Molecular Biology of the Cell
Vol. 10, 1653–1663, May 1999

Coiled Bodies Preferentially Associate with U4, U11, and U12 Small Nuclear RNA Genes in Interphase HeLa Cells but Not with U6 and U7 Genes

Erica Y. Jacobs,* Mark R. Frey,* Wei Wu,* Thomas C. Ingledue,† Thomas C. Gebuhr,* Liming Gao,* William F. Marzluff,† and A. Gregory Matera*‡

*Department of Genetics, Center for Human Genetics and Program in Cell Biology, Case Western Reserve University and University Hospitals of Cleveland, Cleveland, Ohio 44106-4955; and †Program in Molecular Biology and Biotechnology, Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7100

Submitted January 7, 1999; Accepted March 3, 1999
Monitoring Editor: Joseph Gall

Coiled bodies (CBs) are nuclear organelles involved in the metabolism of small nuclear RNAs (snRNAs) and histone messages. Their structural morphology and molecular composition have been conserved from plants to animals. CBs preferentially and specifically associate with genes that encode U1, U2, and U3 snRNAs as well as the cell cycle–regulated histone loci. A common link among these previously identified CB-associated genes is that they are either clustered or tandemly repeated in the human genome. In an effort to identify additional loci that associate with CBs, we have isolated and mapped the chromosomal locations of genomic clones corresponding to bona fide U4, U6, U7, U11, and U12 snRNA loci. Unlike the clustered U1 and U2 genes, each of these loci encode a single gene, with the exception of the U4 clone, which contains two genes. We next examined the association of these snRNA genes with CBs and found that they colocalized less frequently than their multicopy counterparts. To differentiate a lower level of preferential association from random colocalization, we developed a theoretical model of random colocalization, which yielded expected values for χ^2 tests against the experimental data. Certain single-copy snRNA genes (U4, U11, and U12) but not controls were found to significantly ($p < 0.000001$) associate with CBs. Recent evidence indicates that the interactions between CBs and genes are mediated by nascent transcripts. Taken together, these new results suggest that CB association may be substantially augmented by the increased transcriptional capacity of clustered genes. Possible functional roles for the observed interactions of CBs with snRNA genes are discussed.

INTRODUCTION

Coiled bodies (CBs) are nuclear organelles whose precise functions are unknown. They are highly enriched in small nuclear ribonucleoproteins (snRNPs) and several basal transcription factors and have thus received a great deal of recent interest (for review, see Lamond and Earnshaw, 1998; Matera, 1998). Evidence that CBs and their twin structures, called gems, are involved in aspects of snRNP biogenesis and maintenance has intensified this interest, especially given that defects in this pathway result in the genetic disease spinal muscular atrophy (Fischer *et al.*, 1997; Liu *et al.*, 1997; Pellizzoni *et al.*, 1998; for review, see Matera, 1999).

Although CBs are now clearly implicated in snRNP biogenesis, several lines of evidence suggest that they are multifunctional organelles. First, CBs contain many apparently disparate cellular components in addition to the aforementioned snRNAs and associated Sm protein epitopes, including certain nucleolar proteins and small nucleolar RNPs, histone pre-mRNA processing factors, and members of the transcriptional

‡ Corresponding author. E-mail address: gxm26@po.cwru.edu.

apparatus (Matera, 1998). Second, CBs preferentially associate with specific DNA loci within mammalian nuclei and amphibian oocytes, including genes that encode U1, U2, and U3 snRNAs as well as the cell cycle–regulated histone loci (Gall *et al.*, 1981; Callan *et al.*, 1991; Frey and Matera, 1995; Smith *et al.*, 1995; Gao *et al.*, 1997; Schul *et al.*, 1998). If CBs and gems are simply involved in snRNP assembly (and possibly regeneration; Pellizzoni *et al.*, 1998), why do they contain basal transcription factors and associate with snRNA and histone genes?

We suggest that CBs may also participate in regulating snRNA levels by an autogenous feedback loop (Frey and Matera, 1995; Frey *et al.*, 1999) in which a fraction of mature (or partially mature) U snRNPs return, salmon-like, to CBs adjacent to the sites of snRNA synthesis and regulate transcription. Furthermore, we have shown that association of human U2 genes (the *RNU2* locus) with CBs is mediated by nascent U2 RNA (Frey *et al.*, 1999). Based on these data, we hypothesized that when snRNP concentrations are high, the resulting association of the CB with nascent U snRNA would then result in transcriptional attenuation or stalling. Conversely, when snRNP concentrations within the cell are low, CBs would no longer associate with the genes, and transcription might even be up-regulated. Although many mechanisms can be envisioned, it is plausible that excess snRNP-specific proteins (in the absence of the mature snRNAs) might bind directly to the nascent snRNA transcripts and block their interaction with CBs (Frey *et al.*, 1999).

Previously characterized CB-associated loci, including the U1, U2, and U3 genes (*RNU1*, *RNU2*, and *RNU3* loci, respectively) as well as the two major histone gene loci (*HIST1* and *HIST2*), are either tightly clustered or tandemly repeated in the human genome (Frey and Matera, 1995; Gao *et al.*, 1997). In addition to being clustered, these genes share a number of common features. For example, neither snRNA nor histone genes contain introns, and their transcripts are not polyadenylated. The snRNA genes are transcribed from a special class of non-TATA promoters. Furthermore, both histone mRNAs and snRNAs have conserved, 3′-terminal stem–loop structures. But for the presence of an ORF, histone mRNAs are structurally quite similar to snRNAs. In light of these results, we wondered whether genomic organization plays any role in the association frequency of genes with CBs. We therefore wanted to ask whether single-copy snRNA genes could also preferentially associate with CBs.

To investigate this issue, we isolated and mapped the chromosomal locations of bacterial artificial chromosome (BAC) clones containing U4, U6, U7, U11, and U12 snRNA loci by radiation hybrid (RH) and fluorescence in situ hybridization (FISH) analysis and found that they are dispersed among multiple human chromosomes. Each of these loci is single copy, except for U4, which has two copies. We then examined whether these genes associated with CBs and found association values lower than those of multicopy U1 and U2 genes, but higher than other single-copy control loci such as collagen type I (*COL1A1*), c-*myc*, makorin, and nerve growth factor receptor (*NGFR*). To put these observations on a more quantitative basis, we developed a theoretical model that generated random colocalization values in which the frequency of association was solely dependent on the respective areas of the gene signals and those occupied by CBs. This model was used to generate expected values for χ^2 tests against the observed values for each locus. Three of the snRNA loci (*RNU4*, *RNU11*, and *RNU12*) were determined to preferentially associate with CBs, whereas no significant association was found for the *RNU6*, *RNU7*, or the control genes.

MATERIALS AND METHODS

Isolation of BAC Clones and RH Mapping

Primers were constructed and used to screen a human BAC library (Research Genetics, Huntsville, AL) and RH panels (Genebridge 4, Research Genetics; Hudson *et al.*, 1995). Sequence-tagged site mapping data for each of these loci were created and deposited in GenBank using the following primer pairs: *RNU4* (5′-AATACGGCTGGTGGAGTGGGAACA-3′ and 5′-TCGCGCCTCGGATAAACCTCATT-3′), accession number AF114984; *RNU6* (5′-CGAATTTGCGTGTCATC-3′ and 5′-AGGTCGGGCAGGAAGAG-3′), AF114985; *RNU7* (5′-CGCGAACTCTAGAAATGAATGAC-3′ and 5′-TGCTGCGTATGTCTTTGGAG-3′), AF114986; *RNU11* (5′-TGGCTAGGGGTGGCACAAGATACA-3′ and 5′-GTCGATTCCGCACGCAGAGCA-3′), AF114982; and *RNU12* (5′-CCGCTAGGGAGCGACGAACTAC-3′ and 5′-TTCGTGGGTCACAACGTCAATAC-3′), AF114983. The following clone addresses were identified: *RNU4,* 189A14; *RNU6,* 330B12; *RNU11,* 190M11; and *RNU12,* 133E16. Each of the loci was then mapped using the Genebridge 4 RH panel, except the RNU7 primer pair, which was mapped on the higher resolution Stanford G3 RH panel. The complete sequence of a P1 artificial chromosome (PAC) clone containing the U7 gene had already been deposited in the database as part of a large-scale genome-sequencing project (Anisari-Lari *et al.*, 1997).

