TYPE XIII COLLAGEN

Organization and chromosomal localization of the mouse gene, distance between human COL13A1 and prolyl 4-hydroxylase α-subunit genes, and generation of mice expressing an N-terminally altered type XIII collagen

ARI-PEKKA KVIST

Collagen Research Unit, Biocenter Oulu and Department of Medical Biochemistry

OULU 1999

ARI-PEKKA KVIST

TYPE XIII COLLAGEN

Organization and chromosomal localization of the mouse gene, distance between human COL13A1 and prolyl 4-hydroxylase α-subunit genes, and generation of mice expressing an N-terminally altered type XIII collagen

Academic Dissertation to be presented with the assent of the Faculty of Medicine, University of Oulu, for public discussion in the Auditorium of the Department of Medical Biochemistry, on December 2nd, 1999, at 12 noon.

OULUN YLIOPISTO, OULU 1999

Copyright © 1999
Oulu University Library, 1999

Manuscript received 17.9.1999
Accepted 27.9.1999

Communicated by
Doctor Raija Soininen
Doctor Gerard Tromp

ISBN 951-42-5394-9
(URL: http://herkules.oulu.fi/isbn9514253949/)

ALSO AVAILABLE IN PRINTED FORMAT

ISBN 951-42-5393-0
ISSN 0355-3221 (URL: http://herkules.oulu.fi/issn03553221/)

OULU UNIVERSITY LIBRARY
OULU 1999

Kvist, Ari-Pekka, Type XIII collagen. Organization and chromosomal localization of the mouse gene, distance between human COL13A1 and prolyl 4-hydroxylase α-subunit genes, and generation of mice expressing an N-terminally altered type XIII collagen.
Collagen Research Unit, Biocenter Oulu and Department of Medical Biochemistry, University of Oulu, PO Box 5000, FIN-90401 Oulu, Finland
1999
Oulu, Finland
(Received 17 September 1999)

Abstract

The complete exon-intron organization of the gene coding for the mouse α1(XIII) collagen chain, *Col13a1*, was characterized from genomic clones and multiple transcription initiation points were determined. Detailed comparison of the human and mouse genes showed that the exon-intron structures are completely conserved between the species, and both genes have their 5´ untranslated region preceded by a highly conserved putative promoter region. The chromosomal location of the mouse gene was determined to be at chromosome 10, band B4, between markers *D10Mit5* – (2.3± 1.6 cM) – *Col13a1* – (3.4±1.9 cM) – *D10Mit15*.

The location of the genes for both the catalytically important α-subunit of prolyl 4-hydroxylase (P4HA) and human type XIII collagen (COL13A1) were previously mapped to 10q21.3-23.1. Prolyl-4-hydroxylase catalyzes the formation of 4-hydroxyproline in collagens by the hydroxylation of peptide-bound proline and plays a crucial role in the synthesis of these proteins. The order and transcriptional orientation of the COL13A1 and P4HA was determined. These two genes were found to lie at tail to tail orientation on chromosome 10 and the distance between these genes was determined to be about 550 kbp.

To study the function of type XIII collagen we used gene targeting in ES cells to generate a mouse line that carries a mutated type XIII collagen gene. Instead of normal protein, mutant mice express type XIII collagen with an altered amino-terminus in which the cytosolic and the transmembrane domains have been replaced with an unrelated sequence. The homozygous mice are fertile and viable but they show alterations in skeletal muscles, mainly wavy sarcolemma and increased variation in muscle fiber diameter. Ultrastructural studies revealed additional abnormalities such as streaming of z-disks, accumulation and enlargement of mitochondria, and disorganized myofilaments. The basement membranes of the muscle cells showed areas of detachment from the plasma membrane and the fibrillar matrix of the cells was less compact than in control animals. Fibroblasts cultured from mutant mice had normal levels of type XIII collagen but exhibited decreased adhesion to substratum which might be explained by a reduced anchoring strength of the altered protein.

Key words: cre-Lox recombination, extracellular matrix, muscular atrophy, transgenic mice

Acknowledgements

This work was carried out at the Department of Medical Biochemistry, University of Oulu, from 1990 to 1999.

I thank Professor Taina Pihlajaniemi for her guidance and patience throughout the study. I wish to express my respect for Research Professor Kari Kivirikko for his deep insight into research work and providing excellent facilities. My thanks are also due to Professor Ilmo Hassinen for enjoyable coffee room discussions and giving me another view of research work from the perspective of the “energy team”. Docent Leena Ala-Kokko deserves a warm hug for her joyful and positive appearance. The expert contribution of Professor Reinhard Fässler, Docent Raija Sormunen, Doctor Nina Horelli-Kuitunen, and those of all my other co-authors to this work deserve my deepest gratitude.

I wish to acknowledge Doctor Raija Soininen and Docent Gerard Tromp for their precise and constructive comments on the manuscript. Docent Gerard Tromp also made the careful revision of the language for which I am very grateful.

I wish also to thank all my colleagues and staff – both present and past colleagues from the department – for their tenacious attitude and support which guided me through the jungle of never-ending series of disappointments, frustration and failed experiments that seems to be an inseparable part of the type XIII collagen research. Anne Latvanlehto, Malin Sund and Lauri Eklund are due my special thanks because of tireless effort which was invaluable to me. Also, I want to thank all my friends and brass band Teekkaritorvet allowing me the opportunity to escape from the every-day routines.

I express my gratitude to Aila Jokinen, Maija Seppänen, Sirkka Vilmi and Jaana Väisänen for their expert technical assistance, good humour and friendship during all these years.

Above all, I owe my deepest gratitude and love to my wife Laura for her care and understanding. I’m also indebted to her for giving birth to our beloved daughters Reetta and Paula. I wish to express my love and respect to my parents Airi and Paavo as well as to my sisters Arja, Kirsi and Tiina.

This work was supported by grants from the Emil Aaltonen Foundation, the Jenny and Antti Wihuri Foundation and the Research and Science Foundation of Farmos.

Oulu, September 1999 Ari-Pekka Kvist

Abbreviations

$\alpha x(a)$	collagen polypeptide; x: number of chain; a: number of collagen
bp	base pairs
FACIT	fibril associated collagens with interrupted triple helices
cDNA	complementary DNA
COLXAY	Human collagen gene, x:number of collagen; y:number of α chain.
Colxay	Mouse (or chicken) collagen gene, x: number of collagen; y: number of α chain
C-	Carboxy-
ELISA	Enzyme-linked imuunosorbent assay
ES-cells	Embryonic stem cells
FISH	Fluorescense *in situ* hybridization
kbp	kilobase pairs, 1000 base pairs
mRNA	Messenger RNA
NC	Non-collagenous
N-	Amino-
P4HA	Prolyl 4-hydroxylase α-subunit
PCR	Polymerase chain reaction
RACE	Rapid amplification of cDNA ends
RT	Reverse transcriptase
tk	Thymidine kinase

List of original articles

I Kvist A-P, Latvanlehto A, Sund M, Horelli-Kuitunen N, Rehn M, Palotie A, Beier DR & Pihlajaniemi T (1999) Complete Exon-Intron Organization and Chromosomal Location of the Gene for Mouse Type XIII Collagen (*Col13a1*) and Comparison with its Human Homologue. Matrix Biol 18: 261-274.

II Horelli-Kuitunen N, Kvist A-P, Helaakoski T, Kivirikko K, Pihlajaniemi T & Palotie A (1997) The Order and Transcriptional Orientation of the Human COL13A1 and P4HA Genes on Chromosome 10 Long Arm Determined by High-Resolution FISH. Genomics 46: 299-302.

III Kvist A-P, Latvanlehto A, Sund M, Eklund L, Sormunen R, Väisänen T, Fässler R & Pihlajaniemi T. The Expression of N-terminally Altered Type XIII Collagen Causes a New Form of Progressive Muscular Atrophy in Mice. Manuscript.

Contents

Abstract
Acknowledgements
Abbreviations
List of original articles
Contents
1. Introduction..13
2. Review of the literature...14
  2.1. Fibrillar collagens 14
  2.2. The nonfibrillar collagens, types IV, VI-VIII and X..17
  2.3. The FACIT collagens..20
  2.4. Collagens with multiple interruptions 22
  2.5. Collagens that contain transmembrane domains 23
    2.5.1. Type XVII collagen..23
    2.5.2. Type XIII collagen...23
      2.5.2.1. Primary structure..23
      2.5.2.2. Organization and chromosomal localization of the human type XIII gene..24
      2.5.2.2. Alternative splicing 25
      2.5.2.1.3. Tissue distribution..25
      2.5.2.4. Subcellular localization..26
  2.6. Mutations in collagen genes...26
  2.7. Generation of genetically altered mice...27
  2.8. Animal models in collagen research...29
    2.8.1. Spontaneous mutations...29
    2.8.2. Transgenic mice with dominant-negative mutations 30
    2.8.3. Targeted mutations in collagen genes..33
3. Outlines of the present research 36
4. Materials and methods 37
  4.1. Characterization of genomic organization (I, II)...37
  4.2. Isolation of RNA and RT-PCR (I, III) 38
  4.3. Northern analysis (I) 38

4.4. Nuclease S1 and 5'-RACE assays (I) 39
4.5. Chromosomal localization (I,II) 39
4.5.1. Determination of chromosomal location by linkage using single-strand conformational polymorphism 39
4.5.2. Fluorescence in-situ hybridization 39
4.6. Nucleotide sequence analysis (I, III) 40
4.7. Construction of the targeting vector (III) 40
4.8. Generation of targetted ES cell clones and transgenic mice (III) 41
4.9. Adhesion studies (III) 41
4.10. Light microscopic studies (III) 42
4.10.1 Histology 42
4.10.1. Immunofluorescence staining of fibroblasts 42
4.11. Electron microscope studies (III) 42
5. Results 43
5.1. Exon-intron organization of the mouse α1(XIII) collagen gene 43
5.2. Analysis of 5'-flanking regions of mouse and human type XIII collagen genes and sequence comparison 44
5.3. Chromosomal localization of the genes for mouse and human type XIII collagen and human prolyl 4-hydroxylase α-subunit 45
5.4. Generation of mice with a mutated type XIII collagen gene 45
5.5. Analysis of cultured fibroblasts and analysis of mRNA by RT-PCR 46
5.6. Morphological analysis 46
5.7. Electron microscopic analysis 47
6. Discussion 48
7. References 53

1. Introduction

The extracellular matrix is a supporting material that surrounds cells in every tissue. It is composed of elastin, proteoglycans, laminins, fibronectin and collagens, which are its major constituents. Collagens are characteristically of very high tensile strength and form major fibrous components of tissues, such as skin, bone, tendon, cartilage, blood vessels and teeth. Nineteen differerent collagen types, encoded by more than 30 genes dispersed through the genome have been identified to date.

Type XIII collagen is a low molecular weight collagen which contains three collagenous (COL1–COL3) and four noncollagenous domains (NC1–NC4). There are interruptions in the Gly-X-Y sequence in every COL-domain and the NC1 domain contains a highly hydrophobic region indicating that this collagen is a transmembrane protein which localizes at plasma membranes. The primary transcript of type XIII collagen gene is subjected to extensive alternative splicing, resulting hypothetically to 1024 different gene products.

The technical advances and development since the 1980's that accompanied the rapidly expanding field of molecular biology enabled the manipulation of the mouse germline offered a new platform for clinical and basic research namely, animal models with specific genetic alterations. Today, transgenic animals are widely used in biomedical research because they enable detailed studies of specific diseases. Although increasing numbers of analyses defining the effects of the genomic changes are made using cultured cells, transgenic mice yield the most meaningful answers regarding the effects of genes at the level of the organism for multicellular eukaryotes.

When the work on this thesis was initiated the partial organization of the human type XIII cDNA and corresponding gene had been characterized. This work describes the characterization of the entire genomic organization of the mouse type XIII collagen gene and describes those exons of the human type XIII collagen gene which were not published before. The chromosomal localization of and physical distance between the genes for mouse type XIII collagen and prolyl 4-hydroxylase was determined. Moreover, a transgenic mouse strain that lacks the N-terminal and transmembrane portion of type XIII collagen polypeptide was generated. The generation and analysis of the transgenic mouse strain, as well as the conclusions drawn from these studies regarding the function of type XIII collagen, are described.

2. Review of the literature

Collagens are molecules that can be called as “the glue of life” and humans have utilized the physical properties of these unique molecules through the history. According to Prockop & Kivirikko (1995) the superfamily of collagens can be divided into the following classes on the basis of structural features: (a) collagens that form fibrils (types I, II, III, V and XI), (b) collagens that form network-like structures (type IV-family, types VIII and X), (c) collagens that form beaded filaments (type VI), (d) collagens that form anchoring fibrils for basement membranes (type VII), (e) collagens found on the surface of collagen fibrils and known as fibril associated collagens with interrupted triple helices (FACITs) (types IX, XII, XIV, XVI and XIX), (f) collagens that contain transmembrane domains (types XIII and XVII), (g) collagens with multiple triple helical domains and interruptions (MULTIPLEXIN’s, types XV and XVIII) (Oh *et al.* 1994a), and (h) proteins containing triple-helical domains that have not been defined as collagens. For further reading see Chu & Prockop (1993), Fleischmajer *et al.* (1990), Hulmes (1992), Kielty *et al.* (1993), Kivirikko (1995), Vuorio & De Crombrugghe (1990). In this litterature review I will introduce the family of collagens, especially type XIII collagen and the field of transgenic technologies linked to the collagen research.

2.1. Fibrillar collagens

The fibrillar collagen types I, II, III, V and XI are all very similar in structure. Each type is synthesized as a precursor form called procollagen and they are processed in a number of ways prior to secretion into the extracellular space where they interact by forming highly organized fibers and fibrils, thereby providing structural support for the body.

Type I collagen molecules are trimers that consist of one $\alpha 2(I)$ and two $\alpha 1(I)$ chains. It is the most abundant collagen found in the fibrils of tendon, ligaments and bones. Differences in type I collagen isolated from these three tissues are due to differences in the degree of hydroxylation of proline and lysine residues, glycosylation and aldehyde formation, the latter being involved in covalent inter-chain cross-linking. In bones the fibrils are mineralized with calcium hydroxyapatite.

Cartilage and related tissues contain collagen molecules whose structures are presumed to reflect the specialized functions of the proteoglycan-rich extracellular matrices of these tissues. Type II collagen — also called cartilage collagen — is arrayed in quarter-staggered fashion to form fibers similar to those of type I collagen (Linsenmayer 1991). However, type II collagen is an $\alpha 1(II)_3$ homotrimer, consisting of three identical α chains. Despite being found almost exclusively in cartilage, type II collagen also occurs in the vitreous humor.

Type III collagen can be called the fetal and blood vessel collagen but it can also be found in the skin, lung, blood vessels and cornea. Type III collagen is important for the development of skin and the cardiovascular system and for the maintenance of the normal physiological functions of these organs in adult life (Olsen 1995). Type III collagen molecules are α1(III)-homotrimers.

Type V collagen is present in many tissues and organs as a minor collagen component and it can be present as a homotrimer consisting of three identical α1 chains, as a heterotrimer consisting of α1(V), α2(V), and α3(V) chains, or as a heterotrimer consisting of two copies of α1 and one copy of α2 chains (Burgeson *et al.* 1976). The structure of type XI collagen was studied by Morris & Bachinger (1987), who concluded that type XI is a heterotrimer consisting of three different polypeptides, α1(XI), α2(XI), and α3(XI). Fichard *et al.* (1994) have reviewed the structure and function of collagens V and XI and commented on their fundamental role in the control of fibrillogenesis, which they probably achieve by forming a core within the fibrils.

The most characteristic feature of the genes that encode the fibrillar collagens is the unusual pattern of exon sizes that are multiples of 54- and 45 bases in the region of the gene that encodes the triple-helical domain of the polypeptide. This pattern is highly conserved through evolution.

Tromp *et al.* (1988) and Kuivaniemi *et al.* (1988) characterized cDNA clones that represent the full-length mRNAs from the COL1A1 and COL1A2 genes. The length of the COL1A1 gene is known to be over 18 kbp (Körkkö *et al.* 1998). de Wet *et al.* (1987) isolated the entire 38 kbp COL1A2 gene and 22 kbp of its flanking sequences. Retief *et al.* (1985) assigned the α1(I) and α2(I) genes to bands 17q21.31-q22.05 and 7q21.3-q22.1, respectively, by *in situ* hybridization.

Mouse cDNAs encoded by the *Col1a1* and *Col1a2* genes were cloned and the chromosomal locations of these genes were determined to be on chromosomes 11 and 6, respectively. (Irving *et al.* 1989, Li *et al.* 1995a, Munke *et al.* 1986, Phillips *et al.* 1992).

Strom & Upholt (1984) isolated overlapping genomic DNA clones containing most of the coding sequences for chicken type II collagen. Baldwin *et al.* (1989) cloned and characterized the cDNA clones for human type II collagen and the complete genomic structures for the human and mouse type II collagen genes were published by Ala-Kokko *et al.* (1995) and Metsäranta *et al.* (1991), respectively. The overall identity between the coding sequences of the mouse and human type II collagen genes is 89% at the nucleotide level, and this results in only 37 amino acid changes between the mature α1(II) collagen chains. Takahashi *et al.* (1990) localized the COL2A1 gene to 12q13.11-q13.12 and showed that the COL2A1 gene is immediately proximal to the fragile site fra(12)(q13.1). The chromosomal localization of the mouse gene was determined by Cheah *et al.* (1991).

Table 1. Chromosomal locations of human and mouse collagen genes and number of human mutations associated with these genes.

Type	Gene	Human chromosome	Mouse chromosome	Human mutations
I	COL1A1	17q21.31-q22.05	11 (56.0 cM)	164
	COL1A2	7q21.3-q22.1	6 (0.7 cM)	85
II	COL2A1	12q13.11-q13.2	15 (56.8 cM)	40
III	COL3A1	2q32.2	1 (21.1 cM)	69
IV	COL4A1	13q34	8 (5.0 cM)	-
	COL4A2	13q34	8 (5.0 cM)	-
	COL4A3	2q36-q37	ND	5
	COL4A4	2q36-q37	ND	13
	COL4A5	Xq22	X (62.40 cM)	239
	COL4A6	Xq22	X	1
V	COL5A1	9q34.2-q34.3	2 (17.0 cM)	4
	COL5A2	2q32-q33	ND	3
	COL5A3	ND	ND	-
VI	COL6A1	21q22.3	10 (41.0 cM)	1
	COL6A2	21q22.3	10 (41.0 cM)	1
	COL6A3	2q37	1 (53.9 cM)	1
VII	COL7A1	3p21.3	9 (61.0 cM)	99
VIII	COL8A1	3q12-q13.1	ND	-
	COL8A2	1p34.4-p32.3	4 (57.2 cM)	-
IX	COL9A1	6q12-q13	1 (15.0 cM)	-
	COL9A2	1p33-p32.2	4 (53.0 cM)	1
	COL9A3	20q13.3	ND	1
X	COL10A1	6q21-q22.3	10 (22.0 cM)	24
XI	COL11A1	1p21	3 (53.10 cM)	2
	COL11A2	6p21.3	17 (18.51 cM)	3
	COL11A3[a]	12q13.11-q13.2	15 (56.8 cM)	-
XII	COL12A1	6q12-q13	9 (43.0 cM)	-
XIII	COL13A1	10q22	10 (23.1 cM)	-
XIV	COL14A1	8q23[b]	ND	-
XV	COL15A1	9q21-q22	4 (B1-3)	-
XVI	COL16A1	1p34	ND	-
XVII	COL17A1	10q24.3	19 (49.0 cM)	18
XVIII	COL18A1	21q22.3	10 (41.0 cM)	-
XIX	COL19A1	6q12-q13	1 (A3)	-

The human and mouse data and the number of human mutations were collected from the references mentioned in the text and from GeneCards database and Human Gene Mutation Database. The mouse data was obtained from the Mouse Genome Database (http://www.informatics.jax.org/), except for *Col13a1* which is published in paper II in this thesis. The total number of mutations is 774.

[a]The α3(XI) chain of type XI collagen is encoded by the same gene as the α1(II) chain of type II. [b]Location of human undulin gene. ND: not determined.

Ala-Kokko *et al.* (1989) published the full length sequence of the human type III collagen cDNA and the structure of the mouse *Col3a1* gene was published by Toman and de Crombrugghe (1994). The chromosomal localization of the COL3A1 gene was determined by Huerre-Jeanpierre *et al.* (1986), who assigned it to 2q31-q32.3. Tsipouras *et al.* (1988) demonstrated by somatic cell hybrid studies and *in situ* hybridization that the COL3A1 and the COL5A2 loci are very close together, and later Cutting *et al.* (1990) showed by pulsed field gel electrophoresis that the COL3A1 and COL5A2 genes are in the same 35 kbp fragment. Schurr *et al.* (1990) mapped the murine *Col3a1* to mouse chromosome 1.

The organization of the COL5A1 gene has diverged considerably from the conserved organization of the genes for the fibrillar collagen types I-III. COL5A1 has 66 exons, which is greater than the number of exons found in the genes for collagen types I-III. Takahara *et al.* (1991) reported the sequence of the cDNA encoding the complete prepro-α1(V) chain. The collagen-like region and C-terminal noncollagenous region closely resemble the corresponding regions of the α1(XI) chain. Takahara *et al.* (1995) also determined the complete genomic structure of COL5A1 and showed that the gene spans at least 750 kb. Greenspan *et al.* (1992) determined that the COL5A1 gene is located in the segment 9q34.2-q34.3 using cDNA and genomic clones for the COL5A1 gene as probes. Pilz *et al.* (1994) mapped the mouse gene for α1 chain of collagen V to chromosome 2.

The genomic organization of the COL5A2 gene is not known but cDNA sequences derived from this gene is published by Weil *et al.* (1987) and there are also reports of COL5A2-linked diseases (Michalickova *et al.* 1998, Richards *et al.* 1998). No additional information is available about α3(V) chain since its original description (Burgeson *et al.* 1976).

