Am. J. Hum. Genet. 42:435–445, 1988

Cloning and Chromosomal Localization of Human Genes Encoding the Three Chains of Type VI Collagen

Dominique Weil,* Marie-Geneviève Mattei,† Edith Passage,† N'Guyen Van Cong,* Dorothy Pribula-Conway,‡ Karlheinz Mann,§ Rainer Deutzmann,§ Rupert Timpl,§ and Mon-Li Chu‡

*Unité de Recherches de Génétique Médicale, INSERM U12, Paris; †Centre de Génétique Médicale, INSERM U242, Marseille; ‡Department of Biochemistry and Molecular Biology, Jefferson Institute of Molecular Medicine, Thomas Jefferson University, Philadelphia; and §Max-Planck-Institut für Biochemie, Martinsried, Federal Republic of Germany

Summary

Type VI collagen is a heterotrimer composed of three polypeptide chains, α1(VI), α2(VI), and α3(VI). By immunological screening of an expression cDNA library, human cDNAs specific for each chain were isolated and characterized. Major mRNA species encoding these chains have a size of 4.2 kb (α1), 3.5 kb (α2), and 8.5 kb (α3). The cDNA clones were also used to map the genes on human chromosomes by somatic cell hybrid analysis and in situ hybridization. The α1(VI) and α2(VI) collagen genes were both located on chromosome 21, in band q223. This represents a third example of a possible physical proximity of two collagen loci. The α3(VI) collagen gene was localized to chromosome 2, in the region 2q37. The α3(VI) collagen gene is the fifth extracellular matrix gene to be localized to 2q, as four other extracellular matrix genes—i.e., the α1(III) and α2(V) collagen genes, the elastin gene, and the fibronectin gene—have been previously mapped to the distal region of the long arm of chromosome 2.

Introduction

Collagens are a large family of either heterotrimeric or homotrimeric triple-helical proteins present in the extracellular matrices of vertebrates and invertebrates. More than 10 genetically distinct types of collagen have been identified so far (Martin et al. 1985). In recent years, considerable information has been accumulated on the primary structure and genomic organization of several members of this gene family. Those most intensively studied are the fiber-forming collagens, types I, II, and III, which are highly conserved in both their protein and gene structure (Boedtker et al. 1985; de Crombrugghe et al. 1985; Ramirez et al. 1985; Upholt et al. 1985). Closely related is the type V collagen gene, as suggested by recent studies (Myers et al. 1985; Weil et al. 1987). A characteristic feature of this collagen group is a long triple-helical domain with repeating Gly-Xaa-Yaa sequences encoded by short exons consisting of 54 bp or its variations. Other members of this protein family, such as basement-membrane collagen type IV (Kuhn et al. 1985), as well as the cartilage-specific short-chain collagen types IX and X (Olsen et al. 1985), differ considerably in their attachment and macromolecular assembly pattern and show frequent interruptions in the Gly-Xaa-Yaa domain. In addition, their intron/exon arrangement bears little resemblance to that of the fiber-forming collagens (Kurkinen et al. 1985; Lozano et al. 1985; Ninomiya et al. 1986; Sakurai et al. 1986; Soininen et al. 1986).

Type VI collagen is another component of this protein family that is still incompletely characterized at both the structural and genomic level. It is composed of three polypeptides, named α1(VI), α2(VI), and α3(VI) chains (for review, see Timpl and Engel

Received July 27, 1987; revision received October 14, 1987.

Address for correspondence and reprints: Dr. Mon-Li Chu, Department of Biochemistry and Molecular Biology, Thomas Jefferson University, 1020 Locust Street, Philadelphia, PA 19107.

0002-9297/88/4203-0005$02.00

1987). Each chain contains a short triple-helical domain of ~40 kD, which can be isolated after pepsin digestion (Jander et al. 1983). The original chains are, however, much larger, with molecular masses of 140 kD for the α1 and α2 chains and, more strikingly, 250 kD for the α3 chain (Jander et al. 1984; Heller-Harrison and Carter 1984; Hessle and Engvall 1984; Knight et al. 1984; Trüeb and Bornstein 1984; Von der Mark et al. 1984; Gibson and Cleary 1985; Engvall et al. 1986; Trüeb and Winterhalter 1986). Thus, more than two-thirds of each chain consists of globular domains. The unusual proportion of globular and triple-helical domains makes collagen VI a unique member of the collagen family. It is widely distributed in the body and is a component of microfibrils in tissues (Von der Mark et al. 1984; Bruns et al. 1986). The level of collagen VI mRNA is greatly reduced in viral transformed fibroblasts (Trüeb et al. 1985).

Being a ubiquitous structural protein, collagen VI is likely to be involved in several important biological functions. Yet, except for promoting cell attachment in vitro (Carter 1982), which indicates a role as a cell-binding protein, little is known about its other functions. Our understanding of the functional repertoire of collagenous proteins has recently benefited from the molecular characterization of inherited human connective-tissue diseases, which demonstrated the occurrence of several mutations in the collagen I genes in patients with osteogenesis imperfecta and Ehlers-Danlos syndromes (Prockop and Kivirikko 1984; Tsipouras and Ramirez 1987). We adopted a similar approach and have isolated and identified cDNAs coding for α1(VI), α2(VI), and α3(VI) chains as a first step for studying both regulation of collagen VI expression and its possible role in human diseases. These probes were used to accomplish the chromosomal location of these three collagen VI genes.

Material and Methods

Isolation and Identification of cDNA Clones

An expression-λgt11 library from human placenta (Clontech, Palo Alto, CA) was screened with affinity-purified antibodies against collagen VI chains (Von der Mark et al. 1984), and several cDNA clones were plaque-purified. The cDNA inserts (size range 0.5–2.5 kb) were subcloned into the *Eco*RI site of plasmid puc19 (Norrander et al. 1983) and used for the determination of nucleotide sequences of the 5′ and 3′ ends by the dideoxy chain–terminating method (Chen and Seeburg 1985).

Peptide sequences were generated from pepsin-solubilized collagen VI that was isolated from human placenta (Odermatt et al. 1983). The peptide chains of reduced and alkylated material were bound to a Mono Q HR5/5 column (Pharmacia, Uppsala) equilibrated in 4 M urea, 0.02 M Tris-HCl, pH 8.0, and were eluted from the column with a linear NaCl gradient. This allowed the separation of three different chains corresponding to fragments of the α1(VI), α2(VI), and α3(VI) chain (Jander et al. 1983) according to their electrophoretic mobility, their amino acid composition, and high-performance-liquid-chromatography profiles of their tryptic peptides. Partial amino acid sequences were determined for these chains and selected tryptic peptides (0.2–1 nmol) in a gas-phase sequencer (model 470A; Applied Biosystems) or spinning-cup liquid-phase sequencer (model 890C; Beckman) following the manufacturer's protocols.

Northern Blot Hybridization

Total poly(A)$^+$ RNA was prepared from cultured human fibroblasts (GM3349; Institute for Medical Research, Camden, NJ) by the guanidine-thiocyanate method (Chirgwin et al. 1979) and oligo(dT) cellulose chromatography (Aviv and Leder 1972). The RNA was electrophoresed in 1% formaldehyde gels (Lehrach et al. 1977), transferred to nitrocellulose membranes (Thomas 1980), and hybridized with ^{32}P-cDNA probes labeled by nick-translation (Rigby et al. 1977). Hybridization and washing were performed as described elsewhere (Chu et al. 1982).

