Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 5833–5837, August 1987
Genetics

Characterization of the α-tubulin gene family of *Arabidopsis thaliana*

(gene structure/protein homologies/transcription)

STEVEN R. LUDWIG*, DAVID G. OPPENHEIMER†, CAROLYN D. SILFLOW†‡, AND D. PETER SNUSTAD†‡§

Departments of *Biochemistry and †Genetics and Cell Biology and ‡Institute of Plant Molecular Genetics, University of Minnesota, St. Paul, MN 55108

Communicated by Bruce M. Alberts, April 27, 1987

ABSTRACT **The genome of *Arabidopsis thaliana* (Linnaeus) Heynhold was shown to contain an α-tubulin gene family consisting of at least four genes and/or pseudogenes. The primary structure of a transcribed α-tubulin gene was determined. A comparison of the predicted amino acid sequence of the *A. thaliana* α-tubulin with the predicted amino acid sequences of α-tubulins of *Chlamydomonas reinhardtii*, *Stylonychia lemnae*, and *Homo sapiens* reveals a high degree of homology; 90%, 87%, and 83% identity, respectively. Thus, a plant α-tubulin exhibits a high degree of homology to the α-tubulins of protists and animals. The coding sequence of the *A. thaliana* α-tubulin gene is interrupted by four introns, which occur at positions different from those of the less numerous introns of *C. reinhardtii* and rat α-tubulin genes. S1 nuclease mapping data showed that transcription is initiated 99 ± 1 base pairs upstream from the translation initiation codon. Both 5′ and 3′ noncoding gene-specific probes were used to examine the expression of the α-tubulin gene in leaves, roots, and flowers by hybridization to total RNA isolated from these tissues. The results showed that the α-tubulin gene was transcribed in all three tissues.**

Microtubules play central roles in several of the most basic processes of eukaryotic cells: cell division, cell motility, intracellular transport, and the control of cell shape. In plant cells, rigid cell walls obviate the need for direct cytoskeletal maintenance of cell shape. However, wall formation and the division of walled cells require the action of several microtubule arrays unique to plant systems. These arrays include the cortical microtubules involved in orientation of cellulose microfibrils, the preprophase band (which delineates the plane of the ensuing cell division), and the phragmoplast (which forms at the equatorial plane of the spindle to participate in formation of the new cell wall) (see refs. 1–3 for reviews).

Approaches to understanding the diversity of microtubule function in animal and fungal systems have included the molecular and genetic analysis of genes coding for α- and β-tubulin, the major protein components of microtubules (see refs. 4 and 5 for reviews). These studies have shown that many eukaryotes contain multigenic α- and β-tubulin gene families, which include constitutively expressed members as well as genes that exhibit tissue-specific expression. For example, tubulin genes that are preferentially expressed in the testis have been found in mouse, chicken, and *Drosophila* (6–9), whereas tubulin genes that are expressed predominately in neuronal tissue have been identified in chicken, rat, and human (7, 10, 11). The conservation of amino acid sequence between tubulins showing the same tissue-specific expression in different species (12) and the unusual chemical properties of tubulin found in specific microtubules such as the marginal band of avian erythrocytes (13) suggest that the tubulin components contribute to the specificity of microtubule function in some cases. However, a number of studies have also shown that the product of an individual gene can participate in a wide variety of different microtubule types within one cell or tissue (14–17). In these cases, specificity of microtubule function does not derive from the basic tubulin components but from associated proteins, post-translational modification of the tubulin (18, 19), or other cellular factors that control the spatial organization and stability of microtubules.

In contrast to the animal and fungal systems, for higher plants essentially nothing is known about the structure and regulation of the tubulin genes. Multiple α- and β-tubulin isoforms have been demonstrated in *Phaseolus vulgaris* (20); however, no information is available as to whether these isoforms are the products of different genes or are produced by post-translational modification of one or a few gene products.

Arabidopsis thaliana (Linnaeus) Heynhold provides an ideal model system (21–23) with which to dissect the genetic control of α-tubulin synthesis in a higher plant. The key advantages are (*i*) the small genome size, approximately 7×10^4 kilobases (kb) (24), with a minimal amount of repetitive DNA, 1.8×10^4 kb (25); (*ii*) the short generation time, about five weeks (26); and (*iii*) simple transformation techniques using cocultivation of leaf disk explants with *Agrobacterium tumefaciens* harboring disarmed Ti plasmids carrying selectable chimeric genes (27).

In this paper, we present evidence for the existence of an at least four member homologous α-tubulin gene and/or pseudogene family in *A. thaliana*, and we report the primary structure of one member of this gene family. The homologies between the predicted amino acid sequence of this plant α-tubulin and the α-tubulins of animals and protists indicate that this structural protein has been highly conserved throughout evolution. In addition, we show that this α-tubulin gene is transcribed in roots, leaves, and flowers; these results suggest that this gene may code for an α-tubulin that is constitutively synthesized in all tissues of the plant.

MATERIALS AND METHODS

Plant Material. The Landsberg *erecta* and Columbia strains of *A. thaliana* were obtained from E. Meyerowitz (Division of Biology, California Institute of Technology, Pasadena, CA). The plants were grown under a regime of 18 hr of light at 25°C on a mixture of sterile soil, sand, and peat (3:3:2, vol/vol).