The cDNA-derived amino acid sequences of the human α1(XI) and α2(XI) chains were determined and these genes were localized broadly to 1p21 and 6p21.2, respectively (Bernard *et al.* 1988, Henry *et al.* 1988, Kimura *et al.* 1989a). Vuoristo *et al.* (1995) analyzed the human COL11A2 gene confirming the earlier nucleotide sequence analysis of a selected portion of the human gene which showed the characteristic pattern of 54-bp exons (Henry *et al.* 1988). The mouse *Col11a2* gene was isolated and characterized by Vandenberg *et al.* (1996). The α3 chain of type XI collagen is an over-glycosylated variant of the type II mRNA, i.e. it is encoded by the COL2A1 gene (Eyre & Wu 1987).

2.2. The nonfibrillar collagens, types IV, VI-VIII and X

The first collagen found to contain interruptions in the repeating Gly-X-Y sequences was the basement membrane-specific type IV collagen and therefore this molecule is the prototype of nonfibrillar collagens. Type IV collagen does not form ordered fibrillar structures but a complex meshwork of molecules held together at their ends and by lateral aggregations. Type IV collagen is an important component of basement membranes where it associates with laminin, entactin, and heparan sulfate proteoglycans. Basement

membranes compartmentalize tissues, serve as molecular sieves and provide important signals for the differentiation of the cells they support.

Six α(IV) chains are known. The α1(IV) and α2(IV) chains are ubiquitous in tissues, whereas the α3(IV), α4(IV), α5(IV) and α6(IV) chains have more restricted tissue distributions (Butkowski *et al.* 1987, Mariyama *et al.* 1992, Mayne *et al.* 1984, Saus *et al.* 1988, Zhou *et al.* 1993). On the basis of sequence similarities, the chains fall into two classes. α1, α3, and α5 chains make up the α1-like class, and α2 , α4 and α6 chains belong to the α2-like class.

The human cDNAs for all six α(IV) chains have been characterized (Hostikka & Tryggvason 1988, Leinonen *et al.* 1994, Mariyama *et al.* 1994, Myers *et al.* 1990, Oohashi *et al.* 1994, Soininen *et al.* 1987). The complete cDNAs for the α1 and α2 chains, and partial cDNAs for the α3, α4, and α5 chains, have been characterized from the mouse (Miner & Sanes 1994, Muthukumaran *et al.* 1989, Saus *et al.* 1989).

The human COL4A1 and COL4A2 genes are located in a head-to-head configuration on chromosome 13q34 (Emanuel *et al.* 1986, Poschl *et al.* 1988, Soininen *et al.* 1988) and the COL4A3 and COL4A4 genes are similarly arranged on chromosome 2q36-q37 (Boye *et al.* 1998, Kamagata *et al.* 1992, Mariyama *et al.* 1992, Turner *et al.* 1992). COL4A5 is paired with the COL4A6 gene on chromosome Xq22 (Oohashi *et al.* 1994, Zhou *et al.* 1993). Thus, it appears that the type IV collagens evolved by duplication of an ancestral α chain gene, giving rise to a pair of α chain genes with closely apposed 5-prime ends. Koizumi *et al.* (1995) mapped the mouse *Col4a1* and *Col4a2* genes to the centromeric region of mouse chromosome 8 and Brown *et al.* (1993) showed the *Col4a5* gene to lie on chromosome X.

Type VI collagen is unusual among the collagens because of the small size of its collagenous domains and an exceptional supramolecular structure (Engvall *et al.* 1986). Trueb & Winterhalter (1986) showed that type VI collagen consists of two distinct 140-kD subunits (α1(VI) and α2(VI)) and a 200-kD subunit (α3(VI)). Chu *et al.* (1987, 1990) and Chung *et al.* (1976) characterized the three constituent chains by peptide sequences and cDNA clones. Collagen VI is a component of bead-like microfibrillar structures in many tissues (Engel *et al.* 1985). These microfibrils are located close to cells, nerves, blood vessels, and large collagen fibrils and they are considered to have an anchoring function (Kuo *et al.* 1997). Heiskanen *et al.* (1995) determined the the size of the COL6A1 gene to be 29 kb, showed the distance separating the COL6A1 and COL6A2 genes to be 150 kbp and mapped the 5'-3' orientation of these genes. Antonarakis (1993) and Saitta *et al.* (1992) demonstrated that it is 36 kbp long and contains 30 exons. Weil *et al.* (1988) localized the COL6A1 and COL6A2 genes to chromosome 21q22.3 and the COL6A3 gene to chromosome 2q37. The COL6A2 gene is the most telomeric gene on chromosome 21. Justice *et al.* (1990) mapped the mouse homologues of COL6A1 and COL6A2 to chromosome 10.

Type VII collagen is the main constituent of anchoring fibrils, which are located below the basal lamina at the dermal-epidermal basement-membrane zone of the skin. The type VII collagen molecules form disulfide bond-stabilized dimeric aggregates by lateral accretion in a nonstaggered array (Bentz *et al.* 1983, Burgeson 1987).

After the initial cloning of this novel collagen in 1994, it was shown that the type VII collagen mRNA is approximately 9.2 kbp long and it contains an open reading frame of

8833 nucleotides which encodes a polypeptide of 2944 amino acids. The COL7A1 gene was shown to contain 118 exons, more than any previously described gene. Despite this complexity, the COL7A1 gene is compact, consisting of 31 132 bp from the transcription initiation site to the polyadenylation site (Christiano *et al.* 1994a, Christiano *et al.* 1994b, Greenspan 1993, Tanaka *et al.* 1992). The gene was mapped to 3p21.3 by *in situ* hybridization (Greenspan *et al.* 1993, Parente *et al.* 1991). Li *et al.* (1993a) showed that the corresponding mouse gene is located on chromosome 9.

Type VIII collagen was first detected and designated as an endothelial cell collagen (see Sutmuller *et al.* 1997). Type VIII collagen is also a major component of Descemet's membrane, the basement membrane of corneal endothelial cells but COL8A1 mRNA is also expressed in a number of other tissues (Muragaki *et al.* 1992). In-situ hybridization demonstrated COL8A1 mRNA in skin keratinocytes, corneal epithelial and endothelial cells, lens epithelial cells, mesenchymal cells surrounding cartilage and calvarial bone, and in the meninges surrounding the brain. In Descemet's membrane type VIII collagen molecules consist of α1 and α2 subunits in the ratio of 2:1 (Muragaki *et al.* 1991). Muragaki *et al.*(1991a,b) mapped the COL8A1 gene to 3q12-q13.1 and the COL8A2 to 1p34.3-p32.3 by *in-situ* hybridization These authors also demonstrated that the COL8A1 gene consists of only 4 exons, one of which is large and encodes the entire triple-helical and C-terminal non-triple-helical domains. To date these genes have not been described in more detail. The sequence of the α1(VIII) chain is similar to that of the α1 chain of type X collagen, which is produced by hypertrophic chondrocytes. Types VIII and X collagen have been called 'short-chain collagens' because of the relatively small size of their triple-helical domains.

Type X collagen is a product of hypertrophic chondrocytes and has been localized to presumptive mineralization zones of hyaline cartilage (see Sutmuller *et al.* 1997). Kirsch and von der Mark (1991) isolated human type X collagen from human fetal growth-plate cartilage and purified it to homogeneity. They also used an antibody to show the distribution of type X collagen in the growth-plate cartilage and the calcifying zone of the sternum in the human fetus. The distribution of collagen X and its transient expression at sites of calcification suggest that this collagen is associated with events in the later stages of endochondral bone formation.

The triple-helical domain of type X collagen is approximately half the size of that in collagen types I, II, and III. It has been shown that two exons, 169 bp and 2940 bp in length, separated by single 3200-bp intron, encode the complete primary translation product of the COL10A1-gene. Apte *et al.* (1991) and Thomas *et al.* (1991) cloned the human type X gene and assigned it to 6q21-q22.3. The mouse homologue was shown to be located on chromosome 10 (Apte *et al.* 1992).

Table 2. Features of collagen genes and their products

	Gene		Human polypeptide			
	Exons	Size, kbp hu/mo	Mw	Residues	Col/NC Ratio	Accession number/ Reference
COL1A1	51	18/ND	138671	1461	2.37	AF017178
COL1A2	51	38/ND	129423	1366	3.05	AF004877
COL2A1	54	31/29.6	141771	1487	2.68	L10347
COL3A1	51	44/37.6	138555	1466	2.49	P02461
COL4A1	52	>100/ND	160611	1669	3.39	P02462
COL4A2	47	>100/>90	167535	1712	3.37	P08572
COL4A3			161740	1670	5.67	X80031
COL4A4	48	>113/ND	164096	1690	5.50	X81053
COL4A5	51	>140/ND	161044	1685	5.80	SEG_HS4COL5A
COL4A6	46	425/ND	163696	1690	5.47	SEG_HCOL4A6S
COL5A1	66	750/ND	183618	1838	1.67	P20908
COL5A2			144720	1496	4.44	P05997
COL6A1	36	29/ND	108639	1028	0.50	CGHU1A
COL6A2	36	30/ND	108475	1018	0.51	CGHU2A
COL6A3	28		343552	3176	0.12	X52022
COL7A1	118	31.1/31	295220	2944	1.10	Q02388
COL8A1	4	>53/ND	73443	744	1.73	X57527
COL8A2	4		60527	635	2.57	P25067
COL9A1	38	90/100	91892	921	1.86	SEG_HSCOL9AC
COL9A2	32	15/15.9	65131	689	7.88	AF019406
COL9A3	32	23/ND	63742	684	8.28	L41162
COL10A1	3	6.2/ND	66156	680	2.33	Q03692
COL11A1			181156	1806	1.65	P12107
COL11A2	66	28.8/100	171775	1736	2.65	AL031228
COL12A1		>200/ND	333193	3063	0.09	U73778
COL13A1	41/42	140/135	70900	726	2.39	(Hägg 1998)
Col14a1			202667[a]	1888	0.17	X70793
COL15A1	42	145/135	141757	1388	0.42	SEG_HSCOLXV
COL16A1			157693	1603	1.86	Q07092
COL17A1	56	52/ND	150459	1497	0.79	AH005152
COL18A1	43	ND/>102	153841	1516	0.56	AF018081
COL19A1	51	>250/ND	115448	1143	2.89	SEG_AB00468S

Collected from references occurring in the text. [a)]Amino acid sequence from chicken. ND: not determined.

2.3. The FACIT collagens

The collagen types IX, XII, XIV, XVI and XIX are known as Fibril Associated Collagens with Interrupted Triple helices, or FACITs (Fukai *et al.* 1994, Mayne & Brewton 1993,

Ninomiya *et al.* 1990, Shaw & Olsen 1991). Collagen types IX, XII and XIV each contain of one or two collagenous domains that associate with fibrillar collagens, and one collagenous domain and a noncollagenous domain that project out from the fibril. The function of these collagens is not clear, but it is thought that they can serve as interaction components between collagen fibrils and other molecules of the extracellular matrix.

Type IX collagen is composed of three gene products, the α1(IX), α2(IX), and α3(IX) chains and is located on the surface of type II collagen-containing fibrils. Type IX collagen is also a proteoglycan, as it has been shown to contain chondroitin sulfate- and dermatan sulfate-chains covalently linked to the α2(IX) chain (Bruckner *et al.* 1985). The precise function of type IX collagen is not known but the results from certain transgenic mice studies suggest that this collagen is not essential for the assembly of cartilage-fibrils in the extracellular matrix. However, it may have an important role in the maintenance of the structural integrity of the tissue (Mayne *et al.* 1985, Nakata *et al.* 1993). The cDNAs for all three α chains of type IX collagen have been cloned and their chromosomal locations have been determined (Table 2). *Col9a1* and *Col9a2* genes have been localized from the mouse genome (Brewton *et al.* 1995, Kimura *et al.* 1989b, Muragaki *et al.* 1990a, Ninomiya & Olsen 1984, Perälä *et al.* 1993, Perälä 1997, Pihlajamaa *et al.* 1998, Warman *et al.* 1993, Warman *et al.* 1994).

A cDNA clone that encodes a collagen-like polypeptide containing two collagenous and three noncollagenous domains was isolated by Gordon *et al.* (1987) and the novel collagen was called type XII (Gordon *et al.* 1989). The complete cDNA structures of type XII collagen are available from both human and mouse, and both polypeptides have a predicted molecular weight of about 340,000 (Bohme *et al.* 1995, Gerecke *et al.* 1997)(Table 2). Immunohistochemical analyses of mouse embryos showed that type XII collagen is expressed mainly in dense connective tissues of the tendons, ligaments, dermis, cornea, blood vessel walls, meninges, and developing membranous bones (Oh *et al.* 1993). Type XII collagen is thought to act as a crossbridge between fibrils and resist shearing forces caused by tension. The exon-intron organization of the gene that encodes the α1(XII) polypeptide appeared to be similar to those of the genes for the α1 and α2 chains of type IX collagen (Gordon *et al.* 1989). Oh *et al.* (1992) have mapped the mouse *Col12a1* gene to chromosome nine.

Type XIV collagen has been characterized from human and chicken tissues and bovine skin (Dublet & van der Rest 1991, Gordon *et al.* 1991) and it has been shown that these molecules comprise three identical α1(XIV) chains. Structurally, type XIV collagen is very similar to type XII and both of these collagens are also proteoglycans. Type XIV collagen can be found in most of the same tissues as type I collagen, but also in cartilage (Aubert-Foucher *et al.* 1992, Gordon *et al.* 1991, Lunstrum *et al.* 1991, Wälchli *et al.* 1993). Type XIV collagen is found near the surface of collagen fibrils and may be involved in epithelial-mesenchymal interactions as well as in modulation of the biomechanical properties of tissues (Berthod *et al.* 1997).

The nomenclature of the human counterpart of type XIV collagen is ambiguous. The human protein has been isolated and its cDNA has been characterized but the protein is called undulin in the literature and interactions between undulin and other extracellular matrix components have been described (Ehnis *et al.* 1996, Ehnis *et al.* 1997) and the COL14A1 was localized to 8q23 by Schnittger *et al.* (1995). There are also studies that

suggest that the 5'-end of undulin transcript is subject to alternative splicing and also that undulin cDNA is similar to the chicken type XIV collagen (Gerecke *et al.* 1993, Just *et al.* 1991, Schuppan *et al.* 1990, Trueb & Trueb 1992).

Pan *et al.* (1992) and Yamaguchi *et al.* (1992) cloned the cDNA for a new type of collagen which shares certain features with members of the FACIT group. The collagen chain encoded by the cDNA was designated the α1 chain of type XVI collagen and its gene was localized to chromosome 1 by *in situ* hybridization. Fluorescence immunohistochemistry of adult mouse tissues using an affinity-purified antibody revealed a broad distribution of the type XVI collagen (Lai & Chu 1996) and indirect immunofluorescence microscopy studies showed an extracellular distribution of type XVI collagen, which is located close to cells but not associated with fibrillar structures (Grassel *et al.* 1996).

Yoshioka et al. (1992) isolated novel collagen-like cDNA clones from a human rhabdomyosarcoma cDNA library. These clones were predicted to encode a polypeptide containing a distinct collagenous domain with interruptions. This polypeptide was named α1(Y) but the name was later changed to α1(XIX). It has been proposed that type XIX collagen plays a role distinct from that of other FACITs, and may be involved in the assembly of embryonic matrices and maintenance of specific adult tissues (Sumiyoshi *et al.* 1997). The initial analysis of the 12 kbp transcript was done by Myers et. al (1993, 1994) and the complete primary structure was determined by Inoguchi *et al.* (1995). The length of the human COL19A1 gene is over 250 kbp and it is located near the COL9A1 and COL12A1 genes. The gene encoding type XIX collagen has been localized to chromosome 6q12-q13 in the human and to chromosome 1 (band A3) in the mouse (Gerecke *et al.* 1997, Khaleduzzaman *et al.* 1997).

2.4. Collagens with multiple interruptions

Collagen types XV and XVIII form a homologous subgroup within the collagen family (Pihlajaniemi & Rehn 1995, Rehn & Pihlajaniemi 1996). These collagens are characterized by multiple interruptions in their collagenous domains and it has been suggested that they should be termed MULTIPLEXINs (Oh *et al.* 1994a). The complete primary structure of human type XV collagen and the corresponding genomic organization have been determined. The gene is 145 kbp in length and its chromosomal location is 9q21-q22 (Hägg *et al.* 1998a, Kivirikko *et al.* 1994, Muragaki *et al.* 1994, Oh *et al.* 1994b). Type XV collagen has been shown to be expressed in the heart, skeletal muscle, placental tissues, kidney and pancreas (Kivirikko *et al.* 1995, Muragaki *et al.* 1994, Myers *et al.* 1992). The murine gene for type XV collagen was mapped to chromosome 4 (Hägg *et al.* 1997).

Type XVIII collagen was discovered simultaneously by three laboratories in 1994 (Abe *et al.* 1993, Oh *et al.* 1994b, Rehn & Pihlajaniemi 1994). The cDNA derived structure has been determined for mouse and human (Rehn *et al.* 1994, Saarela *et al.* 1998a). The genomic organization of the mouse gene was found to be more than 102 kbp in length (Rehn *et al.* 1996). The chromosomal locations of the human and mouse genes

for type XVIII collagen were determined by Oh *et al.* (1994a) (Table 1). Analyses of mouse and human tissues have shown this collagen to be present along the basement-membrane zones of the blood vessels in intestinal villi, the choroid plexus, skin, liver, kidney, heart, placenta, prostate, ovaries and skeletal muscle (Muragaki *et al.* 1995, Saarela *et al.* 1998a, Saarela *et al.* 1998b).

The functions of types XV and XVIII collagen are currently unknown. In the case of type XVIII collagen, a 20-kDa fragment is derived from its C-terminus by specific proteolytic cleavage. This fragment, called endostatin, has been reported to be antiangiogenic and appears to specifically inhibit endothelial proliferation (O'Reilly *et al.* 1997).

2.5. Collagens that contain transmembrane domains

2.5.1. Type XVII collagen

Type XVII collagen is the only collagen that has been isolated as a direct result of investigating a human disease called bullous pemphigoid (Borradori & Sonnenberg 1999, Fukai *et al.* 1994, Hopkinson & Jones 1996). The work of Li *et al.* (1993a) indicated that the 180-kD bullous pemphigoid antigen is a transmembrane hemidesmosomal collagen, which was designated type XVII. This collagen was first described in 1991 (Li *et al.* 1991) and the complete sequences of the human and mouse cDNAs have been determined (Li *et al.* 1991, Li *et al.* 1993a, Myers *et al.* 1993). Gatalica *et al.* (1997) cloned the entire human COL17A1 gene and elucidated its intron/exon organization. The human and murine genes for the α1 chains of type XVII collagen were mapped to chromosome 10q24.3 and the distal end of chromosome 19, respectively (Copeland *et al.* 1993, Giudice *et al.* 1991, Giudice *et al.* 1992, Li *et al.* 1991, Li *et al.* 1993b, Li *et al.* 1993a).

2.5.2. Type XIII collagen

2.5.2.1. Primary structure

Type XIII collagen was found in a cDNA library derived from the HT-1080 cell line (Pihlajaniemi *et al.* 1987). Four overlapping cDNA clones were characterized and they were predicted to encode a low molecular weight human collagen. The corresponding mouse cDNA was cloned by Hägg *et al.* (1998b), who confirmed that the murine type XIII collagen contains three collagenous (COL1-COL3) and four noncollagenous domains (NC1—NC4). The NC1 domain contains a highly hydrophobic region, which was not originally noticed in the human sequence indicating that this collagen could be a

transmembrane protein. Type XIII collagen molecule was then showed to localize at plasma membranes in such a way that NC1 domain consists of a 36 amino acid-residue intracellular domain, a 23 amino acid-residue transmembrane region and the first 61 residues of the ectodomain. The rest of the molecule belongs to the ectodomain which consists of a 103 residue COL1 domain, a 50 residue NC2 domain, a 172 residue COL2 domain, a 22 residue NC3 domain, a 254 residue COL3 and a 18 residue NC4 domain. There are interruptions in the Gly-X-Y sequence in every COL-domain. To be more precise there is one interruption in the COL1 domain, two in the COL2 domain and one in the COL3 domain. The total lengths of mouse and human type XIII polypeptides are 739 and 731 residues, respectively (Hägg *et al.* 1998b) but these lengths may vary due to alternative splicing of the primary transcript (see below). The cDNA-derived molecular weight for mouse type XIII collagen is 72110 daltons.

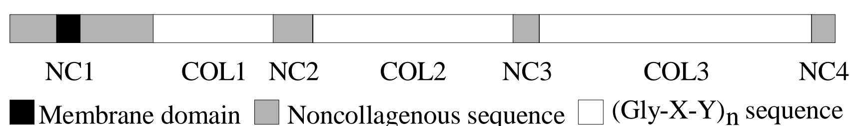

Figure 1. Schematic drawing of type XIII collagen polypeptide

2.5.2.2. Organization and chromosomal localization of the human type XIII gene

Part of the human gene for the α1 chain of type XIII collagen was cloned by Tikka *et al.* (1991) and the COL13A1 was defined to span over 140 kbp. The analysis of this gene revealed an interesting feature, namely the size of its exons. Usually the exons in the genes that encode fibrillar collagens are 54 bp in length but in the gene for type XIII collagen ten out of the 42 exons are 27 bp in length. Tikka *et al.* (1988) defined human type XIII collagen gene to contain 39 exons but further work indicate that the correct number of exons is 42 (Hägg *et al.* 1998b). The sizes of the exons of the human type XIII collagen gene vary between 1088 and eight nucleotides. The largest exon is the first one and the shortest is exon 3. This 8-bp exon is today the shortest exon described in the collagen literature. The locus for COL13A1 was localized to chromosome 10q22 (Pajunen *et al.* 1989, Shows *et al.* 1989). All in all, the gene for the α1 chain of type XIII collagen has some features found in the genes for fibrillar collagens as well as some distinct features (Tikka *et al.* 1988).