Hybrid Cell Lines

A total of 11 human-mouse hybrids and 24 human–Chinese hamster hybrids were prepared from fusions between rodent cell lines and several different human fibroblasts or lymphocytes as described elsewhere (Weil et al. 1980; N'Guyen et al. 1986). The hybrids were characterized by karyotype and isoenzyme analysis.

DNA Isolation and Southern Blot Analysis

High-molecular-weight DNA was isolated from each of the hybrids by using a procedure described elsewhere (Maniatis et al. 1982). Ten micrograms of DNA was digested with restriction enzyme *Taq*I under the condition recommended by the manufacturer (New England Biolabs, Beverly, MA). The di-

gested DNA samples were separated by electrophoresis on 0.7% agarose gels, transferred to nylon membrane by the method of Southern (1975), and hybridized with radiolabeled cDNA probes. The probes were labeled with [α-^{32}P]dCTP by nick-translation (Rigby et al. 1977) to a specific activity of ~5 × 10^8 cpm/μg. Filters were prehybridized at 42 C for 4 h in 50% formamide, 5 × SSC, 5 × Denhardt's solution, and 100 μg denatured salmon-sperm DNA/ml. Hybridization was carried out at 42 C for 20 h in fresh prehybridization solution containing labeled probe (2 × 10^6 cpm/ml). The filters were then rinsed in 2 × SSC at 25 C and washed in 2 × SSC, 0.1% SDS for 10 min each.

In Situ Hybridization

In situ hybridization was performed as described elsewhere (Mattei et al. 1985). Human metaphase chromosomes were prepared from phytohemagglutinin-stimulated lymphocytes of a normal male individual. The cDNA probes were labeled with [^{3}H]dCTP and [^{3}H]dTTP by nick-translation to a specific activity of ~5 × 10^7 cpm/μg and then hybridized to metaphase spreads at a final concentration of 500 ng/ml of hybridization solution. The slides were washed, coated with nuclear track emulsion (Kodak NTB2), and exposed for 11–18 days at 4 C. Following development, chromosome spreads were first stained with buffered Giemsa solution, and then metaphases were photographed. R-banding was then performed using the fluorochrome-photolysis-Giemsa method, and metaphases were rephotographed before analysis.

Results

Properties of cDNA Clones Specific for α1, α2, and α3 Chains of Human Collagen Type VI

Several cDNA clones were isolated from an expression–human placenta library with the aid of antibodies against collagen VI chains. Some of these clones showed a typical fragment pattern after cleavage with *Sau*96 (Lehrach et al. 1978), indicating the presence of multiple Gly-Pro sequences, which are characteristic for the collagen triple helix. This was confirmed by determination of partial DNA sequences, which identified repeating triplet sequences of the type Gly-Xaa-Yaa (fig. 1). Each of these clones was identified by matching peptide sequences, demonstrating that clone P18 (2.1 kb) encodes the α1(VI) chain, clone P8 (1.4 kb) the α2(VI) chain, and clone P24 (1.5 kb) the α3(VI) chain (fig. 1).

Distinct differences were also detected between these clones by Northern hybridization to human fibroblast mRNA (fig. 2). They reacted with a major mRNA species of either 4.2 kb (α1), 3.5 kb (α2), or

clone P18 (α1(VI) chain):

```
1   GGG GAG CGG GGT GGC CCT GGA GAG AGA GGA CCA CGG GGG ACC CCA GGC ACG CGG GGA CCA AGA GGA GAC CCT GGT GAA GCT GGC CCG CAG   90
    G  E  R  G  G  P  G  E  R  G  P  R  G  T  P  G  T  R  G  P  R  G  D  P  G  E  A  G  P  Q

91  GGT GAT CAG GGA AGA GAA GGG CCC GTT GGT GTC CCT GGA GAC CCG GGC GAG GCT GGC CCT ATC GGA CCT AAA GGC TAC CGA GGC GAT GAG  180
    G  D  Q  G  R  E  G  P  V  G  V  P  G  D  P  G  E  A  G  P  I  G  P  K  G  Y  R  G  D  E
```

clone P8 (α2(VI) chain):

```
1   GGT CGC AAG GGG GCC CCT GGC CTG GCT GGC AAG AAC GGG ACC GAT GGA CAG AAG GGC AAG CTG GGG CGC ATC GGA CCT CCT GGC TGC AAG   90
    G  R  K  G  A  P  G  L  A  G  K  N  G  T  D  G  Q  K  G  K  L  G  R  I  G  P  P  G  C  K

91  GGA GAC CCT GGA AAC CGG GGC CCC GAC GGT TAC CCG GGG GAA                                                                132
    G  D  P  G  N  R  G  P  D  G  Y  P  G  E
```

clone P24 (α3(VI) chain):

```
1   GGA GTA AAG GGC TCT CGG GGA TTC CCA GGA GAG AAG GGC GAA GTA GGA GAA ATT GGA CTG GAT GGT CTG GAT GGT GAA GAT GGA GAC AAA   90
    G  V  K  G  S  R  G  F  P  G  E  K  G  E  V  G  E  I  G  L  D  G  L  D  G  E  D  G  D  K

91  GGA TTG CCT GGT TCT TCT GGA GAG AAA GGG AAT CCT GGA AGA AGG GGT GAT AAA GGA CCT CGA GGA GAG AAA GGA GAA AGA GGA GAT GTT  180
    G  L  P  G  S  S  G  E  K  G  N  P  G  R  R  G  D  K  G  P  R  G  E  K  G  E  R  G  D  V
```

Figure 1 Sequence analysis of three cDNA clones encoding the α1(VI), α2(VI), and α3(VI) chains. Nucleotide sequences from the 5′ end of the cDNA clones (first line) and deduced amino acid sequences (second line) in the one-letter code were compared with peptide sequences (underlined residues) determined by Edman degradation. Numbers on both sides indicate nucleotide positions. A full description of the clones will be described elsewhere (Chu et al. 1987).

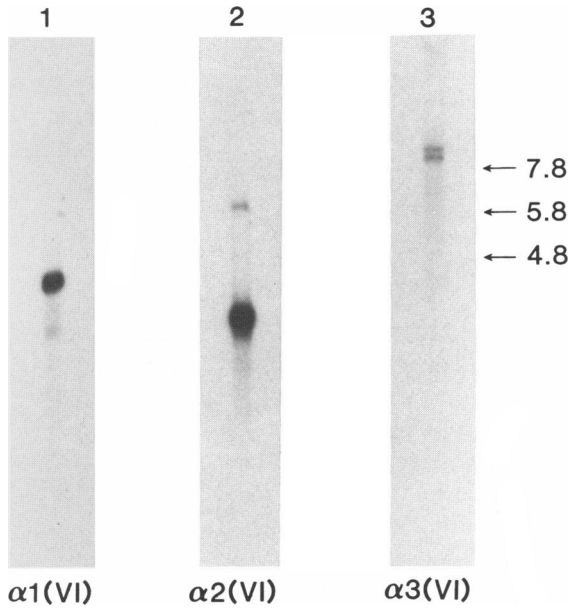

Figure 2 Northern blot analysis of polyadenylated RNA isolated from human fibroblasts with cDNA clones coding for the α1(VI) (lane 1), α2(VI) (lane 2), and α3(VI) (lane 3) chains. Each lane contained 1 μg of RNA. The location of the fibronectin mRNA (7.8 kb) (Bernard et al. 1985) and of the two polymorphic transcripts of α1(I) collagen mRNA (5.8 and 4.8 kb) (Chu et al. 1985) is indicated.

