actions and contests occur (*13*). In view of the wide range of nearest-neighbor distances exhibited by the population observed, it is possible that there are several ecotypes of this species—that is, geographically distinct populations with genetically determined ranges of interindividual spacing, adapted to the environments in which they occur (*14*).

The combination of solitary and communal behavior exhibited by this species suggests that it represents an intermediate stage in the evolution of social behavior in spiders. Perhaps the communal behavior of *M. spinipes* evolved as the result of an increased tolerance of conspecifics in habitats where prey was locally or seasonally abundant. In such habitats, territory size and interindividual distance could at times be reduced, and populations would be large. If web sites with proper architectural support were patchily distributed, as *Agave* and *Opuntia* are, selection might favor individuals capable of tolerating conspecifics and attaching webs together. In such situations, the advantages of group living—exploitation of habitats free of competing species, increased prey capture efficiency, architectural stability of webs, and so on—would usually outweigh the advantages gained by maintaining maximum distances from conspecifics at the cost of aggressive behavior (*3, 4*). Retention of interindividual spacing mechanisms like web defense, however, would ensure survival when prey availability fluctuates. The end result would be a group-living species which displays considerable variation in social grouping tendency, as *M. spinipes* does.

GEORGE W. UETZ
THOMAS C. KANE
GAIL E. STRATTON
Department of Biological Sciences, University of Cincinnati, Cincinnati, Ohio 45221

References and Notes

1. B. Krafft, *Biol. Gabonica* **2**, 235 (1966); E. J. Kullman, *Am. Zool.* **12**, 419 (1972); R. Darchen, *Insectes Soc.* **20**, 379 (1973); Y. D. Lubin, *Zool. J. Linn. Soc.* **54**, 321 (1974); R. E. Buskirk, *Ecology* **56**, 1314 (1975); J. W. Burgess, *Sci. Am.* **234**, 100 (March 1976); V. Brach, *Evolution* **31**, 154 (1977); H. G. Fowler and J. Diehl, *Bull. Br. Arachnol. Soc.* **4**, 241 (1978).
2. R. R. Jackson, *Rev. Arachnol.* **1**, 133 (1978).
3. W. A. Shear, *Bull. Br. Arachnol. Soc.* **1**, 65 (1970).
4. J. W. Burgess, *Symp. Zool. Soc. London* **42**, 69 (1978); R. E. Buskirk, in *Social Insects*, H. R. Hermann, Ed. (Academic Press, New York, in press), vol. 2.
5. G. W. Uetz and J. W. Burgess, *Psyche* **86**, 79 (1979).
6. J. W. Burgess and P. N. Witt, *Interdisc. Sci. Rev.* **1**, 322 (1976).
7. In most localities, colonies represent annual cohorts of individuals hatched at approximately the same time, or successive cohorts hatching from serial egg sacs. Mating and oviposition occur in the late fall; *M. spinipes* overwinters (the dry season) as an egg or spiderling in the egg sac. Only in the moist tropics, where reproduction occurs year-round, is there overlap of a wide variety of contemporary age classes. Marked groups of adults and juveniles showed little change during several months and usually consisted of the offspring of the residents from the previous year.
8. J. W. Burgess, personal communication.
9. Enormous colonies of web-building spiders have been observed in power lines in mountain areas in Central America, but the species was not identified because the power lines were not accessible. Near Cordóba, colonies had spilled over onto the ground and were identified as those of *M. spinipes*. We estimated colony size in several large groups by measuring the volume of the communal web and then extrapolating from density counts. Several of the larger groups contained an estimated 5800 to 6500 individuals.
10. Prey availability was estimated by observation of the number of insects approaching a 1-m^3 colony web for a specified time period and then correction of this estimate by the number of hours of favorable weather conditions in a day.
11. S. E. Riechert, *Symp. Zool. Soc. London* **42**, 211 (1978).
12. ———, *Am. Nat.* **117**, 871 (1981).
13. R. E. Buskirk, *Anim. Behav.* **23**, 560 (1975).
14. It is also possible that the several populations observed represent distinct species or subspecies. Although the genitalia of all specimens match the drawings in the species description of F. O. Pickard-Cambridge [*Biol. Cent. Am.* **2**, 457 (1903)], there are considerable differences in body size, leg length, and life history, as well as social grouping tendency. *Metepeira spinipes* is the only group living spider known in the genus *Metepeira* [H. W. Levi, *Bull. Mus. Comp. Zool. Harv. Univ.* **148**, 185 (1977)], although T. W. Schoener (personal communication) has found groups of *M. daytoni*.
15. We thank R. Hollis and M. K. Uetz for assistance in the field. Supported by the National Geographic Society, the American Philosophical Society, and the University of Cincinnati Research Council.