Genomic Blot Hybridizations. *A. thaliana* DNA was isolated from whole plants as described by Murray and Thompson (28). Digestions with restriction enzymes, gel electrophoresis, transfers to membranes, and hybridizations were

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

§To whom reprint requests should be addressed.

done basically as described by Southern (29) except that alkaline transfers (30) were to Bio-Rad Zeta-Probe nylon membranes. Probes were labeled by primer extension by using the protocol of Hu and Messing (31). Filter hybridizations were done at 65°C in 6× SSC/50 mM Tris·HCl, pH 7.5/10× Denhardt's solution/0.1% $Na_4P_2O_7$/1% $NaDodSO_4$ containing calf thymus DNA at 100 μg/ml. (1× SSC = 0.15 M NaCl/0.015 M sodium citrate; Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone.) The last two washes were carried out in 2× SSC/1% $NaDodSO_4$ at 47°C for 30 min each.

Screening the Genomic Library. An *A. thaliana* Columbia genomic library prepared in λ *sep*6 (25) was screened by the procedure of Benton and Davis (32) using nick-translated (ref. 33, pp. 109–110) *Chlamydomonas reinhardtii* α-tubulin coding-sequence subclone α10-2 (34) as a probe. Hybridizations were carried out at 65°C in 6× SSC/5× Denhardt's solution/20 mM sodium phosphate, pH 7.5/0.5% $NaDodSO_4$ containing calf thymus DNA at 100 μg/ml. The final two filter washes were done in 2× SSC/0.2% $NaDodSO_4$ at 47°C for 30 min each.

Template Preparation and DNA Sequencing. Restriction fragments of the λ α-tubulin clone were subcloned in pUC119, a derivative of pUC19 (35) containing the M13 origin of replication (J. Vieira and J. Messing, personal communication). An overlapping set of deletion subclones suitable for sequencing were generated by the T4 DNA polymerase procedure of Dale *et al.* (36). Packaged single-stranded DNAs of pUC119 derivatives were prepared by using the protocol of McMullen *et al.* (37). Single-stranded template DNAs were isolated as described by Messing (38) and were sequenced by the dideoxy method of Sanger *et al.* (39). Sequence data were analyzed by using the IntelliGenetics software programs on a Sun Microsystems 2/120 computer.

RNA Isolation and Blotting. Total RNA was prepared from *A. thaliana* whole plants or from isolated leaves plus petioles, whole roots, or flowers at various stages of development by the method of Berry *et al.* (40). Poly(A)$^+$ RNA was isolated from total RNA preparations by oligo(dT)-cellulose chromatography as described by Aviv and Leder (41). RNA samples were fractionated on 1.2% agarose/formaldehyde gels and transferred to Schleicher & Schuell Nytran membranes as described by Maniatis *et al.* (ref. 33, p. 203). The probes were labeled by primer extension as described by Hu and Messing (31). Hybridizations were done at 42°C in 50% (vol/vol) formamide/2× Denhardt's solution/0.5% $NaDodSO_4$/5× SSC/0.05% $Na_4P_2O_7$ containing salmon sperm DNA at 200 μg/ml. The last two washes were carried out in 2× SSC/0.5% $NaDodSO_4$ at 47°C for 30 min each.

S1 Nuclease Analysis. The 5′ end of the *A. thaliana* α-tubulin gene transcript was mapped by using the S1 nuclease procedure of Maniatis *et al.* (ref. 33, pp. 207–209). Poly(A)$^+$ RNA (2.5 μg) from *A. thaliana* whole plants was hybridized overnight to end-labeled (10^5 cpm total) probe at 46°C. The hybridization mixture was digested with S1 nuclease (100 units/ml) and analyzed by electrophoresis on a polyacrylamide sequencing gel.

RESULTS

Evidence for at Least Four α-Tubulin Genes and/or Pseudogenes in *A. thaliana*. Genomic DNA blots (Fig. 1) demonstrate the presence of at least four different genomic sequences with homology to a sequenced (see below) *A. thaliana* α-tubulin gene. *Eco*RI and *Hin*dIII genomic digests were hybridized to nonoverlapping 5′ and 3′ α-tubulin genomic subclones in an attempt to determine which might contain complete or nearly complete α-tubulin coding sequences. Although the 5′ probe contained some 5′ noncoding

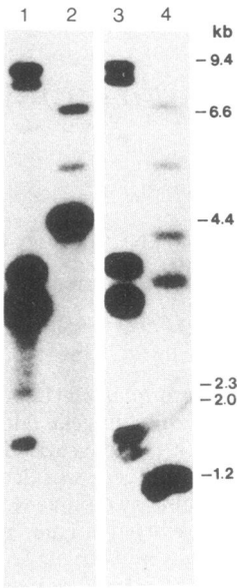

FIG. 1. Analysis of genomic α-tubulin sequences. Lanes 1 and 2, *A. thaliana* Landsberg genomic DNA digested with *Eco*RI and *Hin*dIII, respectively, and probed with a 5′ α-tubulin probe (nucleotides −1000 to 975; see Fig. 3). Lanes 3 and 4, *A. thaliana* Landsberg genomic DNA digested with *Eco*RI and *Hin*dIII, respectively, and probed with a 3′ α-tubulin probe (approximately nucleotide 1034 to nucleotide 2045; see Fig. 3).

