MOLECULAR AND CELLULAR BIOLOGY, May 1983, p. 854–862
0270-7306/83/050854-09$02.00/0

Vol. 3, No. 5

Identification of Two Human β-Tubulin Isotypes

J. L. HALL,[1] L. DUDLEY,[1] P. R. DOBNER,[2] S. A. LEWIS,[1] AND N. J. COWAN[1]*

Department of Biochemistry, New York University Medical Center, New York, New York 10016[1] and Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01605[2]

Received 21 December 1982/Accepted 9 February 1983

The sequence of a human β-tubulin cDNA clone (Dβ-1) is described; our data revealed 95.6% homology compared with the sequence of a human β-tubulin processed pseudogene derived by reverse transcription of a processed mRNA (Wilde et al., Nature [London] **297**:83–84, 1982). However, the amino acid sequence encoded by this cDNA showed less homology with pig and chicken β-tubulin sequences than the latter did to each other, with major divergence within the 15 carboxy-terminal amino acids. On the other hand, an independently isolated, functionally expressed genomic human β-tubulin sequence (5β) possessed a very high degree of homology with chicken and pig β-tubulins in this region. Thus, human cells appear to contain two distinct β-tubulin isotypes. Both the intact β-tubulin cDNA clone and a subclone containing only the 3′ untranslated region detected two mRNA species in HeLa cells; these mRNAs were 1.8 and 2.6 kilobases long and were present in about equal amounts. Two independently subcloned probes constructed from the 3′ untranslated region of the 5β genomic sequence also detected a 2.6-kilobase β-tubulin mRNA. However, the 3′-untranslated-region probes from the cDNA clone and the genomic sequence did not cross-hybridize. Thus, at least two human β-tubulin genes, each specifying a distinct isotype, are expressed in HeLa cells, and the 2.6-kilobase mRNA band is a composite of at least two comigrating β-tubulin mRNAs.

α- and β-tubulins are the major soluble proteins of microtubules. These long filamentous structures are found in all eucaryotic cells and are involved in a remarkable diversity of cellular functions (11). In humans, each of the genes encoding α- and β-tubulins constitutes a large multigene family of about 15 to 20 members (8), the majority of which have been isolated in recombinant fragments (9, 41). Sequence analyses of several of these genes have shown that many of them are pseudogenes (that is, sequences which contain one or more genetic lesions that preclude the synthesis of a functional protein product). Among these pseudogenes is a novel class that is characterized by (i) the absence of intervening sequences, (ii) the presence of a short polyadenylic acid [poly(A)] tract 3′ to the poly(A) signal, and (iii) the existence of short flanking direct repeat sequences (22, 42, 43). These pseudogenes most probably arose by reintegration into the host germ line of a cDNA transcript synthesized on a processed mRNA template.

Because a significant number, perhaps the majority, of human α- and β-tubulin genes are pseudogenes, an important question concerns the number of functionally expressed sequences. One approach to this problem is to investigate the complexity of tubulin mRNAs by constructing cDNA clones. In this paper we describe the sequence of a human β-tubulin cDNA clone that encompasses 98.4% of the coding region. Our data revealed very extensive homology with the sequence of a processed human β-tubulin pseudogene and enabled us to estimate the time of insertion of the mRNA-derived molecule into the germ line. Surprisingly, the amino acid sequence derived from the human β-tubulin cDNA sequence revealed an unexpected degree of divergence compared with known β-tubulin sequences of other species. We present evidence that the differences which we observed reflect the existence of two distinct human β-tubulin isotypes.

MATERIALS AND METHODS

Isolation and sequence analysis of clone Dβ-1. cDNA clone Dβ-1 was isolated by screening a cloned cDNA library prepared from human fetal brain mRNA. RNA was prepared from human fetal brain tissue as described previously (20), except that the tissue was first pulverized in liquid N_2 and then homogenized in a solution containing 0.25 M sucrose, 10 mM Tris-hydrochloride, 10 mM KCl, and 1.5 mM $MgCl_2$ (pH

7.4). Nuclei were removed by centrifugation at 800 × *g* for 5 min, and the RNA was prepared as described previously (23). Poly(A)-containing [poly(A)$^+$] RNA was isolated (1) and used to prepare double-stranded cDNA (5, 40). Terminal deoxynucleotidyl transferase was used to add 15 to 20 deoxycytosine residues to the 3′ termini (10), and an equivalent number of deoxyguanine residues was added to the plasmid vector pUC8 (J. Vieira and J. Messing, Gene, in press) cut with *Pst*I. Equimolar amounts of tailed vector and cDNA were annealed as described previously (39), and the recombinant molecules were used to transform *Escherichia coli* K-12 strain JM83 (24). Colonies were screened by filter hybridization (13), using the excised insert from a chicken β-tubulin cDNA probe (8) labeled with ^{32}P by nick translation (27).

Sequence analysis was performed by the dideoxy chain terminator method (29). Homology with the sequence of the 46β processed pseudogene (43) or chicken β-tubulin cDNA clone pT2 (37) was established by using the computer programs of Staden (32).

Construction and labeling of 3′-untranslated-region subclones. (i) From Dβ-1. A 3′-untranslated-region probe was generated from Dβ-1 by subcloning a fragment extending from a *Hae*III site immediately 5′ to the termination codon to the 3′ flanking *Pst*I site used in the construction of the cDNA clone.

(ii) From 5β. A *Bam*HI fragment containing coding sequences that were C terminal to amino acid 375 and approximately 900 base pairs of noncoding downstream DNA was subcloned from the 5β lambda clone (9) into pBR322. Restriction mapping was used to identify the sites (see Fig. 3B), and the DNA was sequenced and analyzed as described above for Dβ-1.

The single-stranded M13 DNAs from the 3′-end subclones of 5β and Dβ-1 were used as templates for the generation of high-specific-activity probes by primed extension with DNA polymerase I in the presence of ^{32}P-labeled deoxynucleoside triphosphates (16). These probes were used for both Southern blot experiments with digests of genomic DNA and RNA blot transfer experiments (see below).