29 January 1982; revised 13 April 1982

Diverse Mechanisms in the Generation of Human β-Tubulin Pseudogenes

Abstract. *The sequence of two human β-tubulin pseudogenes is described. One contains an intervening sequence but lacks sequences encoding the 55 N-terminal amino acids of the polypeptide chain. A second has no introns but has a polyadenylate signal and an oligoadenylate tract at its 3' end, and it is flanked by a short direct repeat. These sequences have arisen by different mechanisms, including one that probably involves reverse transcription of a processed messenger RNA and reintegration of the complementary DNA copy into the genome.*

With few exceptions, all expressed eukaryotic genes contain intervening sequences (introns), regions of DNA that interrupt the coding sequence and that are spliced out of the primary gene transcript as part of the generation of cytoplasmic messenger RNA (mRNA). Any of a number of genetic events including deletion, insertion, or acquisition of one or more stop codons within the coding sequence can lead to a failure to yield a functional transcript. Such a gene is therefore termed a pseudogene. Among the pseudogenes thus far described, a curious form is that in which the intervening sequences present in the ex-

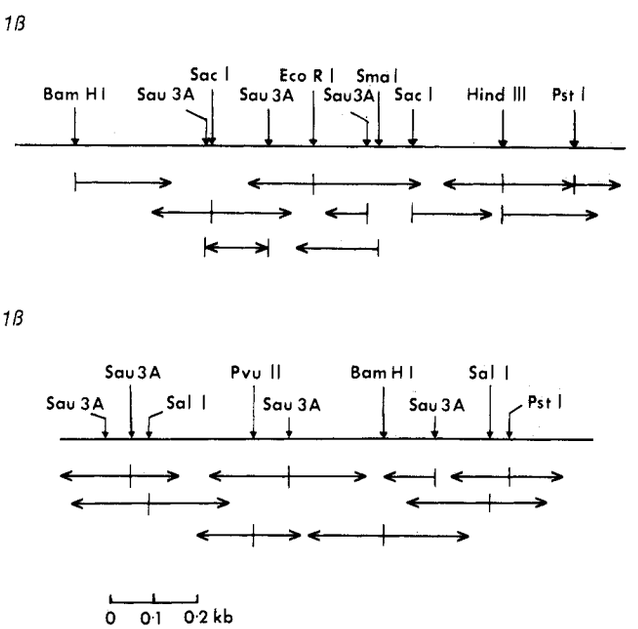


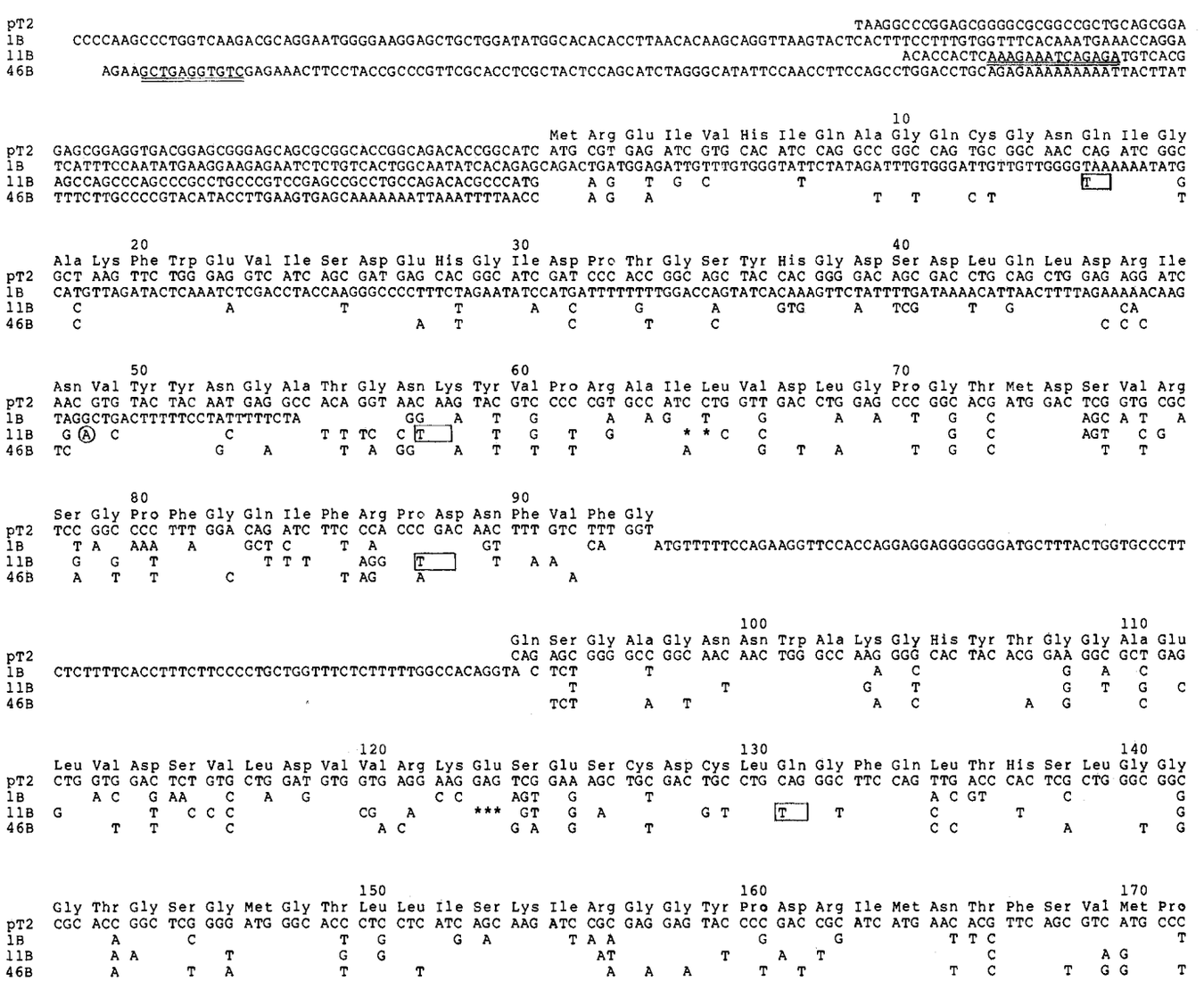
Fig. 1. Sequencing strategies. A 4.1-kb Hind III–Eco RI restriction fragment (11β) and a 1.55-kb Eco RI fragment (1β), each containing β-tubulin sequences, were isolated from the appropriate Charon 4A lambda clones (*3*) by electrophoresis in low melting temperature agarose and extraction with phenol. A further 1.65-kb Bam HI–Hind III fragment containing β-tubulin (1β) was subcloned into pBR322. Restriction endonuclease digestion of these DNA's was performed with the enzymes indicated and the fragments ligated into appropriately cleaved replicative forms of bacteriophage M13 mp8 and mp9. In some cases, plaques containing the DNA to be sequenced were identified on nitrocellulose replicas (*16*) by hybridization with a nick-translated chicken β-tubulin cDNA probe (*17*).

0036-8075/82/0806-0549$01.00/0

