Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 6624–6627, August 1991
Genetics

Stop codon in the procollagen II gene (*COL2A1*) in a family with the Stickler syndrome (arthro-ophthalmopathy)

(type II collagen/vitreous gel/retinal detachment)

NILOFER NINA AHMAD*, LEENA ALA-KOKKO*, ROBERT G. KNOWLTON*, SERGIO A. JIMENEZ*†, ERIC J. WEAVER*, JOSEPH I. MAGUIRE‡, WILLIAM TASMAN‡, AND DARWIN J. PROCKOP*§

*Department of Biochemistry and Molecular Biology, Jefferson Institute of Molecular Medicine, and †Department of Medicine, Jefferson Medical College, Thomas Jefferson University, Philadelphia, PA 19107; and ‡Wills Eye Hospital, Philadelphia, PA 19107

Communicated by Victor A. McKusick, April 22, 1991

ABSTRACT **Linkage analysis with restriction fragment length polymorphisms for the gene for type II procollagen (*COL2A1*) was carried out in a family with the Stickler syndrome, or arthro-ophthalmopathy, an autosomal dominant disorder that affects the eyes, ears, joints, and skeleton. The analysis demonstrated linkage of the disease and *COL2A1* with a logarithm-of-odds score of 1.51 at zero recombination. A newly developed procedure for preparing cosmid clones was employed to isolate the allele for type II procollagen that was linked to the disease. Analysis of over 7000 nucleotides of the gene revealed a single base mutation that altered a CG dinucleotide and converted the codon CGA for arginine at amino acid position α1-732 to TGA, a stop codon. From previous work on procollagen biosynthesis, it is apparent that the truncated polypeptide synthesized from an allele with a stop codon at α1-732 cannot participate in the assembly of type II procollagen, and therefore that the mutation would decrease synthesis of type II procollagen. It was not apparent, however, why the mutation produced marked changes in the eye, which contains only small amounts of type II collagen, but relatively mild effects on the many cartilaginous structures of the body that are rich in the same protein.**

The Stickler syndrome, or arthro-ophthalmopathy (AO), is an autosomal dominant disorder that affects the eyes, ears, joints, and skeleton (1–4). The disorder is highly heterogenous but in many families it produces severe and progressive myopia with vitreal degeneration and retinal detachment, progressive sensorineural hearing loss, cleft palate, and mandibular hypoplasia. Also, it is frequently associated with joint changes that include hypermobility and hypomobility, epiphyseal dysplasia, and degenerative changes similar to osteoarthritis. The incidence of the disease is about 1 in 10,000 (4). Since type II collagen is the most abundant protein of cartilage and is also found in the vitreous gel (5), it has been suspected for some time that mutations in the gene for type II procollagen (*COL2A1*) are the cause of AO. Linkage analysis with restriction fragment length polymorphisms (RFLPs) provided statistical evidence for linkage of *COL2A1* with the syndrome in some but not all families (6–9). Here we identify an allele for type II procollagen that has cosegregated with AO in a large family and demonstrate that the allele has a premature stop codon for translation of the protein.¶

MATERIALS AND METHODS

Genetic Linkage Analysis. Genomic DNA isolated from peripheral blood from each family member was prepared and analyzed for *COL2A1* RFLPs in Southern blot hybridizations as described (8). Alleles of the variable-number tandem repeat at the 3′ end of *COL2A1* (11) were identified as PCR products (Fig. 1). Logarithm-of-odds (lod) scores were calculated with the LIPED program (12).

