SHORT COMMUNICATION

Characterization of the Human *Talin* (*TLN*) Gene: Genomic Structure, Chromosomal Localization, and Expression Pattern

Tamar Ben-Yosef and Clair A. Francomano[1]

Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892

Received August 23, 1999; accepted October 7, 1999

Talin is a high-molecular-weight cytoskeletal protein, localized at cell–extracellular matrix associations known as focal contacts. In these regions, talin is thought to link integrin receptors to the actin cytoskeleton. Talin plays a key role in the assembly of actin filaments and in spreading and migration of various cell types. Talin proteins are found in a wide variety of organisms, from slime molds to humans. The human *Talin* (HGMW-approved symbol *TLN*) gene was previously mapped to chromosome 9p, but little was known of its sequence and genomic structure. To characterize human *TLN* further, we have isolated a single bacterial artificial chromosome clone, harboring the entire gene. The gene extends over more than 23 kb and consists of 57 exons. We have localized *TLN* to human chromosome band 9p13 by both fluorescence *in situ* hybridization and radiation hybrid mapping. Northern blot analysis detected *TLN* expression in various human tissues, including leukocytes, lung, placenta, liver, kidney, spleen, thymus, colon, skeletal muscle, and heart. Based on its chromosomal location, expression pattern, and protein function, we considered *TLN* as a candidate gene for cartilage-hair hypoplasia (CHH), an autosomal recessive metaphyseal chondrodysplasia, previously mapped to 9p13. We sequenced the entire *TLN* coding sequence in several CHH patients, but no functional mutations were detected.

Talin is a high-molecular-weight (~230 kDa) cytoskeletal protein (18). It is localized at focal contacts, specialized cell–cell and cell–extracellular matrix (ECM) associations that are characterized by the presence of filamentous actin at the cytoplasmic face of the junctional complex (2). In these regions, talin is thought to link integrins, a family of transmembrane ECM receptors, to the actin cytoskeleton (8). Talin seems to play a key role in the assembly of actin filaments (2) and is also essential for spreading and migration of various cell types, including fibroblasts (15) and osteoclasts (12). The talin protein N-terminus head contains elements required for localization of talin to cell–ECM junctions (16) and a membrane interaction site (5), while the C-terminus tail contains binding sites for several proteins, including β1-integrin, actin, and vinculin (1, 3, 7).

Talin proteins are found in a wide variety of organisms, from slime molds (11) to humans (4). Murine *Talin* cDNA (Accession No. X56123) is 8233 bp long, encoding a protein of 2541 amino acids. The human *Talin*[2] cDNA sequence was recently published (Accession No. AF078828), and the gene was mapped to human chromosome 9p (6), but little was known of its full sequence and genomic structure. To clone the human *TLN* gene, we screened a human bacterial artificial chromosome (BAC) library (Genome Systems, Inc., St. Louis, MO) using the following PCR primers designed from *TLN* cDNA sequence (Accession No. AF078828): 5′-GATGAGCTCAGTGACGGAAC-3′ and 5′-AAGACTTCATCCGAATGACC-3′; 960 bp; T_m 55°C. A single BAC clone, 280m21, harboring the entire *TLN* coding sequence was identified and partially sequenced using a shotgun sequencing strategy (21). Automated sequencing analysis was performed with dye-terminator chemistry (Perkin–Elmer/Applied Biosystems Division, Foster City, CA). BLAST searches with sequences derived from this BAC revealed that in addition to *TLN*, it harbors the genes *TPM2*, *Luman/LZIP*, and the last exon of the *MN/CA9* gene (Fig. 1A). Comparison of BAC and *TLN* cDNA sequences revealed that the *TLN* gene extends over more than 23 kb and consists of 57 exons and 56 introns (Fig. 1B). Exon boundaries completely conformed to the classical 5′-donor and 3′-acceptor consensus rules (GT . . . AG). The *TLN* coding region starts at base 34 of exon 2 and extends to the first 126 bp of exon 57 (Fig. 1B).

Careful examination of *TLN* genomic sequence dis-

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AF177198, AF178081, and AF178534.