pressed homolog have been precisely eliminated (*1, 2*). Because the mechanism whereby intervening sequences are removed involves splicing at the RNA level, it is logical to ascribe the generation of intron-deficient pseudogenes to a process of reverse transcription, in which a spliced RNA transcript is copied into DNA and the DNA copy is reinserted into the genome. Here we present evidence that two genomic human β-tubulin sequences lack introns but contain 3-polyadenylate [poly (A)] tracts and that each is derived independently by reverse transcription of different spliced mRNA's. A third intron-containing pseudogene which has lost a substantial proportion of its 5′ coding sequences, probably as a consequence of deletion, is also described.

Alpha- and beta-tubulins are the principal proteins of microtubules, filamentous structures found in association with several subcellular structures in eukaryotic cells, including the mitotic apparatus, cilia, flagella, and elements of the cytoskeleton. Evidence based on amino acid sequencing, immunological cross-reactivity, and the ability of cloned tubulin-specific complementary DNA (cDNA) probes to hybridize to DNA fragments from diverse species under stringent conditions points to a close evolutionary conservation among tubulins. In humans, the genes that encode these proteins each constitute a multigene family consisting of about 15 to 20 members, several of which have been isolated as recombinant fragments (*3–5*). These sequences include examples that are short, that is, approximately the size required to encode a human tubulin mRNA. The complete sequence of two such genes was determined by the stratagem depicted in Fig. 1. In Fig. 2, the sequence data is compared with the sequence of a second short human β-tubulin gene (*6*) and a cloned chicken β-tubulin cDNA probe (*7*). Two of the human sequences are devoid of introns, and all three contain lesions (in-frame termination codons and deletions or insertions that cause frameshifts and resultant premature termination signals) that preclude the generation of a functional product. Like 46β, the sequence of 11β includes at its 3′ end a poly(A) signal (AATAAA) and an oligo(A) tract 16 base pairs further downstream. In addition, 11β is flanked by a short direct repeat of sixteen base pairs that differs completely from the direct repeat flanking 46β.