8.5 kb (α3). Some minor bands also hybridized in a characteristic fashion to each specific probe; this may be due to a different length in the 3′-noncoding region or may reflect splicing variants. Since no overlap in the hybridization patterns was observed, the clones appeared suitable for the chromosomal mapping of collagen VI chains.

Chromosomal Mapping Using Somatic Cell Hybrids

Alpha 1(VI) collagen gene.—Hybridization of *Taq*I-digested human genomic DNA to ^{32}P-labeled α1(VI) probe revealed four major bands, ~6.0, 4.4, 4.1, and 1.8 kb in size (fig. 3*A*). Although several distinct cross-hybridizing bands were seen in the rodent DNA with the same probe, the 6.0-, 4.4-, and 4.1-kb bands were unique to the human DNA. The 6.0- and 4.4-kb bands were present in the human-mouse hybrids containing human chromosome 21, whereas either the 6.0- and 4.4-kb bands or the 6.0- and 4.1-kb bands were detected in the human–Chinese hamster hybrids containing human chromosome 21. All three human-specific bands were absent in hybrids lacking human chromosome 21. One exception was hybrid CH 106 IV, where the human-specific bands were detected but chromosome 21 was missing. It was determined that, owing to chromosomal breakage, only part of chromosome 21 was present in this hybrid. Karyotyping could not detect this breakage, but superoxide dismutase (SOD_1), an isozyme already located on chromosome 21, could be detected.

These results strongly support the localization of α1(VI) on chromosome 21. In addition, the data on Southern blotting of human DNA suggested a *Taq*I RFLP. The control human DNA shown in figure 3 is heterozygous for the *Taq*I polymorphism, since both 4.4- and 4.1-kb bands were detected. The hybrid cells positive for α1(VI) can exhibit one or the other band, depending on the presence of the polymorphic site in the cell line used as human parent and on which chromosome 21 (the one bearing the polymorphism or the other) has been retained.

Alpha 2(VI) collagen gene.—When an α2(VI) cDNA probe was used, four *Taq*I fragments, of 4.2, 3.7, 2.7, and 2.4 kb, were detected in human DNA (fig. 3*B*). Two Chinese hamster–specific fragments of 3.7 and 1.8 kb and two mouse-specific fragments of 3.4 and 2.8 kb were present in the corresponding hybrids. Analysis of the chromosome content of all 35 hybrids suggests that the α2(VI) collagen gene also may be located on chromosome 21 (table 1). Again, as noted above in the assignment of α1(VI) gene, a single discordant hybrid (CH 106 VI) was observed. As for the α1(VI) gene, all the other chromosomes could be excluded.

Alpha 3(VI) collagen gene.—When hybridized to human DNA digested with *Taq*I, the α3(VI) cDNA probe detected six fragments, of 5.6, 3.0, 2.5, 2.3, 2.0, and 1.8 kb (fig. 3*C*). The rodent DNA cross-hybridized very strongly with the human probe. However, these rodent-specific bands were clearly different from the human pattern. The presence of the human pattern in the hybrid cell lines is concordant with the presence of human chromosome 2 (table 1). Further regional mapping of the α3(VI) gene was possible because a panel of hybrids was available from a human parental cell strain that carries a translocation of chromosome 2. Examination of the hybrids with derivative chromosome 2 showed no discordancy with 2q321-2qter, whereas 36% discordancy was found for hybrids with 2pter-2q321 (table 1). Therefore, the results suggested strongly

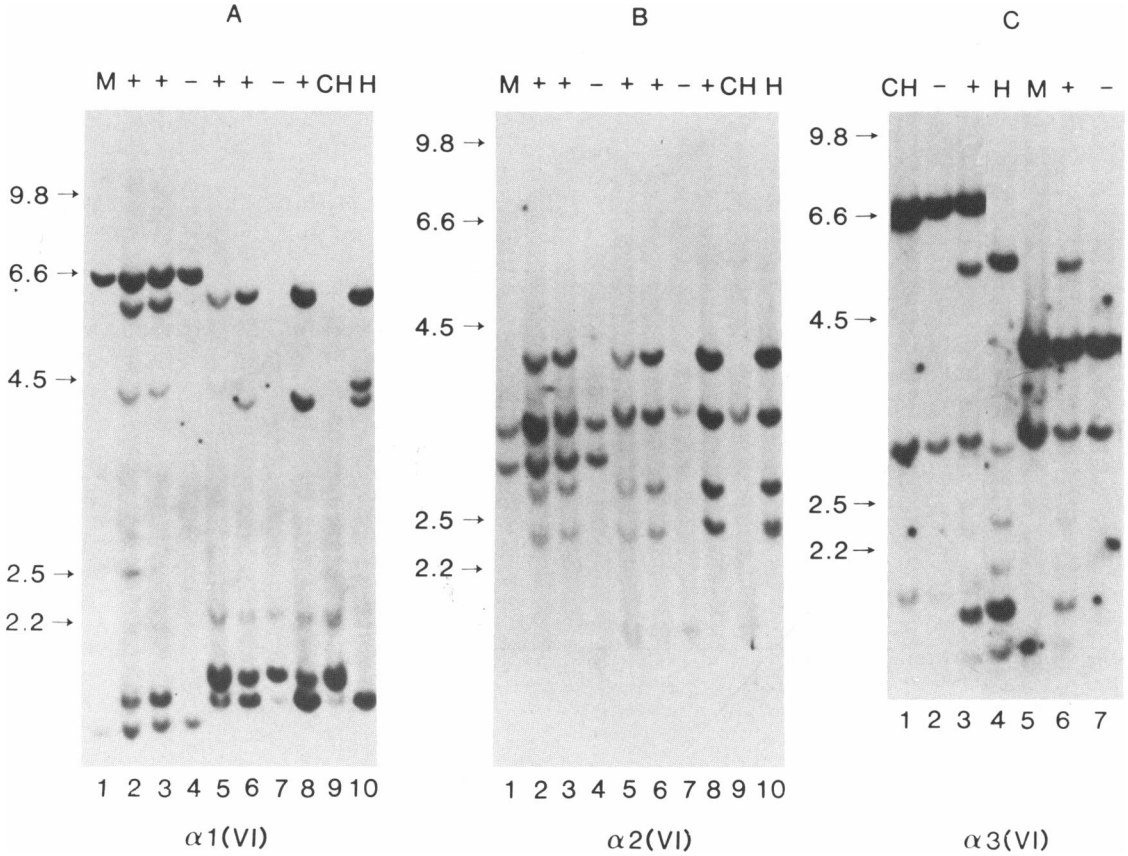

Figure 3 Southern blot analysis of *Taq*I-digested DNA from parental and human-rodent hybrid cells, with the α1(VI) (*A*), the α2(VI) (*B*), or the α3(VI) (*C*) cDNA as a probe. Panels *A* and *B:* lanes 2–3, human-mouse hybrids positive for chromosome 21; lane 4, human-mouse hybrids negative for chromosome 21; lanes 5, 6, and 8, human–Chinese hamster hybrid positive for chromosome 21; lane 7, human–Chinese hamster hybrid negative for chromosome 21. Panel *C:* lane 2, human–Chinese hamster hybrid positive for chromosome 2; lane 3, human–Chinese hamster hybrid negative for chromosome 2; lane 6, human-mouse hybrid positive for chromosome 2; lane 7, human-mouse hybrid negative for chromosome 2. The location of the molecular-weight markers (in kb; *Hin*dIII-digested λphage DNA) is indicated. Patterns produced from mouse (M), Chinese hamster (CH), and human (H) DNA are shown for comparison in panels *A*–*C*.

that the α3(VI) collagen gene is located on the long arm of chromosome 2, in the region q321-qter.