Isolation of Cosmid Clones and DNA Sequencing. The procedure for isolating cosmid clones has been described (13). Genomic DNA was isolated from 2×10^8 cultured human skin fibroblasts and the DNA was partially digested with restriction endonuclease *Sph* I (Boehringer Mannheim). The DNA was fractionated by gel electrophoresis. Fragments ranging from 25 to 35 kilobases (kb) in length were electroeluted from the gel and were cloned into the cosmid vector pJB8 (Amersham), previously engineered to receive *Sph* I–*Sph* I fragments of about 25–35 kb. Clones were screened by filter hybridization using a partial genomic clone of *COL2A1* as a probe. DNA from positive clones was isolated by cesium chloride centrifugation. The clones were analyzed by digestion with several restriction endonucleases and Southern blotting with probes for the gene. For subcloning, the cosmid DNA was digested with *Sph* I, and two fragments of 12 and 14 kb were isolated by agarose gel electrophoresis and electroelution. The two *Sph* I fragments were digested with *Eco*RI and the resulting fragments of 2–7 kb were isolated by agarose gel electrophoresis and then subcloned into the plasmid pUC19 (Boehringer Mannheim). The plasmid subclones were used for sequencing of double-stranded DNA by the dideoxynucleotide method with the use of specifically designed oligonucleotide primers based on the cDNA and genomic DNA sequences (14–16).

Allele-Specific Oligonucleotide Hybridization. Genomic DNA was used as template for the PCR (17). The 5′ primer was an 18-mer with a sequence corresponding to a sequence in intron 39 of *COL2A1*, and the 3′ primer was an 18-mer with a sequence complementary to a sequence in intron 40 (16) (Fig. 2). DNA obtained from the PCR was used for hybridization with allele-specific oligonucleotides (17). DNA blots were prehybridized for 2 hr at 65°C in 6× standard saline citrate (SSC)/5× Denhardt's solution/0.5% SDS containing denatured salmon sperm DNA at 0.25 mg/ml. Hybridization was performed at 37°C for 15 hr. The filters were washed three times in 2× SSC for 15 min at room temperature and then for 2 min at 56°C (ASO-normal probe) or at 57°C (ASO-mutant probe).

RESULTS

Linkage analysis in the family (Fig. 1) was carried out using the PCR (17) to detect polymorphisms in the hypervariable

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.

Abbreviations: AO, arthro-ophthalmopathy; RFLP, restriction fragment length polymorphism; lod, logarithm of odds.

§To whom reprint requests should be addressed.

¶A preliminary abstract was presented at the annual meeting of the American Society for Human Genetics, October 16, 1990, Cincinnati, OH (10).