Preparation of mRNA and blot analyses. Poly(A)$^+$ mRNA was prepared from HeLa cells maintained in Spinner culture as described elsewhere (8). Portions (approximately 5 μg) of this material were run on denaturing 1% agarose gels containing formaldehyde (4), and the gel contents were transferred to nitrocellulose (34). For the analysis of genomic sequences, samples (about 10 μg) of human placental DNA were digested with a threefold excess of restriction endonuclease under the conditions recommended by the manufacturer (New England Biolabs, Inc.). Digests were fractionated on 0.8% agarose gels, and the gel contents were transferred to nitrocellulose (31). RNA and DNA blots were prehybridized and hybridized at 68°C for 12 to 20 h in sealed plastic bags containing 3× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 3× Denhardt solution. The blots were washed twice in 2× SSC containing 0.1% sodium dodecyl sulfate at 68°C and twice in 0.2× SSC containing 0.1% sodium dodecyl sulfate at the same temperature.

Selective hybridization and translation. Probes (approximately 20 μg in either pBR322 or the replicative form of M13) to be tested for the ability to select specific mRNAs were linearized by digestion with *Eco*RI. The digests were extracted with phenol, precipitated with ethanol, and dissolved in 5 μl of 0.2 N NaOH. Then 45 μl of 6× SSC was added, and the mixture was applied directly to 2.3-cm nitrocellulose disks (Schleicher & Schuell Co.). Each wet filter was immersed directly in a vast excess of 6× SSC, air dried, and baked in vacuo at 80°C for 2 h. Prehybridization, hybridization, translation, and analysis of translation products were carried out as described previously (8).

RESULTS

Extensive homology between a human β-tubulin cDNA clone (Dβ-1) and a processed human β-tubulin pseudogene. Southern blot analysis of human genomic DNA in which a chicken β-tubulin cDNA probe was used revealed a total of 15 to 20 bands per restriction digest (8). However, not all of these genes can be expressed, since a significant number are short (9) and contain mutations that render the synthesis of a functional tubulin polypeptide impossible (42, 43). Thus, the number of expressed human β-tubulin genes is unknown. To address this question, we screened a cloned human cDNA library with the chicken β-cDNA probe. A positively hybridizing clone with an insertion about 1.4 kilobases (kb) long was selected for further study. This cDNA, clone Dβ-1, was sequenced essentially by the stratagem described previously for a closely related human β-tubulin processed pseudogene (43). Our data revealed that the cDNA clone included the bulk of the 3′ untranslated region and the entire coding region up to and including amino acid 7 (Fig. 1). A comparison of the cDNA sequence with the sequence of an intronless, poly(A)-containing human β-tubulin pseudogene revealed a high degree of homology (95.6%) that extends throughout the coding and noncoding regions. Within the coding region common to the two sequences, there are 55 base substitutions. Of these differences, 11 are in the first codon position, 16 are in the second, and 28 are in the third. Within the 3′ untranslated region, there are three single-base substitutions, four single-base insertions, and one single-base deletion.

Although the coding regions of tubulin genes are tightly conserved, available sequence evidence suggests that the 3′ untranslated regions are not conserved. Therefore, a high degree of homology in this region suggests a close evolutionary relationship. It seems likely that Dβ-1 represents the expressed product of the gene that, in an ancestral but closely related form, originally gave rise to the 46β processed pseudogene. This interpretation is strengthened by our observation that the human genome contains only a single functionally expressed β-tubulin gene with a 3′ untranslated region ho-