Regional Mapping by In Situ Hybridization

To confirm the chromosomal assignment using somatic cell hybrids described above, and to map the genes more precisely on the chromosomes, we performed in situ hybridization.

Alpha 1(VI) collagen gene.—Eighty-one metaphase cells with a total of 192 silver grains associated with chromosomes were examined (fig. 4). Of the total grains analyzed, 39 (20%) were located on chromosome 21. The distribution of grains on this chromosome was not random. As shown in figure 5, 87% of the total grains on chromosome 21 were in the q222-q223 region, with a maximum in the q223 band. This result confirms assignment of the α1(VI) collagen gene to human chromosome 21, an assignment obtained by analysis of hybrid cells, and further localized the gene to the region 21q223.

Alpha 2 (VI) collagen gene.—Analysis of 96 metaphase cells with 283 silver grains associated with chromosomes revealed that 49 (17%) of the total grains were located on chromosome 21 (figs. 4, 5). The major concentration of grains (89% of the total on chromosome 21) was on the q222-q223 region of this chromosome. Thus, the result was consistent with the assignment made by analysis of somatic cell hybrids, described above. In addition, the α2(VI) collagen gene can now be assigned to a more accurately defined position of 21q223, on which the α(VI) collagen gene is also located.

Table I

Correlation of α1(VI), α2(VI), and α3(VI) Gene Sequences with Human Chromosomes in Somatic Cell Hybrids

Hybrids	Human Chromosomes[a]																						Derivative Chromosomes[b]		α1(VI)/	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	Xp^+	$2q^-$	α2(VI)	α3(VI)
Human-mouse:																												
L 53 C	+	+	+	−	+	+	+	+	−	−	+	−	+	+	−	+	+	+	+	+	+	−	−				+	+
L 53 K	+	/	−	+	+	+	+	+	−	+	+	−	−	−	/	−	+	−	−	+	−	−	+				−	+
L 53 N	−	−	+	/	−	+	−	+	−	−	−	+	+	+	−	/	+	−	+	+	+	−	/				+	−
LA 56 A	+	−	+	−	+	+	+	/	−	/	/	−	−	/	/	/	+	−	/	+	+	+	+				+	−
LA 56 E	−	+	+	/	+	+	+	−	−	−	+	−	−	−	−	−	+	+	+	+	+	+	/				+	+
LA 56 G	/	+	+	−	+	+	+	+	−	+	+	+	+	−	−	+	+	+	+	+	+	+	+				+	+
LA 56 I	−	+	+	/	+	−	+	+	−	−	−	+	+	/	−	+	+	+	−	+	+	+	−				+	+
LA 56 R	−	+	+	−	+	+	+	/	−	−	+	−	−	−	−	−	+	−	/	+	/	−	+				−	+
LA 56 S	−	+	+	−	+	+	+	−	−	−	+	+	−	−	−	−	+	−	−	−	/	−	+				+	+
LA 56 U	+	+	−	+	+	−	+	+	−	/	+	+	−	−	+	+	+	+	+	+	+	+	/				+	+
LA 56 V	/	/	+	−	/	/	/	/	−	−	+	+	−	−	/	/	+	/	/	/	−	+	+				−	+
Human-hamster:																												
V 106	−	−	−	−	+	−	−	−	−	+	−	−	−	−	+	−	−	−	+	/	+	+	+	+			+	−
CH 106 IV	−	−	−	−	−	+	−	−	−	−	+	+	−	−	−	−	−	−	/	/	−	−	−	−			+	−
C 56 C	−	−	−	+	−	+	/	−	−	+	/	+	/	−	−	−	−	−	−	/	+	−	+				+	−
C 56 F	−	−	−	−	−	−	/	+	−	−	−	−	−	+	+	+	−	/	+	−	/	−	+				+	−
C 56 G	−	−	−	+	−	−	−	+	−	−	−	−	+	+	−	−	−	−	−	+	+	−	+				+	−
C 56 L	−	−	+	+	+	−	−	+	+	−	+	+	−	+	−	+	−	/	+	−	+	−	+				+	−
C 56 W	−	−	−	+	−	+	/	+	−	−	+	−	−	−	−	+	−	−	+	−	/	−	−				+	−
C 56 X	+	+	+	/	+	+	−	+	+	−	−	+	−	−	−	−	−	+	+	−	−	+	/				−	+
C 56 Y	−	−	−	−	−	+	−	+	−	+	−	+	−	−	−	−	−	−	+	−	/	−	/				+	−
CHBL E	−	*	+	+	−	+	+	+	/	−	−	−	+	+	+	+	−	−	+	−	−	−	−		+	−	−	+
CHBL D	−	*	−	+	+	+	−	+	+	+	+	−	+	+	−	−	−	−	+	−	+	−	+		+	−	+	+
CHBL H	+	*	+	+	−	+	−	+	+	+	+	+	+	+	+	+	−	/	+	/	+	+	+		+	−	−	+
CHBL I	+	*	−	+	+	−	+	+	−	−	+	+	+	/	+	+	+	/	+	−	−	−	+		+	−	−	+
CHBL J	−	*	−	+	+	+	+	+	+	−	+	+	+	+	+	+	−	−	−	+	+	+	+		+	−	+	+
CHBL K	−	*	+	−	+	+	+	−	−	−	+	−	+	−	+	−	−	−	−	−	+	+	+		+	−	+	+
CHBL N	+	−	−	+	+	+	+	−	+	−	+	−	−	−	−	+	−	−	+	−	+	+	+		−	−	+	−
CH 34 E	−	−	+	+	+	−	−	+	+	−	+	+	−	+	−	+	+	−	+	+	+	−	−	−	−	−	+	−
CH 34 G	−	*	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	+	−	*	−	+	−	+	+
CH 34 S	−	*	−	−	−	+	+	−	/	−	+	+	−	−	−	+	−	−	−	+	/	−	*	−	+	−	+	+
CH 34 V	+	*	+	+	+	+	−	−	+	−	+	+	+	+	+	+	−	+	+	+	−	−	*	−	+	−	−	+
CH 34 X	+	*	−	−	−	/	−	+	+	−	+	+	+	−	+	+	−	+	−	−	+	−	*	−	+	−	+	+
CH 34 Z	+	−	+	+	+	+	−	+	+	+	−	+	+	−	−	+	−	−	+	/	+	/	−	−	−	−	+	−
CH 34 FU	−	*	+	+	−	−	−	+	+	−	−	−	−	−	−	+	−	−	+	+	−	−	*	+	+	/	−	+
CH 34 GT	−	*	−	+	−	+	−	+	+	+	−	+	+	−	/	−	−	+	−	−	+	−	*	−	+	−	+	+
% Discordancy:[c]																												
α1(VI), α2(VI)	69	71	60	55	48	39	55	44	60	54	42	43	50	56	68	47	60	63	48	52	3	53	39	80	46	78		
α3(VI)	45	0	40	51	39	39	32	44	48	60	33	46	38	59	42	47	40	35	58	45	55	47	43	60	0	78		

[a] + = Presence of the human chromosome in 30% of the hybrids; / = presence of the human chromosome in 30% of the hybrids; − = absence of the human chromosome in the hybrids; * = derivative chromosomes.