2.5.2.2. Alternative splicing

The most characteristic feature of type XIII collagen is the extensive alternative splicing of its pre-mRNA (Pihlajaniemi *et al.* 1987). This phenomenon has been studied extensively since its discovery. A total of ten alternatively spliced exons, that apparently combine freely, can be found in this gene (Juvonen *et al.* 1992, Juvonen *et al.* 1993, Juvonen & Pihlajaniemi 1992, Peltonen *et al.* 1997, Pihlajaniemi & Tamminen 1990). Theoretically, 10 exons that combie freely can give rise to a total of 1024 different polypeptides, but only a fraction of this number is found when the mRNAs are analyzed. Analyses covering both the COL1 and NC2 domains of human mRNA demonstrate that at least 12 mRNA species exist through the alternations of exons 3B-5, 12, and 13. In the COL1 domain six of the 16 potential combinations of the exons 3B-5 were found to exist in mRNAs, and as a result, the length of the COL1 domain varied between 57 and 104 amino acid residues. Exons 12 and 13 encoding most of the NC2 domain were found to be alternatively spliced and because these exons never coexisted in the mRNAs the length of NC2 domain varied between 12 and 34 amino acid residues (Juvonen *et al.* 1992, Juvonen & Pihlajaniemi 1992). The length of the COL3 domain varies between 208 and 235 residues, because of alternative splicing of exons 29 and 33. The length of NC4 domain is 18 residues (Juvonen *et al.* 1993). Peltonen *et al.* (1997) found eight common variants and nine rare combinations of alternatively spliced exons in an analysis of long RT-PCR fragments when mouse mRNAs were studied. The lengths of the polypeptides encoded by the alternatively spliced transcripts varies from 651 to 710 residues.

In the case of other collagens alternative splicing of their pre-mRNAs usually takes place at different stages of embryogenesis (Nah *et al.* 1996, Oganesian *et al.* 1997, Pallante *et al.* 1996, Yoshioka *et al.* 1995) or is restricted to the processing of introns at the 5' or 3'-end of pre-mRNA (Feng *et al.* 1994), (Inoguchi *et al.* 1995, Muragaki *et al.* 1990b, Muragaki *et al.* 1995, Rehn & Pihlajaniemi 1995, Saitta *et al.* 1992, Sugimoto *et al.* 1994, Zanussi *et al.* 1992) but an example of alternative splicing of internal introns has also been found in the α1(VI) gene (Stokes *et al.* 1991). The factors that regulate the alternative splicing of collagen pre-mRNAs are not known. For reviews on alternative splicing involving collagens, see Boyd *et al.* (1993) and Sandell (1996).

2.5.2.1.3. Tissue distribution

An interesting feature of type XIII collagen is its expression pattern which includes a wide range of tissues, although the mRNA levels appear to be relatively low. The mRNA for type XIII collagen can be detected in bone, cartilage, intestine, skin and striated muscle by Northern hybridization. In addition, *in situ* hybridization with human fetal tissues has revealed it to be present in the epidermis, cartilage, hair follicles, intestinal mucosal layer cells, placenta and nail root cells of the skin. Furthermore, the positive mRNA signal was detected in the cells in the intestinal mucosal layer, cells forming the reticulin fibers of the bone marrow and endomycium and at the (pre)articular surfaces and at the margins of the epiphyses, whereas the hybridization signal was weaker in the

resting chondrocytes in the middle of the epiphyses. The hybridization signal obtained with the α1(XIII) collagen cDNA probe in cartilaginous areas of the growth plates was similar, but less intense, to that obtained with the type II collagen probe. (Juvonen *et al.* 1993, Sandberg *et al.* 1989). However, the methods used in these studies were not very sensitive and a later study that used an RT-PCR technique to analyze mouse tissues revealed the type XIII collagen mRNA to be present also in the liver, spleen, kidney, small intestine, lung, heart, skeletal muscle, skin, Achilles tendon, sciatic nerve and brain (Peltonen *et al.* 1997). In mouse tissues type XIII collagen was found to co-localize with vinculin in the myotendinous junctions and costameres of skeletal muscle and in the intercalated discs of the heart. Type XIII collagen was also detected on the basal side of the intestinal epithelium, in the epididymis and lung, in some capillaries and in the endoneural sheaths of peripheral nerves. These locations represent a selection of integrin-mediated adherens junctions, which are thought to be the closest equivalent to focal adhesions in tissues (Singer *et al.* 1984).

2.5.2.4. Subcellular localization

Analysis of cultured cells with antibodies against type XIII collagen suggested that it is anchored to the plasma membrane by a transmembrane segment located near its amino terminus (Hägg 1998). All the cell types studied had type XIII collagen in the focal adhesions at the ends of actin stress fibres and this collagen also co-localized with the known focal adhesion components talin and vinculin. This co-localization was also observed in rapidly forming adhesive structures of spreading and moving fibroblasts and in dissociating focal adhesions. These results combined with data on the tissue distribution of type XIII collagen and the studies on the recombinant protein strongly suggest that type XIII collagen has a cell adhesion-associated function in integrin-based cell matrix junctions. This is analogous to the role of type XVII collagen as an integral component of hemidesmosomal junctions in certain stratified squamous epithelia, but the tissue distribution of type XIII collagen of the adherens junctions is much wider (Gatalica *et al.* 1997, Hägg et al., manuscript). Type XIII collagen is thus a novel cell adhesion protein, and is predicted to contain structures that have mechanical strength as well as the potential to provide adhesion surfaces necessary for lateral interactions (Hägg 1998).

2.6. Mutations in collagen genes

Mutations in collagen genes have been characterized in many heritable disorders such as osteogenesis imperfecta, osteoporosis, osteoarthrosis, aortic aneurysms, several chondrodysplasias, several subtypes of the Ehlers-Danlos syndrome, X-linked and autosomal Alport syndrome and dystrophic forms of epidermolysis bullosa. Excessive collagen accumulation also poses a common problem in medicine, as it often leads to

fibrosis and impairment of the normal functioning of the affected tissue. The genetic factors underlying these diseases can be point mutations which cause amino acid substitutions, shifts in the reading frame or premature stop codons. Larger mutations, such as deletions and insertions, have also been found (Balarin *et al.* 1999, Prockop & Kivirikko 1995, Zhou *et al.* 1993). The nature of the diseases caused by mutations in the collagen genes is usually dominant. This dominant-negative effect of these mutations is probably due to the "molecular suicide": the three collagen α chains assemble into a trimeric molecule and if one of the chains is defective the whole trimer is unfunctional and is degraded.

Mutations have been identified in only 13 of the 33 collagen genes that have been described, which indicates that research on the genetic defects that affect collagens is still in its infancy. It is also possible that mutations in the collagen genes are lethal at embryos and fetuses are aborted before detection. The up-dated list of known mutations in collagen genes can be accessed at http://genome-www.stanford.edu/genecards/ (Rebhan *et al.* 1997) or from the Human Gene Mutation Database (Krawczak & Cooper 1997). For literature reviews, see Kivirikko (1993) and Kuivaniemi *et al.* (1997). Dalgleish (1997) described a database of mutations in the COL1A1 and COL1A2 genes, accessible on the World Wide Web (http://www.le.ac.uk/gene-tics/collagen/col1a1.html and http://www.le.ac.uk/genetics/ collagen/col1a2.html).

2.7. Generation of genetically altered mice

The genetically altered mice can be classified into three categories. The first category comprises mice that nature created by spontaneus mutation. In the second category man-made mutations are generated by random insertion of a gene construct into the host genome. The third category of transgenes provides a very powerful tool for investigators, as this methodology enables the targeted introduction of a gene construct that replaces the original gene.

Spontaneus mutations that affect the individual's genotype and possibly also phenotype are the basis of natural selection. Using this natural variation, man has bred all domestic animals and plants. There are also a wide spectrum of mouse lines that are based on spontaneus mutations that were detected. Modern transgenic mouse research began in the 1980's when the first man-made transgenic mice were generated. Those animals were established by microinjection of cloned DNA into the nuclei of fertilized mouse oocytes (Gordon *et al.* 1980), a technology developed from earlier work that used viruses to carry DNA into mouse embryos (Jaenisch 1976). With nuclear injection of DNA, germ-line transmission and expression of transgenes was achieved (Brinster *et al.* 1981, Costantini & Lacy 1981, Gordon & Ruddle 1981, Wagner *et al.* 1981a, Wagner *et al.* 1981b). This was followed by a report of a specific phenotype associated with transgene expression (Palmiter *et al.* 1982). From then on, transgenic mice have been used to address biological questions ranging from the action of oncogenes to the immune system and developmental gene regulation (Jaenisch 1988). The common feature of these studies is the method of integation of DNA into genome. Here the DNA is injected into

the oocyte and random ingration into the genome occurs. In the worst case, the events that are detected are due to inactivation of the another gene, but with careful analysis, this technique can be applied to the wide spectrum of research, e.g. overexpression of the gene, expression of the gene in inappropriate tissues, and expression of mutated DNA (Grosveld & Kollias 1992).

With gene targeting, a any sequence in the genome can specifically be altered, provided that the sequence was cloned and characterized. In this tecnique, DNA is modified and introduced into the pluripotent embryonic stem cells by homologous recombination. The mutated allele is transmitted to offspring through the mouse germline and homozygous mutant animals are derived from intercrosses of heterozygous animals.

The technical development that enabled the manipulation of the mouse germline offered a new platform for clinical and basic research. This technology became available when Evans & Kaufman (1981) and Martin (1981) showed how to grow pluripotential embryonic stem (ES) cells in culture. If the totipotency of ES-cells is maintained, they can be returned into the mouse embryo and contribute to all tissues, including the germ line (Bradley *et al.* 1984). With this technology, DNA under investigation is introduced into the nuclei of the ES cells using electrotransfection. The foreign DNA then integrates into the nuclei of ES cells, and if this DNA carries selectable markers the new phenotype can be observed (Gossler *et al.* 1986, Hooper *et al.* 1987, Kuehn *et al.* 1987, Robertson *et al.* 1986). The most significant development in this technology was the ability to target specific genes in ES-cells (Doetschman *et al.* 1987, Thomas & Capecchi 1987), which enabled the generation of loss-of-function mutations in the mouse (Schwartzberg *et al.* 1989, Zijlstra *et al.* 1989).

The second advance of targeting studies was the use of *Cre* recombinase (for a review, see Galli-Taliadoros *et al.* 1995). The use of *Cre*-enzyme which recognizes the *loxP*-sites of P1-phage enabled also so called “knock-in” and “conditional knock-out” experiments where gene DNA be inserted to the exact position of genome or, as in latter example, the gene can be inactivated at selected tissues. (Gu *et al.* 1994, Kuhn *et al.* 1995).

The development of gene targeting protocols that involve homologous recombination in mouse ES-cells has resulted in the production of a considerable number of mutant mouse lines with specific phenotypes. Several hundred such “targeted” mice have now been created (Brandon *et al.* 1995c, Brandon *et al.* 1995a, Brandon *et al.* 1995b, Jasin *et al.* 1996, Müller 1999). See also TBASE, http://tbase.jax.org/. These animals offer an unparallel tool for detailed studies of specific diseases and for understanding the genetic aspects of our daily life.

2.8. Animal models in collagen research

2.8.1. Spontaneous mutations

Three mouse lines with spontaneous mutations in the collagen genes have been characterized. (Table 3). *cho* mice have chondrodysplasia and mice homozygous for this autosomal recessive mutation die at birth with abnormalities in the cartilage of limbs, ribs, mandible and trachea. The mutation responsible for this phenotype is the deletion of a cytidine residue approximately 570 nucleotides downstream of the translation initiation codon in the α1(XI) mRNA. This causes a shift in the reading frame and a premature stop codon (Li *et al.* 1995b, Seegmiller *et al.* 1971).

The second spontaneous mutation, *Dmm*, results in disproportionate micromelia. The phenotype of the mice homozygous for this autosomal dominant mutation with incomplete penetrance includes disproportionate micromelia, thoracic dysplasia and cleft palate (Brown *et al.* 1981, Pace *et al.* 1997). The chondrocytes of the epiphyseal growth plates are not organized into columns, and ultrastructural analysis reveals excessive dilation of the endoplasmic reticulum and a paucity of collagen fibrils in the extracellular matrix. Mapping studies revealed tight linkage to locus *Col2a1* and sequencing of the α1(II) collagen cDNA from affected animals revealed a three-nucleotide deletion in the region that encodes the globular domain of the C-terminal propeptide. Mice bearing the *Dmm* mutation have served as a useful model not only for the pathogenesis of several human chondrodysplasias but they have also revealed novel insights into normal skeletal morphogenesis.

Table 3. Natural collagen mutations in mice

Locus	Mouse disorder	Mutation	Mutant phenotype	Human disorder
cho	chondrodysplasia	single base deletion in *Col11a1*	absence of α1(XI) collagen in cartilage, abnormally thick collagen fibrils, short spine and long bones cleft palate, completely disorganized growth plate in cho/cho mice	lethal chondrodys-plasia
Dmm	disproportionate micromelia	three nucleotide deletion in *Col2a1*	reduced collagen II content; shortened long bones and spine, small rib cage, growth plate abnormalities	Stickler syndrome
oim	osteogenesis imperfecta murine	single base deletion in *Col1a2*	absence of α2(I) collagen fibrils in skin and bone progressive skeletal deformities, bone fractures, osteopenia	osteogenesis imperfecta type III

The third mutation, *oim,* is a nonlethal recessively inherited mutation which gives rise to phenotypic and biochemical manifestations that are similar to human osteogenesis imperfecta of the moderate to severe types. The phenotype of mice homozygous for *oim* includes skeletal fractures, limb deformities, generalized osteopenia and small body size. Labeling studies have demonstrated an absence of α2(I) collagen chains from the fibroblasts of *oim* mice. Nucleotide sequencing of the cDNA encoding the C-terminal propeptide of the α2(I) chain revealed the deletion of a G residue at nucleotide position 3983, which results in an alteration of the sequence of the last 48 amino acid residues (Chipman *et al.* 1993).

2.8.2. Transgenic mice with dominant-negative mutations

Mov-13, the first transgenic mouse with a dominant-negative mutation in a collagen gene, was generated by retroviral insertion. The integration of a murine retrovirus in the first intron of the *Col1a1* gene resulted in a null allele that is blocked at the level of transcription. Heterozygous mice containing the null allele and wild-type alleles have served as a useful model for osteogenesis imperfecta type I (Bonadio *et al.* 1990, Schnieke *et al.* 1983)(Table 4).

The way the three collagen α chains assemble into a trimeric molecule gives an indispensable tool for collagen researcher allowing a “molecular suicide” strategy to be employed when generating animal models for collagen disorders. Applying this technique the entire collagen molecule can be disrupted in a dominant negative manner by changing only one of the α chains. Khillan *et al.* (1991) used such a strategy to disrupt type I collagen by producing a “mini” version of the gene for the human proα1(I) chain, modeled after a sporadically occurring in-frame deletion that results in a lethal variant of osteogenesis imperfecta in human. The shortened proα1(I) chains were found to associate with the murine wild-type proα1(I) and proα2(I) chains resulting in abnormal fibers, which leads to degradation of both the mutant and the wild-type polypeptides.

Similar studies were conducted on type II collagen by Vandenberg *et al.* (1991) who produced a series of “minigene” versions for human type II procollagen that lacks sections of varying sizes from the molecule. One of the constructs, where a large central region containing 12 of the 52 exons is deleted, was used to generate five lines of transgenic mice expressing the transgene. A large proportion of the mice developed chondrodysplasia with dwarfism and had short and thick limbs, short snout, cranial bulge, cleft palate and delayed mineralization of bone. A number of the animals died shortly after birth. These results were later confirmed by Helminen *et al.* (1993) in independent studies. Metsäranta *et al.* (1992) generated a transgenic mouse line with a 15-amino acid deletion in the *Col2a1* gene. These animals had a severe chondrodysplastic phenotype which included short limbs, a hypoplastic thorax, abnormal craniofacial development and other skeletal deformities. The affected pups died at birth due to respiratory distress apparently caused by a marked reduction in cartilagenous extracellular matrix, disruption of the normal organization of the growth plate, a severe reduction in the amount of cartilage collagen fibrils and abnormalities in their structure.

Another collagen studied by the "molecular suicide" stategy was type IX collagen. This collagen is a heterotrimer made up of three distinct polypeptides, α1(IX), α2(IX) and α3(IX). Nakata *et al.* (1993) generated a mouse line with a truncated α1(IX) chain whose expression was controlled by a tissue-specific promoter/enhancer element. The offspring of two different founders were found to have pathological changes similar to osteoarthritis in the articular cartilage of the knee joints. In addition, mice homozygous for the transgene developed mild chondrodysplasia.

A slightly different approach to procollagen "suicide" was employed by Jacenko *et al.* (1993) and Chung *et al.* (1997), who produced a transgenic mouse line which carried a large in-frame deletion generated into chicken type X collagen. These truncated collagens affected to the formation of native chains by disrupting either the triple-helix formation or fibril formation. The mice developed skeletal and craniofacial deformities which included compression of the hypertrophic growth plate cartilage. They also had decreased formation of new bone, leukocyte deficiency in the bone marrow, lymphopenia and a reduction in the size of the thymus and spleen.

A dominant-negative phenotype can be generated not only with large deletions but also by targeting point mutations to certain regions in collagen molecules. Stacey *et al.* (1988) showed that transgenic mice bearing an altered α1(I) collagen gene with a specific glycine to cysteine substitution has a dominant-lethal phenotype similar to the human disease osteogenesis imperfecta type II. The transgenic mouse line was constructed in a Mov-13 background that may have contributed to the severe phenotype. These authors also demonstrated that only a small amount of the mutant protein (accounting for 10 % of the type I collagen pool in the cells) is needed to disrupt normal collagen function. Liu *et al.* (1995) showed that mice with mutations in the collagenase cleavage site of the α1(I) chain develop marked fibrosis of the dermis similar to that in human scleroderma. The third example of the effects of point mutations is provided by a glycine to cysteine mutation at residue 85 of the mouse type II collagen. The offspring of different founders displayed phenotypes similar to severe chondrodysplasia, characterized by short limbs and trunk, cranio-facial deformities and cleft palate. The affected pups died at birth because they were unable to inflate their lungs. The cartilage anomalies displayed by these transgenic mice are remarkably similar to those of certain human chondrodysplasias (Garofalo *et al.* 1991).

Overexpression of the gene of interest can also be used to model disease. Garofalo *et al.* (1993) showed the applicability of this approach to the study of complex multicomponent protein assemblies. Their results show that overexpression of the wild-type mouse α1(II) collagen led to the development of thick, abnormal collagen fibrils. Furthermore, the highest expression level of the transgene was associated with the highest proportion of abnormal fibrils and the affected pups died at birth. The authors proposed that an imbalance in the constituents of the cartilage collagen fibrils might disrupt the mechanism that controls fibril assembly.

Table 4. Transgenic mice created via retrovirus insertion or pronuclear microinjection

Altered/Expressed locus	Mutant mouse phenotype	Human disease Reference
Retrovirus insertion into the first intron of *Col1a1* gene	Embryonic lethality in homozygous mice; reduced collagen collagen type I content associated with hearing loss and decreased mechanical strength of long bones in heterozygous mice	Osteogenesis imperfecta type I (Bonadio *et al.* 1990)
Glycine to cysteine substitution in *Col1a1* cDNA	Perinatal lethality in founder mice, short and wavy ribs, short and broad long bones, poor mineralization	Osteogenesis imperfecta type II (Stacey *et al.* 1988)
Central in-frame deletion in COL1A1cDNA	Lethal phenotype with with fractures of ribs and long bones in mice expressing high levels of transgene; Fractures and reduced collagen and mineral content in mice expressing moderate levels of transgene	Osteogenesis imperfecta type II (Khillan *et al.* 1991) Osteoporosis (Pereira *et al.* 1993)
Glycine to cysteine substitution in mouse *Col2a1* cDNA	Perinathal lethality, cleft palate, short limbs and trunk, craniofacial abnormalities, disorganization of growth plate	Spondyloepi-physeal dysplasia (Garofalo *et al.* 1991)
15 amino acid deletion in *Col2a1* cDNA	Perinathal lethality and skeletal deformities characterized by short limbs, cartilage hypoplasia and fragility, cleft palate, disorganization of growth plate	Spondyloepiphyseal dysplasia (Metsäranta *et al.* 1992)
Central in-frame deletion in COL2A1 cDNA	Perinathal lethality associated with dwarfism, cranial bulge and cleft palate in the most severely affected mice; Degenerative changes of articular cartilage in viable, older mice	Chondrodysplasia (Vandenberg *et al.* 1991) Osteoarthritis (Helminen *et al.* 1993)
In-frame deletion in *Col9a1* cDNA	Mild proportionate dwarfism, eye abnormalities, degenerative changes in articular cartilage	Chondrodysplasia Osteoarthritis (Nakata *et al.* 1993)
Central in-frame deletions in chicken α1(X) collagen cDNA	Skeletal deformities including mild dwarfism neck lordosis and thoracolumbar kyphosis, compression of the hypertrophic region of growth plates	Spondylometaphyseal dysplasias (Jacenko *et al.* 1993) Metaphyseal chondrodysplasia (Chung *et al.* 1997)

Phenotypic variability and incomplete penetrance of a trait are frequently observed in human monogenic diseases such as osteogenesis imperfecta. Pereira et al. (1993, 1994) demonstrated that phenotypic variability and incomplete penetrance were not explained by variations in genetic background or levels in gene expression by expressing a mutant

COL1A1 gene in an inbred mouse strain. Instead, the authors suggested that phenotypic variation is an inherent feature of the expression of mutated collagen genes.

2.8.3. Targeted mutations in collagen genes

The inactivation of a gene by gene targeting can be used for loss of function studies. In collagen research these kinds of studies have been done for collagen chains α1(II), α1(III), α3(IV), α2(V), α2(VI), α1(IX), α1(X), α1(XV) and α1(XVIII) (Table5).

Homologous recombination in embryonic stem cells was used to generate a transgenic mouse with an inactivated *Col2a1* gene. Homozygous mice developed into fetuses that were born vaginally but died either just before or shortly after birth. The cartilage in these mice consisted of highly disorganized chondrocytes and had a complete lack of extracellular fibrils as discerned by electron microscopy. These results demonstrated that a well-organized cartilage matrix is required as a primary tissue for the development of some components of the vertebrate skeleton, but is not essential for others (Li *et al.* 1995b) (Table 5).

In the experiment where murine *Col3a1* gene was inactivated, about 10% of the animals completely lacking the α1(III)-chain survived to adulthood but had a much shorter lifespan compared with wild-type mice. The major cause of death in the mutant animals was rupture of the major blood vessels, similar to patients with type IV Ehlers-Danlos syndrome. Ultrastructural analysis of tissues from the mutant mice revealed that type III collagen is essential for normal type I collagen fibrillogenesis in the cardiovascular system and other organs (Liu *et al.* 1997).