PCR reactions consisted of 25 ng of each hybrid, 20 mM Tris, pH 8.3, 2.5 mM $MgCl_2$, 50 mM KCl, 0.2 mM dNTPs, 0.8 μM primers, and 0.5 U of *Taq* polymerase in a 10 μl-volume. After a 1-min denaturation at 94°C, 30 cycles of 94°C (20 s), a 30-s annealing (initial annealing at 62°C, final annealing at 47°C), and 72°C extension (45 s, final extension for 5 min), most reactions were performed using a "touchdown" protocol on a PTC-100 thermocycler (MJ Research, Watertown, MA). A 2% agarose gel was used to detect the presence or absence of the respective PCR products. Data vectors for each of the primer pairs are available through their accession numbers.

Histone Processing Assay

Oocytes were injected in the cytoplasm with 35 nl of H_2O or a control oligonucleotide directed against *Xenopus* heterogeneous nuclear RNP (hnRNP) A1 mRNA (5-TCGCCCATCCACTTT-3′) or the anti-frog U7 snRNA (5′-AAGAGCTGTAACACTT-3′) oligonucleotide (2 ng/nl) and incubated for 4 h at 18°C. Oocytes were subsequently injected in the nucleus with a plasmid bearing the mouse H2a614 reporter gene (225 pg of DNA for each plasmid per nucle-

us). All oocytes were coinjected either with the plasmid pTT005 (containing the putative human U7 snRNA gene) or a control pGEM7zF plasmid and incubated at 18°C. Eighteen hours later, RNA was extracted as previously described (Williams *et al.*, 1994) and assayed by S1 nuclease analysis for processing of the mouse H2a614 mRNA (Wang *et al.*, 1999). The percentage of processed histone mRNA was quantified on a PhosphorImager (Storm 840; Molecular Dynamics, Sunnyvale, CA).

Metaphase FISH and Image Acquisition

Human metaphase spreads were prepared from normal peripheral blood according to established procedures. BAC and PAC clones containing elements of human U4, U6, U7, U11, and U12 genes were labeled with biotin-16-dUTP (Boehringer Mannheim, Indianapolis, IN) by nick translation. Approximately 100 ng of each labeled clone were then ethanol precipitated with 2 μg of human Cot1 DNA (Life Technologies, Gaithersburg, MD) and 8 μg of salmon sperm DNA per metaphase slide. The DNA was then dissolved in 50% formamide and 2× SSC; after an overnight incubation at 37°C, slides were washed (three times for 5 min each) in 50% formamide and 2× SSC at 42°C and then in 1× SSC at 60°C. Fluorescein-conjugated avidin (Vector Laboratories, Burlingame, CA) was used to detect the hybridization signals. A DAPI (Boehringer Mannheim) counterstain generated the G/Q banding pattern.

Images were obtained using a Zeiss (Thornwood, NY) Axioplan epifluorescence microscope equipped with a cooled charge-coupled device (CCD) camera (Photometrics, Tucson, AZ). The 16-bit source images were stored as normalized 8-bit gray scale data files by the software program CCD Image Capture (Yale University, New Haven, CT). Highly plane-parallel bandpass filters maintained proper image registration (Ballard and Ward, 1993). Gene Join (Yale University) and an Apple Macintosh computer were used to merge and pseudocolor the images. Finished color prints were produced using Adobe Photoshop 2.5.1 (Adobe Systems, Mountain View, CA) and a dye sublimation printer (Codonics, Middleburg Heights, OH).

Interphase FISH Analysis and Cell Scoring

HeLa cells were grown in a monolayer on chambered slides (Nunc, Rochester, NY), prepermeabilized in Triton X-100, and fixed with paraformaldehyde as described (Frey and Matera, 1995). The cells were then incubated with anti-p80 coilin antibody (to mark CBs), fixed again, and hybridized with a biotinylated (Boehringer Mannheim) gene probe as described (Frey and Matera, 1995).

Scoring of slides for colocalization of genes and CBs was performed using a dual-bandpass filter set (Chroma Technologies, Brattleboro, VT). Signals that were either immediately adjacent or coincident to CBs were scored as positive for association. The number of cells in each set that had none, one, two, or three or more CB–signal associations was recorded. The number of CBs and probe signals per cell was also noted.

Nuclear and nucleolar areas were determined by measuring x and y orthogonal diameters of the DAPI images in the case of nuclei or the DAPI-negative areas in the case of nucleoli and calculating areas according to the formula for an ellipse: $A_{ellipse} = \Pi(D_x/2) \times (D_y/2)$. The resultant values were averaged for >100 nuclei or nucleoli. CB areas and signal diameters were determined similarly using anti-coilin or FISH locus signals.

Statistical Analysis

The number of CBs and probe signals per cell was averaged for each set of experiments. A theoretical model of random colocalization between the CBs and probe signals using these average values was used. Treating the HeLa nucleus as a two-dimensional area within which one or more CB-sized areas and gene signal-sized areas can randomly colocalize yields a colocalization probability proportional to their areas. Our model uses empirical measurements of the areas of HeLa cell nuclei, nucleoli, and CBs to calculate the probability of colocalization within a sample of nuclei.

The probability of a single CB colocalizing with a single FISH signal area is $P_{(\text{colocalization given } c = 1,\ s = 1)} = A_{cb}/A_N$, where c is the number of CBs, and s is the number of locus signals. A_N is equivalent to the nuclear area minus the nucleolar area, and A_{cb} is the effective area available for colocalization between a CB and a signal according to the formula $A_{cb} = \Pi(r_{cb} + d_s)^2$, where r_{cb} is the radius of a CB and d_s is the diameter of a locus signal. This allows the effective area of a CB to include the zone around the CB within which a signal may occur and still colocalize with that CB, permitting the probe signals to be treated as points for the purpose of calculating colocalization probabilities. For more than one CB, the areas can be summed, yielding $P_{(\text{colocalization given } s = 1)} = cA_{cb}/A_N$, as long as $c \leq (A_N/A_{cb})$.

For more than one signal, the random colocalization probablity is more complex. The absence of multiple signals associated with the same CB in our present data set suggests the constraint that only one signal may associate with one CB. The probability of the first gene signal–CB colocalization in a nucleus is as derived above: $P_{(\text{first colocalization})} = cA_{cb}/A_N$. The probability of the second colocalization will be lower, because one fewer CB area is available for colocalization: $P_{(\text{second colocalization})} = (c - 1)A_{cb}/A_N$. For the nth colocalization, $P_{(\text{nth colocalization})} = (c - n)A_{cb}/A_N$, for $0 > c > n$; i.e. the number of gene signals is less that the number of CBs. This is consistent with the observed data.