The sequence of 1β reveals two single-base deletions (in the first codon position of glycines 269 and 400) that lead quickly to in-phase termination codons at amino acid positions 361 and 405. However, in contrast to 46β and 11β, this gene contains a short (120 base pairs) intervening sequence that interrupts the coding sequence at glycine 94 and is flanked by correct consensus splice signals. There



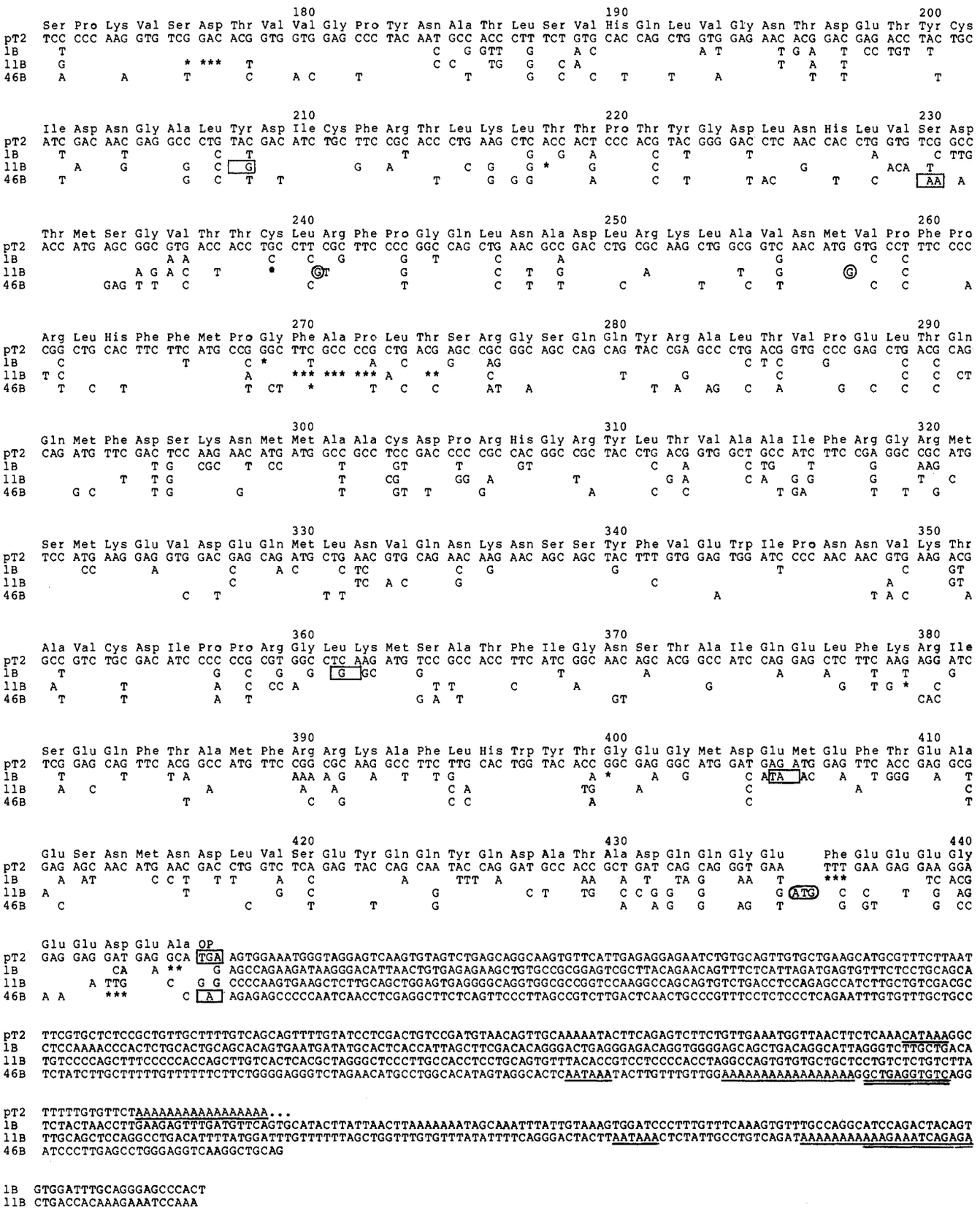


Fig. 2. Sequence of three human β-tubulin pseudogenes. The sequences of 1β and 11β were determined by the dideoxy chain terminator method (*18*) and are compared with the sequence of 46β (*6*). Asterisks denote deleted bases; circles denote inserted bases. Both have been introduced so as to retain maximum homology (represented by gaps within the coding regions) with the prototype chicken β-tubulin cDNA sequence pT2 (*7*). Termination codons are boxed. The poly(A) signals and the oligo(A) tracts are underlined. Flanking direct repeats are doubly underlined. Abbreviations: Ala, alanine; Arg, arginine; Asp, aspartic acid; Asn, asparagine; Cys, cysteine; Gln, glutamine; Glu, glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine, Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; and Val, valine.

is also substantial divergence from the other sequences toward the 3′ end, and 1β contains no discernible poly(A) addition site or significant oligo(A) tract within 700 base pairs further downstream; nor is there evidence of direct or inverted repeat sequences in the regions immediately flanking this gene. Toward the 5′ end (at amino acid 55) an unexpected feature is that all homology, both with respect to the chicken cDNA probe and the intron-lacking genomic sequences, is abruptly lost. Inspection of the region spanning the threonine 55 codon reveals a sequence (TTTCTAC*AG*GT) (T, thymine; C, cytosine; G, guanine; A, adenine) that closely resembles a consensus acceptor splice signal, including the obligatory dinucleotide AG. Indeed, the structure of a 6.8-kb (kilobase) human β-tubulin gene, determined by heteroduplex analysis with the cloned chicken β-tubulin cDNA, shows the presence of a short intervening sequence at approximately this location (*3*). However, Southern's hybridization experiments (*8*) on overlapping recombinant phage that contain the 1β pseudogene and extend 15 kb upstream from the region sequenced show no further evidence of hybridization with human or chicken β-tubulin probes containing 5′ coding sequences. Thus, although the existence of an intervening sequence greater than 15 kb cannot be ruled out, it seems more likely that some DNA rearrangement (deletion, inversion, recombination, or transposition) with a breakpoint upstream from amino acid 55 resulted in a loss of sequences encoding the 5′ end of this gene.