[1] To whom correspondence should be addressed. Telephone: (301) 402-8255. Fax: (301) 496-7157. E-mail: clairf@nhgri.nih.gov.

[2] The HGMW-approved symbol for the gene described in this paper is *TLN*.

0888-7543/99 $30.00

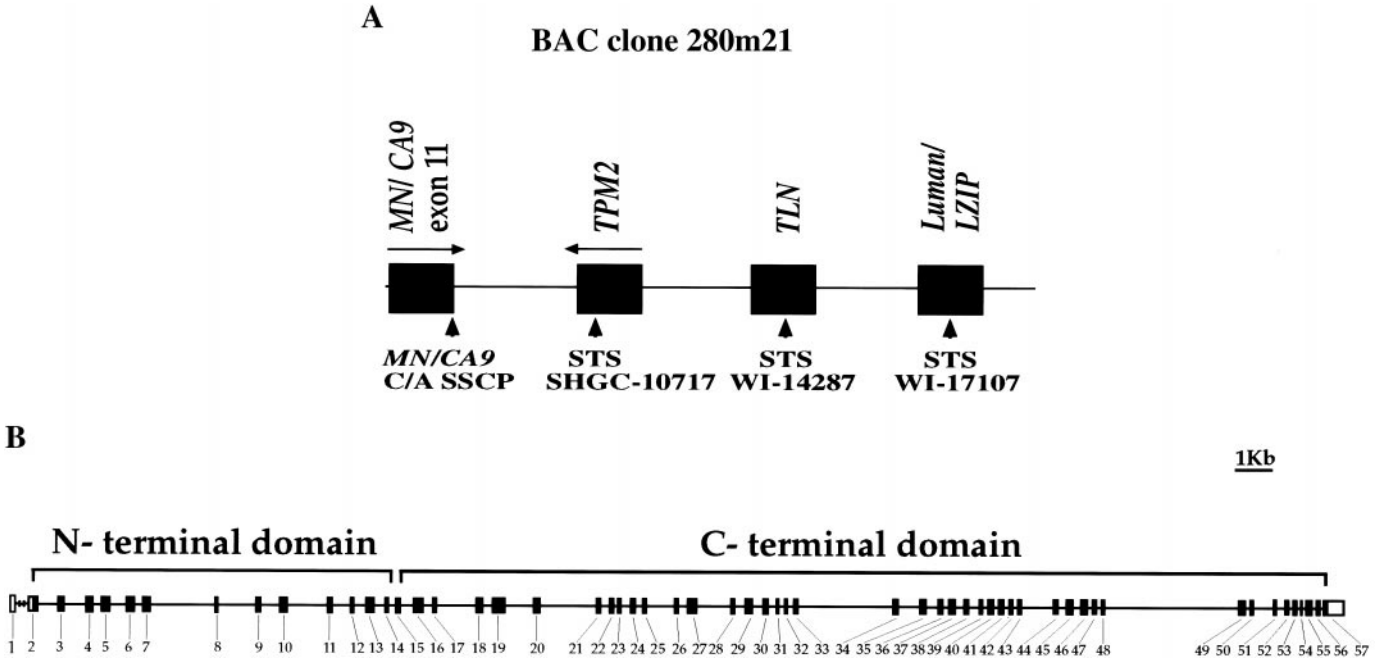

FIG. 1. (**A**) Genes, polymorphic markers, and sequence-tagged sites (STS) in the 280m21 BAC clone. Telomere on left, centromere on right. Relative locations of genes are represented by black boxes. Markers and STS are represented by arrowheads. Orientation of *MN/CA9* and *TPM2* genes is marked by arrows. Orientation of *TLN* and *Luman/LZIP* was not determined. (**B**) Genomic structure of the human *TLN* gene. Coding exons are represented by black boxes; noncoding exons are represented by white boxes. Gene regions coding the N- and C-terminal domains of the talin protein are indicated. Sequences of the exon/intron boundaries are available from the authors upon request.

closed some discrepancies with the published cDNA sequence (GenBank Accession No. AF078828), including the first 39 bp at the cDNA 5′ end and several base changes along the sequence. These were confirmed by sequencing of PCR products derived from the adjacent cDNA regions. cDNA was synthesized from fibroblast total RNA, using random hexamer as primer and SuperScript II Reverse Transcriptase (RT) (Gibco BRL, Gaithersburg, MD). cDNA samples were subjected to PCR amplification with specific *TLN* primers (Table 1). The revised cDNA sequence and the genomic sequence were deposited with GenBank (Accession Nos. AF177198, AF178081, and AF178534).