Instead of calculating the probability of each individual colocalization in this way, the probability of a nucleus having one or more colocalizations between CBs and locus signals was derived as follows. The probability of zero colocalizations given one signal is $P_{(\text{no colocalizations, } s = 1)} = 1 - P_{(\text{first colocalization})}$, and for s signals, $P_{(\text{no colocalizations})} = (1 - P_{(\text{first colocalization})})^s$. For a given nucleus with s signals and c CBs, the probability of that nucleus having one or more colocalizations is therefore $P_{(\text{one or more})} = 1 - P_{(\text{no colocalizations})}$, or $P_{(\text{one or more})} = 1 - (1 - P_{(\text{first colocalization})})^s$. This not only reduces the colocalization probability calculation to a simple binomial distribution but also provides a more conservative test for significance and reduces the effect of the sparse multiple colocalizations in the data set.

A test was constructed to obtain significance levels for the experimental outcomes. The random colocalization probabilities generated according to the formulae above for $P_{(\text{one or more})}$ and $P_{(\text{no colocalizations})}$ were used to generate expected numbers of nuclei with one or more colocalization and expected numbers of nuclei with no colocalizations. For the observed data sets, data were grouped into the number of nuclei per set with one or more colocalizations and the number with no colocalizations. These values were compared with the expected values using a χ^2 test with 1 df. A significance level of 0.01 was chosen to provide a stringent test of significance, indicating that there was a <1% chance that those experimental outcomes found to be significant were actually due to random colocalization. As a further test of the model, we ran a computer simulation of random gene signals and CBs localizing within a nucleus and found that the statistical tests agreed well with the theoretical calculations (see RESULTS).

RESULTS

Isolation and Chromosomal Localization of Dispersed Single-Copy snRNA Loci

To ask whether single-copy snRNA genes also associate with CBs, we first needed to obtain suitable, large-insert FISH probes. To date, only small plasmid clones containing bona fide U4 (Bark *et al.*, 1986), U6 (Kunkel and Pederson, 1988), U11 (Suter-Crazzolara and Keller, 1991), and U12 (Tarn *et al.*, 1995) genes had

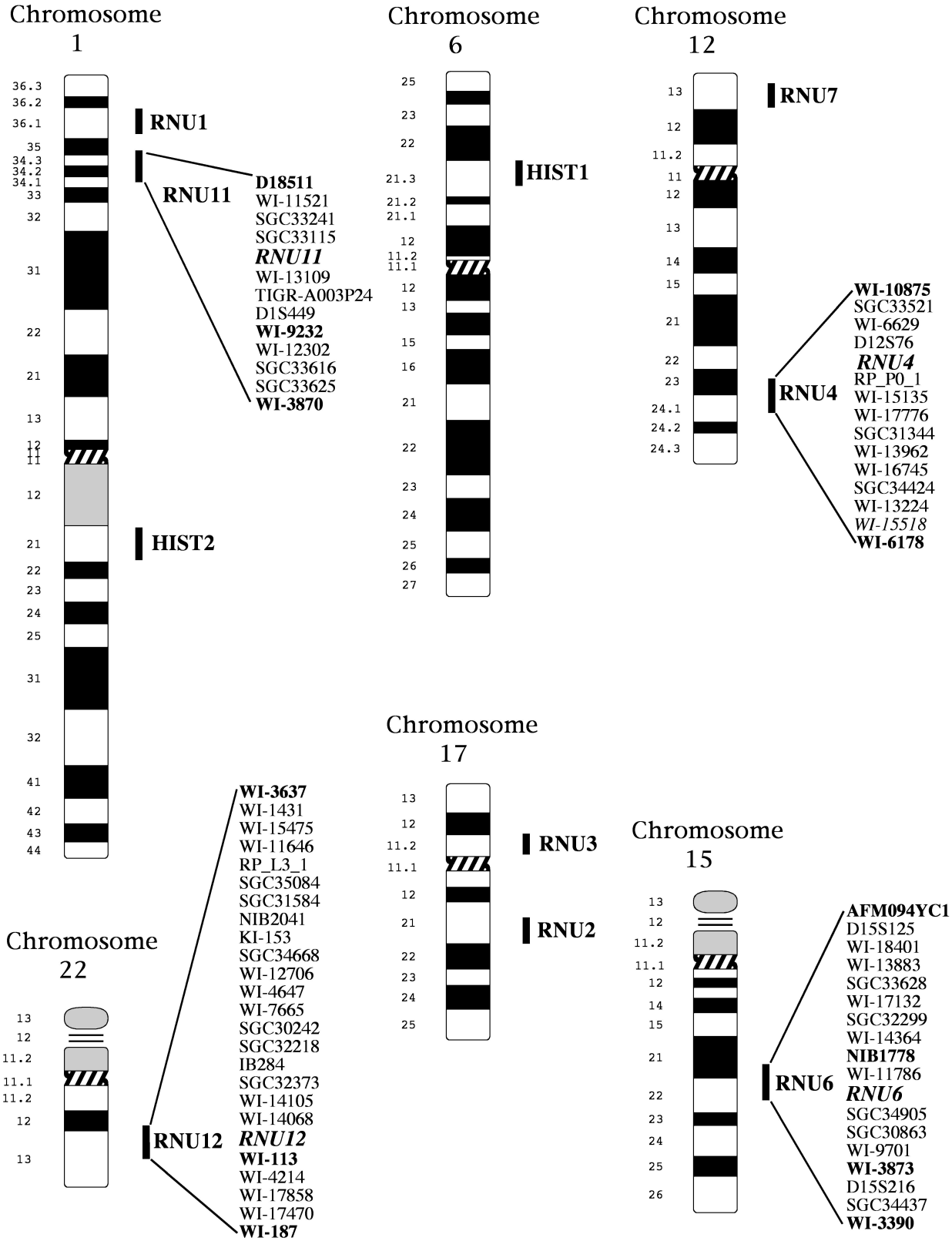
Chromosome 1
36.3
36.2
36.1
35
34.3
34.2
34.1
33
32
31
22
21
13
12
11
12
21
22
23
24
25
31
32
41
42
43
44
RNU1
RNU11
HIST2
D18511
WI-11521
SGC33241
SGC33115
RNU11
WI-13109
TIGR-A003P24
D1S449
WI-9232
WI-12302
SGC33616
SGC33625
WI-3870
Chromosome 6
25
23
22
21.3
21.2
21.1
12
11.2
11.1
12
13
15
16
21
22
23
24
25
26
27
HIST1
Chromosome 12
13
12
11.2
11
12
13
14
15
21
22
23
24.1
24.2
24.3
RNU7
RNU4
WI-10875
SGC33521
WI-6629
D12S76
RNU4
RP_P0_1
WI-15135
WI-17776
SGC31344
WI-13962
WI-16745
SGC34424
WI-13224
WI-15518
WI-6178
Chromosome 17
13
12
11.2
11.1
12
21
22
23
24
25
RNU3
RNU2
Chromosome 22
13
12
11.2
11.1
11.2
12
13
RNU12
WI-3637
WI-1431
WI-15475
WI-11646
RP_L3_1
SGC35084
SGC31584
NIB2041
KI-153
SGC34668
WI-12706
WI-4647
WI-7665
SGC30242
SGC32218
IB284
SGC32373
WI-14105
WI-14068
RNU12
WI-113
WI-4214
WI-17858
WI-17470
WI-187
Chromosome 15
13
12
11.2
11.1
12
14
15
21
22
23
24
25
26
RNU6
AFM094YC1
D15S125
WI-18401
WI-13883
SGC33628
WI-17132
SGC32299
WI-14364
NIB1778
WI-11786
RNU6
SGC34905
SGC30863
WI-9701
WI-3873
D15S216
SGC34437
WI-3390

Figure 1.