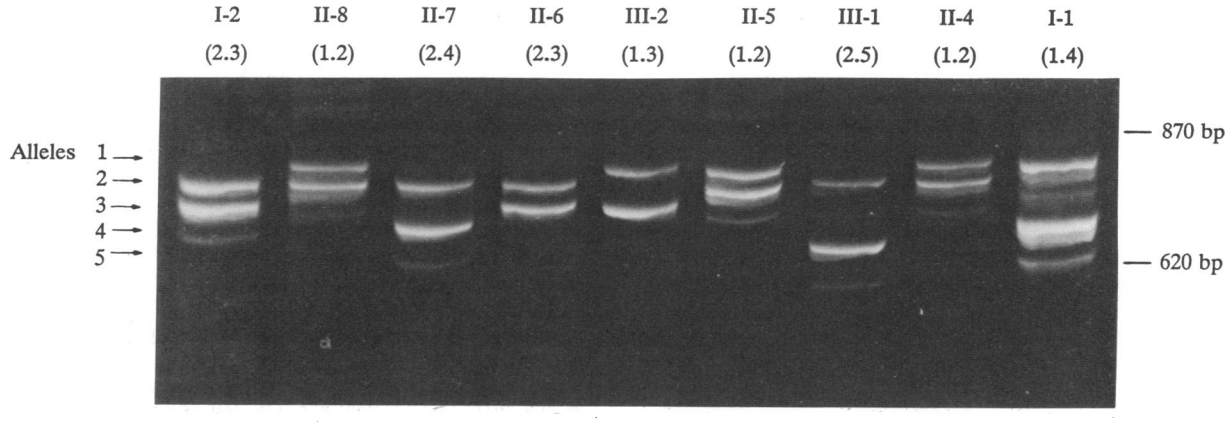

FIG. 1. Variable-number tandem repeat at the 3′ end of *COL2A1*. Numbers in parentheses indicate the five alleles defined by the PCR amplification. Family members (I-1 through III-2) correspond to those indicated in Fig. 2. The oligonucleotide primers used were 5′-CTCCAGGTTAAGGTTGACAGCT-3′ (forward primer) and 5′-CTTTGTCATGAACTAGCTCTGG-3′ (reverse primer). The PCR products were separated by electrophoresis in a 5% polyacrylamide gel and visualized with ethidium bromide. Conditions for the PCR were 30 cycles with 1.5 min at 90°C, 0.45 min at 55°C, and 1.5 min at 72°C. bp, Base pairs.

region at the 3′ end of *COL2A1* (11). Five polymorphic alleles were identified in the family (Fig. 1). Analysis of the data (12) demonstrated linkage of the disease and *COL2A1* with a lod score of 1.51 at zero recombination. Several RFLPs for *COL2A1* (8) were only partially informative but were consistent with linkage at zero recombination.

To sequence the type II procollagen gene, a newly developed procedure for preparing cosmid clones was employed (13). Clones of the gene containing the *Hin*dIII RFLP 14-kb allele (14) that was linked to the disease were digested with *Eco*RI or *Bam*HI, subcloned into pUC19, and sequenced using appropriate primers. Over 7000 nucleotides were analyzed. The data provided the sequences of exons 2–44. Therefore, they included 43 of the 54 exons and 38 of the 43 exons coding for the major triple-helical domain of the protein. They also included 40 bp or more of each of the flanking regions of each exon. With the one exception, all the sequences corresponded to the normal sequence of the gene or to neutral variants that did not change amino acid codons or consensus sequences for RNA splicing (14–16). The one exception was a single base mutation that altered a CG dinucleotide in exon 40 and converted the codon CGA for arginine at amino acid position α1-732 (15) to TGA, a stop codon (Fig. 3).

Genomic DNA from affected and unaffected members of the family was amplified with the PCR and hybridized with allele-specific oligonucleotides (18) containing the normal sequence for codons α1-729 to α1-735 or the same sequence with a single base mutation in the codon for α1-732 that created the stop codon. All affected members of the family had the mutation, whereas unaffected members did not (Fig. 2). Similar assays with genomic DNA from 57 unrelated individuals demonstrated that none had the mutation. Also, similar assays with genomic DNA demonstrated that the mutation was not present in affected members of four additional families with the Stickler syndrome in two of which there was genetic linkage to *COL2A1* with lod scores of 3.54 and 1.20 at zero recombination (8).