The human *TLN* gene was previously mapped to chromosome 9p (6). To refine its chromosomal localization further, the 280m21 BAC clone was nick-translated (17) and used for fluorescence *in situ* hybridization (FISH) on a metaphase preparation from

TABLE 1

Synthetic Oligonucleotide Primers Used for Amplification and Sequencing of *TLN* cDNA

Nucleotides[a]	5′-Primer sequence	3′-Primer sequence	Product size (bp)	Annealing temperature (°C)
48–663	5′-GCA GCG CAG GTA TCG CCA G-3′	5′-GCA GCA GCG TCT CGT GCT CC-3′	615	60
612–1091	5′-GAC CAT GGT CGG ACA CTG AG-3′	5′-CTT GGT GAT GCC CAG AAG CCT G-3′	479	55
965–1732	5′-GTC CGC TAC GTG AAG CTA GC-3′	5′-TCT ACC TGA GAG TGG ATC TC-3′	767	55
1413–2012	5′-GTG GAG CAC GGC TCT GTG GC-3′	5′-AGC TGC TTG CAG CAG GTT CTG-3′	599	60
1944–2691	5′-GTG TCG GAA CTG CTT CGC AG-3′	5′-CTG TGG CAT CAG CTA GGA TC-3′	747	55
2598–3285	5′-GTC AAT GCC ATC AAG GCT GAT G-3′	5′-CCT GTA GGT CTT TCT CTA G-3′	687	55
3234–3724	5′-GAG ATG GAT TCT GCA CTG AG-3′	5′-GCA TTA TCC ACA TCG CGC TG-3′	490	55
3503–4150	5′-GGC TGC GCT GAC ATC AGA TC-3′	5′-GTG CAC ATG GTG ATG AGC TG-3′	647	55
4009–4731	5′-GCA TCT CCA TGT CTT CAA GC-3′	5′-CAT CTA GCG CCT TGA TGG TC-3′	722	55
4447–5232	5′-ATA GCC AAG CTG GAC AGC AAG-3′	5′-GCT CAA TGA GAT GGG AGA TC-3′	785	60
5182–5931	5′-TGC ACA CTG AGA TGC TGA CTG-3′	5′-TCT TGG TGT AGG CAT CAC TG-3′	749	50
5752–6469	5′-AGC TGA CCA GTG ACT ATG GC-3′	5′-GAT GTC ACA TTG GTC ACC ATC-3′	717	55
6321–6941	5′-GTG GTG CTG ATC AAT GCA G-3′	5′-GAT GAG CTC AGT GAC GGA AC-3′	620	60
6607–7137	5′-AAG ACT TCA TCC GAA TGA CC-3′	5′-TGG CAG CTT CTA GTA TCT GC-3′	530	60
6996–7705	5′-GTC ATT GCT GAG AAT GAG CTC C-3′	5′-TCT CGA AGC TCT GAA GGC AG-3′	709	60

Note. The sequences of the synthetic oligonucleotides are tabulated, along with the product sizes and annealing temperatures used for each PCR.

[a] Nucleotide position in Accession No. AF177198.

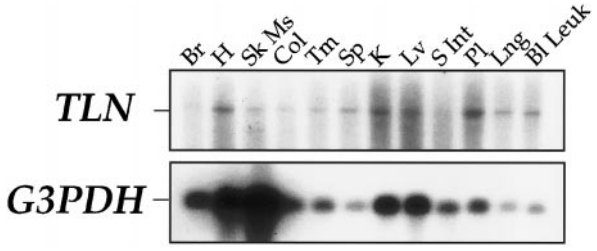

FIG. 2. *TLN* expression pattern in human tissues. A multiple-tissue Northern blot (Clontech, Palo Alto, CA) of poly(A)$^+$ mRNA extracted from 12 human tissues (peripheral blood leukocytes (**Bl Leuk**), lung (**Lng**), placenta (**Pl**), small intestine (**S Int**), liver (**Lv**), kidney (**K**), spleen (**Sp**), thymus (**Tm**), colon (**Col**), skeletal muscle (**Sk Ms**), heart (**H**), and brain (**Br**)) was hybridized with a ^{32}P-labeled 700-bp *TLN*-specific probe. *TLN* expression (8.2-kb transcript) was detected in most tissues examined. *G3PDH* served as an internal control.