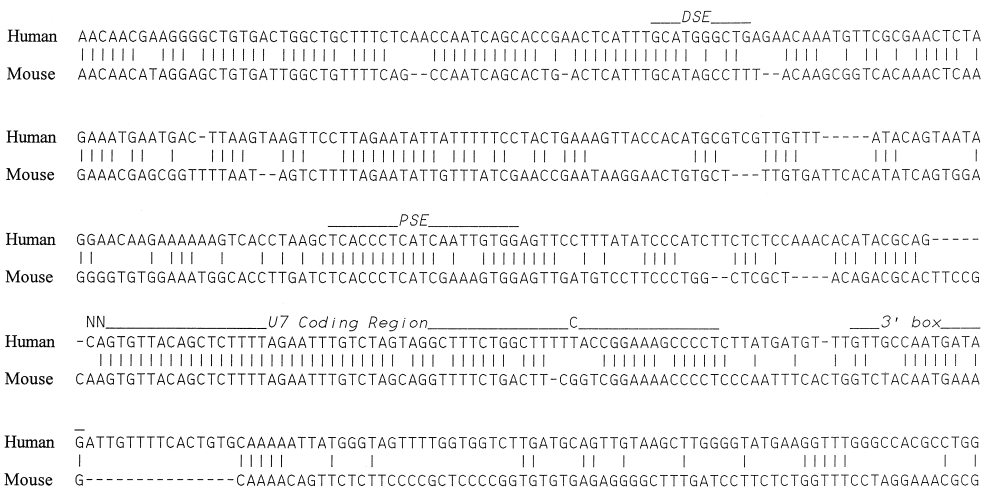

Figure 2. Sequence comparison of human and mouse *RNU7* loci. The U7 genes at human 12p13 (accession number U474924) and mouse chromosome 6F (X54165) along with flanking regions are compared. Note the relatively high degree of upstream similarity, including the functionally conserved control regions called the proximal and distal sequence elements (PSE and DSE, overlined). The human U7 snRNA coding region is also overlined. A single-nucleotide difference between the published RNA sequence (accession number M17910) and the human 12p13 sequence is illustrated. Downstream regions, with the exception of a putative 3′ box terminator sequence (overlined), are not conserved.

been reported in the literature. We therefore used sequence information from these clones to generate PCR primer pairs and screened a human BAC library. These same primer pairs were used to perform RH mapping with the Genebridge 4 panel (Hudson *et al.*, 1995). The results of these RH and cytogenetic mapping experiments (performed in parallel) are shown in Figure 1 and are completely consistent with each other. We found that the two U4 genes mapped to 12q23–24.1, the U6 gene to 15q21–22, U11 to 1p34.2–34.3, and U12 to 22q13. The cytogenetic locations of previously identified CB associated loci are shown for comparison.

Although clones corresponding to functional human U4, U6, U11, and U12 genes have been described in the literature (see above), no such genomic sequence has been reported for U7 genes. Nagafuchi *et al.* (1994) noted that a nearly identical copy of a U7 snRNA sequence was located ~1.3 kb downstream of the dentatorubral and pallidolusian atrophy (DRPLA) gene on chromosome 12p13. This sequence differed by only a single base pair from the reported U7 RNA sequence (Mowry and Steitz, 1987), and its genomic location is syntenic to the region of mouse chromosome 6 that is known to contain the bona fide mouse U7 gene (Turner *et al.*, 1996). Subsequently, we noted that a large-scale genome-sequencing project had sequenced this region of chromosome 12 (Anisari-Lari *et al.*, 1997). Alignment of the regions surrounding the human (accession number U47924) and mouse (X54165) U7 sequences (Figure 2) revealed that the human sequence was unlikely to be a pseudogene, because it contains significant upstream similarity, including recognizable snRNA transcriptional control elements (proximal and distal sequence elements). Thus using a subcloned 3.4-kb *Bam*HI fragment from this region (a kind gift from M. Yamada, National Children's Medical Research Center, Tokyo, Japan), we decided to test whether the human U7 gene was functional in a histone pre-mRNA processing assay.

The Human U7 Gene on 12p13 Encodes a Functional snRNA

Xenopus oocytes were used as a model system for in vivo processing of histone pre-mRNA (Williams *et al.*, 1994; Wang *et al.*, 1999). Unprocessed RNAs are stable in oocytes (Bentley and Groudine, 1988; Middleton and Morgan, 1990), and quantitative results from a processing assay can be generated. Previous results demonstrate that oocytes are competent to process histone pre-mRNA transcribed from the mouse histone H2a-614 gene when injected into the nucleus (Williams *et al.*, 1994). In addition, antisense oligonucleotides directed against *Xenopus* U7 snRNA have been shown to effectively reduce histone mRNA processing in the oocyte (Scharl and Steitz, 1996). Antisense oligonucleotides will bind to the target snRNAs, and the DNA–RNA hybrid is destroyed by endogenous RNase H (Pan and Prives, 1988).

Oocytes were injected with water, a U7 snRNA antisense oligonucleotide, or a control oligonucleotide targeted to hnRNP A1 mRNA. Subsequently, oocytes were injected with the mouse H2a-614 gene and either a control plasmid (pGEM7zf) or plasmid pTT005, containing the putative human U7 snRNA gene. Eighteen

Figure 1 (facing page). Chromosomal locations of snRNA and histone genes used in this study. Map locations of the single-copy snRNA genes (determined by RH and cytogenetic mapping) as well as those of previously characterized snRNA and histone loci are shown. Bars to the right of the ideograms show the range of FISH signals. The nearest markers from the RH mapping are shown to the right of the chromsomes; MIT framework markers are shown in bold.