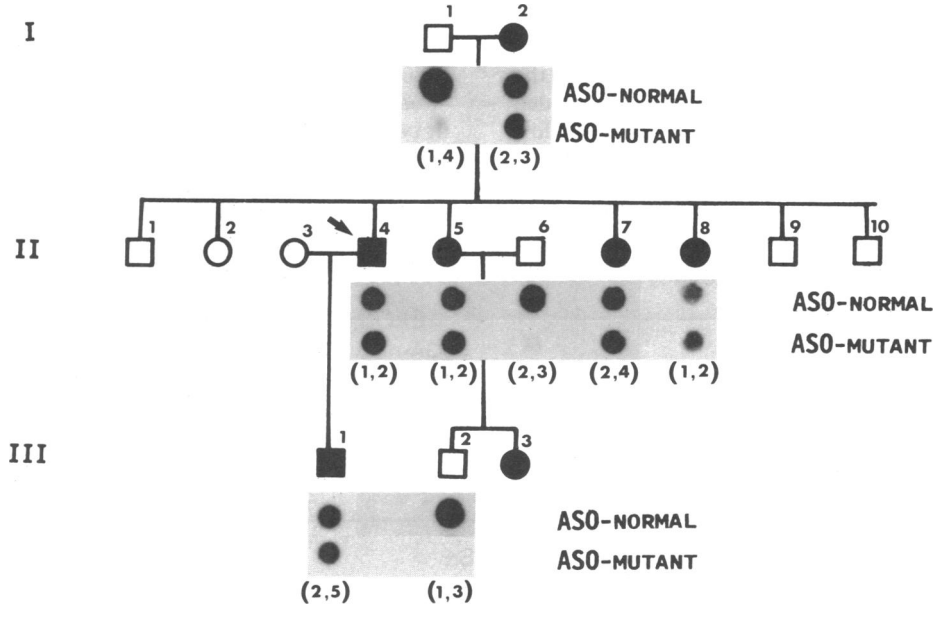

FIG. 2. Pedigree of the family with AO. Arrow indicates the proband. Numbers in parentheses indicate five alleles defined by analysis of PCR products from the hypervariable region at the 3′ end of *COL2A1* (11). Dots below the family members are blots of products of the PCR hybridized with allele-specific oligonucleotides (ASO) (18). The oligonucleotide primers for the PCR were 5′-AAACTGCTGTCACTGAG-3′ (forward primer) and 5′-CGTTCAGTTCTCATCTCC-3′ (reverse primer). The allele-specific oligonucleotides were 5′-AAAGGTGCTCGAG-GAGACA-3′ (normal allele) and 5′-AAAGGTGCTTGAGGAGACA-3′ (mutant allele).

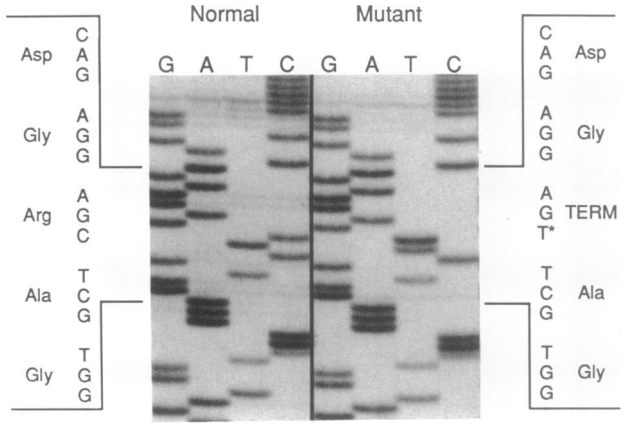

FIG. 3. DNA sequence of a normal allele from an unrelated individual and the allele from the proband that was linked to AO in the proband's family. The sequence shown is for codons of amino acids 730–734 of the α1(II) chain. There is a single base substitution of a T for a C that converts the codon for arginine α1-732 to a termination codon.

DISCUSSION

Type II collagen is a homotrimer that is similar in structure to heterotrimeric type I collagen and homotrimeric type III collagen (5). A series of mutations in the genes for type I and type III procollagens have been shown to produce heritable disorders of connective tissue (19–23). With a few exceptions, the mutations do not decrease expression of the gene. Instead, most of the mutations cause synthesis of an abnormal but partially functioning proα chain of the procollagen. The partially functioning proα chains interfere either with folding of the triple-helical domain of the protein, with processing of the procollagen to collagen, or with self-assembly of the collagen into fibrils (20, 21). Because both the folding of the collagen triple helix and the self-assembly of the protein into fibrils are highly cooperative processes, the deleterious effects of synthesis of an abnormal proα chain are greatly amplified. These generalizations have been shown to apply to >70 mutations in the two structural genes for type I procollagen (*COL1A1* and *COL1A2*) that produce either lethal or severely debilitating variants of osteogenesis imperfecta (19–21). They also apply to at least 6 mutations in the gene for type III procollagen (*COL3A1*) that cause sudden death because of rupture of large arteries and other hollow organs (see refs. 20, 22, and 23). The exceptions to the generalizations are informative, however, because they suggest that mutations decreasing expression of the same genes do not produce severe phenotypes. In particular, mutations decreasing expression of an allele for proα(I) chains were found in probands with type I variants of osteogenesis imperfecta, the mildest forms of the disease (24, 25), and a mutation causing synthesis of nonfunctional proα2(I) chains was found in asymptomatic parents of a child with osteogenesis imperfecta (26).