Cosgrove *et al.* (1996) produced a mouse model for the autosomal form of Alport syndrome by generating a *Col4a3* knockout. The mice developed progressive glomerulonephritis with microhematuria and proteinuria and end-stage renal disease could be seen at about 14 weeks. Immunofluorescence analysis of the glomerular basement membrane showed the absence of type IV collagen α3(IV), α4(IV) and α5(IV) chains as well as the persistence of α1(IV) and α2(IV) chains, which normally locate only in the mesangial matrix. Later the authors reported ultrastructural, physiological and molecular defects in the inner ear of a knockout mouse model for autosomal Alport syndrome (Cosgrove *et al.* 1998).

In order to test how type V collagen controls the physicomechanical properties of collagen fibrils, Andrikopoulos *et al.* (1995) generated mice with a structurally abnormal α2(V) collagen chain by deleteting exon 6 (Table5). Mice homozygous for the mutation survived poorly, possibly because of complications from spinal deformities. In addition, these mice exhibited skin and eye abnormalities caused by disorganized type I collagen fibrils. These results demonstrate that type V collagen is a key determinant in the assembly of tissue-specific matrices and provide an animal model for human connective tissue disorders.

A mouse model for human disease Bethlem myopathy is a mouse line that lacks type VI collagen (Bonaldo *et al.* 1998). The homozygous mutant mice showed histological features of myopathy such as fiber necrosis, phagocytosis and a pronounced variation in fiber diameter. Their muscles also showed signs of stimulation of fiber regeneration. An

interesting feature of this study is that it also indicates haploinsufficiency of the *Col6a1* gene function because similar, although milder, alterations were detected in heterozygous animals. The haploinsufficiency of COL6A1 gene is found also in human patients (Lamandé *et al.* 1998, Pepe *et al.* 1999). These data show that collagen VI is necessary for the maintenance of the integrity of muscle fibers. The phenotypes of the knock-out mice can be quite surprising, as in the case of type IX and X knockouts.

Table 5. Mice with targeted mutations in collagen genes

Inactivated/Altered locus	Mutant mouse phenotype	Human disease Reference
Null mutation for *Col2a1*	Perinatal lethality of homozygous mice, cleft palate, abnormal cartilage, disorganized growth plate, lack of bone marrow	Achondrogenesis type II (Li *et al.* 1995b)
Null mutation for *Col3a1*	Much shorter life span compared with wild-type mice, rupture of the major blood vessels, disturbed fibrillogenesis	Ehlers-Danlos syndrome type IV (Liu *et al.* 1997)
Null mutation for *Col4a3*	Postnathal lethality of null mutants due to renal failure; progressive glomeluronephritis with proteinuria and microhematuria	Autosomal Alport syndroma (Cosgrove *et al.* 1996)
Deletion of exon 6 in *Col5a2*	Spinal deformities, skin and eye abnormalities associated with the disorganization of dermal and corneal collagen fibrils	Ehlers-Danlos syndrome (Andrikopoulos *et al.* 1995)
Null mutation for *Col6a1*	Histological features of myopathy such as fiber necrosis and phagocytosis and a pronounced variation in the fiber diameter, stimulated regeneration of fibers.	Bethlem myopathy (Bonaldo *et al.* 1998)
Null mutation for *Col9a1*	Degenerative changes of articular cartilage of older homozygous mice	Multiple epiphy-seal dysplasia (Hagg *et al.* 1997)
Null mutation for *Col10a1*	No gross alterations in skeletal development; abnormal distribution of cartilage matrix components, altered bone content, *coxa vara*	Schmid metaphyseal chondrodysplasia (Rosati *et al.* 1994)
Null mutation for *Col15a1*	Mild muscle disorder	(Eklund *et al.* 1998)
Null mutation for *Col18a1*	No phenotype observed	(Fukai,N. unpublished)

The mouse strain lacking type IX collagen viable and shows no detectable abnormalities at birth but later develops a severe degenerative joint disease that resembles human osteoarthritis (Hagg *et al.* 1997).

The knock-out for type X collagen had no phenotypic effects at all, the mice were viable and fertile and had no gross abnormalities in the growth or development of the long bones (Rosati *et al.* 1994). This lack of phenotype is surprising because the phenotype of the dominant-negative mouse was quite severe. However, in 1997 Kwan *et al.* reported the phenotype of the type X collagen null-mouse to be similar to human Schmid metaphyseal chondrodysplasia (SMCD). This case shows that the development of the analytical tools can help uncover phenotypes which were classified as normal before. To date, null mutations have also been generated for collagen types XV and XVIII but they have not been associated with pronounced phenotypes (Eklund *et al.* 1998, Fukai,N. unpublished).

3. Outlines of the present research

When this work started type XIII collagen was characterized partially. Later the primary structure of this protein was elucidated but the genomic organization of the gene remained unclear. The apparent co-localization of genes for human type XIII collagen and prolyl 4-hydroxylase α-subunit raised the question of the physical distance between these genes. The most important question to be answered was that concerning the function of type XIII collagen. To find out the genomic organization and facilitate the functional studies this gene, the following aims were set:

1. to isolate and characterize the genomic clones for the mouse *Col13a1* gene, to determine the genomic organization, transcriptional initiation sites and chromosomal localization of this gene and to compare it with the human homologue,
2. to determine the relative transcriptional orientation of the human COL13A1 and P4HA genes, and the physical distance between them; and
3. to produce mice with an altered type XIII collagen gene, and to analyze the consequences of the alteration.

4. Materials and methods

The detailed descriptions of the materials and methods are explained in original articles I-III.

4.1. Characterization of genomic organization (I, II)

A mouse genomic library (951303; Stratagene, La Jolla, CA) in the cosmid vector pWE15 was screened with a mouse type XIII collagen cDNA clone according to standard protocols (Sambrook *et al.* 1989). In addition, a 129Svj mouse genomic library (961301; Stratagene) was screened to obtain the 5' end of the type XIII collagen gene using a 5' PCR fragment as a probe since positive clones were not found from the other library. The same library was also screened for the 3' end using an oligonucleotide from exon 41 as the probe. Positive clones were picked out and DNAs were isolated by the alkaline lysis method. P1 clones were obtained from Genome Systems Inc. (St. Louis, MO) by PCR screening of P1 libraries with oligonucleotides corresponding to intron sequences adjacent to exons 1, 7 and 15. A 5'-end genomic clone of human type XIII collagen was obtained by screening a human genomic library (951202; Stratagene) using an *Eco* RI-fragment from clone CL412 as a probe. This clone, and the 3'-end genomic clone, D2, were generous gifts from Dr. Karl Tryggvason (University of Oulu, Oulu, Finland) (Tikka *et al.* 1991).

Inserts were released from the cosmid clones by *Not* I digestion and from the P1 clones by *Not* I and *Sfi* I digestions for restriction mapping, after which the DNAs were partially digested with the restriction enzymes *Bam* HI, *Eco* RI, *Hind* III or *Xba* I. The partially digested fragments were then separated by pulse field electrophoresis. The size and order of the fragments were defined by Southern blotting and hybridization with the insert-end-specific oligonucleotides T7, T3 or SP6 (Evans *et al.* 1989). Complete digestion of the released inserts was used to certify the restriction fragment sizes. The final restriction map was achieved by subcloning and sequencing the regions of the gene which could not be sequenced directly using cosmid or P1-clones as a template. The approximate exon locations were determined by Southern blotting and hybridization of

complete digestions with exon-specific oligonucleotides. More accurate intron sizes were determined by amplifying the intron areas by PCR using exon-specific primers and separating the products by electrophoresis.

The intron-exon boundaries were sequenced directly from the cosmid and P1 clones using an ABI automatic sequencer (Perkin-Elmer, Norwalk, CT), the ^{T7}Sequencing™ Kit (Amersham Pharmacia Biotech, Uppsala, Sweden) and the Cycle Sequencing Kit (Amersham Pharmacia Biotech) (Sanger *et al.* 1977). The gene areas that could not be sequenced directly from the P1 clones were examined by subcloning *Hind* III and *Eco* RI fragments to the Bluescript™ (Stratagene) vector and sequencing them with the ^{T7}Sequencing™ Kit (Amersham Pharmacia Biotech).

DNA from tail, liver or spleen was extracted according to standard protocols to genotype the transgenic mice (Sambrook *et al.* 1989). The PCR amplifications were performed using three primers, one common primer pairing with two distinct primers, in a single reaction. Each pair was specific for either the wild-type (+) or the transgene (-), making it possible to distinguish -/-, +/- and +/+ genotypes in a single reaction. For Southern hybridizations liver or spleen DNA was digested with *Xba* I, after which the hybridizations were done as described in Sambrook *et al.* (1989).

4.2. Isolation of RNA and RT-PCR (I, III)

Total RNA was isolated from cultured skin fibroblasts by acid guanidinium isothiocyanate-chloroform-phenol extraction (Chomczynski & Sacchi 1987). The RNA (6 µg) was transcribed into single-stranded DNA using a type XIII collagen-specific oligonucleotide primer. One microliter of the RT reaction was used as a template in PCR amplification using *Col13a1*–specific sense and antisense oligonucleotide primers. In negative control samples the template was omitted. The PCR products were separated electrophoretically on agarose gels and the DNA was recovered from agarose by phenol extraction and sequenced.

4.3. Northern analysis (I)

A mouse Multi-Tissue Northern blot (Clontech Laboratories, Inc. Palo Alto, CA) prepared by gel electrophoresis of 2 µg/lane of poly A^+ RNA isolated from various adult mouse tissues was hybridized with a α-^{32}P-labeled 756-bp *Pvu* II fragment of mouse type XIII cDNA with ExpressHyb™ (Clontech) hybridization solution according to protocols from the manufacturer's manual (Yang & Kain 1995).

4.4. Nuclease S1 and 5'-RACE assays (I)

S1 nuclease probe protection experiments were performed as described previously using a 5' end-labeled PCR fragment flanking 5'-end of the gene as a probe (Pihlajaniemi & Myers 1987, Sambrook *et al.* 1989). The double-stranded probe was hybridized to 20 µg of total RNA from mouse lung. After hybridization, S1 buffer was added, and the mixture was digested with S1 nuclease (Boehringer Mannheim GmbH, Mannheim, Germany) at room temperature. The protected fragments were analyzed on a polyacrylamide gel. The total RNA from adult mouse lung was a generous gift from Dr. Sirkku Peltonen (University of Turku, Turku, Finland) and was isolated by guanidium isothiocyanate-chloroform-phenol extraction (Chomczynski & Sacchi 1987).

To prepare 5'RACE cDNA clones corresponding to the initiation of transcription, blunt-end cDNA was generated using a 5' untranslated region primer, poly A^+ RNA from a 18.5 d mouse embryo and the TimeSaver cDNA synthesis kit (Amersham Pharmacia Biotech) as described in Rehn and Pihlajaniemi (1995). After linker ligation and two successive rounds of PCR amplification, the product was digested and cloned. Positive colonies were lysed and their DNA was transferred to nitrocellulose filters, which were then screened using a genomic sequence corresponding to the 5' untranslated region of the mouse type XIII collagen gene as the probe. Positive colonies were picked and the isolated DNA was sequenced using the ^{T7}Sequencing™ Kit (Amersham Pharmacia Biotech).

4.5. Chromosomal localization (I,II)

4.5.1. Determination of chromosomal location by linkage using single-strand conformational polymorphism

Primers were designed to amplify a region corresponding to the intron 25 sequences of *Col13a1* in order to test for single strand conformation polymorphisms (SSCPs) between the mouse strains. These were analyzed as described in Beier (1993). Briefly, genomic DNAs from a series of mouse strains were amplified with radiolabeled oligonucleotides, denatured and electrophoresed in a 6 % non-denaturing acrylamide sequencing gel. One set of primers were used to identify a series of polymorphisms between the C57BL/6J and *M. spretus*, which were then used to analyze DNA prepared from the BSS backcross (Rowe *et al.* 1994). The strain distribution pattern was analyzed using the Map Manager Program (Manly 1993).

4.5.2. Fluorescence in-situ hybridization

The mouse fibroblast cell line L-929 (DSM, Germany) was used as a source of metaphase chromosomes. Cells were cultured according to standard protocols and treated with 5-

bromodeoxyuridine (BrdU) to induce a banding pattern (Lemieux *et al.* 1992, Matsuda *et al.* 1992).

Two genomic clones specific to the mouse *Col13a1* gene were labeled by nick translation and the FISH procedure was carried out as described earlier (Lichter *et al.* 1988, Pinkel *et al.* 1988, Tenhunen *et al.* 1995). Specific hybridization signals were visualized using FITC-conjugated avidin (Vector Laboratories, Burlingame, CA) and the slides were counterstained with DAPI (4'-6'-diamino-2-phenylindole, 0.025 µg/ml). Multicolor digital image analysis system was used to acquire, display and quantify the hybridization signals from the metaphase chromosomes (Heiskanen *et al.* 1996a).

The transcriptional orientation of the human type XIII collagen and prolyl 4-hydroxylase α subunit genes was studied using mechanically stretched chromosomes and probes specific for the 5' and 3' ends of both of these genes (Laan *et al.* 1995). Determination of the clone order was based on stretched chromosomes, interphase nuclei and fiber FISH (Heiskanen *et al.* 1996b).

4.6. Nucleotide sequence analysis (I, III)

Nucleotide sequence data were analyzed using either DNASIS (Amersham Pharmacia Biotech) or GCG Sequence Analysis Software Package (Genetics Computer Group Inc, Madison, WI) (Devereux *et al.* 1984). Consensus sites for the binding of transcription factors were searched for in the Transcription Factor Database using the GCG's FINDPATTERNS and GAP, BESTFIT, FASTA and PLOTSIMILARITY programs were used for alignments and homology searches. The human and mouse promoters were predicted with the PROSCAN (Prestridge 1995, see http://bimas.dcrt. nih.gov/molbio/proscan/) and TSSG programs (Solovyev V.V., Salamov A.A., Lawrence C.B unpublished data, see http://dot.imgen.bcm.tmc.edu:9331 /gene-finder/gf.html). A primer program was used to select the primer pairs (Rozen & Skaletsky 1997).

4.7. Construction of the targeting vector (III)

A genomic fragment which carries the promoter, the transcription initiation sequences, the first exon and a part of the first intron of the mouse gene for type XIII collagen was isolated and subcloned into the pSP72-vector (Promega, Madison, WI). The *loxP* sequence was amplified from the pBS*loxP*-vector (constructed by R. Fässler) by PCR using primers flanking the *Not* I recognition sequences. The PCR product was inserted into the unique *Not* I site at the 5' untranslated region. Subsequently, the cassette containing the neomycin resistance gene (*neo*), the herpes simplex virus tymidine kinase gene (*HSV-tk*) and the *loxP* sequences (a plasmid containing all these genes and sequences was constructed by R. Fässler) surrouding the marker genes was inserted into the unique *Sfi* I site which is located in the first intron of the *Col13a1* gene.

4.8. Generation of targetted ES cell clones and transgenic mice (III)

For ES cell targeting the *Cla* I-linearized targeting vector was electroporated into R1 ES cells (ES cells originally described in Nagy *et al.* 1993). After electroporation the cells were grown on a G418-resistant feeder layer of mouse embryonic fibroblasts in a medium containing 400 mg/l of G418. The colonies which were shown to have undergone homologous recombination were picked, recovered and DNA was characterized by Southern blots.

Two correctly targeted clones were chosen for *Cre* recombinase treatment. The cells were electroporated with the pIC-Cre plasmid (a generous gift from Dr. Werner Mueller, University of Cologne, Germany), which encodes the Cre-protein. Cells were cultured for four days, after which Ganciclovir (Cymevene, Syntex Nordica AB, Södertälje, Sweden) selection was started and continued for five days. Positive colonies were picked and characterized by Southern blots in order to identify altered and conditional types of recombinants.

Two correctly targeted clones in which the marker cassette was deleted by *Cre*, were re-expanded and injected into C57BL/6J blastocysts. Five chimeric males were obtained, and two of them transmitted the mutation into the offspring as mated with C57BL/6J and 129Sv females. Homozygous mice carrying two copies of the mutated alleles were produced by brother-sister mating of the heterozygous mice. The mice were bred and maintained in the animal facility of the University of Oulu.

4.9. Adhesion studies (III)

Fibroblast cultures were established from skin biopsies or 10.5 d embryos taken from normal 129Sv and homozygous mutant mice. The cells were grown under standard conditions.

One subconfluent plate from each of the five separate cell lines was used for adhesion studies. The fibroblasts were washed, detached, collected, centrifuged and diluted to 150 000 cells/ml. A 0.2 ml aliquot of this solution was transferred to a cell culture plates and the cells were allowed to attach for 20, 30, 40 or 60 minutes. After the attachment period the medium was sucked out and the plate was frozen and assayed with the CyQuant Cell Proliferation Assay Kit according to the manufacturers instructions (Molecular Probes, Eugene, Oregon, USA). The wells were counted with the Victor ELISA-plate reader (Wallac, Turku, Finland).

4.10. Light microscopic studies (III)

4.10.1 Histology

Heart, brain, muscles, lung, skin, liver, spleen, kidney and testis from 17 week old wild-type and mutant mice were isolated and frozen in liquid nitrogen. 10 µm sections were cut from the frozen tissues and stained with anti-XIII/NC3 antibodies and hematoxylin-eosin. For the muscle histology studies, both gastrocnemius and quadriceps muscles from 17 and 42 week old animals were prepared, mounted in embedding medium and frozen immediately in isopentane precooled in liquid nitrogen. 10 µm sections were cut and stained with anti-collagen IV (Chemicon International, Inc., Temecula, CA), anti-merosin (ProGen,Tustin, CA), anti-vinculin (Sigma), anti-tenascin-C (Sigma), anti-α5-integrin (Pharmingen, San Diego, CA) and anti-collagen XIII/NC3 antibodies (Hägg *et al.* 1998b).

4.10.1. Immunofluorescence staining of fibroblasts

For immunofluorescent staining, wild-type and type XIII collagen mutated fibroblasts were seeded and cultured to the desired density on sterilized glass coverslips for 1, 2, 3, 4, 9 and 12 hours. The cells were fixed and the anti-XIII/NC3, anti-talin and anti-vinculin antibodies were applied at the appropriate dilutions and incubated for 30 minutes at room temperature, followed by extensive washing with PBS. The porcine anti-rabbit fluoresceine (FITC) secondary antibodies were diluted according to the manufacturer's (DAKO A/S, Glostrup, Denmark) instructions and allowed to bind to the specimens for one hour at room temperature. After washing with PBS, the coverslips were mounted on microscope slides, viewed and photographed using a Leica Aristoplan microscope with the appropriate filter units.

4.11. Electron microscope studies (III)

For electron microscopic studies the gastrocnemius and quadriceps muscles of wild-type mice and mice homozygous for N-terminally altered type XIII collagen were used. The muscles were fixed, dehydrated in acetone and embedded in Epon LX112. Thin-sections were cut with a Reichert Ultracut E-ultramicrotome (Reichert-Jung, Wien, Austria) and examined in a Philips CM100 transmission electron microscope (Philips Export B.V., Eindhoven, The Netherlands) using an acceleration voltage of 80 kV.

5. Results

5.1. Exon-intron organization of the mouse α1(XIII) collagen gene

A total of seven overlapping genomic clones covering the entire mouse type XIII collagen gene were obtained. The first three clones were identified by screening a genomic cosmid library with a mouse cDNA fragment and they covered exons 14 – 41. Then, a cDNA-derived PCR fragment was used as a probe to screen for the 5' end of the gene and resulted in the isolation of a clone that contained the first exon, part of the intron 1 and the 5' flanking sequences. The last exon of this gene was obtained by screening the same library with an exon 41-specific oligonucleotide. This clone included exons 41-42 and the 3' flanking sequences. The rest of the type XIII collagen gene was obtained using a commercial screening service. Clone P1-K1, which was obtained by using oligonucleotides from the first exon, covered exons 1–7, and clone P1-K4 was obtained by PCR screening with oligonucleotides detecting exons 7 and 15. All in all, these clones cover over 180 kbp of the mouse genomic DNA, including 135 kbp corresponding to the type XIII collagen gene.

The mouse type XIII collagen gene was found to consist of 42 exons that range in size from 8 to 836 bp and the exon-intron boundaries conformed well to the consensus sequence AG-exon-GT. A Notable feature of this gene was the abundance of 27-bp exons, as these comprise 10 of the 42 exons. Five of the others were of 36 bp, another five of 45 bp, three of 54 bp and two 87 bp, while the size of remaining of the exons varied between 8 and 836 bp. The lengths of the introns were between 90 bp and 40 kbp. Nucleotide sequences were determined for all the exons, their intron junctions and intronic sequences for introns 10 – 15, 22 – 26 and 28 – 30. For introns 4 – 6, 8, 9, 16, 17, 19 – 21, 27, 31 – 38, 40 and 41 the sizes were determined by amplifying the intron areas by PCR using exon-specific primers. Introns 1, 2, 3, 7, 18 and 39 were too large to be amplified by PCR, so that their sizes were determined from the restriction map. Domain transitions from NC1 to COL1 domain occurred within the exon 2 and the subsequent transitions took place at the exons 13, 16, 26, 27 and 40, respectively. Exon 42 consists only of the termination codon and 3' untranslated sequences. With the exception of exons 3 and 33, the exons encoding only collagenous sequences begin with a

complete codon for glycine, which is characteristic for fibrillar collagen genes. Exon 3 begins with a split codon which codes for glycine and exon 33 begins with an interruption in the collagenous sequence where a codon for glycine is missing.

Human genomic clones covering the previously incompletely characterized region encompassing exons 3a – 5 were found when scanning the GenBank data bank for sequence similarities (GenBank accession number U82211) (Benson *et al.* 1998, Tikka *et al.* 1991). The GenBank accession numbers for the mouse type XIII collagen gene are AF063666-AF063693 and the new GenBank accession number for the 5' end of the human type XIII collagen gene is AF071009.