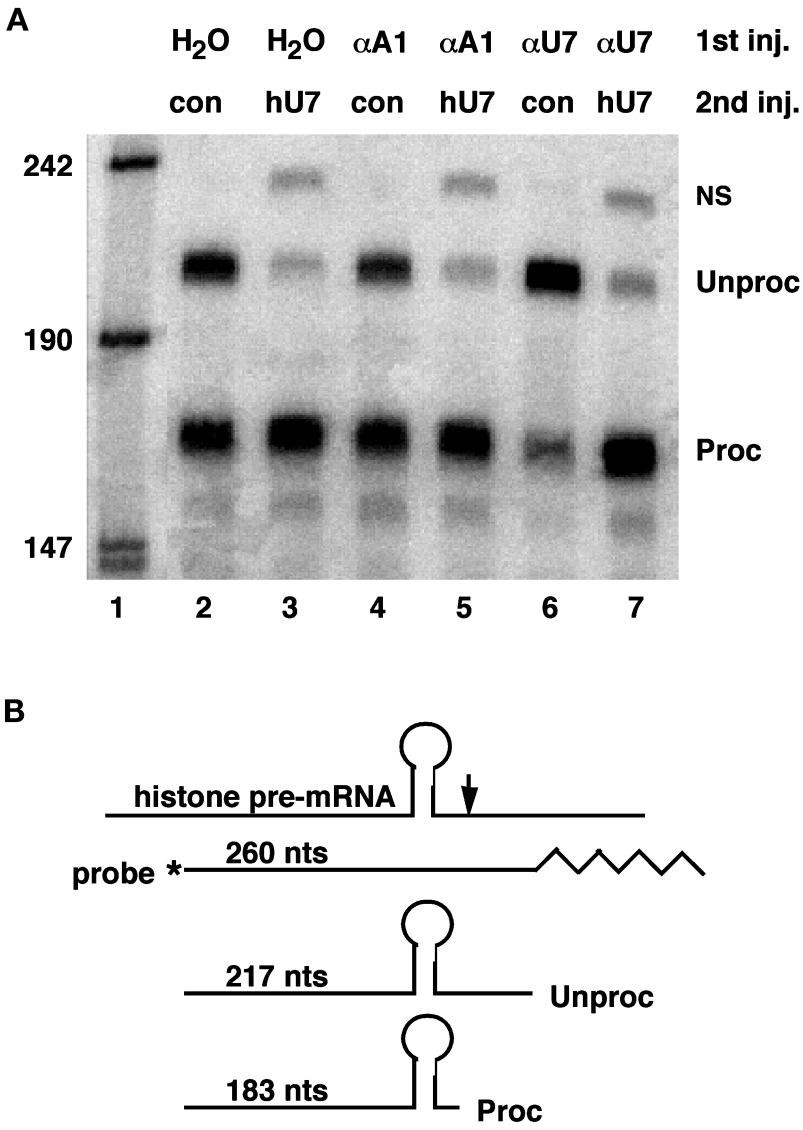

Figure 3. The human U7 sequence at 12p13 encodes a functional snRNA. (A) *Xenopus* oocytes were injected with H_20 (lanes 2 and 3), anti-hnRNP A1 mRNA control oligonucleotide (lanes 4 and 5), or anti-U7 snRNA oligonucleotide (lanes 6 and 7). Four hours later, oocytes were injected with either a control plasmid and the mouse H2a-614 gene (lanes 2, 4, and 6) or plasmid pTT005, containing the putative human U7 gene (hU7) and the mouse H2a-614 gene (lanes 3, 5, and 7). Eighteen hours later, total RNA was prepared from oocytes and analyzed by S1 nuclease mapping using an assay that detects both processed (Proc) and unprocessed (Unproc) histone mRNA. NS, nonspecific transcript derived from the hU7 plasmid. Lane 1, pUC18 DNA digested with *Hpa*II. (B) Schematic of the S1 nuclease assay used in A. The probe contains 34 nucleotides beyond the end of the histone mRNA followed by plasmid sequences that are not present in the transcript from the mouse histone H2a-614 gene.

hours later, oocytes were analyzed for the unprocessed and processed mouse H2a-614 RNA using an S1 nuclease assay. The efficiency of processing of histone pre-mRNA in this batch of oocytes was 60–65% (Figure 3A, lanes 2 and 4). Oocytes treated with anti-U7 oligonucleotides processed histone pre-mRNA much less efficiently (Figure 3A, lane 6). In contrast, oocytes injected with the putative human U7 gene processed histone pre-mRNA with >90% efficiency (Figure 3A, lanes 3, 5, and 7). Even the oocytes treated with U7 antisense oligomers were able to process histone pre-mRNA efficiently when supplemented by injection of the human U7 snRNA gene (Figure 3A, compare lanes 6 and 7). Similar results were obtained with oocytes from several frogs (our unpublished results). We conclude that a U7 snRNA is transcribed from pTT005 and that this snRNA assembles with endogenous proteins in the oocyte to form a functional U7 snRNP.

Single-Copy snRNA Genes Associate with CBs Less Frequently than Their Multicopy Counterparts

Using an interphase FISH and cell-scoring protocol developed previously (Frey and Matera, 1995), we analyzed the distributions of U4, U6, U7, U11, and U12 snRNA genes with respect to CBs in unsynchronized HeLa-ATCC cells (Table 1). If the CB and gene signals were found to overlap, they were scored as a colocalization event. Examples of typical CB colocalizations with single-copy snRNA gene loci are shown in Figure 4. We also scored a number of control loci, including the *NGFR*, *HPRT*, c-*myc*, *COL1A1*, makorin, and 5S rRNA (*RN5S*) genes (Table 1). Previous studies

Table 1. Coiled body association data and statistical analysis

Cytogenetic location	Locus name	Nuclei (n)	Average signals (n)	Average CBs (n)	Nuclei ≥1 associations (n)	Observed (%)	Expected (%)	χ^2	p
17q21.3	*RNU2*	129	2.7	4.8	93	72.1	6.6	893.5	2.5E–196
17p11.2	*RNU3*	202	2.9	5.3	110	54.4	7.9	605.1	1.3E–133
6p21.3	*HIST2*	124	2.2	5.8	54	43.5	6.7	270.1	1.1E–60
1q21	*HIST1*	125	1.4	5.9	38	30.6	4.4	204.4	2.3E–46
1p36.1	*RNU1*	109	2.9	5.0	41	37.6	7.4	144.2	3.1E–33
12q23–24.1	*RNU4*	200	3.0	4.9	57	28.5	7.7	122.9	1.5E–28
1p34.2–34.3	*RNU11*	150	2.9	4.5	26	17.3	6.8	26.0	3.5E–07
22q13	*RNU12*	200	3.0	4.6	33	16.5	7.3	25.2	5.1E–07
17q21.3–22	*NGFR*	135	2.5	4.6	15	11.0	6.1	6.0	0.014
12p13	*RNU7*	140	3.0	5.1	15	10.7	8.0	1.4	0.23
15q21–22	*RNU6*	150	3.0	4.5	14	9.3	7.0	1.3	0.26
17q21.3–22	*COL1A1*	122	3.0	4.4	7	6.0	7.0	0.3	−0.60
Xq26.1	*HPRT*	101	2.1	4.5	3	4.9	3.0	0.8	−0.38
8q24	c-*myc*	84	1.3	4.4	0	0	3.1	2.7	−0.10
7q34	*makorin*	100	2.8	4.6	1	1.0	6.8	5.3	−0.022
1q42.1	*RN5S*	121	3.8	4.5	3	2.5	9.0	6.3	−0.012
NA[a]	Random	200	3.0	4.8	14	7.0	7.5	8.1E–02	0.78

[a] NA, not applicable.

have shown that only transcriptionally active U2 gene constructs associate with CBs (Frey *et al.*, 1999). Therefore, the only stipulation we made regarding choice of the control genes is that they are single copy (with the exception of *RN5S*) and that their transcripts are expressed in HeLa cells (with the exception of NGFR).

To develop a locus association test, empirical measurements of the areas of HeLa cell nuclei, nucleoli, and CBs were used to calculate the probability of colocalization within a sample of nuclei. The probability of a single CB colocalizing with a single FISH signal is proportional to the respective areas of the two structures and the nuclear area, as diagrammed in Figure 5. In our data set, all the FISH signals were excluded from nucleoli. Thus the effective nuclear area is equivalent to the area of the nucleus minus that of the nucleoli. Conversely, the effective area of the CB is actually larger by one gene signal diameter (Figure 5), because the two objects (CBs and gene signals) must touch to be scored positively for association (see MATERIALS AND METHODS for additional details). Therefore, the probability that a random gene would colocalize with at least one of n CBs is proportional to the effective nuclear areas of all the CBs in the cell, divided by the effective nuclear area (i.e., the nucleoplasm).

Using these parameters, we generated expected values of nuclei with one or more random associations and used them for a χ^2 test against experimental data. As expected, the association of CBs with the previously characterized histone and snRNA loci was highly significant (with p values in the 1.0E-30–1.0E-200 range; see Table 1). Interestingly, we found that among our test probes, the *RNU4*, *RNU11*, and *RNU12* loci also displayed significant association frequencies (Table 1). However, neither the *RNU6* and *RNU7* loci nor the control gene loci met the 0.01 stringency threshold. Taken together with the *RN5S* experiments, the *RNU6* data imply that pol III genes do not preferentially associate with CBs (albeit with several caveats; see DISCUSSION).