To date, only four mutations in the human gene for type II procollagen have been reported. One was a deletion of exon 48 (27) and another was an insertion of 45 bp into exon 48 (28) in two probands with spondyloepiphyseal dysplasia, a disease characterized by disproportionate dwarfism, myopia, and other changes in tissues rich in type II collagen. A third mutation was a substitution of serine for a glycine residue of exon 46 (29) in a proband with achondrogenesis, a usually lethal disorder of cartilage. The fourth mutation was a

Table 1. Clinical features of affected members of the family with AO

Family member (age in years)	Myopia (OD/OS)	Lattice degeneration (OD/OS)	Retinal detachment (OD/OS)	Arthropathy*
I-2 (54)	+/+	+/+	+/+	Moderate
II-4 (30)	+/+	+/+	+/+	Mild
II-5 (32)	+/+	+/+	+/+	Mild
II-7 (29)	+/+	+/+	+/−	Mild
II-8 (26)	+/+	+/+	+/+	Mild
II-1 (8)	+/+	+/+	+/+	None

Family members are identified as in Fig. 2. The proband (II-4) had sudden loss of vision in the left eye at the age of 13. In other family members, ocular changes were noted as early as age 5 (II-8) or as late as age 15 (II-5). OD, right eye; OS, left eye.

*The proband (II-4) complained of chronic neck stiffness. X-rays revealed loss of the normal cervical lordosis, reduction of the disc space of cervical vertebrae 5 and 6, scalloping of the anterior endplates of thoracic vertebrae 4 through 9 consistent with a developmental anomaly, and mild degenerative changes in the mid-dorsal segment. The mother (I-2) had Heberden's nodes, as well as pain and stiffness of many joints, especially the knees. X-ray photographs of the mother showed hypertrophic spurs of cervical vertebrae 4 to 7, osteoarthritic and osteoporotic changes throughout the mid-dorsal and lower dorsal spine, moderate degenerative changes with osteophyte formation in the knees, and moderate to severe degeneration of interphalangeal and carpal–metacarpal joints. In addition, the proband's mother had a congenitally shortened fourth metatarsal. She stated that her own mother had a similar deformity of the feet, together with loss of vision. The three other affected members of the proband's generation (II-5, II-7, and II-8) complained of chronic neck stiffness. X-rays revealed findings similar to those of the proband except that two (II-5 and II-8) also had anterior discal herniation in the lumbar spine. The proband's 8-year-old son (III-1) had no joint symptoms, but x-rays revealed flattening of cervical vertebrae 3 through 6. None of the affected family members had a Marfanoid habitus or apparent hearing loss. There was disagreement among several examiners as to whether affected family members had facial dysmorphism.

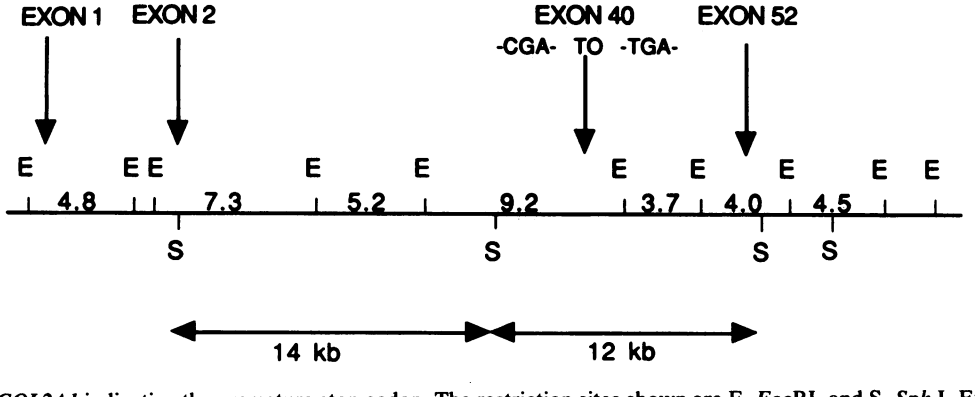

FIG. 4. Map of *COL2A1* indicating the premature stop codon. The restriction sites shown are E, *Eco*RI, and S, *Sph* I. Exons are numbered as suggested by Baldwin *et al.* (15). The map is taken from Sangiorgi *et al.* (32). Numbers indicate length in kilobases.

substitution of cysteine for arginine at α1-519 in a family with primary generalized osteoarthritis associated with a mild chondrodysplasia (30, 31). The mutation described here, in contrast, introduces a premature stop codon into a collagen gene (Fig. 4).