sequence, this apparently did not affect the results, since all of the nonsequenced fragments that hybridized to the 5′ probe also hybridized to the 3′ probe. *Hin*dIII fragments of approximately 6.5, 5.2, and 4.2 kb were shown to contain sequences homologous to both the 5′ and 3′ α-tubulin gene probes (Fig. 1, lanes 2 and 4). On the autoradiograph shown, the 4.2-kb fragment is partially masked by strong hybridization of the 5′ probe to a 4.4-kb *Hin*dIII fragment; however, its presence is clear after a shorter exposure (data not shown). The 4.4-kb *Hin*dIII fragment contains the 5′ end of the sequenced gene; the strong hybridization thus results from the perfect homology between this fragment and the probe. The 3′ end of the sequenced gene is located on a 1.2-kb *Hin*dIII fragment (Fig. 1, lane 4). These *Hin*dIII results and the supporting *Eco*RI data (Fig. 1, lanes 1 and 3) indicate the presence of four different α-tubulin genes and/or pseudogenes. In addition, a *Hin*dIII fragment approximately 3 kb in size is homologous to the 3′ probe but not to the 5′ probe. This fragment probably contains the 3′ distal end of a gene or pseudogene, the remainder of which resides on the 6.5-kb, the 5.2-kb, or the 4.2-kb *Hin*dIII fragment. Alternatively, one of these three large *Hin*dIII fragments may contain two 5′ α-tubulin gene segments, one of which is joined to the 3-kb *Hin*dIII fragment that exhibits homology to the 3′ α-tubulin gene probe. In addition, one or more of the large *Hin*dIII and *Eco*RI fragments could contain more than one α-tubulin gene and/or pseudogene. In fact, in *A. thaliana*, such clustering of related genes is known to occur for the three chlorophyll-*a*/*b*-binding protein coding sequences (42) and for two β-tubulin genes and/or pseudogenes (D.G.O. and N. Haas, unpublished data). Thus, our data indicate that the *A. thaliana* genome contains at least four, but possibly more, α-tubulin genes and/or pseudogenes. The above genomic blot analyses have been carried out with DNA from both the Columbia and the Landsberg *erecta* strains of *A. thaliana*, and no difference in gene number was observed.

Primary Structure of an *A. thaliana* α-Tubulin Gene. A λ *A. thaliana* Columbia genomic library (kindly provided by E. Meyerowitz) was screened for cross-hybridization to a *Chlamydomonas reinhardtii* α-tubulin gene probe. Blot hybridization analyses of DNA from a positive clone showed that a 1.2-kb *Hin*dIII fragment and a 3-kb *Eco*RI fragment hybridized strongly to the *C. reinhardtii* α-tubulin gene probe. The fragments were subcloned in pUC119 and both strands of the gene were sequenced in a series of overlapping clones (Fig. 2) generated by the unidirectional deletion method of Dale *et al.* (36). The nucleotide sequence of the α-tubulin gene is

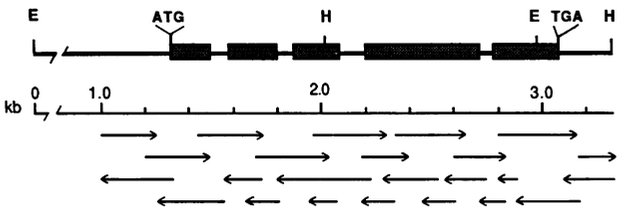

FIG. 2. DNA sequencing strategy. The top line shows the positions of the *Eco*RI and *Hin*dIII fragments that were subcloned and used to generate deletion subclones for sequencing, and it shows the α-tubulin coding region, with the boxes representing the coding sequences. E, *Eco*RI site; H, *Hin*dIII site. The arrows at the bottom indicate the lengths of the DNA segments that were sequenced and the direction in which they were sequenced.

presented in Fig. 3. The coding region of the *A. thaliana* α-tubulin gene contains a 1353-base-pair (bp) coding region that is 72% identical to the coding sequences of the *C. reinhardtii* α-tubulin genes. The α-tubulin coding sequence is interrupted by four sequences presumed to be introns. All four introns contain consensus splice junctions (5′-$\frac{A}{T}$G/GT...CAG/$\frac{TT}{GG}$-3′) similar to the splice junctions of other eukaryotic genes. All of the putative introns are small, ranging in size from 77 to 111 bp. This is consistent with the small intron size in the alcohol dehydrogenase gene of *A. thaliana* (43) and with the small size of the *Arabidopsis* genome. The introns vary in G+C content from 34% to 39%, while the G+C content of the coding region is 47%.

Homology of the *A. thaliana* α-Tubulin to α-Tubulins of Animals and Protists. The predicted product of the sequenced *A. thaliana* α-tubulin gene is a polypeptide of 450 amino acids with a molecular weight of 49,660. This polypeptide exhibits a high degree of homology with the α-tubulins of other species (Fig. 4). The amino acid sequence of the *A. thaliana* α-tubulin is 90%, 87%, and 83% identical to the amino acid sequences of the α-tubulins of *C. reinhardtii* (34), *Stylonichia lemnae* (a ciliated protozoan; ref. 45), and *Homo sapiens* (44), respectively. Although the carboxyl termini of α-tubulins are highly divergent, *A. thaliana*, *C. reinhardtii*, and human α-tubulins all contain a carboxyl-terminal tyrosine.