5.2. Analysis of 5'-flanking regions of mouse and human type XIII collagen genes and sequence comparison

S1 nuclease probe-protection mapping of the mouse gene showed several transcription initiation sites between 470 and 525 bp upstream of the translation initiation codon, but the main start site of transcription is 500 nucleotides upstream from the initiation ATG codon. Six 5'-RACE cDNA clones corresponding to the 5´ end of the mouse mRNAs were isolated, and all of them started 548 bp before the translation initiation codon.

The 5'-flanking areas of the mouse and human type XIII collagen genes were sequenced and compared. Analysis of 2642 bp of both flanking regions shows an overall identity of 58.3%, To see how the identity between the 5' untranslated and 5' flanking regions of the human and mouse genes is distributed, the optimally aligned sequences were plotted with computer programs. This showed the identity between these two 5' regions to be quite low but to increase just before the putative promoter areas. A search for potential promoters with two different computer programs, PROSCAN and TSSG, pointed to an identical position in both genes, which is the predicted direction of the promoter. 77.13% identity was found in the computationally proposed promoter and the adjacent 5'-untranslated areas but no significant homologies were detected upstream of the putative promoters. Collectively, these data suggest that the promoter of the mouse type XIII collagen gene is located between -864 and -568 bp upstream of the initiation ATG. The 5'-flanking region of the mouse type XIII collagen gene before the initiation of transcription does not contain a complete TATA box but contains a modified TATA-like box in the position -601 bp before the initiation codon. Lack of a real TATA box is in agreement with the presence of multiple transcription start sites observed downstream of this site. The 5'-most cDNA clone found here, starts 53 bp downstream of the modified TATA box. In the case of the human type XIII collagen gene, a proposed promoter area was to lie at area from -905 to -605 bp in the 5´ direction with respect to the initiation ATG, and programs proposed a modified TATA box to lie 636 bp before the initiation ATG. Furthermore, the 5'-flanking region contains numerous consecutive GC boxes both in the human and mouse genes. This result changes the location of human promoter suggested by Tikka *et al.* (1991) and places the promoter further in the 5' direction.

The putative promoters were also analyzed for the presence of transcription factor binding sites using the Transcription Factor Data Base (TFDB). A number of consensus motifs for several transcription factors were identified.

5.3. Chromosomal localization of the genes for mouse and human type XIII collagen and human prolyl 4-hydroxylase α-subunit

A PCR product corresponding to intron 25 of *Col13a1* was analyzed and found to identify a polymorphism between inbred mouse strains. *Col13a1* was found to map to chromosome 10 with a LOD score of 25. The position of *Col13a1* with respect to the flanking microsatellite markers is: *D10Mit5* – (2.3±1.6 cM) – *Col13a1* – (3.4±1.9 cM) – *D10Mit15*. Thus SSCP analysis localized the *Col13a1* to mouse chromosome 10, tightly linked to the microsatellite marker *D10Mit5*.

As an alternative approach, the location of the mouse *Col13a1* gene was also determined by fluorescent *in situ* hybridization. All in all, 68 out of 78 hybridizations showed a specific location on chromosome 10, band B4.

The genes for human type XIII collagen and prolyl 4-hydroxylase α-subunit were superimposed when hybridized on metaphase chromosomes. When mechanically stretched chromosomes were used as a target for FISH, the gene order was determined to be centromere–COL13A1–P4HA–telomere. As only rough estimations of the distances separating these genes could be made based on mechanically stretched chromosomes, more precise distance measurements between COL13A1 and P4HA genes were performed using fiber FISH. Based on the the known clone sizes, the distance between the COL13A1 and the P4HA genes was estimated to be 550 kb and the order of these tail to tail oriented genes along the chromosome was determined to be centromere–COL13A1–P4HA–telomere.

5.4. Generation of mice with a mutated type XIII collagen gene

A targeting vector containg a 8 kbp *Bam* I-fragment of type XIII collagen gene, *loxP*-sequences and selection marker genes was constructed. The targeting vector was introduced into ES cells and 360 colonies which survived the selection, and had thus the targeting construct in their genome, were recovered and DNA of those clones was analyzed by Southern blotting. In 14 clones the homologous recombination had taken place, and four clones had obtained all three *loxP* sites. Two of these clones were exposed to the *Cre* treatment and 164 clones that survived Ganciclovir selection were characterized.

Two clones in which the deletion of a 1200 bp genomic fragment between *Sfi* I and *Not* I restriction sites was detected were re-expanded and injected into the C57/B6 blastocysts to obtain chimeric males. From five chimeric males tested by breeding with

C57BL/6J females, two produced heterozygous offspring. A homozygous line carrying the mutation was established by brother-sister mating of heterozygous mice.

Genotyping of the offspring derived from intercrosses of heterozygotes, pups at various embryonic stages and aborted embryos revealed that pups were born in the expected ratios. The homozygous mutant mice appeared normal in their development, gross anatomy and behaviour, and they were fertile and produced offspring at the same rate as the heterozygotes. The phenotype of mice with heterogenous background, obtained from mating C57Bl/6 and 129SvJ mice, did not differ detectably from mice with pure 129Sv genetic background.

5.5. Analysis of cultured fibroblasts and analysis of mRNA by RT-PCR

Fibroblasts from both wild-type and mutant mice expressing N-terminally altered type XIII collagen were cultured and stained with anti-XIII/NC3, anti-phosphotyrosin, anti-paxillin, anti-talin and anti-vinculin antibodies. No differences in the staining pattern or morphology was observed. Since the original plan was to produce a type XIII knock-out mouse line, it came as a surprise that the cells from the transgenic mice showed staining similar to that in wild-type cells. Therefore, we performed RT-PCR analysis, and a type XIII collagen transcript was detected. Closer analysis revealed however that a new 5' end originating from the *loxP* site and extending to the first intron had been acquired. The first intron contains a cryptic splice site which serves as a donor site for splicing to the second exon in frame. This results in an altered type XIII collagen molecule in which the coding region of the first exon is replaced by a totally unrelated sequence. The first 96 amino acid residues of the wild-type protein are replaced potentially by either 65 or 11 residues in the altered protein, depending on which two potential initiation methionines is used. An important change in the new N-terminus is the absence of the cytosolic and transmembrane domain.

Analysis of cultured skin fibroblasts revealed that the cells from the transgenic mice did not adhere as well as wild-type cells. The same phenomenon could also be observed in the test tube after centrifugation, when a pellet of fibroblast cells obtained from the mice expressing N-terminally altered type XIII collagen detached from the bottom of the tube more quickly than wild-type cells did.

5.6. Morphological analysis

All tissues of the transgenic mice, except for skeletal muscle, appeared normal when examined under the light microscope. The skeletal muscle tissues were further characterized by immunofuorescense staining with anti-collagen IV, anti-merosin, anti-vinculin, anti-tenascin-C, anti-α5-integrin and anti-collagen XIII/NC3 antibodies. In general, the diameter and appearance of the muscle fibers was variable and fibers had

rough edges. The same differences were also detected with anti-collagen IV and anti-XIII/NC3 antibodies. Staining with the anti-collagen IV antibodies also revealed an altered appearance of the basement membranes. An irregular banding of the muscle fibers could be seen after staining with the anti-XIII/NC3 antibodies. An elevated expression of the α5-integrin was seen in the blood vessels of the muscle as demonstrated with staining with the anti-α5-integrin antibody. The muscle sections were also stained with an anti-tenascin-C antibody, but this revealed no apparent differences between the mutated and wild-type muscles.

5.7. Electron microscopic analysis

Muscle for electron microscopy were collected from 17 and 42 week-old wild-type and mutated mice. In the 17 week-old mutated mice the most common finding was the accumulation and enlargement of mitochondria in the muscle cells. Abnormalities in the z-bands and myofilament organization were also detected. The z-bands showed streaming and disintegration and the sarcolemmal basement membrane showed a thinner and looser structure than that seen in the corresponding wild-type mice. In the mutated muscle, accumulation of unorganized debris near the basement membrane was seen, while collagen fibers were seen close to the basement membrane in the wild-type animals. The described abnormalities became increasingly common in older mice. In the 42 week-old mice expressing N-terminally altered type XIII collagen, some of the muscle fibers were abnormal in shape and contained deep invaginations of the sarcolemma. Disorganization of the myofilaments and disruption of the z-bands were also seen. The sarcolemmal basement membrane was in many cases detached or even totally absent. Deep papillary projections of the sarcoplasm were also detected.

6. Discussion

Type XIII collagen is transmembrane collagen found in a wide variety of tissues. In cultured cells it is found concentrated in the focal adhesion plaques. The genomic organization of *Col13a1* gene was analyzed. To study the function *in vivo*, mice with an altered gene were generated.

The present research was initiated by searching and characterizing the genomic clones for mouse type XIII collagen gene. The gene encoding the mouse type XIII collagen was found to be about 135 kbp in size and the ensuing mRNAs are about 3000 nucleotides in length. *Col13a1* contains 42 exons, which vary in size from 8 to 836 bp. The exon-intron boundaries conform well to the consensus sequence. The coding sequences are unevenly distributed in the gene, as the first two exons and introns encompass half of the gene. The length of the exons which encode noncollagenous parts of type XIII collagen were found to vary between 8 and 144 bp. The prominent feature of *Col13a1* gene is the abundance of 27-bp exons. In fibrillar collagens the triple-helix is encoded predominantly by exons of 54 bp or derivatives of this. The common feature for both 27 bp and 54 bp exons is that exon begins with codon for glycine. The exons of *Col13a1* show also features characteristic of non-fibril-forming collagens because of the presence of exons of more variable in size (Christiano *et al.* 1994b, Hayman *et al.* 1991, Li *et al.* 1991, Lozano *et al.* 1985, Perälä *et al.* 1994, Saitta *et al.* 1991, Soininen *et al.* 1989, Tikka *et al.* 1991, Wälchli *et al.* 1992, Zhou *et al.* 1994). The evolutionary origin of this kind of arrangement is unknown and leads to the question about the origin of introns but addressing that question is beyond the scope of this thesis.

A gap in the previously characterized clones encoding the exons 3a – 4b of human type XIII collagen (Tikka *et al.* 1991) was covered by a clone identified when scanning of the GenBank data bank. This search for sequence similarities revealed one sequence, GenBank accession number U82211, which contained the human type XIII collagen gene region covering the missing exons (Benson *et al.* 1998). The 99-bp exon 31 in mouse has not been reported in the human species but is likely to exist there too. The 8-bp exon 3 found both in the mouse clones and in the human high-throughput genomic sequencing clone is the shortest exon known in collagen genes to date, superseding the 9-bp exon found in the human gene for the α4 chain of type IV collagen (Boye *et al.* 1998).

Characterization of mouse and human cDNA clones predicted a longer N-terminal noncollagenous NC1 domain than had previously been thought, which prompted us to study further the 5´-flanking sequences of the mouse and human type XIII collagen genes (Hägg *et al.* 1998b, Pihlajaniemi & Tamminen 1990). The study on the human gene (Tikka *et al.* 1991) has suggested that the promoter lies within the sequences that is shown in our studies to represent 5´-untranslated sequences, and thus the data point to a revised understanding of the promoter and 5´-untranslated sequences of the human type XIII collagen gene. 2642 bp of sequences preceding the initiation methionine for this extended NC1 domain were compared between the human and mouse genes and their homology was found to be low with the exception of the 5´-untranslated region of about 550 bp and an adjacent apparent promoter region of about 350 bp containing a modified TATAA motif, several GC boxes and other conserved putative cis-acting elements.

The restriction map of the mouse type XIII collagen genomic area, with four commonly cutting enzymes and two rarely cutting enzymes, *Not* I and *Sfi* I was constructed. Surprisingly, a number of rare sites found in the gene, close to the noncollagenous sequences encoding exon 1. The mouse type XIII collagen gene includes six *Sfi* I restriction sites in total, three of which were in exon areas coding for collagenous sequences, which is not surprising since a basis for the *Sfi* I recognition sequence is a glycine codon, and glycine is every third amino acid in collagen sequence. The restriction map obtained here was used as a tool in gene targeting studies.

The gene encoding mouse type XIII collagen was found to be located in band B4, at *D10Mit5* – (2.3±1.6 cM) – *Col13a1* – (3.4±1.9 cM) – *D10Mit15* in chromosome 10. According to the consensus genetic map data maintained in the Mouse Genome Database at the Jackson Laboratory (http://www.informatics.jax.org), sub-chromosomal linkage relationships in this region have been conserved on the human chromosome 10q21-22. This result is consistent with the evidence that the human homologue of *Col13a1* has been mapped to chromosome 10q22 (Shows *et al.* 1989). We confirmed this location and determined also that the gene encoding the catalytically active α-subunit of prolyl 4-hydroxylase is located about 550 kbp from the type XIII collagen gene. The distance between COL13A1 and P4HA is so large that several genes can lie between them. However, these two genes are involved in the same metabolic cascade and little, if anything, is known about chromosome wide gene regulation. Therefore the possibility that these two genes share some transcriptional elements can not be entirely excluded.

The genomic clones obtained were used to construct a vector for targeting the type XIII collagen gene by homologous recombination. The targeting construct contained three *loxP* sites, one in the 5' untranslated region, and two flanking the *neo*-TK marker cassette. Homologous recombination and, subsequently, *Cre* recombination generated ES cells where, in one allele, 1.2 kb in the 5' region of the Col13a1 gene was deleted. Mice carrying the targeted mutation (which was presumed to cause inactivation of the gene) were produced, and by mating of heterozygous mice, homozygous mutant mice were obtained. When the homozygous mutant mice were born, they seemed healthy, had a normal life span and were fertile. On later studies we found out that instead of total inactivation, an alteration in the gene was observed, so that these mice express an altered type XIII collagen with a new N-terminal domain. Results of the RT-PCRs and subsequent sequencing of these products showed that these chimeras had a new

translation start codon inside of the *loxP*-sequence, and the transcript was then correctly spliced using cryptic donor site located at the beginning of first intron keeping the same reading frame, therefore the rest of protein was unaltered. Gene targeting has been used to generate mutations where only part of the protein is altered (Andrikopoulos *et al.* 1995, White *et al.* 1997) but the unique feature in generating mice expressing N-terminally altered type XIII collagen is that the translation starting codon is changed and cryptic splice sequence of the first intron is utilized.

The most obvious phenotypic manifestation of mice expressing N-terminally altered type XIII collagen was seen in skeletal muscle, which show a number of atrophic abnormalities. We also examined fibroblasts from these mice and found that the cultured cells had decreased adhesion to culture plates.

Considering the results of our previous studies which suggested that type XIII collagen is located in focal adhesions and adhesive tissue structures (Hägg *et al.*, manuscript), the observation that the loss of the transmembrane domain did not affect the development or postnatal life of the mice expressing the N-terminally altered protein was contrary to the hypothesized mode of action. This lack of a pronounced phenotype is surprising, considering that deficiencies or mutations in other collagens usually have severe consequences (Bonadio *et al.* 1990, Chung *et al.* 1997, De Paepe *et al.* 1997, Hagg *et al.* 1997, Ikegawa *et al.* 1998, Li *et al.* 1995b, Liu *et al.* 1997). On the other hand, there are studies where the knock-out mice of various collagens were made but only a mild phenotype was observed (Bonaldo *et al.* 1998, Eklund *et al.* 1998, Rosati *et al.* 1994).

It was also surprising that the altered collagen is correctly transported into focal adhesion plaques, as it does not carry a transmembrane domain or a signal sequence. Apparently the N-terminally altered type XIII collagen is translocated into the lumen of the endoplasmic reticulum and transported to the plasma membrane of the cells, possibly by interacting with other plasma membrane molecules and thereby being carried to the correct location. Despite its altered stucture the protein seems to be stable enough at physiological temperatures to be detected with an antibodies, as shown by fibroblast studies. Also, the function of the protein seem not to be severely disturbed.

In order to find potential alterations caused by the expression of the N-terminally altered protein several tissues were sectioned, stained and analyzed by light microscopy. The only abnormalities revealed by this analysis were seen in skeletal muscle. To further define the structural defects, both light and electron microscopic analyses of the gastrocnemius and quadriceps muscles were carried out. At the light microscopic level the staining with the hematoxylin-eosin, anti-collagen-IV, anti-merosin, anti-α5-integrin and anti-collagen XIII/NC3 antibodies revealed changes in the overall structure of the muscles, as the diameter and appearance of the muscle fibers was different than in wild-type muscles and differences were also detected with anti-collagen IV and anti-XIII/NC3 antibodies. Staining with the anti-collagen IV antibodies revealed an altered appearance of the basement membranes and an uneven structure of the muscle fibers could be seen after staining with the anti-XIII/NC3 antibodies. Staining with the anti-α5-integrin antibody showed an elevated expression of the α5-integrin in the blood vessels of the muscle. No apparent differences between the mutated and wild-type muscles were seen when stained with an anti-tenascin-C antibody.

In electron microscopic studies the skeletal muscles of seventeen week-old mutant mice showed considerable accumulation and enlargement of mitochondria, streaming of the z-disks, loosened structure of the basement membrane, and papillary projections on the surface of abnormal fibers. These changes seemed to be focal and progressive because the phenotype is more severe in 42 week-old mice than in younger animals. All these changes are unspecific and typical to muscles affected by different myopathies (Engel & Banker 1986). A notable feature of the mice expressing N-terminally altered type XIII collagen is that they do not show necrosis or regeneration of the abnormal fibers.

The degenerative changes seen in the muscles of mice expressing N-terminally altered type XIII collagen type could be explained by considering these changes as a consequence of decreased adhesion. As type XIII collagen is located in focal adhesion plaques, it is likely that changes in the structure of this molecule can cause changes in cellular adhesion. Indeed, decreased adhesion was measured in fibroblasts cultured from the transgenic mice, and it is likely that this finding extends to other tissues, because type XIII collagen is expressed in almost all tissues examined so far. Weakened linkage is likely to initiate the separation of the contracting muscle from the the surrounding basement membrane and further contraction is likely to aggravate the initial disruption. When the separation has occurred, the muscle fiber can twist and the internal organelles can be damaged causing the streaming of z-bands, accumulation of mitochondria, generation of papillary projections and, finally, the destruction of the structure of the myofilaments. These observations can be considered as atrophic changes of the muscle. This hypothesis is also represented by Taverna *et al.* (1998) to explain their observation on mutated α5-integrin in chimeric mice. A larger number of mitochondrial transections per unit area seen in the samples can also result from elongation or increase in complexity of individual mitochondria. Moreover, an apparent increase in the size of the organelle in longitudinal sections can be due to reorientation of the organelle in space, as seen in recently denervated muscle (Engel & Banker 1986). As the denervation of muscle fibers induces the accumulation of tenascin-C in the muscle (Erickson & Bourdon 1989) and this kind of accumulation was not seen in the mice expressing N-terminally altered type XIII collagen. The changes seen in these mice are probably mechanistically different from those in denervated muscle.

Muscles are attached to the extracellular matrix by several additional molecules, such as several integrins and the dystrophin-dystroglycan complex, which is considered to be the main mediator of muscle attachment (Campbell 1995, Ibraghimov-Beskrovnaya *et al.* 1992, Mayer *et al.* 1997). There is also a report indicating that type VI collagen interacts with type IV collagen at basement membranes and with several other extra cellular matrix components thereby providing a machinery for transmembrane signalling (Kuo *et al.* 1997). It is noteworthy that the proposed functions of types VI and XIII collagens are on opposite sides of the muscular basement membrane; it is thought that type VI collagen acts to support the basement membrane on the side opposite the sarcolemma, whereas type XIII acts at the interface between the sarcolemma and the basemement membrane.

Two muscle dystrophies, a subtype of Miyoshi-type distal muscular dystrophy (MMD) and facioscapulohumeral muscular dystrophy have been mapped to chromosome 10. MMD is an autosomal recessively inherited progressive disorder and its putative locus

has been linked to the limb-girdle muscular dystrophy 2B locus on chromosome 2p12-14. Linssen *et al.* (1998) reported genetic heterogenity within this disease and screening of two non-chromosome 2-linked families revealed that a 23 cM region on chromosome 10 segregated with MMD (Linssen *et al.* 1998). This region is far from the locus of type XIII collagen gene, which has been mapped to 10q22 (Shows *et al.* 1989). Facioscapulohumeral muscular dystrophy is linked to region 10qter (Cacurri *et al.* 1998), which is also distant from the gene for type XIII collagen. Therefore, it is unlikely that the changes seen in type XIII collagen mutant mice are related to the loci mentioned above.

In this thesis the genomic organization of the murine type XIII collagen gene, its regulatory elements, the chromosomal localization and comparison to the human type XIII collagen gene is demonstrated. It also shows the physical distance and transciptional orientation of the human type XIII collagen and prolyl 4-hydroxylase α-subunit genes. In addition, the generation of transgenic mice and that the expression of N-terminally altered type XIII collagen causes a new form of progressive muscular atrophy in mice is described. I hypothesize that the cause of this atrophy is the weakened linkage between the sarcolemma and the basement membrane. I therefore postulate that type XIII collagen participates along with integrins and the dystrophin-dystroglycan complex in mediating the linkage between the muscle fiber and the extracellular matrix. It is also likely that type XIII collagen is a potential candidate in some inherited muscular atrophies for which the cause has not been identified.

7. References

Abe N, Muragaki Y, Yoshioka H, Inoue H & Ninomiya Y (1993) Identification of a novel collagen chain represented by extensive interruptions in the triple-helical region. Biochem. Biophys. Res. Commun. 196: 576-582.

Ala-Kokko L, Kontusaari S, Baldwin CT, Kuivaniemi H & Prockop DJ (1989) Structure of cDNA clones coding for the entire prepro α1(III) chain of human type III procollagen. Differences in protein structure from type I procollagen and conservation of codon preferences. Biochem. J. 260: 509-516.

Ala-Kokko L, Kvist A-P, Metsäranta M, Kivirikko KI, de Crombrugghe B, Prockop DJ & Vuorio E (1995) Conservation of the sizes of 53 introns and over 100 intronic sequences for the binding of common transcription factors in the human and mouse genes for type II procollagen (COL2A1). Biochem. J. 308: 923-929.

Andrikopoulos K, Liu X, Keene DR, Jaenisch R & Ramirez F (1995) Targeted mutations in the col5a2 gene reveal a regulatory role for type V collagen during matrix assembly. Nat. Genet. 9: 31-36.