Cartilage can only rarely be obtained from patients, and the cells proliferate poorly and frequently dedifferentiate in culture. Therefore, it will be difficult to demonstrate directly the effects of the premature stop codon in *COL2A1*. However, considerable information is available about the biosynthesis of fibrillar procollagens such as type II (33). The folding of procollagen differs from the folding of other proteins in that it involves formation of a nucleus of triple helix near the COOH terminus of the molecule and then propagation of the triple-helical structure in a zipper-like manner from the COOH to the NH_2 terminus (21). However, the nucleus of triple helix cannot form until after the three proα chains associate and become disulfide-linked through their globular COOH-terminal propeptides. The stop codon at α1-732 defined here will cause synthesis of a truncated polypeptide that lacks a COOH-terminal propeptide and, therefore, cannot participate in the assembly of type II procollagen. A similar effect was seen with a frameshift mutation in a proband with osteogenesis imperfecta that altered the last 33 amino acids in the COOH-terminal propeptide of proα2(I) chains and thereby caused synthesis of proα2(I) chains that were not assembled into type I procollagen (see refs. 20 and 26). For these reasons, it is very likely that the premature stop codon at α1-732 would decrease synthesis of type II procollagen.

The mutation defined here produced early loss of vision because of retinal detachment but, surprisingly, only mild malformation or degeneration of cartilage (Table 1). Therefore, the expected reduction in synthesis of type II procollagen produced by the mutation apparently had more effect on the eye, which contains only small amounts of type II collagen (5, 34), than on the many cartilaginous structures in the body that are rich in the protein.

We thank Drs. Clinton Baldwin, Ian Hopkinson, and Deike Strobel for advice and help in the work.