S1 Nuclease Mapping of the Transcription Start Site. The 5′ end of the α-tubulin gene transcript was defined by S1 nuclease mapping (Fig. 5) using a 375-bp *Cla* I–*Bam*HI fragment (nucleotides 47 to 421; see Fig. 3). The 5′ untranslated region extends 99 ± 1 nucleotides upstream from the start of translation. The S1 nuclease data indicate that a secondary transcription start site may exist 7 bp upstream from the major transcription start site. The comparisons of predicted amino acid sequences (Fig. 4) indicate that translation starts at the first AUG downstream from the start of transcription, in agreement with the Kozak "scanning" model (46). A putative TATA box, TAAATAA, lies upstream of the transcription start point at position −34 to −28, and the sequence CGAATCT, which is similar to the consensus "CAAT box" GGCCAATCT (47), is present at position −122 to −116.

The *A. thaliana* α-Tubulin Gene Is Transcribed in Roots, Leaves, and Flowers. The S1 nuclease result shows that the sequenced α-tubulin gene is transcribed in tissue from whole plants. As a first step in investigating whether this α-tubulin gene is expressed constitutively or in a tissue-specific manner, we have determined the relative levels of tubulin transcripts in total RNA isolated from roots, leaves, and flowers. Our approach was to prepare two identical total RNA blots and to probe one with an α-tubulin coding sequence subclone that should hybridize to all α-tubulin mRNAs and to probe the other with a 5′ or 3′ noncoding sequence subclone that was shown to be gene specific by genomic blot analysis (data not shown). The 5′ gene-specific probe contained the DNA sequence from nucleotide 47 to nucleotide 371; the 3′ gene-specific probe carried the DNA segment from nucleotide 2167 to nucleotide 2274 (see Fig. 3). Both gene-specific probes yielded the same results (Fig. 6). As expected, the coding-sequence probe hybridized to RNA from all three tissues (Fig. 6, lanes 1–3). Since a single band was observed on the blot for the RNAs from all three tissues,

```
1     caatagttat ttaaaccaat aaaccctaat ttaacctata atatcgattt aggtcggtta
61    agctattatt tacctcactc atactaacga aatacctaaa ttccggtaag ttatttaaac
121   cggataaaaa tcgacacctt ccagatctag atccagatcg aatctcacgt attttgtcta
181   tgcccagtcc caaaggcagg ctgtcacaac ttggcggacc ccaaccaatt atcaaatctc
241   attcaataaa taatcaattt cccacaaaaa tcatggaaat taggaaaaaa agaagatctc
301   gctcaccgct ctcatttcct caattctccg tccgtcgaag aagcgaatcg ctcgaaatta
                          M  R  E  I     I  S  I     H  I  G     Q  A  G  I
361   gggtttctac tgagagaagA TGCGAGAAAT CATAAGCATT CACATCGGAC AAGCTGGGAT
       Q  V  G     N  S  C     W  E  L  Y     C  L  E     H  G  I     Q  P  D  G
421   CCAGGTCGGA AATTCGTGTT GGGAGCTTTA CTGTCTCGAA CATGGAATCC AGCCGGACGG
       M  M  P     S
481   AATGATGCCG AGgtataaac cctaattcac tctttgtact ctgtttctgt ggatcggaag
                                                   D  T     T  V  G     V  A  H
541   atttttgaga tacgagtctt ttgcgttgtt gcagTGATAC TACAGTTGGT GTTGCACACG
       D  A  F  N     T  F  F     S  E  T     G  A  G  K     H  V  P     R  A  V
601   ATGCGTTCAA TACTTTCTTT AGCGAGACTG GTGCTGGGAA GCATGTCCCT AGGGCTGTCT
       F  V  D  L     E  P  T     V  I  D     E  V  R  T     G  T  Y     R  Q  L
661   TCGTTGATCT CGAGCCTACC GTTATCGACG AAGTTCGTAC TGGTACTTAC CGTCAACTTT
       F  H  P  E     Q  L  I     S  G  K     E  D  A  A     N  N  F     A  R  G