[b] Translocated chromosome Xp^+ = 2qter-2q321::Xp22-Xqter; $2q^-$ = 2pter-2q321::Xp22-Xqter.

[c] Discordancy indicates the presence of the gene sequence but the absence of specific chromosome, or the reverse. % Discordancy = the percentage of discordant hybrids of the total hybrids.

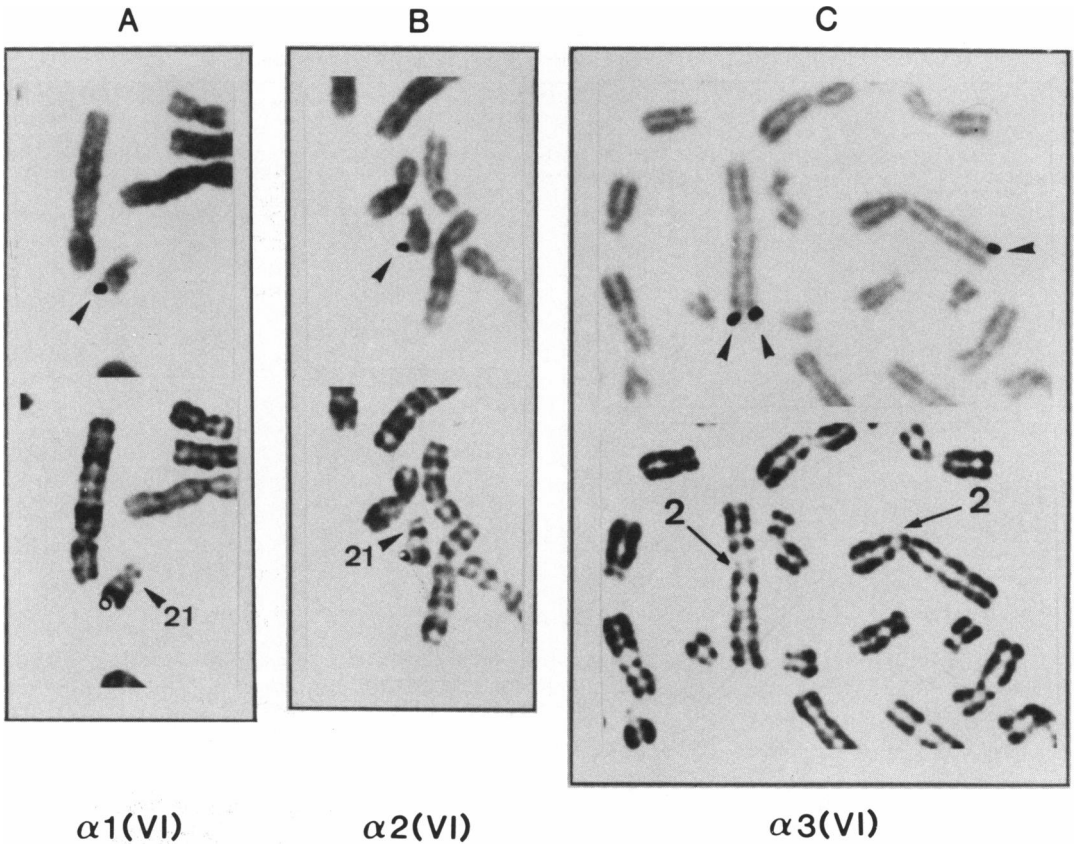

Figure 4 Examples of in situ hybridization of human metaphase chromosomes with tritium-labeled α1(VI) (*A*), α2(VI) (*B*), and α3(VI) (*C*) cDNA probes. Top panels: Silver grains (arrowheads) on chromosomes 21 and 2 after Giemsa staining. Bottom panels: Identification of chromosomes with silver grains by R-banding.

Alpha 3(VI) collagen gene.—Hybridization of a total of 98 metaphase cells with α1(VI) probe yielded 264 silver grains associated with chromosomes, of which 74 grains (28%) were on chromosome 2 (figs. 4, 5). The grains on this chromosome were concentrated on the region of q36-q37 (81% of the total grains on chromosome 2), with a peak of grains in the q37 band. The data were in good agreement with the chromosomal and regional assignment made by analysis of somatic cell hybrids. Thus, the α3(VI) gene maps to the long arm of chromosome 2, in the region q37.

Discussion

Extensive screening of a human-expression library with affinity-purified antibodies against collagen VI chains allowed us to identify several specific cDNA clones, from which three were selected for protein characterization. Each clone showed at the 5′ end a sequence encoding triple-helical Gly-Xaa-Yaa repeats, which, by matching peptide sequences, could be assigned to the α1(VI), α2(VI), and α3(VI) chains, respectively. These clones reacted specifically in Northern hybridization with a major mRNA species of either 3.5 kb (α2), 4.2 kb (α1), or 8.5 kb (α3). The sizes of the mRNAs are in good agreement with a molecular mass of ~140 kD for α1(VI) and α2(VI) and of 250 kD determined for the precursor form of α3(VI) chain (Trüeb and Winterhalter 1986). These differences, together with those observed by Southern hybridization (fig. 3), clearly indicate that each collagen VI chain is the product of a separate gene.

In addition, we have employed two independent procedures—somatic cell hybrid analysis and in situ hybridization—to map the genes for type VI collagen. Our results indicated that the α1(VI) and α2(VI) genes are syntenic on human chromosome 21q223, whereas the α3(VI) gene is located on chromosome 2q37.