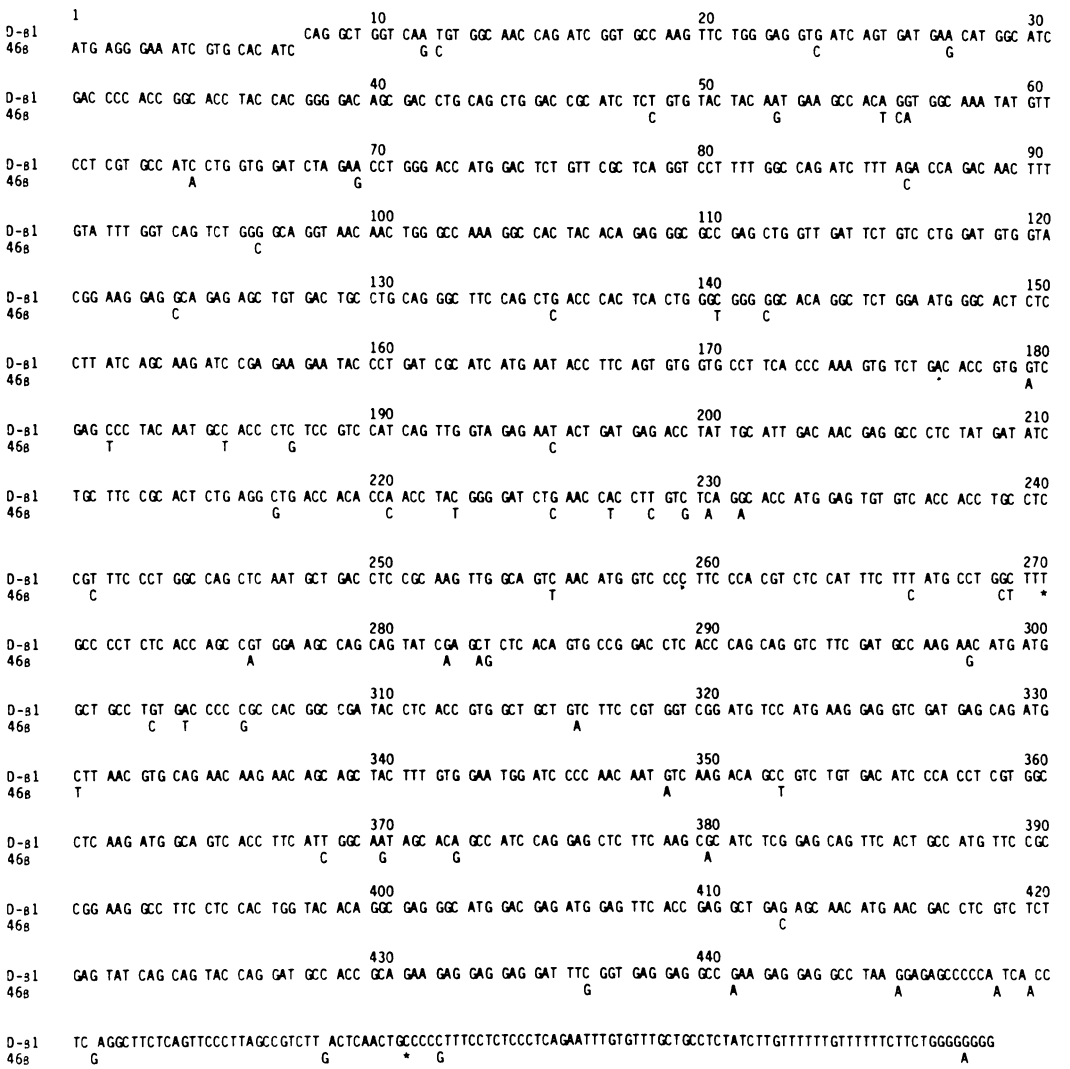

FIG. 1. Sequence of the human β-tubulin cDNA clone Dβ-1. The sequence of the 46β processed pseudogene (43) is shown for comparison. The asterisks indicate deleted bases.

mologous to Dβ-1 (M. Gwo-Shu Lee and N. J. Cowan, unpublished data).

A 3′-untranslated-region probe from Dβ-1 detects two sizes of human β-tubulin mRNA. Analysis of total HeLa poly(A)$^+$ RNA by blot hybridization using the intact Dβ-1 cDNA clone as a probe revealed the presence of two bands of about equal intensity at 1.8 and 2.6 kb (Fig. 2, track A). The large (2.6-kb) species is not a precursor of the smaller (1.8-kb) mRNA, since (i) the poly(A)$^+$ preparation was purified from cytoplasmic polyribosomes and (ii) poly(A)$^+$ RNA from nuclei showed no enrichment of the 2.6-kb band (data not shown). Surprisingly, a corresponding experiment in which the subcloned 3′ untranslated region of the cDNA clone was used as the probe gave results very similar to those obtained when the intact cDNA clone was used (Fig. 2, track B). Because only a single functional gene bears 3′-untranslated-region homology with Dβ-1 (see above), two mRNA products must be transcribed from this single sequence. This could occur either by a single gene transcript being subject to alternative splicing pathways or by the acquisition of a 3′ poly(A) tract at two alternative addition sites.

A 3′-untranslated-region probe subcloned from a genomic β-tubulin sequence detects only a 2.6-kb mRNA. We previously described the structure of a 6.8-kb human β-tubulin gene (5β) that contains three intervening sequences (9). To determine the nature of the mRNA (if any)

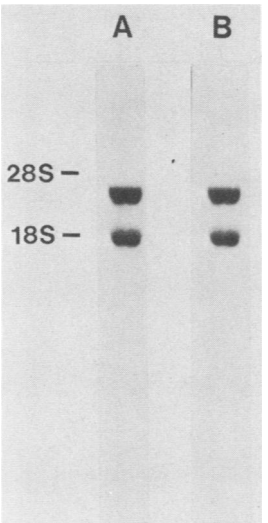

FIG. 2. Blot analysis of HeLa poly(A)$^+$ RNA. Approximately 5 μg of HeLa cell poly(A)$^+$ RNA was resolved in adjacent lanes of a 1% agarose gel containing formaldehyde (4). The contents of the gel were transferred to a nitrocellulose filter (34), and the filter was hybridized with about 10^7 cpm of either of the following probes labeled by nick translation. Track A. Intact Dβ-1; track B, subcloned 3'-untranslated-region probe from Dβ-1.

transcribed from this gene, two subcloned fragments were prepared, each spanning regions downstream from the termination codon (Fig. 3B). Probe A extended from an *Xho*I site 31 base pairs downstream from the TAA signal to a *Sac*I site 181 base pairs further downstream. Probe B extended from an *Xba*I site 364 base pairs downstream from the TAA signal to a *Bam*HI site 430 base pairs further downstream. When we used either probe A or probe B in RNA blots of HeLa poly(A)$^+$ RNA, we obtained the results shown in Fig. 3C. Both probes detected a single RNA species 2.6 kb long. This mRNA is not related to the 2.6-kb mRNA detected with the 3' probe from cDNA clone Dβ-1, since the sequence of the 3' untranslated regions of Dβ-1 (Fig. 1) and 5β showed no significant homology (data not shown). To demonstrate that the 2.6-kb mRNA detected by probe B was the transcription product of the 5β gene and not a different gene bearing a similar 3' untranslated region, genomic Southern blot experiments were performed with four different restriction digests of human DNA. Figure 4 shows that there was a single major band per restriction digest. The size of each major band corresponded precisely to the size of a fragment predicted on the basis of restriction maps of overlapping phages containing the 5β gene (Fig. 3A). Therefore, it seems likely that the 2.6-kb mRNA detected by probe B is indeed transcribed from 5β. However, the presence of a second, weaker band in three of the four restriction digests suggests the presence of a second genomic sequence related to 5β, and the possibility that the 2.6-kb transcript detected by probe B might be transcribed from another gene that bears a 3' untranslated region closely related to 5β cannot be ruled out.

To show that the 2.6-kb transcript detected by probe B indeed encoded human β-tubulin, a selective hybridization-translation experiment was carried out. Clone Dβ-1, its subcloned 3' end, and probe B from the 3' end of 5β (Fig. 3B) were each linearized, denatured, and bound to nitrocellulose filters. Each filter was hybridized in the presence of HeLa poly(A)$^+$ mRNA, and the selected material was eluted and translated in a cell-free system (25) containing [^{35}S]methionine. A sodium dodecyl sulfate-polyacrylamide gel analysis (21) of the translation products is shown in Fig. 5; track 1 shows the endogenous activity of the reticulocyte lysate, and track 2 reveals the translation products of total HeLa poly(A)$^+$ RNA. Both filters containing the entire Dβ-1 clone and filters containing its 3' untranslated region alone selected material that translated to give a pronounced β-tubulin band (Fig. 5, tracks 3 and 4). The filter containing probe B from the 3' end of 5β (Fig. 3B) also selected material that translated to yield β-tubulin (Fig. 5, track 5), although at much lower abundance. This observation is consistent with the weaker signals produced when this probe was used in RNA blots (Fig. 3C) compared with the signals produced when either Dβ-1 or its subcloned 3' untranslated region was used (Fig. 2) and is presumably a consequence of the relatively low abundance of mRNA transcribed from the 5β gene in HeLa cells.