721   TCCATCCAGA GCAGCTCATT TCTGGGAAAG AAGATGCTGC TAACAACTTC GCTAGAGGAC
       H  Y  T
781   ATTACACTGg tgagaatcca ctttctgctt gttctaactt gttagggttt ttgaaatggt
                                      V  G     K  E  I     V  D  L     C  L  D  R
841   ccttagaatg cttcttgtaa ttgcagTTGG GAAGGAAATT GTGGATCTAT GTCTTGACCG
       V  R  K     L  A  D     N  C  T  G     L  Q  G     F  L  V     F  N  A  V
901   TGTGAGGAAA CTTGCCGACA ACTGTACTGG CTTACAAGGG TTTTTGGTGT TCAATGCTGT
       G  G  G     T  G  S     G  L  G  S     L  L  L     E  R  L     S  V  D  Y
961   TGGTGGTGGA ACTGGTTCTG GATTGGGTTC TCTGTTGCTA GAGCGTTTGT CTGTAGATTA
       G  K  K     S  K  L     G  F  T  I     Y  P  S     P  Q
1021  TGGAAAGAAG TCTAAGCTTG GTTTTACCAT ATACCCTTCT CCTCAGgtac tctttctttt

1081  gctagacact gaagtagaaa ctcttacatg acatgtttct atagtggtct aatttggatg
                                                   V  S  T  A     V  V  E  P
1141  aaacttggtt ttaatgcata ttcctctgtg atgtcagGTT TCTACTGCTG TTGTAGAACC
       Y  N  S     V  L  S     T  H  S  L     L  E  H     T  D  V     A  V  L  L
1201  TTACAACAGT GTTCTCTCAA CGCATTCCCT TCTTGAACAT ACGGATGTAG CTGTCCTCTT
       D  N  E     A  I  Y     D  I  C  R     R  S  L     D  I  E     R  P  T  V
1261  GGATAACGAA GCCATCTATG ACATTTGCCG CAGATCCCTA GATATCGAGA GGCCAACCTA
       T  N  L     N  R  L     I  S  Q  I     I  S  S     L  T  T     S  L  R  F
1321  CACAAACTTG AACAGGTTGA TATCACAGAT CATTTCATCC TTGACAACAT CTTTGAGGTT
       D  G  A     I  N  V     D  I  T  E     F  Q  T     N  L  V     P  Y  P  R
1381  TGATGGTGCC ATCAACGTGG ATATCACTGA GTTCCAGACC AATCTTGTCC CATATCCCCG
       I  H  F     M  L  S     S  Y  A  P     V  I  S     A  A  K     A  Y  H  E
1441  TATCCATTTC ATGCTGTCAT CTTATGCACC AGTCATCTCA GCCGCCAAGG CTTACCACGA
       Q  L  S     V  P  E     I  T  N  A     V  F  E     P  A  S     M  M  A  K
1501  GCAGCTATCA GTCCCTGAGA TCACCAATGC CGTTTTTGAG CCAGCTAGCA TGATGGCAAA
       C  D  P     R  H  G     K  Y  M  A     C  C  L     M  Y  R     G  D  V  V
1561  GTGTGACCCG AGACACGGAA AGTACATGGC CTGTTGTTTG ATGTACCGAG GAGATGTTGT
       P  K  D     V  N  A     A  V  G  T     I  K  T     K  R  T     V  Q  F  V
1621  TCCCAAAGAT GTTAATGCTG CTGTTGGCAC CATCAAGACA AAGAGGACTG TTCAGTTTGT
       D  W
1681  TGACTGgtat gatctagaaa tggataaagc tgctactgtt tctaccttca tgttgttatt
                                                   C  P  T  G     F  K  C
1741  atataataac tcctttggca tgggttttgt tttgcagGTG CCCAACTGGG TTCAAATGTG
       G  I  N  Y     Q  P  P     T  V  V     P  G  G  D     L  A  K     V  Q  R
1801  GAATCAACTA CCAACCTCCA ACAGTTGTTC CAGGTGGTGA CCTCGCTAAG GTTCAGAGAG
       A  V  C  M     I  S  N     N  T  A     V  A  E  V     F  S  R     I  D  H
1861  CTGTATGTAT GATCAGTAAC AACACAGCAG TTGCAGAGGT GTTCTCACGG ATCGACCACA
       K  F  D  L     M  Y  A     K  R  A     F  V  H  W     Y  V  G     E  G  M
1921  AGTTTGATCT CATGTATGCG AAGAGGGCAT TTGTCCACTG GTACGTTGGT GAAGGAATGG
       E  E  G  E     F  S  E     A  R  E     D  L  A  A     L  E  K     D  Y  E
1981  AGGAAGGTGA ATTCTCTGAG GCACGTGAAG ACTTGGCCGC ACTGGAGAAA GATTACGAAG
       E  V  G  A     E  G  G     D  D  E     E  D  E  G     E  D  Y
2041  AAGTTGGTGC TGAAGGTGGA GACGATGAAG AAGATGAAGG TGAAGACTAT TGAcaccttt
2101  tctcaaaaaa actgcaacca aaaaaagttt tttcgtttat ctgttttcac ttttgattct
2161  gttgattgtt taaatcctgt aaaatattcg ttcggtcttc gtgttttatc tcggttggat
2221  acattttcat tggtagagcg accattgcga gttcccgagt tcattatcaa gctt
```