An additional feature of 1β is that the largest (4.9 kb) intron contained in the 6.8-kb gene (*3*) is absent. One possibility is that this intervening sequence was eliminated by reverse transcription of an RNA processing intermediate, and the cDNA copy reintegrated into the genome. This mechanism would appear unlikely, however, in view of the absence of a poly(A) signal, poly(A) tract, or flanking direct repeats. A more plausible explanation is that 1β was generated solely by mutational events from a functional β-tubulin gene that lacks the 4.9-kb intron. In contrast to the known intron-containing pseudogenes of the globin gene families (*9*), 1β does not appear to be closely linked to other β-tubulin sequences. A transposition event may have conveyed 1β to its present location, or it may represent a member of the β-tubulin multigene family that was already dispersed, and subsequently mutated so as to be no longer functional.

The addition of A residues to the 3′ ends of mRNA molecules is a post-transcriptional event. Hence, the occurrence of 3′-poly(A) tracts in the two intron-lacking human β-tubulin pseudogenes (Fig. 2) is suggestive of a mechanism involving the reverse transcription of processed mRNA. How could such molecules be generated and, once generated, how could they become integrated into the genome? Among the requirements would be a reverse transcriptase and a functional primer for the synthesis of plus and minus strands. Although no obvious molecules exist that might serve as primers, the rarity of successful synthesis could be a consequence of improbable events. A possible integration mechanism might then involve recombination with a transposable element, for example, retrovirus DNA. Indeed, recombination of retroviral sequences with a fragment of genomic DNA and loss of intervening sequences during the RNA phase of the virus life cycle—a process known to occur with incorporated host genes (*10*)—can satisfactorily explain the observation that two intron-lacking mouse α-globin pseudogenes both retain homology with respect to functionally expressed genes 5′ to the mRNA capping site (*1, 2*). This mechanism is also consistent with the detection of flanking retrovirus-like elements (*11*) and the observation that at least one of the pseudogenes is dispersed relative to its expressed homolog (*12*). However, neither of these pseudogenes contains a 3′-oligo(A) tract and, although a spliced mRNA or its cDNA cognate could conceivably recombine directly with a retrovirus sequence, a much simpler mechanism would involve the direct integration of a cDNA copy into a staggered host chromosomal break. An aberrant human immunoglobulin λ light chain transcript (*13*) may also have been generated in this fashion. Direct repeats could then arise as a consequence of DNA repair at the integration site (*14*).