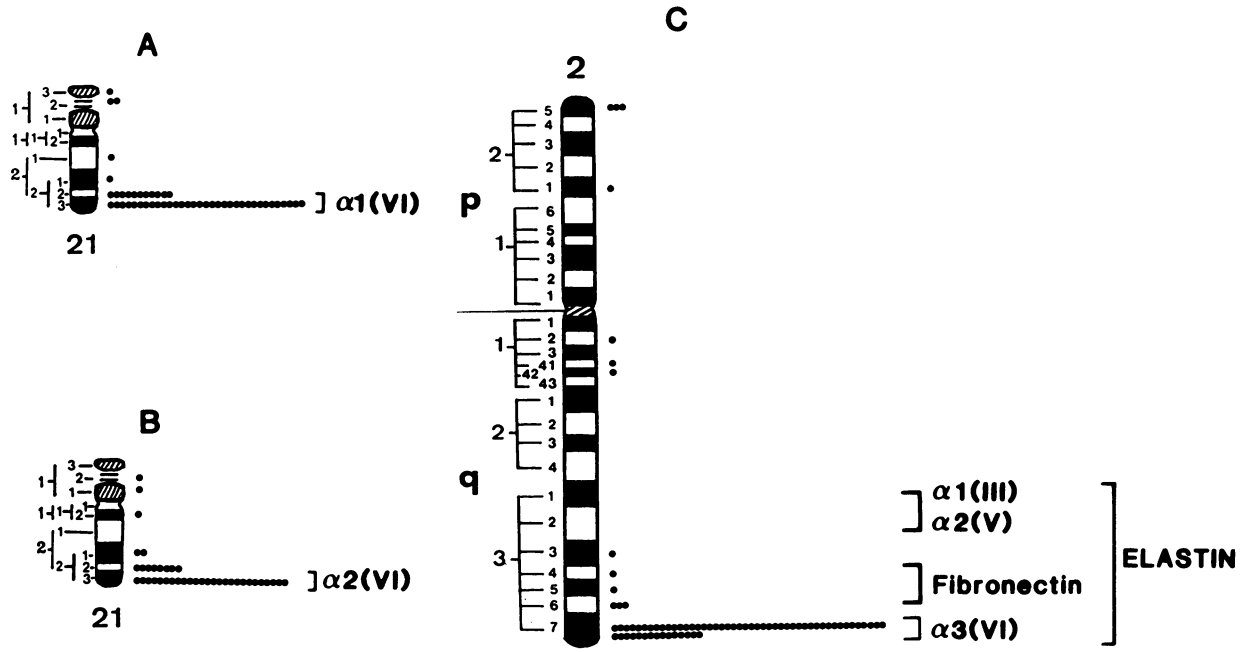

Figure 5 Distribution of silver grains on metaphase chromosomes for the α1(VI) probe (*A*) and the α2(VI) probe (*B*) on chromosome 21 and the α3(VI) probe (*C*) on chromosome 2. The location of the genes for collagen α1(III) chain, collagen α2(V) chain, elastin, and fibronectin (see text for references) is indicated by the bracket.

The family of collagen α-chain genes is known to be dispersed in the human genome. The highly conserved α1(I), α2(I), and α1(II) genes were mapped to three separate human chromosomes, namely 17, 7, and 12, respectively (Huerre et al. 1982; Junien et al. 1982; Solomon et al. 1983, 1985; Huerre-Jeanpierre et al. 1986*b*). Recently, other collagen genes have been mapped to two additional chromosomal loci (see below). These studies revealed two exceptions for the dispersion of collagen genes. The genes for α1(III) and α2(V) were assigned to the same segment of chromosome 2 (Emanuel et al. 1985*b;* Huerre-Jeanpierre 1986*a,* 1986*b*); and the two genes for α1(IV) and α2(IV) both were mapped to the distal long arm of chromosome 13 (Griffin et al. 1987). Our assignment of the α1(VI) and α2(VI) genes, therefore, represents another case of possible close linkage between two genes encoding two of the constituent chains of one collagen type. At the same time, our data also provide evidence for the dispersion of collagen genes, because the gene encoding the third chain of collagen VI is located on a separate chromosome. The fact that the α1(VI) and α2(VI) genes are located close to each other implies that these two genes may have arisen by duplication of an ancestral gene. This is consistent with our current knowledge of the structure of α1(VI) and α2(VI) chains. As mentioned above, these two polypeptides are similar in size and their mRNAs are also comparable in size. The lack of apparent structural homology of the α3 chain with the αl and α2 chains is in accordance with the presence of the α3(VI) gene on a different chromosome. This finding suggests that the α3(VI) gene may have evolved separately. Alternatively, the duplication and divergence of the α3 gene from the common ancestor may have occurred early in evolutionary history. As the genes for structurally related fiber-forming collagens are dispersed on different chromosomes, it has been proposed by several investigators that such separation of the genetic locus may be a consequence of the favorable evolutionary events, which prevent these homologous genes from crossing over (Solomon et al. 1985; Griffin et al. 1987). According to this hypothesis, one would predict that, despite the apparent similarity between the α1(VI) and α2(VI), these two chains may in fact be quite divergent in their fine structure so that crossing-over is less likely to occur. The elucidation of the origin of the type VI collagen chains awaits a more detailed analysis of the primary structure of these chains, as well as further characterization of their genomic organization.

The close linkage of the α1(VI) and α2(VI) gene loci may be reflected in their expression, which appeared to be coordinately regulated in several normal and tumor cells we have examined (Chu et al. 1987). Similarly, the variable expression of the α3(VI) gene in these same cells is likely the result of its distinct chromosomal location. Although current data support the heterotrimeric configuration of α1:α2:α3 = 1:1:1, we cannot exclude the possibility that trimers composed of one or two constituent chains exist in different tissues.

Several other genes coding for extracellular matrix proteins have been assigned to the long arm of chromosome 2 in the region of 2q24 to 2qter. Specifically, α1(III) and α2(V) collagen genes on 2q24-3-2q31 (Emanuel et al. 1985*b*), the elastin gene on 2q31-2qter (Emanuel et al. 1985*a*), and the fibronectin gene on 2q34-2q36 (Henry et al. 1985; Jhanwar et al. 1986; Prowse et al. 1986). Our assignment of the α3(VI) gene to 2q37 represents the fifth extracellular matrix gene mapped to the same region. The predominance of genes that are involved in related functions in the long arm of chromosome 2 is quite striking. It would be of interest to further investigate whether these proteins are evolutionarily related and whether these genes are coordinately regulated.

Using DNA probes of known chromosomal origin for RFLP analysis has been proved to be extremely powerful in studying inherited human diseases such as Huntington disease, cystic fibrosis, and others (for review, see Gusella 1986). This approach has been recently applied to the study of heritable connective-tissue disorders that are quite heterogeneous in nature (Tsipouras et al. 1983, 1986; Grobler-Rabie 1985; Wenstrup et al. 1986). The molecular genetic information on type VI collagen provided here probably will be an important tool in our further understanding of this group of diseases. The DNA probes will also be valuable in studying other diseases that have been mapped to distinct genetic loci on the long arm of chromosome 21, such as Down syndrome and Alzheimer disease (Williams et al. 1975; St. George-Hyslop et al. 1987). Distinct differences in the regional maps precludes, however, the possibility that these genes are identical or closely linked to those encoding α1(VI) and α2(VI) chains. Down syndrome is consistently correlated to a trisomy of chromosome 21. It remains, therefore, an interesting question whether this anomaly is also correlated with an altered expression of collagen type VI.

Acknowledgments

We thank Loretta Renkart for her technical assistance and Sandy Parsons for typing of the manuscript. This work was supported by research grants AR-28450, AR-38188, AR-38912, and AR-38923 from the National Institutes of Health; grant 86-51 from the American Heart Association, New Jersey affiliate; and the Deutsche Forschungsgemeinschaft.

References

Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymadylic acid–cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.

Bernard, M. P., M. Kolbe, D. Weil, and M.-L. Chu. 1985. Human cellular fibronectin: comparison of the carboxyl-terminal portion with rat identifies primary structural domains separated by hypervariable regions. Biochemistry **24**:2698–2704.

Boedtker, H., M. Finer, and S. Aho. 1985. The structure of the chicken α2 collagen gene. Ann. NY Acad. Sci. **460**:85–116.