FIG. 3. Nucleotide sequence of the *A. thaliana* α-tubulin gene and the predicted amino acid sequence of the encoded α-tubulin. The amino acid sequence is shown in the one-letter code above the corresponding codons. The coding sequence is presented in uppercase letters. The putative "TATA box" is underlined and the start of transcription is indicated with an asterisk.

```
                                    30                                  60
A.thal.  MREIISIHIG QAGIQVGNSC WELYCLEHGI QPDGMMPSDT TVGVAHDAFN TFFSETGAGK
Chlamy.     V               A                Q    K  I GGD
human       C    V     V I  A                Q    K  I GGD S
Stylon.     V    V       I  A     F          Q    K  I GGD             E

                                    90                                 120
A.thal.  HVPRAVFVDL EPTVIDEVRT GTYRQLFHPE QLISGKEDAA NNFARGHYTV GKEIVDLCLD
Chlamy.     CI L       V                                    I        A
human                                       T         Y      I     I  V
Stylon.     C  L                                             I

                                   150                                 180
A.thal.  RVRKLADNCT GLQGFLVFNA VGGGTGSGLG SLLLERLSVD YGKKSKLGFT IYPSPQVSTA
Chlamy.   I                                                    V
human     I     Q           HS F       FT    M             E S    A
Stylon.   I     Q            S                                 V

                                   210                                 240
A.thal.  VVEPYNSVLS THSLLEHTDV AVLLDNEAIY DICRRSLDIE RPTYTNLNRL ISQIISSLTT
Chlamy.                        M                                A V     A
human          I T   TT   S C  FMV           N                   V  I A
Stylon.                       M     V       N                 A V     A

                                   270                                 300
A.thal.  SLRFDGAINV DITEFQTNLV PYPRIHFMLS SYAPVISAAK AYHEQLSVPE ITNAVFEPAS
Chlamy.         L                              I   E          A      A
human           L    L                 P A T       E          AD     C    N
Stylon.         L    V            SVIIRP  TP E-----          A     SA

                                   330                                 360
A.thal.  MMAKCDPRHG KYMACCLMYR GDVVPKDVNA AVGTIKTKRT VQFVDWCPTG FKCGINYQPP
Chlamy.    V                             S A        I
human    Q V               L              IA        I            V
Stylon.                                    A        I

                                   390                                 420
A.thal.  TVVPGGDLAK VQRAVCMISN NTAVAEVFSR IDHKFDLMYA KRAFVHWYVG EGMEEGEFSE
Chlamy.                        S  IG I    L
human                       L  T  I  AWA  L
Stylon.      S       M         S  I

                                   450
A.thal.  AREDLAALEK DYEEVGAEGG-DDEEDEGEDY.
Chlamy.             F      SAEGAG G   E
human       M            VDSVEGEG E   E
Stylon.  V               I IVEGEG E  ME.-
```

FIG. 4. Homology of the *A. thaliana* α-tubulin to other α-tubulins. The predicted amino acid sequence of the α-tubulin of *A. thaliana* is shown on the top line. The predicted amino acid sequences of α-tubulins of *Chlamydomonas reinhardtii* α-1 (34), cultured human keratinocytes (44), and *Stylonychia lemnae* (45) are given on the three lines below the *A. thaliana* sequence, showing only those amino acids that differ from the corresponding amino acids in the *A. thaliana* sequence. The dashes indicate the absence of amino acids corresponding to those in the *A. thaliana* sequence.

the mRNAs that hybridize with the α-tubulin coding sequence probe must have approximately the same size distribution in all of these tissues. The gene-specific probes also hybridized, and to approximately the same degree, to the RNAs from leaves, roots, and flowers (Fig. 6, lanes 4–6). Thus, the sequenced α-tubulin gene is transcribed in all three tissues.

DISCUSSION

The results presented in this paper demonstrate the existence in *A. thaliana* of an α-tubulin gene family containing at least four members. This α-tubulin gene family may include pseudogenes, which are well documented in mammalian tubulin gene families (48, 49). However, a tubulin gene family containing four or five functional genes would not be novel, since at least six functional α-tubulin genes exist in the mouse genome (6) and seven different β-tubulin genes function in chicken (12). We have cloned one of the α-tubulin genes from *A. thaliana* and determined its nucleotide sequence and the predicted amino acid sequence of its gene product. The *A. thaliana* α-tubulin shows strong homology to the α-tubulins of animals and protists. This homology indicates that the structural tubulin gene predates the plant–animal evolutionary split, more than one billion years ago.

Little and colleagues (50, 51) have examined partial proteolytic cleavage patterns of α-tubulins from 28 different sources. Their results indicate that plant and *Chlamydomonas* α-tubulins are very similar by this criterion. Our data

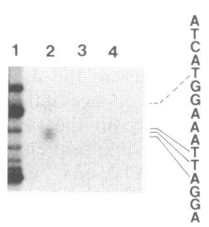

FIG. 5. S1 nuclease mapping of the transcription start site. Autoradiograph of a polyacrylamide sequencing gel showing the dC dideoxy sequencing lane for a known pUC119 ladder (lane 1, dG, dA, and dT lanes on the left not shown), the S1 nuclease digestion products of a reaction mixture containing 2.5 μg of *A. thaliana* whole plant poly(A)$^+$ RNA hybridized to an α-tubulin gene-specific probe (see text) and digested with S1 nuclease (100 units/ml) (lane 2), and the products of control reaction mixtures containing 2.5 μg of tRNA instead of poly(A)$^+$ RNA (lane 3) or no RNA or S1 nuclease (lane 4).