Antonarakis SE (1993) Human chromosome 21: genome mapping and exploration, circa 1993. Trends. Genet. 9: 142-148.

Apte S, Mattei MG & Olsen BR (1991) Cloning of human α1(X) collagen DNA and localization of the COL10A1 gene to the q21-q22 region of human chromosome 6. FEBS Lett. 282: 393-396.

Apte SS, Seldin MF, Hayashi M & Olsen BR (1992) Cloning of the human and mouse type X collagen genes and mapping of the mouse type X collagen gene to chromosome 10. Eur. J. Biochem. 206: 217-224.

Aubert-Foucher E, Font B, Eichenberger D, Goldschmidt D, Lethias C & van der Rest M (1992) Purification and characterization of native type XIV collagen. J. Biol. Chem. 267: 15759-15764.

Balarin MA, da SL, V & Varella-Garcia M (1999) A dup(17)(p11.2p11.2) detected by fluorescence in situ hybridization in a boy with Alport syndrome. Am. J. Med. Genet. 82: 183-186.

Baldwin CT, Reginato AM, Smith C, Jimenez SA & Prockop DJ (1989) Structure of cDNA clones coding for human type II procollagen. The α1(II) chain is more similar to the α1(I) chain than two other α chains of fibrillar collagens. Biochem. J. 262: 521-528.

Beier DR (1993) Single-strand conformation polymorphism (SSCP) analysis as a tool for genetic mapping. Mamm. Genome. 4: 627-631.

Benson DA, Boguski MS, Lipman DJ, Ostell J & Ouellette BF (1998) GenBank. Nucleic Acids Res. 26: 1-7.

Bentz H, Morris NP, Murray LW, Sakai LY, Hollister DW & Burgeson RE (1983) Isolation and partial characterization of a new human collagen with an extended triple-helical structural domain. Proc. Natl. Acad. Sci. USA. 80: 3168-3172.

Bernard M, Yoshioka H, Rodriguez E, van der Rest M, Kimura T, Ninomiya Y, Olsen BR & Ramirez F (1988) Cloning and sequencing of pro-α1 (XI) collagen cDNA demonstrates that type XI belongs to the fibrillar class of collagens and reveals that the expression of the gene is not restricted to cartilagenous tissue. J. Biol. Chem. 263: 17159-17166.

Berthod F, Germain L, Guignard R, Lethias C, Garrone R, Damour O, van der Rest M & Auger FA (1997) Differential expression of collagens XII and XIV in human skin and in reconstructed skin. J. Invest. Dermatol. 108: 737-742.

Bohme K, Li Y, Oh PS & Olsen BR (1995) Primary structure of the long and short splice variants of mouse collagen XII and their tissue-specific expression during embryonic development. Dev. Dyn. 204: 432-445.

Bonadio J, Saunders TL, Tsai E, Goldstein SA, Morris-Wiman J, Brinkley L, Dolan DF, Altschuler RA, Hawkins JE, Jr. & Bateman JF (1990) Transgenic mouse model of the mild dominant form of osteogenesis imperfecta. Proc. Natl. Acad. Sci. USA. 87: 7145-7149.

Bonaldo P, Braghetta P, Zanetti M, Piccolo S, Volpin D & Bressan GM (1998) Collagen VI deficiency induces early onset myopathy in the mouse: an animal model for bethlem myopathy. Hum. Mol. Genet. 7: 2135-2140.

Borradori L & Sonnenberg A (1999) Structure and function of hemidesmosomes: more than simple adhesion complexes. J. Invest. Dermatol. 112: 411-418.

Boyd CD, Pierce RA, Schwarzbauer JE, Doege K & Sandell LJ (1993) Alternate exon usage is a commonly used mechanism for increasing coding diversity within genes coding for extracellular matrix proteins. Matrix. 13: 457-469.

Boye E, Mollet G, Forestier L, Cohen-Solal L, Heidet L, Cochat P, Grunfeld JP, Palcoux JB, Gubler MC & Antignac C (1998) Determination of the genomic structure of the COL4A4 gene and of novel mutations causing autosomal recessive Alport syndrome. Am. J. Hum. Genet. 63: 1329-1340.

Bradley A, Evans M, Kaufman MH & Robertson E (1984) Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature. 309: 255-256.

Brandon EP, Idzerda RL & McKnight GS (1995a) Knockouts. Targeting the mouse genome: a compendium of knockouts (Part I). Current Biology. 5: 625-634.

Brandon EP, Idzerda RL & McKnight GS (1995b) Targeting the mouse genome: a compendium of knockouts (Part II). Current Biology. 5: 758-765.

Brandon EP, Idzerda RL & McKnight GS (1995c) Targeting the mouse genome: a compendium of knockouts (Part III). Current Biology. 5: 873-881.

Brewton RG, Wood BM, Ren ZX, Gong Y, Tiller GE, Warman ML, Lee B, Horton WA, Olsen BR & Baker JR (1995) Molecular cloning of the α3 chain of human type IX collagen: linkage of the gene COL9A3 to chromosome 20q13.3. Genomics. 30: 329-336.

Brinster RL, Chen HY, Trumbauer M, Senear AW, Warren R & Palmiter RD (1981) Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. Cell. 27: 223-231.

Brown KS, Cranley RE, Greene R, Kleinman HK & Pennypacker JP (1981) Disproportionate micromelia (Dmm): an incomplete dominant mouse dwarfism with abnormal cartilage matrix. J. Embryol. Exp. Morphol. 62: 165-182.

Brown SD, Avner P, Boyd Y, Chapman V, Rastan S, Sefton L, Thomas JD & Herman GE (1993) Encyclopedia of the mouse genome III. October 1993. Mouse X chromosome. Mamm. Genome. 4 Spec No: 269-281.

Bruckner P, Vaughan L & Winterhalter KH (1985) Type IX collagen from sternal cartilage of chicken embryo contains covalently bound glycosaminoglycans. Proc. Natl. Acad. Sci. USA. 82: 2608-2612.

Burgeson RE (1987) Type VII collagen. In:Mayne, R & Burgeson, R (eds) Structure and Function of Collagen Types 145-172. Academic Press, New York.

Burgeson RE, El Adli FA, Kaitila II & Hollister DW (1976) Fetal membrane collagens: identification of two new collagen α chains. Proc. Natl. Acad. Sci. USA. 73: 2579-2583.

Butkowski RJ, Langeveld JP, Wieslander J, Hamilton J & Hudson BG (1987) Localization of the Goodpasture epitope to a novel chain of basement membrane collagen. J. Biol. Chem. 262: 7874-7877.

Cacurri S, Piazzo N, Deidda G, Vigneti E, Galluzzi G, Colantoni L, Merico B, Ricci E & Felicetti L (1998) Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric KpnI repeat units implicated in facioscapulohumeral muscular dystrophy. Am. J. Hum. Genet. 63: 181-190.

Campbell KP (1995) Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. Cell. 80: 675-679.

Cheah KS, Au PK, Lau ET, Little PF & Stubbs L (1991) The mouse Col2a-1 gene is highly conserved and is linked to Int- 1 on chromosome 15. Mamm. Genome. 1: 171-183.

Chipman SD, Sweet HO, McBride DJ, Jr., Davisson MT, Marks SC, Jr., Shuldiner AR, Wenstrup RJ, Rowe DW & Shapiro JR (1993) Defective pro α2(I) collagen synthesis in a recessive mutation in mice: a model of human osteogenesis imperfecta. Proc. Natl. Acad. Sci. USA. 90: 1701-1705.

Chomczynski P & Sacchi N (1987) Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162: 156-159.

Christiano AM, Greenspan DS, Lee S & Uitto J (1994a) Cloning of human type VII collagen. Complete primary sequence of the α1(VII) chain and identification of intragenic polymorphisms. J. Biol. Chem. 269: 20256-20262.

Christiano AM, Hoffman GG, Chung-Honet LC, Lee S, Wen C, Uitto J & Greenspan DS (1994b) Structural organization of the human type VII collagen gene (COL7A1), composed of more exons than any previously characterized gene. Genomics. 21: 169-179.

Chu ML, Mann K, Deutzmann R, Pribula-Conway D, Hsu-Chen CC, Bernard MP & Timpl R (1987) Characterization of three constituent chains of collagen type VI by peptide sequences and cDNA clones. Eur. J. Biochem. 168: 309-317.

Chu ML & Prockop D (1993) Collagen: gene structures. In:Royce, P M & Steinmann, B (eds) Connective Tissue and Its Heritable Disorders. Molecular, Genetics, and Medical Aspects 149-165. Wiley-Liss, Inc., New York.

Chu ML, Zhang R-Z, Pan T-C, Stokes D, Conway D, Kuo H-J, Glanville R, Mayer U, Mann K, Deutzmann R & Timpl R (1990) Mosaic structure of globular domains in the type VI collagen α3(VI) chain: similarity to von Willebrand factor, fibronectin, actin, salivary proteins and aprotinin type protease inhibitors. EMBO J. 9: 385-393.

Chung E, Rhodes K & Miller EJ (1976) Isolation of three collagenous components of probable basement membrane origin from several tissues. Biochem. Biophys. Res. Commun. 71: 1167-1174.

Chung KS, Jacenko O, Boyle P, Olsen BR & Nishimura I (1997) Craniofacial abnormalities in mice carrying a dominant interference mutation in type X collagen. Dev. Dyn. 208: 544-552.

Copeland NG, Gilbert DJ, Li K, Sawamura D, Giudice GJ, Chu ML, Jenkins NA & Uitto J (1993) Chromosomal localization of mouse bullous pemphigoid antigens. BPAG1 and BPAG2: identification of a new region of homology between mouse and human chromosomes. Genomics. 15: 180-181.

Cosgrove D, Meehan DT, Grunkemeyer JA, Kornak JM, Sayers R, Hunter WJ & Samuelson GC (1996) Collagen COL4A3 knockout: a mouse model for autosomal Alport syndrome. Genes Dev. 10: 2981-2992.

Cosgrove D, Samuelson G, Meehan DT, Miller C, McGee J, Walsh EJ & Siegel M (1998) Ultrastructural, physiological, and molecular defects in the inner ear of a gene-knockout mouse model for autosomal Alport syndrome. Hear. Res. 121: 84-98.

Costantini F & Lacy E (1981) Introduction of a rabbit β-globin gene into the mouse germ line. Nature. 294: 92-94.

Cutting GR, McGinniss MJ, Kasch LM, Tsipouras P & Antonarakis SE (1990) Physical mapping by PFGE localizes the COL3A1 and COL5A2 genes to a 35-kb region of human chromosome 2. Genomics. 8: 407-410.

Dalgleish R (1997) The human type I collagen mutation database. Nucleic Acids Res. 25: 181-187.

De Paepe A, Nuytinck L, Hausser I, Anton-Lamprecht I & Naeyaert JM (1997) Mutations in the COL5A1 gene are causal in the Ehlers-Danlos syndromes I and II. Am. J. Hum. Genet. 60: 547-554.

de Wet W, Bernard M, Benson-Chanda V, Chu ML, Dickson L, Weil D & Ramirez F (1987) Organization of the human pro-α2(I) collagen gene. J. Biol. Chem. 262: 16032-16036.

Devereux J, Haeberli P & Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12: 387-395.

Doetschman T, Gregg RG, Maeda N, Hooper ML, Melton DW, Thompson S & Smithies O (1987) Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature. 330: 576-578.

Dublet B & van der Rest M (1991) Type XIV collagen, a new homotrimeric molecule extracted from foetal bovine skin and tendon, with a triple helical disulfide- bonded domain homologous to type IX and type XII collagens. J. Biol. Chem. 266: 6853-6858.

Ehnis T, Dieterich W, Bauer M, Kresse H & Schuppan D (1997) Localization of a binding site for the proteoglycan decorin on collagen XIV (undulin). J. Biol. Chem. 272: 20414-20419.

Ehnis T, Dieterich W, Bauer M, Lampe B & Schuppan D (1996) A chondroitin/dermatan sulfate form of CD44 is a receptor for collagen XIV (undulin). Exp. Cell Res. 229: 388-397.

Eklund L, Fässler R, Muona A, Väisänen T, Fukai N, Olsen B & Pihlajaniemi T (1998) Lack of Type XV Collagen Cause a Mild Late Onset Skeletal Myopathy with Variability in Different Muscle Groups in Mice. In: ABSTRACTS - XVIth Meeting of the Federation of the European Connective Tissue Societies D6.Uppsala.

Emanuel BS, Sellinger BT, Gudas LJ & Myers JC (1986) Localization of the human procollagen α1(IV) gene to chromosome 13q34 by in situ hybridization. Am. J. Hum. Genet. 38: 38-44.

Engel AG & Banker BQ (1986) Ultrastructural Changes in Diseased Muscle. In:Engel, A G & Banker, B Q (eds) Myology 909-1044. McGraw-Hill, Inc., USA.

Engel J, Furthmayr H, Odermatt E, von der Mark H, Aumailley M, Fleischmajer R & Timpl R (1985) Structure and macromolecular organization of type VI collagen. Ann. N. Y. Acad. Sci. 460:25-37: 25-37.

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

Erickson HP & Bourdon MA (1989) Tenascin: an extracellular matrix protein prominent in specialized embryonic tissues and tumors. Annu. Rev. Cell Biol. 5:71-92: 71-92.

Evans GA, Lewis K & Rothenberg BE (1989) High efficiency vectors for cosmid microcloning and genomic analysis. Gene. 79: 9-20.

Evans MJ & Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. Nature. 292: 154-156.

Eyre D & Wu J-J (1987) Type XI or 1α2α3α collagen. In:Mayne, R & Burgeson, R E (eds) Structure and Function of Collagen Types 261-281. Academic Press, Inc., Orlando, FL.

Feng L, Xia Y & Wilson CB (1994) Alternative splicing of the NC1 domain of the human α3(IV) collagen gene. Differential expression of mRNA transcripts that predict three protein variants with distinct carboxyl regions. J. Biol. Chem. 269: 2342-2348.

Fichard A, Kleman J-P & Ruggiero F (1994) Another look at collagen V and XI molecules. Matrix Biol. 14: 515-531.

Fleischmajer R, Perlish JS, Burgeson RE, Shaikh-Bahai F & Timpl R (1990) Type I and type III collagen interactions during fibrillogenesis. Ann. N. Y. Acad. Sci. 580: 161-175.

Fukai N, Apte S & Olsen BR (1994) Nonfibrillar collagens. Meth. Enzymol. 245: 3-28.

Galli-Taliadoros LA, Sedgwick JD, Wood SA & Korner H (1995) Gene knock-out technology: a methodological overview for the interested novice. J. Immunol. Methods. 181: 1-15.

Garofalo S, Vuorio E, Metsäranta M, Rosati R, Toman D, Vaughan J, Lozano G, Mayne R, Ellard J & Horton W (1991) Reduced amounts of cartilage collagen fibrils and growth plate anomalies in transgenic mice harboring a glycine-to-cysteine mutation in the mouse type II procollagen α1-chain gene. Proc. Natl. Acad. Sci. USA. 88: 9648-9652.

Gatalica B, Pulkkinen L, Li K, Kuokkanen K, Ryynänen M, McGrath JA & Uitto J (1997) Cloning of the human type XVII collagen gene (COL17A1), and detection of novel mutations in generalized atrophic benign epidermolysis bullosa. Am. J. Hum. Genet. 60: 352-365.

Gerecke DR, Foley JW, Castagnola P, Gennari M, Dublet B, Cancedda R, Linsenmayer TF, van der Rest M, Olsen BR & Gordon MK (1993) Type XIV collagen is encoded by alternative transcripts with distinct 5' region and is a multidomain protein with homologies to von Willebrand's factor, fibronectin, and other matrix proteins. J. Biol. Chem. 268: 12177-12184.

Gerecke DR, Olson PF, Koch M, Knoll JHM, Taylor R, Hudson DL, Champliaud M-F, Olsen BR & Burgeson RE (1997) Complete primary structure of two splice variants of collagen XII, and assignment of α1(XII) collagen (COL12A1), α1(IX) collagen (COL9A1), and α1(XIX) collagen (COL19A1) to human chromosome 6q12-q13. Genomics. 41: 236-242.

Giudice GJ, Emegy DJ & Diaz LA (1992) Cloning and primary structural analysis of the bullous pemphigoid autoantigen BP180. J. Invest. Dermatol. 99: 243-250.

Giudice GJ, Squiquera HL, Elias PM & Diaz LA (1991) Identification of two collagen domains within the bullous pemphigoid autoantigen, BP180. J. Clin. Invest. 87: 734-738.

Gordon JW & Ruddle FH (1981) Integration and stable germ line transmission of genes injected into mouse pronuclei. Science. 214: 1244-1246.

Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA & Ruddle FH (1980) Genetic transformation of mouse embryos by microinjection of purified DNA. Proc. Natl. Acad. Sci. USA. 77: 7380-7384.

Gordon MK, Castagnola P, Dublet B, Linsenmayer TF, van der Rest M, Mayne R & Olsen BR (1991) Cloning of a cDNA for a new member of the class of fibril- associated collagens with interrupted triple helices. Eur. J. Biochem. 201: 333-338.

Gordon MK, Gerecke DR, Dublet B, van der Rest M & Olsen BR (1989) Type XII collagen. A large multidomain molecule with partial homology to type IX collagen. J. Biol. Chem. 264: 19772-19778.

Gordon MK, Gerecke DR & Olsen BR (1987) Type XII collagen: Distinct extracellular matrix component discovered by cDNA cloning. Proc. Natl. Acad. Sci. USA. 84: 6040-6044.

Gossler A, Doetschman T, Korn R, Serfling E & Kemler R (1986) Transgenesis by means of blastocyst-derived embryonic stem cell lines. Proc. Natl. Acad. Sci. USA. 83: 9065-9069.

Grassel S, Timpl R, Tan EM & Chu ML (1996) Biosynthesis and processing of type XVI collagen in human fibroblasts and smooth muscle cells. Eur. J. Biochem. 242: 576-584.

Greenspan DS (1993) The carboxyl-terminal half of type VII collagen, including the non-collagenous NC-2 domain and intron/exon organization of the corresponding region of the COL7A1 gene. Hum. Mol. Genet. 2: 273-278.

Greenspan DS, Byers MG, Eddy RL, Cheng W, Jani-Sait S & Shows TB (1992) Human collagen gene COL5A1 maps to the q34.2–q34.3 region of chromosome 9, near the locus for nail-patella syndrome. Genomics. 12: 836-837.

Greenspan DS, Byers MG, Eddy RL, Hoffman GG & Shows TB (1993) Localization of the human collagen gene COL7A1 to 3p21.3 by fluorescence in situ hybridization. Cytogenet. Cell Genet. 62: 35-36.

Grosveld F & Kollias G (1992) Transgenic Animals. Academic Press Ldt, London.

Gu H, Marth JD, Orban PC, Mossmann H & Rajewsky K (1994) Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. Science. 265: 103-106.

Hagg R, Hedbom E, Mollers U, Aszodi A, Fässler R & Bruckner P (1997) Absence of the α1(IX) chain leads to a functional knock-out of the entire collagen IX protein in mice. J. Biol. Chem. 272: 20650-20654.

Hayman AR, Koppel J & Trueb B (1991) Complete structure of the chicken α2(VI) collagen gene. Eur. J. Biochem. 197: 177-184.

Heiskanen M, Kallioniemi O & Palotie A (1996a) Fiber-FISH: experiences and a refined protocol. Genet. Anal. 12: 179-184.

Heiskanen M, Peltonen L & Palotie A (1996b) Visual mapping by high resolution FISH. Trends. Genet. 12: 379-382.

Heiskanen M, Saitta B, Palotie A & Chu ML (1995) Head to tail organization of the human COL6A1 and COL6A2 genes by fiber-FISH. Genomics. 29: 801-803.