DISCUSSION

Comparison of β-tubulin sequences from chickens, pigs, and humans suggests the existence of two β-tubulin isotypes. Apart from the human cDNA sequence described above, the only available β-tubulin sequence data are from chickens (37) and pigs (19); in the latter case, the data are restricted to the amino acid sequence. A comparison of the amino acid sequences (Fig. 6) revealed that the pig and chicken β-tubulins are significantly more homologus to one another (98.5%) than either is to the human sequence (approximately 95%). This surprising result contrasts sharply with the expectation that the two mammalian species would be more similar to each other than to the avian species. In particular, there is dramatic divergence within the carboxy-terminal 15 amino acids, where the sequences of Dβ-1 and either chicken or pig β-tubulin differ in six positions. This observation is

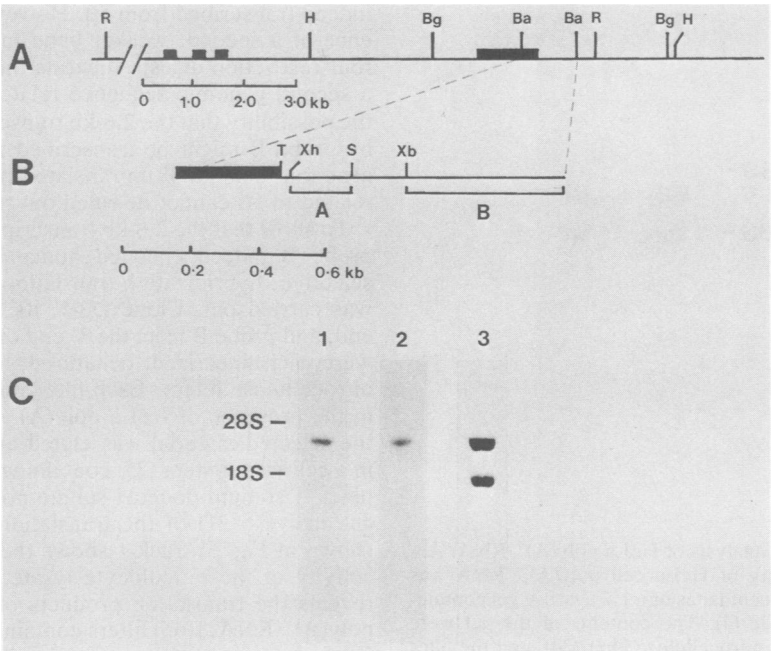

FIG. 3. Two 3′-untranslated-region probes from a human β-tubulin genomic sequence (5β) detect a 2.6-kb mRNA. (A) Restriction map of a human genomic β-tubulin sequence (5β) (9). The bars indicate coding regions, as determined by electron microscopic heteroduplex analysis. (B) Expanded view of a 1.1-kb *Bam*HI-*Bam*HI subclone encoding the 100 C-terminal amino acids and approximately 800 base pairs of downstream sequence. The regions labeled A and B were subcloned into the replicative form of bacteriophage M13 using the restriction sites shown. Ba, *Bam*HI; Bg, *Bgl*II; H, *Hin*dIII; R, *Eco*RI; S, *Sac*I; T, position of termination codon (TAA), as determined by sequence analysis; Xb, *Xba*I; Xh, *Xho*I. (C) Blot analysis of HeLa poly(A)$^+$ RNA. Track 1, Probe A; track 2, probe B; track 3, intact human β-tubulin cDNA cloned probe (Dβ-1). The positions of rRNA markers are shown.

not a consequence of sequence error, since, with the exception of two single-base substitutions, the sequence of the cDNA clone and that of the 46β processed pseudogene (Fig. 1) are identical in this region. In contrast, a partial sequence analysis of a second expressed human β-tubulin gene (5β) (Fig. 3B) revealed that, within the 15 C-terminal amino acids, the sequence more closely resembled the sequences of chicken and pig β-tubulins, differing from the latter in only three positions (Fig. 6). Thus, these data constitute strong evidence for the existence of two major human β-tubulin isotypes rather than an unexpected divergence of β-tubulin sequences between humans and pigs. Curiously, both human sequences are 444 amino acids long, in contrast to the 445 amino acids of pig and chicken sequences. The data in Fig. 6 are displayed in such a way as to retain maximum homology among all four β-tubulin polypeptides. The functional significance of the single amino acid difference between human and nonhuman β-tubulins is not clear.

Of the three complete sequences compared in Fig. 6, two (chicken and human) were obtained by analyzing cDNA clones, whereas the third (pig) was determined by direct amino acid sequence analysis of purified β-tubulin. In all cases, brain was used as the tissue of origin. Our experiments show that two β-tubulin genes yielding proteins with substantially different C termini are expressed in HeLa cells; a similar situation exits in human neural tissue (Dudley and Cowan, unpublished data). Therefore, it is surprising that the amino acid sequence analysis of β-tubulin purified from porcine brain revealed the existence of only limited internal microheterogeneity (19), with no heterogeneity toward the C terminus. One possible explanation for this observation is that the β-tubulin isotype represented by Dβ-1 is restricted in its expression to a defined developmental stage and therefore may be absent from nonembryonic porcine brain. Alternatively, it is conceivable that the purification procedure used to prepare the porcine tubulin (i.e., from a post-100,000 × *g* supernatant) could have selectively enriched for a limited subset of microtubules assembled from a distinct β-tubulin isotype. There have been several reports describing tubulins in membrane