peripheral blood (13). Signals were clearly visible, after antibody detection and inverted DAPI banding, as symmetrical spots on chromosomal bands 9p21–p13, on both copies of chromosome 9 (data reviewed but not shown). No specific signal was obtained in any other chromosomal region. For further assessment of *TLN* chromosomal localization, we screened 83 cell lines from the Stanford Human Genome medium-resolution (G3) human/rodent radiation hybrid (RH) panel, with the same PCR primers used for BAC library screening. Analysis with the RH mapper server (Stanford Human Genome Center: http://www-shgc.stanford.edu/RHindex.html) revealed that the gene is located on human chromosomal band 9p13, tightly linked to the marker SHGC-36744 (lod score = 12.38), between the markers D9S163 and D9S1874.

The *TLN* expression pattern in human tissues was analyzed by Northern blot hybridization. *TLN* expression was detected in most tissues examined (Fig. 2).

We mapped *TLN* to an interval on chromosomal region 9p13 in which the cartilage-hair hypoplasia (CHH) locus was previously located (19). CHH is an autosomal-recessive disorder, prevalent among the Old Order Amish in the United States and in the Finnish population, characterized by short stature, hypoplastic hair growth, T-cell-mediated immunodeficiency and additional hematological abnormalities, and increased risk of malignancies (14). *TLN* plays essential roles in both lymphocytes and bone (12). Moreover, Wistar Furth rats, which have a mutation in the *Tln* gene (9, 10), present a defect in platelet formation and an increased malignancy risk (10). We thus considered *TLN* as a candidate gene for CHH. To test this hypothesis, the entire *TLN* coding sequence of two unrelated CHH patients was sequenced. cDNA derived from cultured fibroblasts from the two patients served as a template for PCR with *TLN*-specific primer pairs (Table 1). Sequencing of the PCR products detected no functional mutations in the *TLN* coding sequence of these patients, while Northern and Western blot analyses indicated that both had normal levels of *TLN* mRNA and talin protein (data not shown), thus excluding the possibility of a mutation affecting a *TLN* promoter or a regulatory element, which would result in an altered expression level. Studies of Amish, Finnish, and sporadic CHH families demonstrated evidence for genetic homogeneity of CHH (20), suggesting that all CHH patients harbor mutations at the same locus, located on human chromosome 9p13. We thus conclude that *TLN* is not a likely candidate for CHH.

ACKNOWLEDGMENTS

We thank Dr. David Valle for the CHH cell lines, Dr. Amalia Dutra for assistance with FISH analysis, and Holly Root for assistance with RH mapping. This work was supported by the Division of Intramural Research (DIR) of the National Human Genome Research Institute.