Bruns, R. R., W. Press, E. Engvall, R. Timpl, and J. Gross. 1986. Type VI collagen in extracellular, 100-nm periodic filaments and fibrils: identification by immunoelectron microscopy. J. Cell. Biol. **103**:393–404.

Carter, W. G. 1982. Transformation-dependent alterations in glycoproteins of the extracellular matrix of human fibroblasts. J. Biol. Chem. **257**:13805–13815.

Chen, E. Y., and P. H. Seeburg. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.

Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.

Chu, M.-L., W. de Wet, M. Bernard, and F. Ramirez. 1985. Fine structural analysis of the human proα1(I) collagen gene. J. Biol. Chem. **260**:2315–2320.

Chu, M.-L., K. Mann, R. Deutzmann, D. Pribula-Conway, C. C. Hsu-chen, M. P. Bernard, and R. Timpl. 1987. Characterization of three constituent chains of collagen type VI by peptide sequences and cDNA clones. Eur. J. Biochem. **168**:309–317.

Chu, M.-L, J. C. Myers, M. P. Bernard, J. F. Ding, and F. Ramirez. 1982. Cloning and characterization of five overlapping cDNAs specific for the human proα1(I) collagen chain. Nucleic Acids Res. **10**:5925–5934.

de Crombrugghe, B., A. Schmidt, G. Liau, C. Setoyama, M. Mudryj, Y. Yamada, and C. McKeon. 1985. Structural and functional analysis of the genes for α2(I) and α1(III) collagens. Ann. NY Acad. Sci. **460**:142–162.

Emanuel, B. S., L. Cannizzaro, N. Ornstein-Goldstein,

Z. K. Indik, K. Yoon, M. May, L. Oliver, C. Boyd, and J. Rosenbloom. 1985*a*. Chromosomal localization of the human elastin gene. Am. J. Hum. Genet. 37:873–882.

Emanuel, B. S., L. A. Cannizzaro, J. M. Seyer, and J. C. Myers. 1985*b*. Human α1(III) and α2(V) procollagen genes are located on the long arm of chromosome 2. Proc. Natl. Acad. Sci. USA 82:3385–3389.

Engvall, E., H. Hessle, and G. Klier. 1986. Molecular assembly, secretion, and matrix deposition of type VI collagen. J. Cell. Biol. 102:703–710.

Gibson, M. A., and E. G. Cleary. 1985. CL glycoprotein is the tissue form of type VI collagen. J. Biol. Chem. 260:11149–11159.

Griffin, C. A., B. S. Emanuel, J. R. Hansen, W. K. Cavenee, and J. C. Myers. Human collagen genes encoding basement membrane α1(IV) and α2(IV) chains map to the distal long arm of chromosome 13. 1987. Proc. Natl. Acad. Sci. USA 84:512–516.

Grobler-Rabie, A. F., G. Wallis, D. K. Brebner, P. Beighton, A. J. Bester, and C. G. Mathew. 1985. Detection of a high-frequency *Rsa*I polymorphism in the human proα1(I) collagen gene which is linked to an autosomal dominant form of osteogenesis imperfecta. EMBO J. 4:1745–1748.

Gusella, J. F. 1986. DNA polymorphism and human disease. Annu. Rev. Biochem. 55:831–854.

Heller-Harrison, R. A., and W. G. Carter. 1984. Pepsin-generated type VI collagen is a degradation product of GP140. J. Biol. Chem. 259:6858–6864.

Henry, I., M. Jeanpierre, M. Bernard, D. Weil, K. H. Grzeschik, F. Ramirez, M.-L. Chu, and C. Junien. 1985. The structural gene for fibronectin (FN) maps to 2p323-qter (abstract). Cytogenet. Cell Genet. 40:650.

Hessle, H., and E. Engvall. 1984. Type VI collagen: studies on its localization, structure, and biosynthetic form with monoclonal antibodies. J. Biol. Chem. 259:3955–3961.

Huerre C., C. Junien, D. Weil, M.-L. Chu, M. Morabito, V. C. N'Guyen, J. C. Myers, C. Foubert, M. S. Gross, D. J. Prockop, A. Boué, J. C. Kaplan, A. de La Chapelle, and F. Ramirez. 1982. Human type I procollagen genes are located on different chromosomes. Proc. Natl. Acad. Sci. USA 79:6627–6630.

Huerre-Jeanpierre, C., I. Henry, P. Bernard, P. Galano, D. Weil, K. H. Grzeschik, F. Ramirez, and C. Junien. 1986*a*. The proα2(V) collagen gene (COL5A2) maps to 2q14-2q32, syntenic to the proα1(III) collagen locus (COL3A1). Hum. Genet. 73:64–67.

Huerre-Jeanpierre, C., M.-G. Mattei, D. Weil, K. H. Grzeschik, M.-L. Chu, F. O. Sangiorgi, M. E. Sobel, F. Ramirez, and C. Junien. 1986*b*. Further evidence for the dispersion of the human fibrillar collagen genes. Am. J. Hum. Genet. 38:26–37.

Jander, R., J. Rauterberg, and R. W. Glanville. 1983. Further characterization of the three polypeptide chains of bovine and human short-chain collagen (intima collagen). Eur. J. Biochem. 133:39–46.

Jander, R., D. Troyer, and J. Rauterberg. 1984. A collagen-like glycoprotein of the extracellular matrix is the undegraded form of type VI collagen. Biochemistry 23:3675–3681.

Jhanwar, S. C., J. T. Jensen, M. Kaelbling, R. S. K. Chaganti, and H. P. Klinger. 1986. In situ localization of human fibronectin (FN) genes to chromosome regions 2p14-p16, 2q34-q36, and 11q12.1-q13.5 in germ line cells, but to chromosome 2 sites only in somatic cells. Cytogenet. Cell Genet. 41:47–53.

Junien, C., D. Weil, J. C. Myers, V. C. N'Guyen, M.-L. Chu, C. Foubert, M. S. Gross, D. J. Prockop, J. C. Kaplan, and F. Ramirez. 1982. Assignment of the human proα1(I) collagen structural gene (COLIA2) to chromosome 7 by molecular hybridization. Am. J. Hum. Genet. 34:381–387.

Knight, K. R., S. Ayad, C. A. Shuttleworth, and M. E. Grant. 1984. A collageneous glycoprotein found in dissociative extracts of foetal bovine nuchal ligament: evidence for a relationship with type VI collagen. Biochem. J. 220:395–401.

Kuhn, K., R. W. Glanville, W. Babel, R.-G. Qian, H. Dieringer, T. Voss, B. Siebold, I. Oberbaumer, U. Schwarz, and Y. Yamada. 1985. The structure of type IV collagen. Ann. NY Acad. Sci. 460:14–24.

Kurkinen, M., M. P. Bernard, D. P. Barlow, and L. T. Chow. 1985. Characterization of 64-, 123- and 182-base-pair exons in the mouse α2(IV) collagen gene. Nature 317:177–179.

Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions: a critical reexamination. Biochemistry 16:4743–4747.

Lehrach, H., A. M. Frischauf, D. Hanahan, J. Wozney, F. Fuller, R. Czkvenjokov, H. Boedtker, and P. Doty. 1978. Construction and characterization of a 2.5-kilobase procollagen clone. Proc. Natl. Acad. Sci. USA 75:5417–5421.