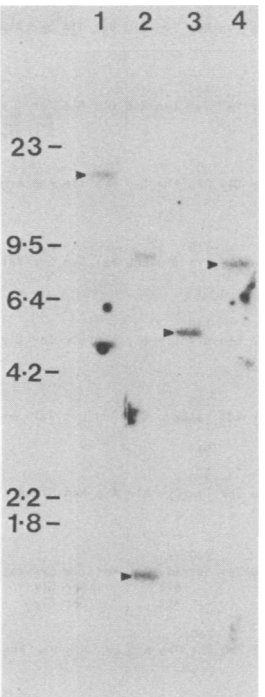

FIG. 4. Genomic Southern blot of human DNA when probe B was used (see Fig. 3B). Samples (10 μg) of human placental DNA were digested with one of four restriction endonucleases, and the mixtures were resolved on an 0.8% agarose gel. After transfer to nitrocellulose (31), the blots were hybridized with approximately 2×10^7 cpm of probe B (Fig. 3B) labeled by primed extension in the presence of α-^{32}P-labeled deoxynucleoside triphosphates (17). Track 1, *Eco*RI; track 2, *Bam*HI; track 3, *Bgl*II; track 4, *Hin*dIII. The arrowheads indicate bands corresponding to 5β (Fig. 3A).

fractions, particularly those from brain (2, 3, 33). Indeed, tubulin-specific mRNA has been found in association with the membrane-bound polyribosomes of embryonic chicken brain (7). This raises the intriguing possibility that functionally different microtubules are polymerized from distinctly different tubulin proteins.

Evolutionary origin of the 46β processed pseudogene. Restriction mapping and Southern blot analysis of several human β-tubulin gene-containing fragments have revealed regions of tubulin-specific DNA that are short (that is, no larger than the minimum size required to encode a β-tubulin mRNA) (9). A similar situation exists in the case of human α-tubulin genes (F. C. Wefald and N. J. Cowan, unpublished data). Sequence N. J. Cowan, unpublished data). A sequence analysis of several of the short β-tubulin genes has revealed some remarkable features (42, 43). The coding regions are devoid of intervening sequences but include one or more genetic lesions (insertions, deletions, or single-base changes that result in in-frame termination codons) that preclude the formation of a functional β-tubulin protein. In addition, these sequences contain a short tract of adenine residues downstream from the poly(A) addition sequence (AATAAA). Finally, immediately downstream from the oligoadenylic acid tract, there is a short sequence that is repeated directly upstream to the putative cap site. Because these molecules lack intervening sequences and contain a feature characteristic of mature cytoplasmic mRNAs [i.e., a 3′-terminal poly(A) tract], they have been termed processed pseudogenes (15).

Among proteins that are universally expressed in eucaryotic cells, such as the actins and tubulins, there is a high degree of evolutionary conservation (8). However, the 3′ untranslated regions of the genes encoding these proteins appear not to be under such evolutionary constraints; for example, the 3′ untranslated regions of the α-tubulin genes of *Drosophila* have been used to distinguish various gene prod-

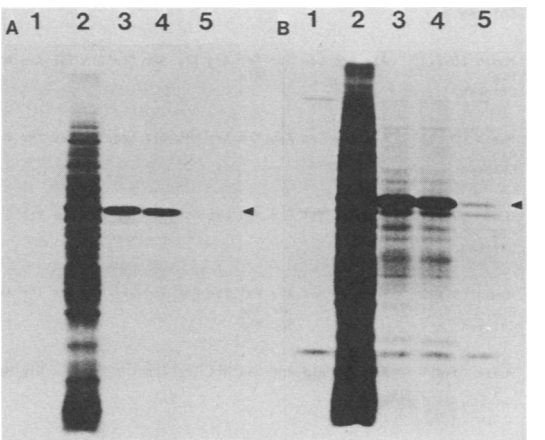

FIG. 5. Selective hybridization and translation using intact human β-tubulin cDNA clone Dβ-1, its subcloned 3′ end, and the subcloned 3′ end of the genomic β-tubulin sequence 5β. Hybrid selections were performed, using linearized, denatured, filter- and translation in a micrococcal nuclease-treated rabbit reticulocyte cell-free system (25) containing [^{35}S]methionine, the products were analyzed directly on stacking sodium dodecyl sulfate-polyacrylamide gels (21). Track 1, Endogenous activity of lysate (no added mRNA); track 2, translation products from 2 μg of total HeLa poly(A)$^+$ mRNA; track 3, mRNA selected by the intact Dβ-1 clone; track 4, mRNA selected by the subcloned 3′ untranslated region of Dβ-1; track 5, mRNA selected by probe B from the 3′ end of genomic β-tubulin sequence 5β (Fig. 3B). The two panels (A and B) are identical except that B was exposed for three times as long as A to visualize the translation products in track 5. The arrowheads indicate the position of (unlabeled) β-tubulin marker.

										10										20										30
Human (D-β1)								Gln	Ala	Gly	Gln	Cys	Gly	Asn	Gln	Ile	Gly	Ala	Lys	Phe	Trp	Glu	Val	Ile	Ser	Asp	Glu	His	Gly	Ile
Pig	Met	Arg	Glu	Ile	Val	His	Ile																							
Chicken																														

										40										50										60
Human (D-β1)	Asp	Pro	Thr	Gly	Thr	Tyr	His	Gly	Asp	Ser	Asp	Leu	Gln	Leu	Asp	Arg	Ile	Ser	Val	Tyr	Tyr	Asn	Glu	Ala	Thr	Gly	Gly	Lys	Tyr	Val
Pig					Ser		/Val								Glu			/Asn							Ala/Ser	Gly/Ser	Asn/His			
Chicken					Ser										Glu			Asn								Gly	Asn			

										70										80										90
Human (D-β1)	Pro	Arg	Ala	Ile	Leu	Val	Asp	Leu	Glu	Pro	Gly	Thr	Met	Asp	Ser	Val	Arg	Ser	Gly	Pro	Phe	Gly	Gln	Ile	Phe	Arg	Pro	Asp	Asn	Phe
Pig																														
Chicken																														