The sequences corresponding to the 3′ untranslated and 5′ untranslated regions of the two intron-lacking human β-tubulin pseudogenes are of different length and show no significant homology. It seems likely, therefore, that 46β and 11β arose by reverse transcription of two independently expressed β-tubulin genes. If we assume that reintegration of cDNA copies can occur anywhere in the genome, the propagation of intron-lacking pseudogenes in the germ line would require that there be no interference with the expression of other essential genes. The accumulation of a number of mutations in these integrated cDNA's suggests that they were probably nonfunctional. Eventually, such progressive changes might result in a region of DNA that is unrecognizable with respect to the original transcript. That such a degenerative process indeed occurs is suggested by the greater number of insertions, deletions, and amino acid substitutions (compared with the functional chicken cDNA sequence) in 11β, and implies that the integration of this reverse transcript significantly predates the integration of 46β. The sequence of a third intron-lacking human β-tubulin pseudogene that also contains multiple point mutations, small deletions, and insertions throughout its length is consistent with this conclusion (*15*). It remains to be seen to what extent the genomes of higher vertebrates are populated by dispersed reintegrated cDNA molecules generated by reverse transcription of spliced mRNA's. In any event, it is clear that several distinct and independent pathways can lead to the existence of homologous but apparently functionless genelike sequences.

C. D. WILDE
C. E. CROWTHER
N. J. COWAN*
Department of Biochemical Sciences,
Princeton University,
Princeton, New Jersey 08544

References and Notes

1. Y. Nishioka, A. Leder, P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2806 (1980).
2. E. F. Vanin, G. I. Goldberg, P. W. Tucker, O. Smithies, *Nature (London)* **286**, 222 (1980).
3. N. J. Cowan, C. D. Wilde, L. T. Chow, F. C. Wefald, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4877 (1981).
4. C. D. Wilde, L. T. Chow, F. C. Wefald, N. J. Cowan, *ibid.* **79**, 96 (1982).
5. C. D. Wilde, C. E. Crowther, N. J. Cowan, *J. Mol. Biol.* **155**, 533 (1982).
6. C. D. Wilde, C. E. Crowther, T. P. Cripe, M. G.-S. Lee, N. J. Cowan, *Nature (London)* **297**, 83 (1982).
7. P. Valenzuela, M. Quiroba, J. Zaldivar, W. J. Rutter, M. W. Kirschner, D. W. Cleveland, *ibid.* **289**, 650 (1981).
8. E. Southern, *J. Mol. Biol.* **98**, 503 (1975).
9. P. F. R. Little, *Cell* **28**, 683 (1982).
10. S. P. Goff, E. Gilboa, O. N. Witte, D. Baltimore, *ibid.* **22**, 777 (1980).
11. K. Lueders, A. Leder, P. Leder, E. Kuff, *Nature (London)* **295**, 426 (1982).
12. A. Leder, D. Swan, F. Ruddle, P. D'Eustachio, P. Leder, *ibid.* **293**, 196 (1981).
13. G. F. Hollis, P. A. Hieter, O. W. McBride, D. Swan, P. Leder, *ibid.* **296**, 321 (1982).
14. S. W. Van Arsdell, R. A. Denison, L. B. Bernstein, A. M. Weiner, T. Manser, R. F. Gesteland, *Cell* **26**, 11 (1981).
15. T. P. Cripe and N. J. Cowan, unpublished data.
16. W. D. Benton and R. W. Davis, *Science* **196**, 180 (1977).
17. D. W. Cleveland, M. A. Lopata, R. J. McDonald, N. J. Cowan, W. J. Rutter, M. W. Kirschner, *Cell* **20**, 95 (1980).
18. F. Sanger, A. R. Coulson, B. G. Barrell, A. J. H. Smith, B. Roe, *J. Mol. Biol.* **143**, 161 (1980).
19. Supported by NIH grant GM26456. C.D.W. acknowledges the award of a Muscular Dystrophy postdoctoral fellowship.
* Present address: Department of Biochemistry, New York University Medical Center, New York 10016.

9 April 1982; revised 24 May 1982