Technical Considerations

Note that our theoretical colocalization model treats the nucleus and the structures within as two-dimensional areas. This two-dimensional model closely approximates the actual nonconfocal image analysis of the experimental data and is appropriate for relatively flattened adherent cells such as HeLa. Taking into account the depth of field of our objective (~1 μm), the volume sampled is approximately one-third to one-half the volume of a HeLa cell. Although area is not generally a good approximation of volume, by constraining the statistical analysis to two dimensions, we actually make the association test *more* stringent. Random (expected) colocalizations would actually be less frequent if the genes were given more room to roam and would lower the threshold of significance. Thus the two-dimensional approach gives a more conservative picture of CB associations, raising the threshold and detecting only the most significant interactions. We, therefore, cannot completely exclude the possibility that other, less frequent, associations are not of biological interest. It is also noteworthy that if volume measurements are used instead of area, the model can be adapted both to other cell shapes and to

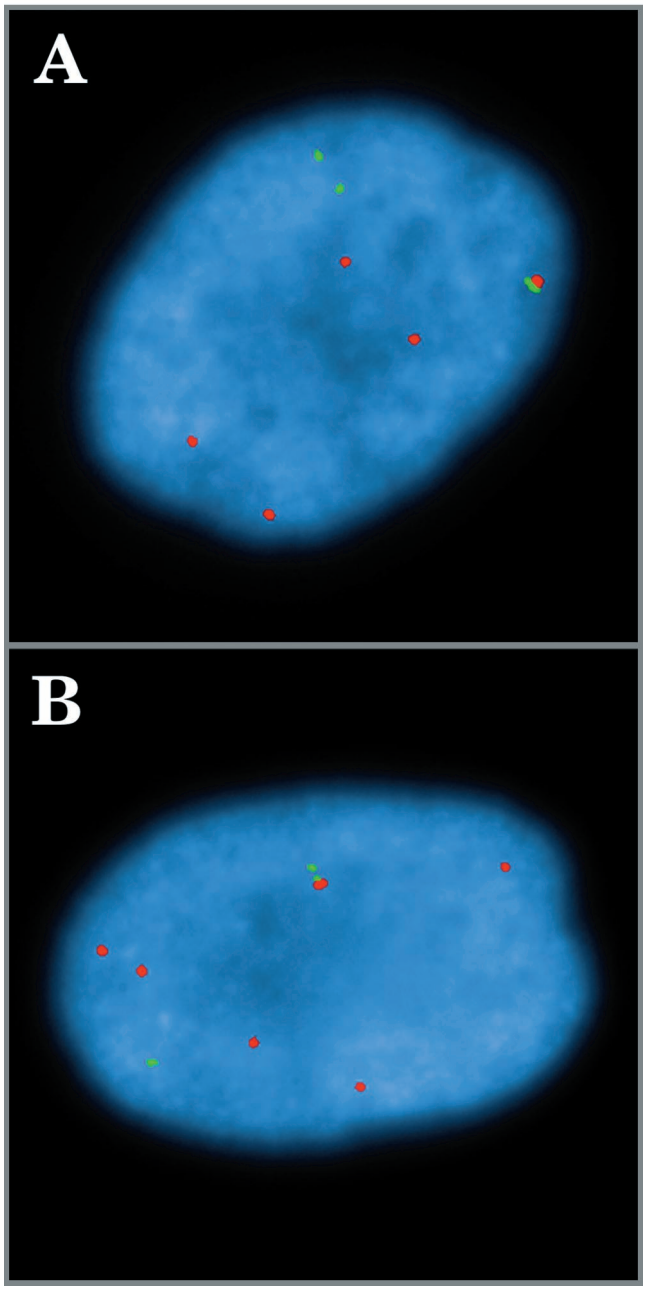

Figure 4. Single-copy snRNA loci also colocalize with CBs in interphase HeLa cells. Nuclei were stained with DAPI (blue); CBs are displayed in red; and snRNA gene signals are in green. In A (*RNU11*) and B (*RNU12*), BAC clones corresponding to the appropriate locus were used to generate the FISH signals; CB immunofluorescence was performed, and cells were scored as described previously (Frey and Matera, 1995). In each panel, a typical colocalization event is shown; the full data set is presented in Table 1.

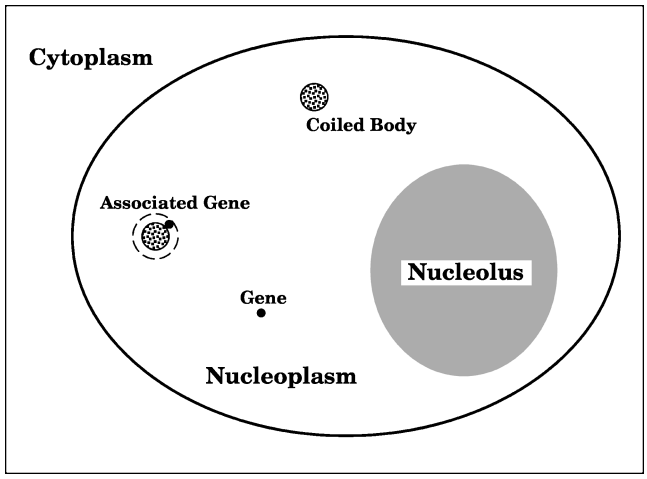

Figure 5. Diagram showing a nucleus with CBs, locus signals, a nucleolus, and a CB–locus association. The dashed line around the CB represents the effective area of a CB that is available for colocalization if locus signals are considered points. The nucleoplasmic area is essentially equal to the area of the nucleus minus the nucleolar area. Average values were nuclear area, 133.7 μm^2; nucleolar area, 10.9 μm^2; nucleoplasmic area (A_N), 123 μm^2; and effective CB area (A_{cb}), 0.66 μm^2.

alternate image analysis methods such as three-dimensional confocal microscopy.

Significant deviations from the random model could theoretically include both cases of more and fewer nuclei with colocalizations than expected. We therefore used a two-tailed χ^2 distribution but found no cases of significantly fewer colocalizations than expected at the 0.01 stringency level. Makorin and *RN5S* genes, however, met the 0.05 stringency level with "negative" p values of 0.022 and 0.012, respectively. These genes may be partitioned into regions of chromatin that do not usually come into contact with CBs. Similar negative results have been obtained with various centromeric (alphoid) DNA probes (Smith *et al.*, 1995; Schul, 1998).

The *NGFR* locus did not meet our stringent criteria for association but was significant at the 0.05 level with a p value of 0.014. As mentioned previously (Frey and Matera, 1995), the elevated CB association frequency of *NGFR* may simply be attributable to its close proximity (~6–7 Mb; Deloukas *et al.*, 1998) to the *RNU2* locus. However, the *COL1A1* locus is located ~200 kb proximal to *NGFR* (i.e., slightly closer to *RNU2*), and this gene did not significantly associate with CBs (Table 1).

DISCUSSION

We wanted to determine whether, like their multicopy U snRNA brethren (Frey and Matera, 1995; Smith *et al.*, 1995; Gao *et al.*, 1997; Schul *et al.*, 1998), single-copy snRNA genes might also preferentially associate with CBs in interphase HeLa cells. To answer this question we first had to isolate BAC clones and map the chro-

mosomal locations of bona fide human U4, U6, U11, and U12 snRNA genes. We found these genes to be dispersed among four different chromosomes (*RNU4*, 12q23–24.1; *RNU6*, 15q21–22; *RNU11*, 1p34.2–34.3; and *RNU12*, 22q13). Furthermore, database analysis identified a putatively functional U7 gene on chromosome 12p13 (Anisari-Lari *et al.*, 1997), whose map location is syntenic to the known functional *RNU7* locus in mouse (Turner *et al.*, 1996). We therefore tested the ability of transcripts from this gene to process histone messages in a *Xenopus* oocyte expression system and found that this gene is completely functional. An analysis of the interphase distributions of these gene loci with respect to CBs revealed that only the CB colocalization frequencies of the U4, U11, and U12 genes were significant when compared with a random model (Table 1).