										100										110										120
Human (D-β1)	Val	Phe	Gly	Gln	Ser	Gly	Ala	Gly	Asn	Asn	Trp	Ala	Lys	Gly	His	Tyr	Thr	Glu	Gly	Ala	Glu	Leu	Val	Asp	Ser	Val	Leu	Asp	Val	Val
Pig																														
Chicken																														

										130										140										150
Human (D-β1)	Arg	Lys	Glu	Ala	Glu	Ser	Cys	Asp	Cys	Leu	Gln	Gly	Phe	Gln	Leu	Thr	His	Ser	Leu	Gly	Gly	Gly	Thr	Gly	Ser	Gly	Met	Gly	Thr	Leu
Pig				Ser																										
Chicken				Ser																										

										160										170										180
Human (D-β1)	Leu	Ile	Ser	Lys	Ile	Arg	Glu	Glu	Tyr	Pro	Asp	Arg	Ile	Met	Asn	Thr	Phe	Ser	Val	Val	Pro	Ser	Pro	Lys	Val	Ser	Asp	Thr	Val	Val
Pig																														
Chicken																				Met										

										190										200										210
Human (D-β1)	Glu	Pro	Tyr	Asn	Ala	Thr	Leu	Ser	Val	His	Gln	Leu	Val	Glu	Asn	Thr	Asp	Glu	Thr	Tyr	Cys	Ile	Asp	Asn	Glu	Ala	Leu	Tyr	Asp	Ile
Pig																														
Chicken																														

										220										230										240
Human (D-β1)	Cys	Phe	Arg	Thr	Leu	Arg	Leu	Thr	Thr	Pro	Thr	Tyr	Gly	Asp	Leu	Asn	His	Leu	Val	Ser	Gly	Thr	Met	Glu	Cys	Val	Thr	Thr	Cys	Leu
Pig						Lys															Ala			Ser	Gly					
Chicken						Lys															Ala			Ser	Gly					

										250										260										270
Human (D-β1)	Arg	Phe	Pro	Gly	Gln	Leu	Asn	Ala	Asp	Leu	Arg	Lys	Leu	Ala	Val	Asn	Met	Val	Pro	Phe	Pro	Arg	Leu	His	Phe	Phe	Met	Pro	Gly	Phe
Pig																														
Chicken																														

										280										290										300
Human (D-β1)	Ala	Pro	Leu	Thr	Ser	Arg	Gly	Ser	Gln	Gln	Tyr	Arg	Ala	Leu	Thr	Val	Pro	Asp	Leu	Thr	Gln	Gln	Val	Phe	Asp	Ala	Lys	Asn	Met	Met
Pig					/Ala													Glu					Met							
Chicken																		Glu					Met			Ser				

										310										320										330
Human (D-β1)	Ala	Ala	Cys	Asp	Pro	Arg	His	Gly	Arg	Tyr	Leu	Thr	Val	Ala	Ala	Val	Phe	Arg	Gly	Arg	Met	Ser	Met	Lys	Glu	Val	Asp	Glu	Gln	Met
Pig																														
Chicken																Ile														

										340										350										360
Human (D-β1)	Leu	Asn	Val	Gln	Asn	Lys	Asn	Ser	Ser	Tyr	Phe	Val	Glu	Trp	Ile	Pro	Asn	Asn	Val	Lys	Thr	Ala	Val	Cys	Asp	Ile	Pro	Pro	Arg	Gly
Pig																														
Chicken																														

										370										380										390
Human (D-β1)	Leu	Lys	Met	Ala	Val	Thr	Phe	Ile	Gly	Asn	Ser	Thr	Ala	Ile	Gln	Glu	Leu	Phe	Lys	Arg	Ile	Ser	Glu	Gln	Phe	Thr	Ala	Met	Phe	Arg
Pig				Ser	Ala																									
Chicken				Ser	Ala																									

										400										410										420
Human (D-β1)	Arg	Lys	Ala	Phe	Leu	His	Trp	Tyr	Thr	Gly	Glu	Gly	Met	Asp	Glu	Met	Glu	Phe	Thr	Glu	Ala	Glu	Ser	Asn	Met	Asn	Asp	Leu	Val	Ser
Pig																														
Chicken																														

										430										440					
Human (D-β1)	Glu	Tyr	Gln	Gln	Tyr	Gln	Asp	Ala	Thr	Ala	Glu	Glu	Glu	Glu	Asp	Phe	Gly	Glu	Glu	Ala	Glu	Glu	*	Glu	Ala
Pig											Asp		Gln	Gly	Glu		Glu			Gly			Asp		
Chicken											Asp		Gln	Gly	Glu		Glu			Gly			Asp		
Human (5β)											*		Gln	Gly	Glu		Glu						Glu	Val	

FIG. 6. Comparison of the amino acid sequences of β-tubulins from humans, pigs, and chickens. The data are derived from Fig. 1 (humans) and the studies of Krauhs et al. (19) (pigs) and Valenzuela et al. (37) (chickens). The C-terminal amino acids encoded by 5β (Fig. 3B), as determined from the nucleotide sequence of the 1.1-kb *Bam*HI subclone (Fig. 3B), are included for comparison. The asterisks at position 443 in Dβ-1 and at position 431 in the 5β sequence have been introduced as gaps so as to retain maximum homology among the β-tubulin sequences.

ucts (17), and even such closely related sequences as the α- and β-globin genes of humans show significant divergence in these regions (12). Thus, extensive homology between 3′ untranslated regions implies a close evolutionary relationship. Indeed, dissection of the human β-tubulin gene family with the subcloned 3′ end of Dβ-1 has identified a subfamily consisting of a single functional β-tubulin gene and three processed pseudogenes, including 46β. Therefore, it seems reasonable to conclude that Dβ-1 represents the expressed counterpart of the 46β pseudogene and that differences between the two sequences are a consequence of evolutionary drift since the original 46β integration event. In this context it is interesting to note that the number of third-codon-position differences between the cDNA clone and 46β (i.e., 28) is roughly double the number of first- or second-codon-position differences (11 and 16, respec-

tively). The greater number of third-position differences can be explained by the fact that the pseudogene acquires mutations that are not subject to selective pressure and are therefore random. In contrast, the functional gene can in general only acquire changes that are silent (i.e., changes that do not affect the amino acid sequence). Therefore, changes in the third codon position represent the cumulative effects of evolutionary drift in both functional and nonfunctional sequences. Assuming that approximately one-half of the third-base differences are due to changes in 46β and a rate of divergence of 0.7% per 10^6 years for neutral mutations (26), the 46β pseudogene may be estimated to have originated about 4.4×10^6 years ago.