1. Stickler, G. B., Belau, P. G., Farrell, F. J., Jones, J. D., Pugh, D. G., Steinberg, A. G. & Ward, L. E. (1965) *Mayo Clin. Proc.* **40,** 433–455.
2. Opitz, J. M., France, T., Herrmann, J. & Spranger, J. W. (1972) *N. Engl. J. Med.* **286,** 546–547.
3. Blair, N. P., Albert, D. M., Liberfarb, R. M. & Hirose, T. (1979) *Am. J. Ophthalmol.* **88,** 876–888.
4. Pyeritz, R. E. (1989) in *Arthritis and Allied Conditions: A Textbook of Rheumatology*, ed. McCarty, D. J. (Lea & Febiger, Philadelphia), p. 1343.
5. Kühn, K. (1987) in *Structure and Function of Collagen Types*, eds. Mayne, R. & Burgeson, R. E. (Academic, New York), pp. 1–42.
6. Francomano, C. A., Liberfarb, R. M., Hirose, T., Maumenee, I. H., Streeten, E. A., Meyers, D. A. & Pyeritz, R. E. (1987) *Genomics* **1,** 293–296.
7. Francomano, C. A., Rowan, B. G., Liberfarb, R. M., Hirose, T., Maumenee, I. H., Stoll, H. U. & Pyeritz, R. E. (1988) *Am. J. Hum. Genet.* **43,** A83 (abstr.).
8. Knowlton, R. G., Weaver, E. J., Struyk, A. F., Knobloch, W. H., King, R. A., Norris, K., Shamban, A., Uitto, J., Jimenez, S. A. & Prockop, D. J. (1989) *Am. J. Hum. Genet.* **45,** 681–688.
9. Fryer, A. E., Upadhyaya, M., Littler, M., Bacon, P., Watkins, D., Tsipouras, P. & Harper, P. S. (1990) *J. Med. Genet.* **27,** 91–93.
10. Ahmad, N. N., Ala-Kokko, L., Knowlton, R. G., Weaver, E. J., Maguire, J. I., Tasman, W. & Prockop, D. J. (1990) *Am. J. Hum. Genet.* **47,** A206 (abstr.).
11. Wu, S., Seino, S. & Bell, G. I. (1990) *Nucleic Acids Res.* **18,** 3102.
12. Ott, J. (1974) *Am. J. Hum. Genet.* **26,** 588–597.
13. Ala-Kokko, L. & Prockop, D. J. (1990) *Matrix. Col. Rel. Res.* **10,** 279–284.
14. Nunez, A. M., Kohno, K., Martin, G. R. & Yamada, Y. (1985) *Biochemistry* **24,** 6343–6348.
15. Baldwin, C. T., Reginato, A. M., Smith, C., Jimenez, S. A. & Prockop, D. J. (1989) *Biochem. J.* **262,** 521–528.
16. Ala-Kokko, L. & Prockop, D. J. (1990) *Genomics* **8,** 454–460.
17. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230,** 1350–1354.
18. Studencki, A. B. & Wallace, R. B. (1984) *DNA* **3,** 7–15.
19. Byers, P. H., Bonadio, J. F., Cohn, D. H., Starman, B. J., Wenstrup, R. J. & Willing, N. C. (1988) *Ann. N.Y. Acad. Sci.* **543,** 117–128.
20. Kuivaniemi, H., Tromp, G. & Prockop, D. J. (1991) *FASEB J.* **5,** 2052–2060.
21. Prockop, D. J. (1990) *J. Biol. Chem.* **265,** 15349–15353.
22. Superti-Furga, A., Gugler, E., Gitzelmann, R. & Steinmann, B. (1988) *J. Biol. Chem.* **263,** 6226–6232.
23. Tromp, G., Kuivaniemi, H., Stolle, C., Pope, F. M. & Prockop, D. J. (1989) *J. Biol. Chem.* **264,** 19313–19317.
24. Wenstrup, R. J., Willing, M. C., Starman, B. J. & Byers, P. H. (1990) *Am. J. Hum. Genet.* **46,** 975–982.
25. Willing, M. C., Cohn, D. H. & Byers, P. H. (1990) *J. Clin. Invest.* **85,** 282–290.
26. Dickson, L. A., Pihlajaniemi, T., Deak, S., Pope, F. M., Nicholls, A., Prockop, D. J. & Myers, J. C. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4524–4528.
27. Lee, B., Vissing, H., Ramirez, F., Rogers, D. & Rimoin, D. (1989) *Science* **244,** 978–980.
28. Tiller, G., Rimoin, D. L., Murray, L. W. & Cohn, D. H. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 3889–3893.
29. Vissing, H., D'Alessio, M., Lee, B., Ramirez, F., Godfrey, M. & Hollister, D. W. (1989) *J. Biol. Chem.* **264,** 18265–18267.
30. Knowlton, R. G., Katzenstein, P. L., Moskowitz, R. W., Weaver, E. J., Malemud, C. J., Pathria, N., Jimenez, S. A. & Prockop, D. J. (1990) *N. Engl. J. Med.* **322,** 526–530.
31. Ala-Kokko, L., Baldwin, C. T., Moskowitz, R. W. & Prockop, D. J. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6565–6568.
32. Sangiorgi, F. O., Benson-Chanda, V., de Wet, W. J., Sobel, M. E., Tsipouras, P. & Ramirez, F. (1985) *Nucleic Acids Res.* **13,** 2207–2225.
33. Prockop, D. J. & Kivirikko, K. I. (1984) *N. Engl. J. Med.* **311,** 376–386.
34. Thorogood, P., Bee, J. & van der Mark, K. (1986) *Dev. Biol.* **116,** 497–509.