REFERENCES

1. Bass, M. D., Smith, B. J., Prigent, S. A., and Critchley, D. R. (1999). Talin contains three similar vinculin-binding sites predicted to form an amphipathic helix. *Biochem. J.* **341:** 257–263.
2. Burridge, K., Fath, K., Kelly, T., Nuckolls, G., and Turner, C. (1988). Focal adhesions: Transmembrane junctions between the extracellular matrix and the cytoskeleton. *Annu. Rev. Cell Biol.* **4:** 487–525.
3. Burridge, K., and Mangeat, P. (1984). An interaction between vinculin and talin. *Nature* **308:** 744–746.
4. Collier, N. C., and Wang, K. (1982). Human platelet P235: A high M_r protein which restricts the length of actin filaments. *FEBS Lett.* **143:** 205–210.
5. Dietrich, C., Goldmann, W. H., Sackmann, E., and Isenberg, G. (1993). Interaction of NBD-talin with lipid monolayers. A film balance study. *FEBS Lett.* **324:** 37–40.
6. Gilmore, A. P., Ohanian, V., Spurr, N. K., and Critchley, D. R. (1995). Localisation of the human gene encoding the cytoskeletal protein talin to chromosome 9p. *Hum. Genet.* **96:** 221–224.
7. Gilmore, A. P., Wood, C., Ohanian, V., Jackson, P., Patel, B., Rees, D. J., Hynes, R. O., and Critchley, D. R. (1993). The cytoskeletal protein talin contains at least two distinct vinculin binding domains. *J. Cell Biol.* **122:** 337–347.
8. Hynes, R. O. (1992). Integrins: Versatility, modulation, and signaling in cell adhesion. *Cell* **69:** 11–25.
9. Jackson, C. W., Hutson, N. K., Steward, S. A., and Rees, D. J. G. (1993). Detection of a mutation in the cytoskeletal protein talin in the Wistar Furth rat: A rat strain with defective platelet formation and a high tumor incidence. *Blood* **82:** 340a.
10. Jackson, C. W., Hutson, N. K., Steward, S. A., and Stenberg, P. E. (1992). A unique talin antigenic determinant and anomalous megakaryocyte talin distribution associated with abnormal platelet formation in the Wistar Furth rat. *Blood* **79:** 1729–1737.
11. Kreitmeier, M., Gerisch, G., Heizer, C., and Muller-Taubenberger, A. (1995). A talin homologue of *Dictyostelium* rapidly assembles at the leading edge of cells in response to chemoattractant. *J. Cell Biol.* **129:** 179–188.
12. Lakkakorpi, P. T., and Vaananen, H. K. (1991). Kinetics of the osteoclast cytoskeleton during the resorption cycle in vitro. *J. Bone Miner. Res.* **6:** 817–826.
13. Lundsteen, C., and Lind, A. M. (1985). A test of a climate room for preparation of chromosome slides. *Clin. Genet.* **28:** 260–262.
14. McKusick, V., Eldridge, R., Hostetler, J., Ruangwit, U., and Egeland, J. (1965). Dwarfism in the Amish. II. Cartilage-hair hypoplasia. *Bull. Johns Hopkins Hosp.* **116:** 285–326.
15. Nuckolls, G. H., Romer, L. H., and Burridge, K. (1992). Microinjection of antibodies against talin inhibits the spreading and migration of fibroblasts. *J. Cell Sci.* **102:** 753–762.

16. Nuckolls, G. H., Turner, C. E., and Burridge, K. (1990). Functional studies of the domains of talin. *J. Cell Biol.* **110:** 1635–1644.
17. Pinkel, D., Gray, J. W., Trask, B., van den Engh, G., Fuscoe, J., and van Dekken, H. (1986). Cytogenetic analysis by in situ hybridization with fluorescently labeled nucleic acid probes. *Cold Spring Harbor Symp. Quant. Biol.* **51 Pt 1:** 151–157.
18. Rees, D. J., Ades, S. E., Singer, S. J., and Hynes, R. O. (1990). Sequence and domain structure of talin. *Nature* **347:** 685–689.
19. Sulisalo, T., Klockars, J., Makitie, O., Francomano, C. A., de la Chapelle, A., Kaitila, I., and Sistonen, P. (1994). High-resolution linkage-disequilibrium mapping of the cartilage-hair hypoplasia gene. *Am. J. Hum. Genet.* **55:** 937–945.
20. Sulisalo, T., van der Burgt, I., Rimoin, D. L., Bonaventure, J., Sillence, D., Campbell, J. B., Chitayat, D., Scott, C. I., de la Chapelle, A., Sistonen, P., and Kaitila, I. (1995). Genetic homogeneity of cartilage-hair hypoplasia. *Hum. Genet.* **95:** 157–160.
21. Wilson, R. K., and Mardis, E. R. (1997). Shotgun sequencing. *In* "Analyzing DNA" (B. Birren, E. D. Green, S. Klapholz, R. M. Myers, and J. Roskams, Eds.), pp. 397–454, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.