Origin of the 1.8- and 2.6-kb mRNAs. The detection of two major β-tubulin mRNA species of different size in HeLa cells is an unexpected finding. The 3'-untranslated-region probes from both the genomic sequence (5β) and the cDNA clone detected a 2.6-kb mRNA (Fig. 3). These probes do not cross-hybridize; furthermore, overlapping phage containing 5β and including regions far downstream from the coding region show no homology with the cDNA 3'-untranslated-region probe. Thus, the 2.6-kb mRNA band detected by a coding region probe must be a composite of at least two comigrating species transcribed from different genes. The amount of mRNA transcribed from one of these genes (5β) is apparently relatively low, since in both RNA blot experiments (Fig. 3C) and selective hybridization-translation experiments (Fig. 5), the signals were weak compared with the corresponding signals obtained when the cDNA 3' probe was used.

Surprisingly, the two mRNA bands detected in blot experiments when either the intact cDNA clone or its subcloned 3' end (Fig. 2) was used are essentially indistinguishable. Because there is only a single functional gene that shares a homologous 3' untranslated region with Dβ-1, mRNAs of two different sizes must be transcribed from the same promoter. This may be a consequence either of alternative splicing pathways or of poly(A) addition at different sites. Alternative splicing has been shown to be the mechanism used in immunoglobulin heavy-chain synthesis as a means of transcribing mRNAs encoding both secreted and membrane-bound μ and δ chains from a single gene (6, 28). However, no functional requirement for alternative C-terminal domains is known to exist in the case of β-tubulin. On the other hand, addition of poly(A) at more than one site has been shown to be a feature of a number of eucaryotic genes (14, 30, 35, 36).

Concluding remarks. The sequence data reported above unequivocally demonstrate the expression in HeLa cells of a β-tubulin gene which is significantly divergent (especially within its 15 C-terminal amino acids) from β-tubulin sequences reported previously. This divergence does not seem to be a consequence of evolutionary drift, since a second expressed gene, 5β, has been identified that yields a polypeptide much more closely homologous to previously described β-tubulin polypeptides in other species. Thus, in humans there appear to be two distinct β-tubulin isotypes. It remains to be determined whether these two isotypes are capable of contributing to all microtubules or whether there are functional constraints such that certain microtubules are restricted to a single isotype content.

Some 15 to 20 β-tubulin-like sequences have been detected in the human genome by a β-tubulin cDNA probe (8). Sequence analysis of nine of these genes has shown that only two are capable of producing a translatable polypeptide; of the seven pseudogenes examined thus far, five are of the processed type (M. Gwo-Shu Lee and N. J. Cowan, Cell, in press). Thus, the periodic reintegration of reverse transcripts into germ line DNA, although rare, is nonetheless significant in evolutionary terms. Processed pseudogenes have also been described for rat α-tubulin (22) and human metalothionein (18). This passage of information from RNA to DNA is not restricted to the transcription products of *pol*II genes; a significant proportion of the abundant U_2 pseudogenes appear to have been generated by a similar mechanism (38). Despite their gradual evolutionary drift, these integrated transcripts retain many of their original features for significant periods of time and might be recruited, in whole or in part, for the generation of new functional sequences.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grants GM 31740 and GM 26028 from the National Institutes of Health to N.J.C. and to L. Villa-Komaroff, respectively, and by a Basic Research Grant from the Muscular Dystrophy Association. P.R.D. acknowledges the award of an American Cancer Society Postdoctoral Fellowship.

We thank Barbara Stebbins for technical assistance.

ADDENDUM IN PROOF

We have recently isolated a second human β-tubulin cDNA clone, DB-2. Sequence analysis of this clone reveals complete homology with corresponding coding region sequences of 5β.

LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid cellulose. Proc. Natl. Acad. Sci. U.S.A. **69**:1408–1412.
2. **Bhattacharrya, B., and J. Wolff.** 1975. Membrane-bound tubulin in brain and thyroid tissue. J. Biol. Chem. **250**:7639–7646.
3. **Bhattacharrya, B., and J. Wolff.** 1976. Polymerization of