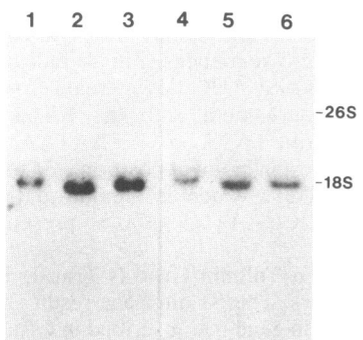

FIG. 6. RNA blot analyses. Total RNA isolated from roots (lanes 1 and 4), leaves (lanes 2 and 5), and flowers (lanes 3 and 6) was electrophoretically fractionated on a 1.2% agarose/formaldehyde gel. Lanes 1, 2, and 3 were probed with a primer-extended pUC119 subclone containing part of the α-tubulin coding sequence (the 1.2-kb *Hin*dIII fragment shown in Fig. 2). Lanes 4, 5, and 6 were probed with a primer-extended pUC119 subclone containing a DNA sequence specific to the 3′ region (segment from nucleotide 2167 to nucleotide 2274; see Fig. 3) of the cloned α-tubulin gene.

show that the *Arabidopsis* α-tubulin is indeed more closely related to the *Chlamydomonas* α-tubulins than to the α-tubulins of other protists and animals. Of the amino acid differences between the α-tubulins of *A. thaliana* and *C. reinhardtii*, over one-third (15/43) are of the very conservative isoleucine ↔ valine ↔ leucine type. Only 7 involve changes in charge at pH 7, and 4 of these are within the carboxyl-terminal 10 amino acids. Eight of the 10 carboxyl-terminal amino acids in the *A. thaliana* α-tubulin are aspartic acid or glutamic acid; this includes 4 more acidic residues than occur in the corresponding segment of the *C. reinhardtii* α-tubulin. Although the carboxyl-terminal regions of α-tubulins are highly divergent, the α-tubulins of *A. thaliana*, *C. reinhardtii*, and *H. sapiens* all contain a carboxyl-terminal tyrosine. Evidence from studies on African green monkey kidney TC-7 cells indicates that the carboxyl-terminal tyrosine of α-tubulins in these cells is involved in the establishment of separate populations of microtubules with distinct functions (52). Whether or not an analogous mechanism operates in plant cells is unknown.

On the basis of the conserved positions of five introns in the triosephosphate isomerase genes of maize and chicken, Marchionni and Gilbert (53) concluded that introns were in place before the divergence of the plant and animal kingdoms. The α-tubulin genes sequenced to date provide no evidence for conserved intron positions across kingdoms, although intron positions are conserved within kingdoms. The intron positions (codons interrupted) for sequenced α-tubulin genes are as follows: *A. thaliana*, 38, 109, 176–177, and 346; human (54) and rat (55), 1–2, 76, and 125–126; *C. reinhardtii*, 15–16 and 90 (34); and *Schizosaccharomyces pombe* α1, 19 (56). The α-tubulin genes of *Stylonychia lemnae* (45) and *Trypanosoma brucei* (57) and the *S. pombe* α2-tubulin gene (56) contain no introns.

The roles that multiple tubulin genes play in plant cells will probably deviate to some degree from their roles in animal cells since some animal tubulin genes are expressed primarily in the testes and brain. The cloned *A. thaliana* α-tubulin gene described in this paper was shown to be transcribed in roots, leaves, and flowers. Our present goal is to clone the other α-tubulin genes of *A. thaliana* and to characterize their patterns of expression during plant development. Whether plants, like animals, regulate tubulin genes in a tissue-specific manner remains to be determined.

We gratefully acknowledge the assistance of Elliot M. Meyerowitz, who provided the *A. thaliana* genomic library used in this work. This research was supported by U.S. Department of Agriculture Biotechnology Competitive Research Grant 85-CRCR-1-1754. D.G.O. was supported by a University of Minnesota Graduate School Doctoral Dissertation Fellowship. The Sun Microsystems 2/120 computer and the IntelliGenetics software used in this work were made available by the University of Minnesota Molecular Biology Computer Facility.