Lozano, G., Y. Ninomiya, H. Thompson, and B. R. Olsen. 1985. A distinct class of vertebrate collagen genes encodes chicken type IX collagen polypeptides. Proc. Natl. Acad. Sci. USA 82:4050–4054.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Isolation of high-molecular-weight eukaryotic DNA from cells grown in tissue culture. Pp. 280–281 *in* Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Martin, G. R., R. Timpl, P. Muller, and K. Kuhn. 1985. The genetically distinct collagens. Trends Biochem. Sci. 10:285–287.

Mattei, M. G., N. Philip, E. Passage, J. P. Moisan, J. L. Mandel, and J. F. Mattei. 1985. DNA probe localization at 18p113 band by in situ hybridization and identification of a small supernumerary chromosome. Hum. Genet. 69:268–271.

Myers, J. C., H. R. Loidl, C. A. Stolle, and J. M. Seyer. 1985. Partial covalent structure of the human α2 type V collagen chain. J. Biol. Chem. 260:5533–5541.

N'Guyen, V. C. N., D. Weil, C. Finaz, O. Cohen-Haguehauer, M. S. Gross, J. C. Jejou-Foubert, M. F. deTand, C. Cochet, J. deGrouchy, and J. Frezol. 1986. Panel of twenty-five independent man-rodent hybrids for human genetic marker mapping. Anal. Genet. **29**:20–26.

Ninomiya, Y., M. Gordon, M. VanderRest, T. Schmidt, T. Linsenmayer, and B. R. Olsen. 1986. The developmentally regulated type X collagen gene contains a long open reading frame without introns. J. Biol. Chem. **26**:5041–5050.

Norrander, J., T. Kempe, and J. Messing. 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.

Odermatt, E., J. Risteli, V. van Delden, and R. Timpl. 1983. Structural diversity and domain composition of a unique collagenous fragment (intima collagen) obtained from human placenta. Biochem. J. **211**:295–302.

Olsen, B. R., Y. Ninomiya, G. Lozano, H. Konomi, M. Gordon, G. Green, J. Parsons, J. Seyer, H. Thompson, and G. Vasios. 1985. Short-chain collagen genes and their expression in cartilage. Ann. NY Acad. Sci. **460**: 141–153.

Prockop, D. J., and K. I. Kivirikko. 1984. Heritable diseases of collagen. N. Engl. J. Med. **311**:376–386.

Prowse, K. R., J. U. Tricoli, R. J. Klebe, and T. B. Shows. 1986. Assignment of the human fibronectin structural gene to chromosome 2. Cytogenet. Cell Genet. **41**:42–46.

Ramirez, F., M. Bernard, M.-L. Chu, L. Dickson, F. Sangiorgi, D. Weil, W. de Wet, C. Junien, and M. Sobel. 1985. Isolation and characterization of human fibrillar collagen genes. Ann. NY Acad. Sci. **460**:117–129.

Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.

St. George-Hyslop and 22 others. 1987. The genetic defect causing familial Alzheimer's disease maps on chromosome 21. Science **235**:885–890.

Sakurai, Y., M. Sullivan, and Y. Yamada. 1986. Alpha1 type IV collagen gene evolved differently from fibrillar collagen genes. J. Biol. Chem. **261**:6654–6657.

Soininen, R., L. Chow, M. Kurkinen, K. Tryggvason, and D. J. Prockop. 1986. The gene for the alpha1(IV) chain of human type IV procollagen: the exon structures do not coincide with the two structural subdomains in the globular carboxy-terminus of the protein. EMBO J. **5**:2821–2823.

Solomon, E., L. Hiorns, R. Dalgleish, P. Tolstoshev, R. Crystal, and B. Sykes. 1983. Regional localization of the human α2(I) collagen gene on chromosome 7 by molecular hybridization. Cytogenet. Cell Genet. **35**:64–66.

Solomon, E., L. R. Hiorns, N. Spurr, M. Kurkinen, D. Barlow, B. L. M. Hogan, and R. Dalgleish. 1985. Chromosomal assignments of the genes coding for human type II, III, and IV collagen: a dispersed gene family. Proc. Natl. Acad. Sci. USA **82**:3330–3334.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.

Thomas, P. S. 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5206.

Timpl, R., and J. Engel. 1987. Collagen type VI. Pp. 105–143 *in* R. Mayne and R. E. Burgeson, eds. Structure and function of collagen types: biological and extracellular matrix. Academic Press, New York.

Trüeb, B., and P. Bornstein. 1984. Characterization of the precursor form of type VI collagen. J. Biol. Chem. **259**:8597–8604.

Trüeb, B., J. B. Lewis, and W. G. Carter. 1985. Translatable mRNA for GP140 (a subunit of type VI collagen) is absent in SV40 transformed fibroblasts. J. Cell. Biol. **100**:638–641.

Trüeb, B., and K. H. Winterhalter. 1986. Type VI collagen is composed of a 200 kd subunit and two 140 kd subunits. EMBO J **5**:2815–2819.

Tsipouras, P., J. C. Myers, F. Ramirez, and D. J. Prockop. 1983. Restriction fragment length polymorphism associated with the proα2(I) gene of human type I procollagen. J. Clin. Invest. **72**:1261–1267.

Tsipouras, P., P. H. Byers, R. C. Schwartz, M.-L. Chu, D. Weil, G. Pepe, S. B. Cassidy, and F. Ramirez. 1986. Ehlers-Danlos syndrome type IV: cosegregation of the phenotype to a COL3A1 allele of type III procollagen. Hum. Genet. **74**:41–46.

Tsipouras, P., and F. Ramirez. 1987. Genetic disorders of collagen. J. Med. Genet. **24**:2–8.

Upholt, W. B., C. M. Storm, and L. J. Sandell. 1985. Structure of the type II collagen gene. Ann. NY Acad. Sci. **46**:130–140.

Von der Mark, H., M. Aumailley, G. Wick, R. Fleischmajer, and R. Timpl. 1984. Immunochemistry, genuine size and tissue localization of collagen VI. Eur. J. Biochem. **142**:493–502.

Weil, D., M. Bernard, S. Gargano, and F. Ramirez. 1987. The proα2(V) collagen gene is evolutionarily related to the major fibrillar-forming collagens. Nucleic Acids Res. **12**:9383–9394.

Weil, D., D. Cottreau, V. C. Nguyen, R. Rebourcet, C. Foubert, M. S. Gross, J. C. Dreyfus, and A. Kahn. 1980. Assignment of the gene for F type phosphofunctokinase to human chromosome 10 by somatic cell hybridization and specific immunoprecipitation. Ann. Hum. Genet. **44**:11–16.

Wenstrup, R. J., P. Tsipouras, and P. H. Byers. 1986. Osteogenesis imperfecta type IV; biochemical conformation of genetic linkage to the proα2(I) gene of type I collagen. J. Clin. Invest. **78**:1449–1455.

Williams, J., R. L. Sumit, P. R. Martens, and R. A. Kimbrell. 1975. Familial Down syndrome due to t(10;21) translocation: evidence that the Down phenotype is related to trisomy of a specific segment of chromosome 21. Am. J. Hum. Genet. **27**:478–485.