Helminen HJ, Kiraly K, Pelttari A, Tammi MI, Vandenberg P, Pereira R, Dhulipala R, Khillan JS, Ala-Kokko L & Hume EL (1993) An inbred line of transgenic mice expressing an internally deleted gene for type II procollagen (COL2A1). Young mice have a variable phenotype of a

chondrodysplasia and older mice have osteoarthritic changes in joints. J. Clin. Invest. 92: 582-595.
Henry I, Bernheim A, Bernard M, van der Rest M, Kimura T, Jeanpierre C, Barichard F, Berger R, Olsen BR & Ramirez F (1988) Mapping of a human fibrillar collagen gene, pro α1(XI) (COL11A1), to the p21 region of chromosome 1. Genomics. 3: 87-90.
Hooper M, Hardy K, Handyside A, Hunter S & Monk M (1987) HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. Nature. 326: 292-295.
Hopkinson SB & Jones JCR (1996) Type XVII collagen and collagen-like molecules: related by more than a common motif. Cell Dev. Biol. 7: 659-666.
Hostikka SL & Tryggvason K (1988) The complete primary structure of the α2 chain of human type IV collagen and comparison with the α1(IV) chain. J. Biol. Chem. 263: 19488-19493.
Huerre-Jeanpierre C, Henry I, Bernard M, Gallano P, Weil D, Grzeschik KH, Ramirez F & Junien C (1986) The pro α2(V) collagen gene (COL5A2) maps to 2q14–2q32, syntenic to the pro α1(III) collagen locus (COL3A1). Hum. Genet. 73: 64-67.
Hulmes DJS (1992) The collagen superfamily - diverse structures and assemblies. Essays Biochem. 27: 49-67.
Hägg P (1998) Characterization of Collagen Types XIII and XV. Acta Univ Oul D 466.
Hägg P, Rehn M, Huhtala P, Väisänen T, Tamminen M & Pihlajaniemi T (1998b) Type XIII collagen is identified as a plasma membrane protein. J. Biol. Chem. 273: 15590-15597.
Hägg PM, Horelli-Kuitunen N, Eklund L, Palotie A & Pihlajaniemi T (1997) Cloning of mouse type XV collagen sequences and mapping of the corresponding gene to 4B1-3. Comparison of mouse and human α1(XV) collagen sequences indicates divergence in the number of small collagenous domains. Genomics. 45: 31-41.
Hägg PM, Muona A, Lietard J, Kivirikko S & Pihlajaniemi T (1998a) Complete exon-intron organization of the human gene for the α1 chain of type XV collagen (COL15A1) and comparison with the homologous COL18A1 gene. J. Biol. Chem. 273: 17824-17831.
Ibraghimov-Beskrovnaya O, Ervasti JM, Leveille CJ, Slaughter CA, Sernett SW & Campbell KP (1992) Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. Nature. 355: 696-702.
Ikegawa S, Nishimura G, Nagai T, Hasegawa T, Ohashi H & Nakamura Y (1998) Mutation of the Type X Collagen Gene (COL10A1) Causes Spondylometaphyseal Dysplasia. Am. J. Hum. Genet. 63: 1659-1662.
Inoguchi K, Yoshioka H, Khaleduzzaman M & Ninomiya Y (1995) The mRNA for α1(XIX) collagen chain, a new member of FACITs, contains a long unusual 3' untranslated region and displays many unique splicing variants. J. Biochem. 117: 137-146.
Irving NG, Hardy JA, Bahary N, Friedman JM & Brown SD (1989) The α2 chain of type 1 collagen does not map to mouse chromosome 16 but maps close to the Met proto-oncogene on mouse chromosome 6. Cytogenet. Cell Genet. 50: 121-122.
Jacenko O, LuValle PA & Olsen BR (1993) Spondylometaphyseal dysplasia in mice carrying a dominant negative mutation in a matrix protein specific for cartilage-to- bone transition. Nature. 365: 56-61.
Jaenisch R (1976) Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. Proc. Natl. Acad. Sci. USA. 73: 1260-1264.
Jaenisch R (1988) Transgenic animals. Science. 240: 1468-1474.
Jasin M, Moynahan ME & Richardson C (1996) Targeted transgenesis. Proc. Natl. Acad. Sci. USA. 93: 8804-8808.
Just M, Herbst H, Hummel M, Durkop H, Tripier D, Stein H & Schuppan D (1991) Undulin is a novel member of the fibronectin-tenascin family of extracellular matrix glycoproteins. J. Biol. Chem. 266: 17326-17332.
Justice MJ, Siracusa LD, Gilbert DJ, Heisterkamp N, Groffen J, Chada K, Silan CM, Copeland NG & Jenkins NA (1990) A genetic linkage map of mouse chromosome 10: localization of eighteen molecular markers using a single interspecific backcross. Genetics. 125: 855-866.
Juvonen M & Pihlajaniemi T (1992) Characterization of the spectrum of alternative splicing of α1(XIII) collagen transcripts in HT-1080 cells and calvarial tissue resulted in identification of

two previously unidentified alternatively spliced sequences, one previously unidentified exon, and nine new mRNA variants. J. Biol. Chem. 267: 24693-24699.

Juvonen M, Pihlajaniemi T & Autio-Harmainen H (1993) Location and alternative splicing of type XIII collagen RNA in the early human placenta. Lab. Invest. 69: 541-551.

Juvonen M, Sandberg M & Pihlajaniemi T (1992) Patterns of expression of the six alternatively spliced exons affecting the structures of the COL1 and NC2 domains of the α1(XIII) collagen chain in human tissues and cell lines. J. Biol. Chem. 267: 24700-24707.

Kamagata Y, Mattei MG & Ninomiya Y (1992) Isolation and sequencing of cDNAs and genomic DNAs encoding the α4 chain of basement membrane collagen type IV and assignment of the gene to the distal long arm of human chromosome. J. Biol. Chem. 267: 23753-23758.

Khaleduzzaman M, Sumiyoshi H, Ueki Y, Inoguchi K, Ninomiya Y & Yoshioka H (1997) Structure of the human type XIX collagen (COL19A1) gene, which suggests it has arisen from an ancestor gene of the FACIT family. Genomics. 45: 304-312.

Khillan JS, Olsen AS, Kontusaari S, Sokolov B & Prockop DJ (1991) Transgenic mice that express a mini-gene version of the human gene for type I procollagen (COL1A1) develop a phenotype resembling a lethal form of osteogenesis imperfecta. J. Biol. Chem. 266: 23373-23379.

Kielty CM, Hopkinson I & Grant ME (1993) Collagen: The collagen family: Structure, assembly and organization in the extracellular matrix. In:Royce, P M & Steinmann, B (eds) Connective Tissue and Its Heritable Disorders. Molecular, Genetics, and Medical Aspects 103-147. Wiley-Liss, Inc., New York.

Kimura T, Cheah KS, Chan SD, Lui VC, Mattei MG, van der Rest M, Ono K, Solomon E, Ninomiya Y & Olsen BR (1989a) The human α2(XI) collagen (COL11A2) chain. Molecular cloning of cDNA and genomic DNA reveals characteristics of a fibrillar collagen with differences in genomic organization. J. Biol. Chem. 264: 13910-13916.

Kimura T, Mattei MG, Stevens JW, Goldring MB, Ninomiya Y & Olsen BR (1989b) Molecular cloning of rat and human type IX collagen cDNA and localization of the α1(IX) gene on the human chromosome 6. Eur. J. Biochem. 179: 71-78.

Kirsch T & von der Mark K (1991) Isolation of human type X collagen and immunolocalization in fetal human cartilage. Eur. J. Biochem. 196: 575-580.

Kivirikko KI (1995) Posttranslational processing of collagens. In:Bittar, E E, Bittar, N, & Greenwich, C T (eds) Principles of Medical Biology, Vol 3. Cellular organelles the Extracellular Matrix 233-254. JAI Press, Greenwich, Conn.

Kivirikko S, Heinämäki P, Rehn M, Honkanen N, Myers JC & Pihlajaniemi T (1994) Primary structure of the α1 chain of human type XV collagen and exon-intron organization in the 3' region of the corresponding gene. J. Biol. Chem. 269: 4773-4779.

Kivirikko S, Saarela J, Myers JC, Autio-Harmainen H & Pihlajaniemi T (1995) Distribution of type XV collagen transcripts in human tissue and their production by muscle cells and fibroblasts. Am. J. Pathol. 147: 1500-1509.

Koizumi T, Hendel E, Lalley PA, Tchetgen MB & Nadeau JH (1995) Homologs of genes and anonymous loci on human chromosome 13 map to mouse chromosomes 8 and 14. Mamm. Genome. 6: 263-268.

Krawczak M & Cooper DN (1997) The human gene mutation database. Trends Genet. 13: 121-122.

Kuehn MR, Bradley A, Robertson EJ & Evans MJ (1987) A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. Nature. 326: 295-298.

Kuhn R, Schwenk F, Aguet M & Rajewsky K (1995) Inducible gene targeting in mice. Science. 269: 1427-1429.

Kuivaniemi H, Tromp G, Chu ML & Prockop DJ (1988) Structure of a full-length cDNA clone for the prepro α2(I) chain of human type I procollagen. Comparison with the chicken gene confirms unusual patterns of gene conservation. Biochem. J. 252: 633-640.

Kuivaniemi H, Tromp G & Prockop DJ (1997) Mutations in fibrillar collagens (types I, II, III, and XI), fibril-associated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels. Human Mutation. 9: 300-315.

Kuo HJ, Maslen CL, Keene DR & Glanville RW (1997) Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen. J. Biol. Chem. 272: 26522-26529.

Kwan KM, Pang MK, Zhou S, Cowan SK, Kong RY, Pfordte T, Olsen BR, Sillence DO, Tam PP & Cheah KS (1997) Abnormal compartmentalization of cartilage matrix components in mice lacking collagen X: implications for function. J. Cell Biol. 136: 459-471.

Körkkö J, Ala-Kokko L, De Paepe A, Nuytinck L, Earley J & Prockop DJ (1998) Analysis of the COL1A1 and COL1A2 genes by PCR amplification and scanning by conformation-sensitive gel electrophoresis identifies only COL1A1 mutations in 15 patients with osteogenesis imperfecta type I: identification of common sequences of null-allele mutations. Am J Hum. Genet. 62: 98-110.

Laan M, Kallioniemi OP, Hellsten E, Alitalo K, Peltonen L & Palotie A (1995) Mechanically stretched chromosomes as targets for high-resolution FISH mapping. Genome Res. 5: 13-20.

Lai CH & Chu ML (1996) Tissue distribution and developmental expression of type XVI collagen in the mouse. Tissue Cell. 28: 155-164.

Lamandé SR, Bateman JF, Hutchison W, McKinlay GR, Bower SP, Byrne E & Dahl HH (1998) Reduced collagen VI causes Bethlem myopathy: a heterozygous COL6A1 nonsense mutation results in mRNA decay and functional haploinsufficiency. Hum. Mol. Genet. 7: 981-989.

Leinonen A, Mariyama M, Mochizuki T, Tryggvason K & Reeders ST (1994) Complete primary structure of the human type IV collagen α4(IV) chain. Comparison with structure and expression of the other α(IV) chains. J. Biol. Chem. 269: 26172-26177.

Lemieux N, Dutrillaux B & Viegas-Pequignot E (1992) A simple method for simultaneous R- or G-banding and fluorescence in situ hybridization of small single-copy genes. Cytogenet. Cell Genet. 59: 311-312.

Li K, Christiano AM, Copeland NG, Gilbert DJ, Chu ML, Jenkins NA & Uitto J (1993b) cDNA cloning and chromosomal mapping of the mouse type VII collagen gene (Col7a1): evidence for rapid evolutionary divergence of the gene. Genomics. 16: 733-739.

Li K, Sawamura D, Giudice GJ, Diaz LA, Mattei M-G, Chu ML & Uitto J (1991) Genomic organization of collagenous domains and chromosomal assignment of human 180-kDa bullous pemphigoid antigen-2, a novel collagen of stratified squamous epithelium. J. Biol. Chem. 266: 24064-24069.

Li K, Tamai K, Tan EM & Uitto J (1993a) Cloning of type XVII collagen. Complementary and genomic DNA sequences of mouse 180-kilodalton bullous pemphigoid antigen (BPAG2) predict an interrupted collagenous domain, a transmembrane segment, and unusual features in the 5'-end of the gene and the 3'-untranslated region of the mRNA. J. Biol. Chem. 268: 8825-8834.

Li SW, Khillan JS & Prockop DJ (1995a) The complete cDNA coding sequence for the mouse pro α1(I) chain of type I procollagen. Matrix Biol. 14: 593-595.

Li SW, Prockop DJ, Helminen H, Fässler R, Lapvetelainen T, Kiraly K, Peltarri A, Arokoski J, Lui H & Arita M (1995b) Transgenic mice with targeted inactivation of the Col2 α1 gene for collagen II develop a skeleton with membranous and periosteal bone but no endochondral bone. Genes Dev. 9: 2821-2830.

Lichter P, Cremer T, Tang CJ, Watkins PC, Manuelidis L & Ward DC (1988) Rapid detection of human chromosome 21 aberrations by in situ hybridization. Proc. Natl. Acad. Sci. USA. 85: 9664-9668.

Linsenmayer TF (1991) Collagen. In:Hay, E D (ed) Cell Biology of Extracellular Matrix 7-44. Plenum Press, New York.

Linssen WH, de Visser M, Notermans NC, Vreyling JP, Van Doorn PA, Wokke JH, Baas F & Bolhuis PA (1998) Genetic heterogeneity in Miyoshi-type distal muscular dystrophy. Neuromuscul. Disord. 8: 317-320.

Liu X, Wu H, Byrne M, Jeffrey J, Krane S & Jaenisch R (1995) A targeted mutation at the known collagenase cleavage site in mouse type I collagen impairs tissue remodeling. J. Cell Biol. 130: 227-237.

Liu X, Wu H, Byrne M, Krane S & Jaenisch R (1997) Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development. Proc. Natl. Acad. Sci. USA. 94: 1852-1856.

Lozano G, Ninomiya Y, Thompson H & Olsen BR (1985) A distinct class of vertebrate collagen genes encodes chicken type IX collagen polypeptides. Proc. Natl. Acad. Sci. USA. 82: 4050-4054.

Lunstrum GP, Morris NP, McDonough AM, Keene DR & Burgeson RE (1991) Identification and partial characterization of two type XII-like collagen molecules. J. Cell Biol. 113: 963-969.

Manly KF (1993) A Macintosh program for storage and analysis of experimental genetic mapping data. Mamm. Genome. 4: 303-313.

Mariyama M, Leinonen A, Mochizuki T, Tryggvason K & Reeders ST (1994) Complete primary structure of the human α3(IV) collagen chain. Coexpression of the α3(IV) and α 4(IV) collagen chains in human tissues. J. Biol. Chem. 269: 23013-23017.

Mariyama M, Zheng K, Yang-Feng TL & Reeders ST (1992) Colocalization of the genes for the α3(IV) and α4(IV) chains of type IV collagen to chromosome 2 bands q35-q37. Genomics. 13: 809-813.

Martin GR (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. USA. 78: 7634-7638.

Matsuda Y, Harada YN, Natsuume-Sakai S, Lee K, Shiomi T & Chapman VM (1992) Location of the mouse complement factor H gene (cfh) by FISH analysis and replication R-banding. Cytogenet. Cell Genet. 61: 282-285.

Mayer U, Saher G, Fässler R, Bornemann A, Echtermeyer F, von der M, Miosge N & Poschl E (1997) Absence of integrin α7 causes a novel form of muscular dystrophy. Nat. Genet. 17: 318-323.

Mayne R & Brewton RG (1993) New members of collagen superfamily. Curr. Opin. Cell Biol. 5: 883-890.

Mayne R, van der Rest M, Ninomiya Y & Olsen BR (1985) The structure of type IX collagen. Ann. N. Y. Acad. Sci. 460: 38-46.

Mayne R, Wiedemann H, Irwin MH, Sanderson RD, Fitch JM, Linsenmayer TF & Kuhn K (1984) Monoclonal antibodies against chicken type IV and V collagens: electron microscopic mapping of the epitopes after rotary shadowing. J. Cell Biol. 98: 1637-1644.

Metsäranta M, Garofalo S, Decker G, Rintala M, de Crombrugghe B & Vuorio E (1992) Chondrodysplasia in transgenic mice harboring a 15-amino acid deletion in the triple helical domain of pro α1(II) collagen chain. J. Cell Biol. 118: 203-212.

Metsäranta M, Toman D, de Crombrugghe B & Vuorio E (1991) Mouse type II collagen gene. Complete nucleotide sequence, exon structure, and alternative splicing. J. Biol. Chem. 266: 16862-16869.

Michalickova K, Susic M, Willing MC, Wenstrup RJ & Cole WG (1998) Mutations of the α2(V) chain of type V collagen impair matrix assembly and produce ehlers-danlos syndrome type I. Hum. Mol. Genet. 7: 249-255.

Miner JH & Sanes JR (1994) Collagen IV α3, α4, and α5 chains in rodent basal laminae: sequence, distribution, association with laminins, and developmental switches. J. Cell Biol. 127: 879-891.

Morris NP & Bachinger HP (1987) Type XI collagen is a heterotrimer with the composition (1α, 2α, 3α) retaining non-triple-helical domains. J. Biol. Chem. 262: 11345-11350.

Munke M, Harbers K, Jaenisch R & Francke U (1986) Chromosomal mapping of four different integration sites of Moloney murine leukemia virus including the locus for α1(I) collagen in mouse. Cytogenet. Cell Genet. 43: 140-149.

Muragaki Y, Abe N, Ninomiya Y, Olsen BR & Ooshima A (1994) The human α1(XV) collagen chain contains a large amino-terminal non-triple helical domain with tandem repeat structure and homology to the α1(XVIII) collagen. J. Biol. Chem. 269: 4042-4046.

Muragaki Y, Jacenko O, Apte S, Mattei MG, Ninomiya Y & Olsen BR (1991) The α2(VIII) collagen gene. A novel member of the short chain collagen family located on the human chromosome 1. J. Biol. Chem. 266: 7721-7727.

Muragaki Y, Kimura T, Ninomiya Y & Olsen BR (1990a) The complete primary structure of two distinct forms of human α1(IX) collagen chains. Eur. J. Biochem. 192: 703-708.

Muragaki Y, Mattei MG, Yamaguchi N, Olsen BR & Ninomiya Y (1991) The complete primary structure of the human α1(VIII) chain and assignment of its gene (COL8A1) to chromosome 3. Eur. J. Biochem. 197: 615-622.

Muragaki Y, Nishimura I, Henney A, Ninomiya Y & Olsen BR (1990b) The α1(IX) collagen gene gives rise to two different transcripts in both mouse embryonic and human fetal RNA. Proc. Natl. Acad. Sci. USA. 87: 2400-2404.

Muragaki Y, Shiota C, Inoue M, Ooshima A, Olsen BR & Ninomiya Y (1992) α1(VIII)-collagen gene transcripts encode a short-chain collagen polypeptide and are expressed by various epithelial, endothelial and mesenchymal cells in newborn mouse tissues. Eur. J. Biochem. 207: 895-902.

Muragaki Y, Timmons S, Griffith CM, Oh SP, Fadel B, Quetermous T & Olsen BR (1995) Mouse Col18a1 is expressed in a tissue-specific manner as three alternative variants and is localized in the basement membrane zone. Proc. Natl. Acad. Sci. USA. 92: 8763-8767.

Muthukumaran G, Blumberg B & Kurkinen M (1989) The complete primary structure for the α1-chain of mouse collagen IV. Differential evolution of collagen IV domains. J. Biol. Chem. 264: 6310-6317.

Myers JC, Jones TA, Pohjolainen ER, Kadri AS, Goddard AD, Sheer D, Solomon E & Pihlajaniemi T (1990) Molecular cloning of α5(IV) collagen and assignment of the gene to the region of the X chromosome containing the Alport syndrome locus. Am. J. Hum. Genet. 46: 1024-1033.

Myers JC, Kivirikko S, Gordon MK, Dion AS & Pihlajaniemi T (1992) Identification of a previously unknown human collagen chain, α1(XV), characterized by extensive interruptions in the triple-helical region. Proc. Natl. Acad. Sci. USA. 89: 10144-10148.

Myers JC, Sun MJ, D'Ippolito JA, Jabs EW, Neilson EG & Dion AS (1993) Human cDNA clones transcribed from an unusual high-molecular- weight RNA encode a collagen chain. Gene. 123: 211-217.

Myers JC, Yang H, D'Ippolito JA, Presente A & Miller MKDAS (1994) The triple-helical region of human type XIX collagen consists of multiple collagenous subdomains and exhibits limited sequence homology to α1(XVI). J. Biol. Chem. 269: 18549-18557.

Müller U (1999) Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. Mech. Dev. 82: 3-21.

Nagy A, Rossant J, Nagy R, Abramow-Newerly W & Roder JC (1993) Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc. Natl. Acad. Sci. USA. 90: 8424-8428.

Nah HD, Bennett VD, Niu Z & Adams SI (1996) Alternative transcript of the chick α2(I) collagen gene is transiently expressed during endochondral bone formation and during development of the central nervous system. Dev. Dyn. 206: 146-158.

Nakata K, Ono K, Miyazaki J, Olsen BR, Muragaki Y, Adachi E, Yamamura K & Kimura T (1993) Osteoarthritis associated with mild chondrodysplasia in transgenic mice expressing α1(IX) collagen chains with a central deletion. Proc. Natl. Acad. Sci. USA. 90: 2870-2874.

Ninomiya Y, Castagnola P, Gerecke D, Gordon MK, Jacenko O, LuValle P, McCarthy M, Muragaki Y, Nishimura I, Oh S, Rosenblum N, Sato N, Sugrue S, Taylor R, Vasios G, Yamaguchi N & Olsen BR (1990) The Molecular Biology of Collagens with Short Triple-helical Domains. In:Sandell, L J & Boyd, C D (eds) Extracellular Matrix Genes 79-114. Academic Press, Inc., San Diego.

Ninomiya Y & Olsen BR (1984) Synthesis and characterization of cDNA encoding a cartilage-specific short collagen. Proc. Natl. Acad. Sci. USA. 81: 3014-3018.

O'Reilly MS, Boehm T, Shing Y, Fukai N, Vasios G, Lane WS, Flynn E, Birkhead JR, Olsen BR & Folkman J (1997) Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. Cell. 88: 277-285.

Oganesian A, Zhu Y & Sandell LJ (1997) Type IIA procollagen amino propeptide is localized in human embryonic tissues. J. Histochem. Cytochem. 45: 1469-1480.

Oh SP, Griffith CM, Hay ED & Olsen BR (1993) Tissue-specific expression of type XII collagen during mouse embryonic development. Dev. Dyn. 196: 37-46.

Oh SP, Kamagata Y, Muragaki Y, Timmons S, Ooshima A & Olsen BR (1994b) Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa-Yaa repeats identify a distinct family of collagenous proteins. Proc. Natl. Acad. Sci. USA. 91: 4229-4233.

Oh SP, Taylor RW, Gerecke DR, Rochelle JM, Seldin MF & Olsen BR (1992) The mouse α1(XII) and human α1(XII)-like collagen genes are localized on mouse chromosome 9 and human chromosome 6. Genomics. 14: 225-231.

Oh SP, Warman ML, Seldin MF, Cheng SD, Knoll JH, Timmons S & Olsen BR (1994a) Cloning of cDNA and genomic DNA encoding human type XVIII collagen and localization of the α1(XVIII) collagen gene to mouse chromosome 10 and human chromosome 21. Genomics. 19: 494-499.

Olsen BR (1995) The roles of collagen genes in skeletal development and morphogenesis. Experientia. 51: 194-195.

Oohashi T, Sugimoto M, Mattei M-G & Ninomiya Y (1994) Identification of a new collagen chain, α6(IV), by cDNA isolation and assignment of the gene to chromosome Xq22, which is the same locus for COL4A5. J. Biol. Chem. 269: 7520-7526.

Pace JM, Li Y, Seegmiller RE, Teuscher C, Taylor BA & Olsen BR (1997) Disproportionate micromelia (Dmm) in mice caused by a mutation in the C-propeptide coding region of Col2a1. Dev. Dyn. 208: 25-33.

Pajunen L, Tamminen M, Solomon E & Pihlajaniemi T (1989) Assignment of the gene coding for the α1 chain of collagen type XIII (COL13A1) to human chromosome region 10q11-qter. Cytogenet. Cell Genet. 52: 190-193.

Pallante KM, Niu Z, Zhao Y, Cohen AJ, Nah HD & Adams SL (1996) The chick α2(I) collagen gene contains two functional promoters, and its expression in chondrocytes is regulated at both transcriptional and post-transcriptional levels. J. Biol. Chem. 271: 25233-25239.