We conclude that although clustered snRNA genes colocalize more frequently with CBs than do their single-copy counterparts, the interaction does not depend on the genomic organization of the gene. However, the interaction does appear to depend on the type of RNA polymerase. Both U6 snRNA and 5S rRNA genes (transcribed by pol III) do not associate with CBs, whereas those for U1, U2, U3, U4, U11, and U12 (pol II) are each preferentially localized adjacent to CBs. Taken together with our previous findings that the association of snRNA genes with CBs is mediated by the nascent snRNA transcripts (Frey *et al.*, 1999), these data highlight the possibility that CBs may associate with snRNA genes through interactions with the RNA polymerase II complex bearing the nascent transcript and not simply with the nascent snRNA itself. Clearly, additional experiments will be required to answer this question definitively.

Finally, we are left with a dilemma. Does a statistical preference for colocalization of a given gene with CBs reflect a functional interaction or is it merely fortuitous? The diversity of genes that exhibit such associations (Figure 1), the functional similarities between the products of these genes (they encode small RNAs involved in RNA processing or intronless mRNA precursors that are processed by small RNAs), and the phylogenetic conservation of these interactions from amphibians to mammals (Frey *et al.*, 1999; our unpublished observations) all argue that CBs talk to these genes for a biologically important reason.

Genomic Caveats

Although uncovering the precise nature of these CB/gene associations will be the subject of future investigations, there are also a number of unknown factors to be considered when interpreting these statistical data. For example, we do not know whether the U6 gene we have identified on chromosome 15q21–22 is the major site of U6 synthesis, even though constructs of the gene are active in vitro (Kunkel and Pederson, 1988), and it remains the only bona fide U6 gene in the database. Thus the reason that this gene does not associate with CBs may not be due to the type of polymerase involved but because it is not highly transcribed. Similarly, the U7 gene on 12p13 also may not be the major site of U7 synthesis. However, several lines of evidence suggest that, in fact, this sequence is the major U7 gene. First, its genomic location is syntenic with the mouse U7 gene (Turner *et al.*, 1996; Nadeau and Sankoff, 1998). Second, despite the fact that the coding region contains a single-nucleotide difference (Figure 2) from the published U7 RNA sequence (Mowry and Steitz, 1987), a subsequent study found that this nucleotide position (49) was a U in five isolated U7 cDNAs (Yu *et al.*, 1996). The original sequencing of the 3′ end of U7 was done by RNase A and T_1 mapping and could easily be erroneous. Nevertheless, although there may be multiple active U7 genes in the human genome, the locus at 12p13 does not significantly associate with CBs.

There are ~100 copies of the U4 sequence (including pseudogenes) in the human genome (Bark *et al.*, 1986). Although the two U4 genes contained within our BAC clone are the only known active members of the U4 gene family in the database, the locus on 12q23–24.1 does not correspond to the major U4 locus in human genomic DNA (Bark *et al.*, 1986). The major locus remains uncloned, although based on the conservation of flanking sequences detected by genomic blotting, it seems probable that U4 genes are also tightly clustered (Bark *et al.*, 1986). Based on the CB association frequency observed for the *RNU4* locus on 12q (Table 1), it seems likely that these genes are active. It would also be worthwhile to isolate the major U4 cluster and assay both its level of expression and frequency of colocalization with CBs.

CB Biogenesis and Functions

CBs and interchromatin granule clusters both contain mature U snRNPs, yet there are many interchromatin granule clusters but only a few CBs. What limits the number of CBs, and where do they form? One attractive hypothesis is that the CB-associated genes nucleate CBs, just as rDNA (the "nucleolus organizer region") nucleates nucleoli. If CB-associated genes nucleate CBs, why are some CBs attached to specific loci, whereas others appear to be free in the nucleoplasm? The facts that CBs do not accumulate newly synthesized RNA (Fakan and Bernhard, 1971; Callan and Gall, 1991; Jordan *et al.*, 1997; Schul *et al.*, 1998) and that they can form in the complete absence of cognate genomic DNA (Bell *et al.*, 1992; Bauer *et al.*, 1994) argue against this model. Our data demonstrating association of CBs with multiple different snRNA loci are more compatible with spontaneous assembly of CBs, perhaps in nucleoli (Lyon *et al.*, 1997; Issac *et*

al., 1998; Sleeman *et al*., 1998) followed by RNA-mediated association with the genes.

The dynamic nature of CBs is well described (Andrade *et al*., 1993; Carmo-Fonseca *et al*., 1993; Ferreira *et al*., 1994). One interesting idea that has yet to be explored is that association with the genes may actually *stabilize* CBs. We suggest that active snRNA genes may be tethered to CBs through multiple weak interactions with their nascent transcripts. Cognate "receptor" molecules would necessarily be localized within or adjacent to CBs. Presumably, the higher transcriptional capacities of multicopy snRNA and histone gene clusters would thereby increase the chances of association with CBs relative to single-copy genes. Furthermore, CBs are capable of simultaneously associating with multiple chromosomal loci (Frey *et al*., 1999). Regardless of whether CBs are stabilized by interaction with nearby genes, the fact that CBs accumulate high concentrations of mature snRNPs supports the existence of an autogenous feedback loop (Frey and Matera, 1995) through which CBs could participate in coordinate regulation of the expression of several different gene families.

ACKNOWLEDGMENTS

We are indebted to M. Yamada for the U7 plasmid, M. Anisari-Lari for the U7 PAC clone, and E. Chan for anti-coilin antibodies. We thank L. Carrel and T. Gray for the collagen and makorin probes, respectively. We also thank K. Jacobs for helpful suggestions regarding the statistical analysis. This work was supported by National Institutes of Health grants GM-53034 (to A.G.M.) and GM-29832 (to W.F.M.). M.R.F. was supported in part by National Institutes of Health predoctoral traineeship T32-GM-08056, and T.C.I. was supported by a postdoctoral fellowship from the Mellon Foundation.

REFERENCES

Andrade, L.E.C., Tan, E.M., and Chan, E.K.L. (1993). Immunocytochemical analysis of the coiled body in the cell cycle and during cell proliferation. Proc. Natl. Acad. Sci. USA *90*, 1947–1951.

Anisari-Lari, M.A., Shen, Y., Muzny, D.M., Lee, W., and Gibbs, R.A. (1997). Large-scale sequencing in human chromosome 12p13: experimental and computational gene structure determination. Genome Res. *7*, 268–280.

Ballard, S.G., and Ward, D.C. (1993). Fluorescence in situ hybridization using digital imaging microscopy. J. Histochem. Cytochem. *41*, 1755–1759.

Bark, C., Weller, P., Zabielski, J., and Pettersson, U. (1986). Genes for human U4 small nuclear RNA. Gene *50*, 333–344.

Bauer, D.W., Murphy, C., Wu, Z.W., Wu, C.H.H., and Gall, J.G. (1994). In vitro assembly of coiled bodies in *Xenopus* egg extract. Mol. Biol. Cell *5*, 633–644.

Bell, P., Dabauvalle, M.-C., and Scheer, U. (1992). In vitro assembly of prenucleolar bodies in *Xenopus* egg extract. J. Cell Biol. *118*, 1297–1304.

Bentley, D.L., and Groudine, M. (1988). Sequence requirements for premature termination of transcription in the human c-myc gene. Cell *53*, 245–256.