membrane tubulin. Nature (London) **264**:576–577.
4. **Boedtker, H.** 1971. Conformation-independent molecular weight determination of RNA by gel electrophoresis. Biochim. Biophys. Acta **240**:448–453.
5. **Buell, G. N., M. P. Wickens, F. Payvar, and R. T. Schimke.** 1978. Synthesis of full length cDNAs from four partially purified oviduct mRNAs. J. Biol. Chem. **253**:2471–2482.
6. **Cheng, H.-L., F. Blattner, L. Fitzmaurice, J. F. Mushinski, and P. W. Tucker.** 1982. Structure of genes for membrane and secreted murine IgD chains. Nature (London) **296**:410–415.
7. **Cleveland, D. W., M. W. Kirschner, and N. J. Cowan.** 1978. Isolation of separate mRNAs for α- and β-tubulin and characterization of the corresponding *in vitro* translation products. Cell **15**:1021–1031.
8. **Cleveland, D. W., M. Lopata, R. J. McDonald, N. J. Cowan, W. J. Rutter, and M. W. Kirschner.** 1980. Number and evolutionary conservation of α- and β-tubulin and cytoplasmic β- and γ-actin genes using specific cloned cDNA probes. Cell **20**:95–205.
9. **Cowan, N. J., C. D. Wilde, L. T. Chow, and F. C. Wefald.** 1981. Variation in the structure of human β-tubulin genes. Proc. Natl. Acad. Sci. U.S.A. **78**:4877–4881.
10. **Dang, G. R., and R. Wu.** 1981. An improved procedure for utilizing terminal transferase to add homopolymers to 3′ termini of DNA. Nucleic Acids Res. **9**:4173–4188.
11. **Dustin, P.** 1978. Microtubules. Springer-Verlag, Berlin.
12. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. DeRiel, B. G. Forget, S. M. Weissman, J. K. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Shoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human β-globin gene family. Cell **21**:653–668.
13. **Grunstein, M., and D. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. U.S.A. **72**:3961–3965.
14. **Hellig, R., F. Perrin, F. Gannon, J. L. Mandel, and P. Chambon.** 1980. The ovalbumin gene family: structure of the X gene and evolution of duplicated split genes. Cell **20**:625–637.
15. **Hollis, G. F., P. A. Hieter, O. W. McBride, D. Swan, and P. Leder.** 1982. Processed genes: a dispersed human immunoglobulin gene bearing evidence of RNA-type processing. Nature (London) **296**:321–325.
16. **Hu, N., and J. Messing.** 1982. The making of strand-specific M13 probes. Gene **17**:271–277.
17. **Kalfayan, L., and P. C. Wensink.** 1982. Developmental regulation of *Drosophila* α-tubulin genes. Cell **29**:91–98.
18. **Karin, M., and R. I. Richards.** 1982. Human metallothionein genes—primary structure of the metallothionein-II gene and a related processed gene. Nature (London) **299**:797–802.
19. **Krauhs, E., M. Little, T. Kempf, R. Hofer-Warbinek, W. Ade, and H. Ponstingl.** 1981. Complete amino acid sequence of β-tubulin from porcine brain. Proc. Natl. Acad. Sci. U.S.A. **78**:4156–4160.
20. **Kurnit, D. M., B. M. Wentworth, L. DeLong, and L. Villa-Komaroff.** 1982. Construction of cloned libraries from RNA of human fetal tissues. Cytogenet. Cell Genet. **34**:193–203.
21. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
22. **Lemischka, I., and P. Sharp.** 1982. The sequence of an expressed rat α-tubulin gene and a pseudogene with an inserted repetitive element. Nature (London) **300**:330–335.
23. **McKnight, G. S.** 1978. The induction of ovalbumin and conalbumin mRNA by oestrogen and progesterone in chick oviduct explant cultures. Cell **14**:403–413.
24. **Messing, J.** 1979. A multipurpose cloning system based on the single stranded bacteriophage M13. Recombinant DNA Tech. Bull. **2**:43–48.
25. **Pelham, H. R. B., and R. J. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
26. **Perler, F., A. Efstratiadis, P. Lomedico, W. Gilbert, R. Kolodner, and J. Dodgsen.** 1980. The evolution of genes: the chicken preproinsulin gene. Cell **20**:555–566.
27. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling DNA to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
28. **Rogers, J., P. Early, C. Carter, K. Calame, M. Bond, L. Hood, and R. Wall.** 1980. Two mRNAs with different 3′-ends encode membrane-bound and secreted forms of immunoglobulin μ-chain. Cell **20**:303–312.
29. **Sanger, F., A. R. Coulsen, B. G. Barrell, A. J. H. Smith, and B. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
30. **Setzer, D. R., M. McGrogan, and R. T. Schimke.** 1982. Nucleotide sequence surrounding multiple polyadenylation sites in the mouse dihydrofolate reductase gene. J. Biol. Chem. **257**:5143–5147.
31. **Southern, E.** 1975. Detection of specific sequences among fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
32. **Staden, R.** 1980. A new computer method for the storage and manipulation of DNA gel reading data. Nucleic Acids Res. **8**:3673–3694.
33. **Stadler, J., and W. W. Franke.** 1972. Colchicine-binding proteins in chromatin and membranes. Nature (London) New Biol. **237**:237–238.
34. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. U.S.A. **77**:5201–5205.
35. **Tosi, M., R. A. Young, O. Hagenbuchle, and U. Schibler.** 1981. Multiple polyadenylation sites in a mouse α-amylase gene. Nucleic Acids Res. **9**:2313–2323.
36. **Unterman, R. D., K. R. Lynch, H. K. Nakhasi, K. P. Dolan, J. W. Hamilton, D. V. Cohn, and P. Feigelson.** 1981. Cloning and sequence of several α_{2u}-globulin cDNAs. Proc. Natl. Acad. Sci. U.S.A. **78**:3478–3482.
37. **Valenzuela, P., M. Quiroga, J. Zaldivar, J. W. Rutter, M. W. Kirschner, and D. W. Cleveland.** 1981. Nucleotide and corresponding amino acid sequences encoded by α- and β-tubulin mRNAs. Nature (London) **289**:650–655.
38. **Van Arsdell, S. W., R. A. Denison, L. B. Bernstein, A. M. Weiner, T. Manser, and R. F. Gesteland.** 1981. Direct repeats flank three small nuclear RNA pseudogenes in the human genome. Cell **26**:11–17.
39. **Villa-Komaroff, L., A. Efstratiadis, S. Broome, P. Lomedico, R. Tizard, B. Naber, W. L. Chick, and W. Gilbert.** 1978. A bacterial clone synthesizing proinsulin. Proc. Natl. Acad. Sci. U.S.A. **75**:3727–3731.
40. **Wickens, M. P., G. N. Buell, and R. T. Schimke.** 1978. Synthesis of double stranded DNA complementary to lysozyme ovomucoid and ovalbumin mRNAs. J. Biol. Chem. **253**:2483–2495.
41. **Wilde, C. D., L. T. Chow, F. C. Wefald, and N. J. Cowan.** 1982. Structure of two human α-tubulin genes. Proc. Natl. Acad. Sci. U.S.A. **79**:96–100.
42. **Wilde, C. D., C. E. Crowther, and N. J. Cowan.** 1982. Diverse mechanisms in the generation of human β-tubulin pseudogenes. Science **217**:549–552.
43. **Wilde, C. D., C. E. Crowther, T. P. Cripe, M. Gwo-Shu Lee, and N. J. Cowan.** 1982. Evidence that a human β-tubulin pseudogene is derived from its corresponding mRNA. Nature (London) **297**:83–84.