Palmiter RD, Brinster RL, Hammer RE, Trumbauer ME, Rosenfeld MG, Birnberg NC & Evans RM (1982) Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. Nature. 300: 611-615.

Pan T-C, Zhang R-Z, Mattei M-G, Timpl R & Chu ML (1992) Cloning and chromosomal location of human α1(XVI) collagen. Proc. Natl. Acad. Sci. USA. 89: 6565-6569.

Parente MG, Chung LC, Ryynänen J, Woodley DT, Wynn KC, Bauer EA, Mattei MG, Chu ML & Uitto J (1991) Human type VII collagen: cDNA cloning and chromosomal mapping of the gene. Proc. Natl. Acad. Sci. USA. 88: 6931-6935.

Peltonen S, Rehn M & Pihlajaniemi T (1997) Alternative splicing of mouse α1(XIII) collagen RNAs results in at least 17 different transcripts, predicting α1(XIII) collagen chains with length varying between 651 and 710 amino acid residues. DNA Cell Biol. 16: 227-234.

Pepe G, Giusti B, Bertini E, Brunelli T, Saitta B, Comeglio P, Bolognese A, Merlini L, Federici G, Abbate R & Chu ML (1999) A heterozygous splice site mutation in COL6A1 leading to an in-frame deletion of the alpha1(VI) collagen chain in an italian family affected by bethlem myopathy. Biochem. Biophys. Res. Commun. 258: 802-807.

Pereira R, Halford K, Sokolov BP, Khillan JS & Prockop DJ (1994) Phenotypic variability and incomplete penetrance of spontaneous fractures in an inbred strain of transgenic mice expressing a mutated collagen gene (COL1A1). J. Clin. Invest. 93: 1765-1769.

Pereira R, Khillan JS, Helminen HJ, Hume EL & Prockop DJ (1993) Transgenic mice expressing a partially deleted gene for type I procollagen (COL1A1). A breeding line with a phenotype of spontaneous fractures and decreased bone collagen and mineral. J. Clin. Invest. 91: 709-716.

Perälä M (1997) Molecular Genetics of the α 2(IX) Collagen Chain. Turun yliopisto, Turku Ser. D, Tom. 273.

Perälä M, Elima K, Metsäranta M, Rosati R, de Crombrugghe B & Vuorio E (1994) The exon structure of the mouse α2(IX) collagen gene shows unexpected divergence from the chick gene. J. Biol. Chem. 269: 5064-5071.

Perälä M, Hanninen M, Hastbacka J, Elima K & Vuorio E (1993) Molecular cloning of the human α2(IX) collagen cDNA and assignment of the human COL9A2 gene to chromosome 1. FEBS Lett. 319: 177-180.

Phillips CL, Morgan AL, Lever LW & Wenstrup RJ (1992) Sequence analysis of a full-length cDNA for the murine pro α2(I) collagen chain: comparison of the derived primary structure with human pro α2(I) collagen. Genomics. 13: 1345-1346.

Pihlajamaa T, Vuoristo MM, Annunen S, Perälä M, Prockop DJ & Ala-Kokko L (1998) Human COL9A1 and COL9A2 genes. Two genes of 90 and 15 kb code for similar polypeptides of the same collagen molecule. Matrix Biol. 17: 237-241.

Pihlajaniemi T & Myers JC (1987) Characterization of a pro-α2(I) collagen gene mutation by nuclease S1 mapping. Meth. Enzymol. 145: 213-222.

Pihlajaniemi T, Myllylä.R., Seyer J, Kurkinen M & Prockop DJ (1987) Partial characterization of a low molecular weight human collagen that undergoes alternative splicing. Proc. Natl. Acad. Sci. USA. 84: 940-944.

Pihlajaniemi T & Rehn M (1995) Two new collagen subgroups: Membrane associated collagens and types XV and XVIII. Prog. Nucleic Acid Res. Mol. Biol. 50: 225-262.

Pihlajaniemi T & Tamminen M (1990) The α1 chain of type XIII collagen consists of three collagenous and four noncollagenous domains, and its primary transcript undergoes complex alternative splicing. J. Biol. Chem. 265: 16922-16928.

Pilz A, Prohaska R, Peters J & Abbott C (1994) Genetic linkage analysis of the Ak1, Col5a1, Epb7.2, Fpgs, Grp78, Pbx3, and Notch1 genes in the region of mouse chromosome 2 homologous to human chromosome 9q. Genomics. 21: 104-109.

Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J & Gray J (1988) Fluorescence in situ hybridization with human chromosome- specific libraries: detection of trisomy 21 and translocations of chromosome 4. Proc. Natl. Acad. Sci. USA. 85: 9138-9142.

Poschl E, Pollner R & Kuhn K (1988) The genes for the α1(IV) and α2(IV) chains of human basement membrane collagen type IV are arranged head-to-head and separated by a bidirectional promoter of unique structure. EMBO J. 7: 2687-2695.

Prestridge DS (1995) Predicting Pol II promoter sequences using transcription factor binding sites. J. Mol. Biol. 249: 923-932.

Prockop DJ & Kivirikko KI (1995) Collagens: molecular biology, diseases, and potentials for therapy. Annu. Rev. Biochem. 64: 403-434.

Rebhan M, Chalifa-Caspi V, Prilusky J & Lancet D (1997) GeneCards: integrating information about genes, proteins and diseases. Trends Genet. 13: 163-163.

Rehn M, Hintikka E & Pihlajaniemi T (1994) Primary structure of the α1 chain of mouse type XVIII collagen, partial structure of the corresponding gene, and comparison of the α1(XVIII) chain with its homologue, the α1(XV) collagen chain. J. Biol. Chem. 269: 13929-13935.

Rehn M, Hintikka E & Pihlajaniemi T (1996) Characterization of the mouse gene for the α1 chain of type XVIII collagen (Col18a1) reveals that the three variant N- terminal polypeptide forms are transcribed from two widely separated promoters. Genomics. 32: 436-446.

Rehn M & Pihlajaniemi T (1994) α1(XVIII), a collagen chain with frequent interruptions in the collagenous sequence, a distinct tissue distribution, and homology with type XV collagen. Proc. Natl. Acad. Sci. USA. 91: 4234-4238.

Rehn M & Pihlajaniemi T (1995) Identification of three N-terminal ends of type XVIII collagen chains and tissue-specific differences in the expression of the corresponding transcripts. The longest form contains a novel motif homologous to rat and Drosophila frizzled proteins. J. Biol. Chem. 270: 4705-4711.

Rehn M & Pihlajaniemi T (1996) Type XV and type XVIII collagens, a new subgroup within the family of collagens. Cell Dev. Biol. 7: 673-679.

Retief E, Parker MI & Retief AE (1985) Regional chromosome mapping of human collagen genes α2(I) and α1(I) (COLIA2 and COLIA1). Hum. Genet. 69: 304-308.

Richards AJ, Martin S, Nicholls AC, Harrison JB, Pope FM & Burrows NP (1998) A single base mutation in COL5A2 causes Ehlers-Danlos syndrome type II. J. Med. Genet. 35: 846-848.

Robertson E, Bradley A, Kuehn M & Evans M (1986) Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature. 323: 445-448.

Rosati R, Horan GS, Pinero GJ, Garofalo S, Keene DR, Horton WA, Vuorio E, de Crombrugghe B & Behringer RR (1994) Normal long bone growth and development in type X collagen-null mice. Nat. Genet. 8: 129-135.

Rowe LB, Nadeau JH, Turner R, Frankel WN, Letts VA, Eppig JT, Ko MS, Thurston SJ & Birkenmeier EH (1994) Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. Mamm. Genome. 5: 253-274.

Rozen S & Skaletsky HJ (1997) Primer3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

Saarela J, Rehn M, Oikarinen A, Autio-Harmainen H & Pihlajaniemi T (1998b) The short and long forms of type XVIII collagen show clear tissue specificities in their expression and location in basement membrane zones in humans. Am. J. Pathol. 153: 611-626.

Saarela J, Ylikärppä R, Rehn M, Purmonen S & Pihlajaniemi T (1998a) Complete primary structure of two variant forms of human type XVIII collagen and tissue-specific differences in the expression of the corresponding transcripts. Matrix Biol. 16: 319-328.

Saitta B, Timpl R & Chu ML (1992) Human α2(VI) collagen gene. Heterogeneity at the 5'-untranslated region generated by an alternate exon. J. Biol. Chem. 267: 6188-6196.

Saitta B, Wang Y-M, Renkart L, Zhang R-Z, Pan T-C, Timpl R & Chu ML (1991) The exon organization of the triple-helical coding regions of the human α1(VI) and α2(VI) collagen genes is highly similar. Genomics. 11: 145-153.

Sambrook J, Fritch EF & Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sandberg M, Tamminen M, Hirvonen H, Vuorio E & Pihlajaniemi T (1989) Expression of mRNAs coding for the α1 chain of type XIII collagen in human fetal tissues: comparison with expression of mRNAs for collagen types I, II, and III. J. Cell Biol. 109: 1371-1379.

Sandell LJ (1996) Genes and gene regulation of extracellular matrix proteins: an introduction. Connect. Tissue Res. 35: 1-6.

Sanger F, Nicklen S & Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74: 5463-5467.

Saus J, Quinones S, MacKrell A, Blumberg B, Muthukumaran G, Pihlajaniemi T & Kurkinen M (1989) The complete primary structure of mouse α2(IV) collagen. Alignment with mouse α1(IV) collagen. J. Biol. Chem. 264: 6318-6324.

Saus J, Wieslander J, Langeweld JPM, Quinones S & Hudson PG (1988) Identification of the Goodpasture antigens as the α3(IV) chain of collagen IV. J. Biol. Chem. 263: 13374-13380.

Schnieke A, Harbers K & Jaenisch R (1983) Embryonic lethal mutation in mice induced by retrovirus insertion into the α1(I) collagen gene. Nature. 304: 315-320.

Schnittger S, Herbst H, Schuppan D, Dannenberg C, Bauer M & Fonatsch C (1995) Localization of the undulin gene (UND) to human chromosome band 8q23. Cytogenet. Cell Genet. 68: 233-234.

Schuppan D, Cantaluppi MC, Becker J, Veit A, Bunte T, Troyer D, Schuppan F, Schmid M, Ackermann R & Hahn EG (1990) Undulin, an extracellular matrix glycoprotein associated with collagen fibrils. J. Biol. Chem. 265: 8823-8832.

Schurr E, Skamene E, Morgan K, Chu ML & Gros P (1990) Mapping of Col3a1 and Col6a3 to proximal murine chromosome 1 identifies conserved linkage of structural protein genes between murine chromosome 1 and human chromosome 2q. Genomics. 8: 477-486.

Schwartzberg PL, Goff SP & Robertson EJ (1989) Germ-line transmission of a c-abl mutation produced by targeted gene disruption in ES cells. Science. 246: 799-803.

Seegmiller R, Fraser FC & Sheldon H (1971) A new Chondrodystrophic mutant in mice. Electron microscopy of normal and abnormal chondrogenesis. J. Cell Biol. 48: 580-593.

Shaw LM & Olsen BR (1991) FACIT collagens: diverse molecular bridges in extracellular matrices. Trends Biochem. Sci. 16: 191-194.

Shows TB, Tikka L, Byers MG, Eddy RL, Haley LL, Henry WM, Prockop DJ & Tryggvason K (1989) Assignment of the human collagen α1(XIII) chain gene (COL13A1) to the q22 region of chromosome 10. Genomics. 5: 128-133.

Singer II, Kawka DW, Kazazis DM & Clark RA (1984) In vivo co-distribution of fibronectin and actin fibers in granulation tissue: immunofluorescence and electron microscope studies of the fibronexus at the myofibroblast surface. J. Cell Biol. 98: 2091-2106.

Soininen R, Haka-Risku T, Prockop DJ & Tryggvason K (1987) Complete primary structure of the α1 chain of human basement membrane (type IV) collagen. FEBS Lett. 225: 188-194.

Soininen R, Huotari M, Ganguly A, Prockop DJ & Tryggvason K (1989) Structural organization of the gene for the α1 chain of human type IV collagen. J. Biol. Chem. 264: 13565-13571.

Soininen R, Huotari M, Hostikka SL, Prockop DJ & Tryggvason K (1988) The structural genes for α1 and α2 chains of human type IV collagen are divergently encoded on opposite DNA strands and have an overlapping promoter region. J. Biol. Chem. 263: 17217-17220.

Stacey A, Bateman J, Choi T, Mascara T, Cole W & Jaenisch R (1988) Perinatal lethal osteogenesis imperfecta in transgenic mice bearing an engineered mutant pro-α1(I) collagen gene. Nature. 332: 131-136.

Stokes DG, Saitta B, Timpl R & Chu ML (1991) Human α3(VI) collagen gene. Characterization of exons coding for the amino-terminal globular domain and alternative splicing in normal and tumor cells. J. Biol. Chem. 266: 8626-8633.

Strom CM & Upholt WB (1984) Isolation and characterization of genomic clones corresponding to the human type II procollagen gene. Nucleic Acids Res. 12: 1025-1038.

Sugimoto M, Oohashi T & Ninomiya Y (1994) The genes COL4A5 and COL4A6, coding for basement membrane collagen chains α5(IV) and α6(IV), are located head-to- head in close proximity on human chromosome Xq22 and COL4A6 is transcribed from two alternative promoters. Proc. Natl. Acad. Sci. USA. 91: 11679-11683.

Sumiyoshi H, Inoguchi K, Khaleduzzaman M, Ninomiya Y & Yoshioka H (1997) Ubiquitous expression of the α1(XIX) collagen gene (Col19a1) during mouse embryogenesis becomes restricted to a few tissues in the adult organism. J. Biol. Chem. 272: 17104-17111.

Sutmuller M, Bruijn JA & de Heer E (1997) Collagen types VIII and X, two non-fibrillar, short-chain collagens. Structure homologies, functions and involvement in pathology. Histol. Histopathol. 12: 557-566.

Takahara K, Hoffman GG & Greenspan DS (1995) Complete structural organization of the human α1(V) collagen gene (COL5A1): divergence from the conserved organization of other characterized fibrillar collagen genes. Genomics. 29: 588-597.

Takahara K, Sato Y, Okazawa K, Okamoto N, Noda A, Yaoi Y & Kato I (1991) Complete primary structure of human collagen α1(V) chain. J. Biol. Chem. 226: 13124-13129.

Takahashi E, Hori T & Sutherland GR (1990) Mapping of the human type II collagen gene (COL2A1) proximal to fra(12) (q13.1) by nonisotopic in situ hybridization. Cytogenet. Cell Genet. 54: 84-85.

Tanaka T, Takahashi K, Furukawa F & Imamura S (1992) Molecular cloning and characterization of type VII collagen cDNA. Biochem. Biophys. Res. Commun. 183: 958-963.

Taverna D, Disatnik MH, Rayburn H, Bronson RT, Yang J, Rando TA & Hynes RO (1998) Dystrophic muscle in mice chimeric for expression of α5 integrin. J. Cell Biol. 143: 849-859.

Tenhunen K, Laan M, Manninen T, Palotie A, Peltonen L & Jalanko A (1995) Molecular cloning, chromosomal assignment, and expression of the mouse aspartylglucosaminidase gene. Genomics. 30: 244-250.

Thomas JT, Cresswell CJ, Rash B, Nicolai H, Jones T, Solomon E, Grant ME & Boot-Handford RP (1991) The human collagen X gene. Complete primary translated sequence and chromosomal localization. Biochem. J. 280: 617-623.

Thomas KR & Capecchi MR (1987) Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell. 51: 503-512.

Tikka L, Elomaa O, Pihlajaniemi T & Tryggvason K (1991) Human α1(XIII) collagen gene. Multiple forms of the gene transcripts are generated through complex alternative splicing of several short exons. J. Biol. Chem. 266: 17713-17719.

Tikka L, Pihlajaniemi T, Henttu P, Prockop DJ & Tryggvason K (1988) Gene structure for the α1 chain of a human short-chain collagen (type XIII) with alternatively spliced transcripts and translation termination codon at the 5' end of the last exon. Proc. Natl. Acad. Sci. USA. 85: 7491-7495.

Toman PD & de Crombrugghe B (1994) The mouse type-III procollagen-encoding gene: genomic cloning and complete DNA sequence. Gene. 147: 161-168.

Tromp G, Kuivaniemi H, Stacey A, Shikata H, Baldwin CT, Jaenisch R & Prockop DJ (1988) Structure of a full-length cDNA clone for the prepro α1(I) chain of human type I procollagen. Biochem. J. 253: 919-922.

Trueb B & Winterhalter KH (1986) Type VI collagen is composed of a 200 kd subunit and two 140 kd subunits. EMBO J. 5: 2815-2819.

Trueb J & Trueb B (1992) Type XIV collagen is a variant of undulin. Eur. J. Biochem. 207: 549-557.

Tsipouras P, Schwartz RC, Liddell AC, Salkeld CS, Weil D & Ramirez F (1997) Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located on the long arm of human chromosome 2. Genomics. 3: 275-277.

Turner N, Mason PJ, Brown R, Fox M, Povey S, Rees A & Pusey CD (1992) Molecular cloning of the human Goodpasture antigen demonstrates it to be the α3 chain of type IV collagen. J. Clin. Invest. 89: 592-601.

Vandenberg P, Khillan JS, Prockop DJ, Helminen H, Kontusaari S & Ala-Kokko L (1991) Expression of a partially deleted gene of human type II procollagen (COL2A1) in transgenic mice produces a chondrodysplasia. Proc. Natl. Acad. Sci. USA. 88: 7640-7644.

Vandenberg P, Vuoristo MM, Ala-Kokko L & Prockop DJ (1996) The mouse col11a2 gene. Some transcripts from the adjacent rxr-α gene extend into the col11a2 gene. Matrix Biol. 15: 359-367.

Vuorio E & de Crombrugghe B (1990) The family of collagen genes. Annu. Rev. Biochem. 59: 837-872.

Vuoristo MM, Pihlajamaa T, Vandenberg P, Prockop DJ & Ala-Kokko L (1995) The human COL11A2 gene structure indicates that the gene has not evolved with the genes for the major fibrillar collagens. J. Biol. Chem. 270: 22873-22881.

Wagner EF, Stewart TA & Mintz B (1981a) The human β-globin gene and a functional viral thymidine kinase gene in developing mice. Proc. Natl. Acad. Sci. USA. 78: 5016-5020.

Wagner TE, Hoppe PC, Jollick JD, Scholl DR, Hodinka RL & Gault JB (1981b) Microinjection of a rabbit β-globin gene into zygotes and its subsequent expression in adult mice and their offspring. Proc. Natl. Acad. Sci. USA. 78: 6376-6380.

Warman ML, McCarthy MT, Perälä M, Vuorio E, Knoll JH, McDaniels CN, Mayne R, Beier DR & Olsen BR (1994) The genes encoding α2(IX) collagen (COL9A2) map to human chromosome 1p32.3-p33 and mouse chromosome 4. Genomics. 23: 158-162.

Warman ML, Tiller GE, Polumbo PA, Seldin MF, Rochelle JM, Knoll JH, Cheng SD & Olsen BR (1993) Physical and linkage mapping of the human and murine genes for the α1 chain of type IX collagen (COL9A1). Genomics. 17: 694-698.

Weil D, Bernard M, Gargano S & Ramirez F (1987) The proα2(V) collagen gene is evolutionary related to the fibrillar-forming collagens. Nucleic Acids Res. 15: 181-198.

Weil D, Mattei MG, Passage E, N'Guyen VC, Pribula-Conway D, Mann K, Deutzmann R, Timpl R & Chu ML (1988) Cloning and chromosomal localization of human genes encoding the three chains of type VI collagen. Am. J. Hum. Genet. 42: 435-445.

White JK, Auerbach W, Duyao MP, Vonsattel JP, Gusella JF, Joyner AL & MacDonald ME (1997) Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. Nat. Genet. 17: 404-410.

Wälchli C, Koller E, Trueb J & Trueb B (1992) Structural comparison of the chicken genes for α1(VI) and α2(VI) collagen. Eur. J. Biochem. 205: 583-589.

Wälchli C, Trueb J, Kessler B, Winterhalter KH & Trueb B (1993) Complete primary structure of chicken collagen XIV. Eur. J. Biochem. 212: 483-490.

Yamaguchi N, Kimura S, McBride OW, Hori H, Yamada Y, Kanamori T, Yamakoshi H & Nagai Y (1992) Molecular cloning and partial characterization of a novel collagen chain, α1(XVI), consisting of repetitive collagenous domains and cysteine-containing non-collagenous segments. J. Biochem. 112: 856-863.

Yang TT & Kain SR (1995) Fast hybridization solution for the detection of immobilized nucleic acids. Biotechniques. 18: 498-503.

Yoshioka H, Inoguchi K, Khaleduzzaman M, Ninomiya Y, Andrikopoulos K & Ramirez F (1995) Coding sequence and alternative splicing of the mouse α1(XI) collagen gene (Col11a1). Genomics. 28: 337-340.

Yoshioka H, Zhang H, Ramirez F, Mattei M-G, Moradi-Ameli M, van der Rest M & Gordon MK (1992) Synteny between the loci for a novel FACIT-like collagen locus (D6S228E) and α1(IX) collagen (COL9A1) on 6q12-q14 in humans. Genomics. 13: 884-886.

Zanussi S, Doliana R, Segat D, Bonaldo P & Colombatti A (1992) The human type VI collagen gene. mRNA and protein variants of the α3 chain generated by alternative splicing of an additional 5-end exon. J. Biol. Chem. 267: 24082-24089.

Zhou J, Leinonen A & Tryggvason K (1994) Structure of the human type IV collagen COL4A5 gene. J. Biol. Chem. 269: 6608-6614.

Zhou J, Mochizuki T, Smeets H, Antignac C, Laurila P, De Paepe A, Tryggvason K & Reeders ST (1993) Deletion of the paired α5(IV) and α6(IV) collagen genes in inherited smooth muscle tumors. Science. 261: 1167-1169.

Zijlstra M, Li E, Sajjadi F, Subramani S & Jaenisch R (1989) Germ-line transmission of a disrupted β2-microglobulin gene produced by homologous recombination in embryonic stem cells. Nature. 342: 435-438.