Callan, H.G., and Gall, J.G. (1991). Association of RNA with the B and C snurposomes of *Xenopus* oocyte nuclei. Chromosoma *101*, 69–82.

Callan, H.G., Gall, J.G., and Murphy, C. (1991). Histone genes are located at the sphere loci of *Xenopus* lampbrush chromosomes. Chromosoma *101*, 245–251.

Carmo-Fonseca, M., Ferreira, J., and Lamond, A.I. (1993). Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis—evidence that the coiled body is a kinetic nuclear structure. J. Cell Biol. *120*, 841–852.

Deloukas, P., *et al.* (1998). A physical map of 30,000 human genes. Science *282*, 744–746.

Fakan, S., and Bernhard, W. (1971). Localization of rapidly and slowly labeled nuclear RNA as visualized by high resolution autoradiography. Exp. Cell Res. *67*, 129–141.

Ferreira, J.A., Carmo-Fonseca, M., and Lamond, A.I. (1994). Differential interaction of splicing snRNPs with coiled bodies and interchromatin granules during mitosis and assembly of daughter cell nuclei. J. Cell Biol. *126*, 11–23.

Fischer, U., Liu, Q., and Dreyfuss, G. (1997). The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. Cell *90*, 1023–1029.

Frey, M.R., Bailey, A.D., Weiner, A.M., and Matera, A.G. (1999). Association of snRNA genes with coiled bodies is mediated by nascent snRNA transcripts. Curr. Biol. *9*, 126–135.

Frey, M.R., and Matera, A.G. (1995). Coiled bodies contain U7 small nuclear RNA and associate with specific DNA sequences in interphase cells. Proc. Natl. Acad. Sci. USA *92*, 5915–5919.

Gall, J.G., Stephenson, E.C., Erba, H.P., Diaz, M.O., and Barsacchi-Pilone, G. (1981). Histone genes are located at the sphere loci of newt lampbrush chromosomes. Chromosoma *84*, 159–171.

Gao, L., Frey, M.R., and Matera, A.G. (1997). Human genes encoding U3 snRNA associate with coiled bodies in interphase cells and are clustered on chromosome 17p11.2 in a complex inverted repeat structure. Nucleic Acids Res. *25*, 4740–4747.

Hudson, T., Stein, L., Gerety, S., Ma, J., Castle, A., Silva, J., Slonim, D., Baptista, R., Kruglyak, L., and Xu, S. (1995). An STS-based map of the human genome. Science *270*, 1945–1954.

Issac, C., Yang, Y., and Meier, U.T. (1998). Nopp140 functions as a molecular link between the nucleolus and the coiled bodies. J. Cell Biol. *142*, 407–417.

Jordan, P., Cunha, C., and Carmo-Fonseca, M. (1997). The cdk7-cyclin H-MAT1 complex associated with TFIIH is localized in coiled bodies. Mol. Biol. Cell *8*, 1207–1217.

Kunkel, G.R., and Pederson, T. (1988). Upstream elements required for efficient transcription of a human U6 RNA gene resemble those of U1 and U2 genes even though a different polymerase is used. Genes & Dev. *2*, 196–204.

Lamond, A.I., and Earnshaw, W.C. (1998). Structure and function in the nucleus. Science *280*, 547–553.

Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997). The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. Cell *90*, 1013–1021.

Lyon, C.E., Bohmann, K., Sleeman, J., and Lamond, A.I. (1997). Inhibition of protein dephosphorylation results in the accumulation of splicing snRNPs and coiled bodies within the nucleolus. Exp. Cell Res. *230*, 84–93.

Matera, A.G. (1998). Of coiled bodies, gems, and salmon. J. Cell. Biochem. *70*, 181–192.

Matera, A.G. (1999). RNA splicing: more clues from spinal muscular atrophy. Curr. Biol. *9*, R140–R142.

Middleton, K.M., and Morgan, G.T. (1990). Premature termination of transcription can be induced on an injected α-tubulin gene in *Xenopus* oocytes. Mol. Cell. Biol. *10*, 727–735.

Mowry, K.L., and Steitz, J.A. (1987). Identification of the human U7 snRNP as one of several factors involved in the 3′ end maturation of histone premessenger RNAs. Science *238*, 1682–1687.

Nadeau, J.H., and Sankoff, D. (1998). Counting on comparative maps. Trends Genet. *14*, 495–501.

Nagafuchi, S., Yanagisawa, H., Ohsaki, E., Shirayama, T., Tadokoro, K., Inoue, T., and Yamada, M. (1994). Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidolusian atrophy (DRPLA). Nat. Genet. *8*, 177–182.

Pan, Z.-Q., and Prives, C. (1988). Assembly of functional U1 and U2 human-amphibian hybrid snRNPs in *Xenopus laevis* oocytes. Science *241*, 1328–1331.

Pellizzoni, L., Kataoka, N., Charroux, B., and Dreyfuss, G. (1998). A novel function for SMN, the spinal muscular atrophy disease gene product, in premRNA splicing. Cell *95*, 615–624.

Scharl, E.C., and Steitz, J.A. (1996). Length suppression in histone mRNA 3′-end maturation: processing defects of insertion mutant premessenger RNAs can be compensated by insertions into the U7 small nuclear RNA. Proc. Natl. Acad. Sci. USA *93*, 14659–14664.

Schul, W. (1998). The Spatial and Functional Organization of Transcription and RNA Processing in Relation to Nuclear Domains. Ph.D. Thesis. Amsterdam: University of Amsterdam.

Schul, W., van Driel, R., and de Jong, L. (1998). Coiled bodies and U2 snRNA genes adjacent to coiled bodies are enriched in factors required for snRNA transcription. Mol. Biol. Cell *9*, 1025–1036.

Sleeman, J., Lyon, C.E., Platani, M., Kreivi, J.-P., and Lamond, A.I. (1998). Dynamic interactions between splicing snRNPs, coiled bodies and nucleoli revealed using snRNP protein fusions to the green fluorescent protein. Exp. Cell Res. *243*, 290–304.

Smith, K.P., Carter, K.C., Johnson, C.V., and Lawrence, J.B. (1995). U2 and U1 snRNA gene loci associate with coiled bodies. J. Cell. Biochem. *59*, 473–485.

Suter-Crazzolara, C., and Keller, W. (1991). Organization and transient expression of the gene for human U11 snRNA. Gene Expr. *1*, 91–102.

Tarn, W.Y., Yario, T.A., and Steitz, J.A. (1995). U12 snRNA in vertebrates: evolutionary conservation of 5′ sequences implicated in splicing of premRNAs containing a minor class of introns. RNA *1*, 644–656.

Turner, P.C., Whalen, A., Schümperli, D., and Matera, A.G. (1996). The *bona fide* mouse U7 snRNA gene maps to a different chromosome than two U7 pseudogenes. Genomics *31*, 250–252.

Wang, Z.-F., Ingledue, T.C., Dominski, Z., Sanchez, R., and Marzluff, W.F. (1999). Two *Xenopus* proteins that bind the 3′ end of histone mRNA: implications for translational control of histone synthesis during oogenesis. Mol. Cell. Biol. *19*, 835–845.

Williams, A.S., Ingledue, T.C., Kay, B.K., and Marzluff, W.F. (1994). Changes in the stem-loop at the 3′ terminus of histone mRNA affects its nucleocytoplasmic transport and cytoplasmic regulation. Nucleic Acids Res. *22*, 4660–4666.

Yu, Y.T., Tarn, W.Y., Yario, T.A., and Steitz, J.A. (1996). More Sm snRNAs from vertebrate cells. Exp. Cell Res. *229*, 276–281.