1. Lloyd, C. W., ed. (1981) *The Cytoskeleton in Plant Growth and Development* (Academic, New York).
2. Gunning, B. E. S. & Hardham, A. R. (1982) *Annu. Rev. Plant Physiol.* **33**, 651–698.
3. Tiwari, S. C., Wick, S. M., Williamson, R. E. & Gunning, B. E. S. (1984) *J. Cell Biol.* **99**, 63s–69s.
4. Raff, E. C. (1984) *J. Cell Biol.* **99**, 1–10.
5. Cleveland, D. W. & Sullivan, K. F. (1985) *Annu. Rev. Biochem.* **54**, 331–365.
6. Villasante, A., Wang, D., Dobner, P., Dolph, P., Lewis, S. A. & Cowan, N. J. (1986) *Mol. Cell. Biol.* **6**, 2409–2419.
7. Havercroft, J. C. & Cleveland, D. W. (1984) *J. Cell Biol.* **99**, 1927–1935.
8. Raff, E. C., Fuller, M. T., Kaufman, T. C., Kemphues, K. J., Rudolph, J. E. & Raff, R. A. (1982) *Cell* **28**, 33–40.
9. Sullivan, K. F., Machlin, P. S., Ratrie, H., III, & Cleveland, D. W. (1986) *J. Biol. Chem.* **261**, 13317–13322.
10. Bond, J. F., Robinson, G. S. & Farmer, S. R. (1984) *Mol. Cell. Biol.* **4**, 1313–1319.
11. Hall, J. L. & Cowan, N. J. (1985) *Nucleic Acids Res.* **13**, 207–223.
12. Sullivan, K. F. & Cleveland, D. W. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4327–4331.
13. Murphy, D. B. & Wallis, K. T. (1983) *J. Biol. Chem.* **258**, 7870–7875.
14. Oakley, B. R. & Morris, N. R. (1981) *Cell* **24**, 837–845.
15. Kemphues, K. J., Kaufman, T. C., Raff, R. A. & Raff, E. C. (1982) *Cell* **31**, 655–670.
16. Neff, N., Thomas, J. H., Grisafi, P. & Botstein, D. (1983) *Cell* **33**, 211–219.
17. Youngblom, J., Schloss, J. A. & Silflow, C. D. (1984) *Mol. Cell. Biol.* **4**, 2686–2696.
18. L'Hernault, S. W. & Rosenbaum, J. L. (1983) *J. Cell Biol.* **97**, 258–263.
19. Gard, D. L. & Kirschner, M. W. (1985) *J. Cell Biol.* **100**, 764–774.
20. Hussey, P. J. & Gull, K. (1985) *FEBS Lett.* **181**, 113–118.
21. Meyerowitz, E. M. & Pruitt, R. E. (1985) *Science* **229**, 1214–1218.
22. Estelle, M. A. & Somerville, C. R. (1986) *Trends Genet.* **2**, 89–92.
23. Rédei, G. P. (1970) *Biblio. Genet.* **20**, 1–150.
24. Leutwiler, L. S., Hough-Evans, B. R. & Meyerowitz, E. M. (1984) *Mol. Gen. Genet.* **194**, 15–23.
25. Pruitt, R. E. & Meyerowitz, E. M. (1986) *J. Mol. Biol.* **187**, 169–183.
26. Rédei, G. P. (1975) *Annu. Rev. Genet.* **9**, 111–127.
27. Lloyd, A. M., Barnason, A. R., Rogers, S. G., Byrne, M. C., Fraley, R. T. & Horsch, R. B. (1986) *Science* **234**, 464–466.
28. Murray, M. G. & Thompson, W. F. (1980) *Nucleic Acids Res.* **8**, 4321–4326.
29. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
30. Reed, K. C. & Mann, D. A. (1985) *Nucleic Acids Res.* **13**, 7207–7221.
31. Hu, N. T. & Messing, J. (1982) *Gene* **17**, 271–277.
32. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
33. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
34. Silflow, C. D., Chisholm, R. L., Conner, T. W. & Ranum, L. P. W. (1985) *Mol. Cell. Biol.* **5**, 2389–2398.
35. Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **26**, 101–106.
36. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13**, 31–40.
37. McMullen, M. D., Hunter, B., Phillips, R. L. & Rubenstein, I. (1986) *Nucleic Acids Res.* **12**, 4953–4968.
38. Messing, J. (1983) *Methods Enzymol.* **101**, 20–78.
39. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
40. Berry, J. D., Nikolau, B. J., Carr, J. P. & Klessig, D. F. (1985) *Mol. Cell. Biol.* **5**, 2238–2246.
41. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
42. Leutwiler, L. S., Meyerowitz, E. M. & Tobin, E. M. (1986) *Nucleic Acids Res.* **14**, 4051–4064.
43. Chang, C. & Meyerowitz, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1408–1412.
44. Cowan, N. J., Dobner, P. R., Fuchs, E. V. & Cleveland, D. W. (1983) *Mol. Cell. Biol.* **3**, 1738–1745.
45. Helftenbein, E. (1985) *Nucleic Acids Res.* **13**, 415–433.
46. Kozak, M. (1980) *Cell* **22**, 7–8.
47. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50**, 349–383.
48. Wilde, C. D., Crowther, C. E. & Cowan, N. J. (1982) *Science* **217**, 549–552.
49. Gwo-Shu Lee, M., Lewis, S. A., Wilde, C. D. & Cowan, N. J. (1983) *Cell* **33**, 477–487.
50. Little, M., Luduena, R. F., Langford, G. M., Asnes, C. F. & Farrell, K. (1981) *J. Mol. Biol.* **149**, 95–107.
51. Little, M., Luduena, R. F., Morejohn, L. C., Asnes, C. & Hoffman, E. (1984) *Origins Life* **13**, 169–176.
52. Gunderson, G. G., Kalnoski, M. H. & Bulinski, J. C. (1984) *Cell* **38**, 779–789.
53. Marchionni, M. & Gilbert, W. (1986) *Cell* **46**, 133–141.
54. Hall, J. L. & Cowan, N. J. (1985) *Nucleic Acids Res.* **13**, 207–223.
55. Lemischka, I. & Sharp, P. A. (1982) *Nature (London)* **300**, 330–335.
56. Toda, T., Adachi, Y., Hiraoka, Y. & Yanagida, M. (1984) *Cell* **37**, 233–242.
57. Thomashow, L. S., Milhausen, M., Rutter, W. J. & Agabian, N. (1983) *Cell* **32**, 35–43.