INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9431848

Structure, function and expression of type V collagen

Andrikopoulos, Konstantinos Ioannou, Ph.D.

City University of New York, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

STRUCTURE, FUNCTION AND EXPRESSION OF TYPE V COLLAGEN

by

KONSTANTINOS IOANNOU ANDRIKOPOULOS

A dissertation submitted to the Graduate Faculty in Biomedical Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1994

This manuscript has been read and accepted for the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

3-29-94
Date

Robert J. Desnick, Ph.D., M.D.
Chair of Examining Committee

3-27-94
Date

Terry A. Krulwich, Ph.D.
Executive Officer

Christine Eng, M.D.

Yiannis Ioannou, Ph.D.

Rudolf Jaenisch, M.D.

Francesco Ramirez, Ph.D.

Supervisory Committee

The City University of New York

ABSTRACT

STRUCTURE, FUNCTION AND EXPRESSION OF TYPE V COLLAGEN

by

Konstantinos Ioannou Andrikopoulos

Adviser: Professor Francesco Ramirez

The correct assembly of the fibrillar collagen networks plays a critical role in animal morphogenesis. Type V collagen is a minorly represented member of the fibrillar collagen family; it is believed that one of its functions is to regulate type I collagen fibrillogenesis. This study was designed to prove the validity of this hypothesis. We first cloned the mouse pro-α2(V) collagen gene (*col5a2*) and examined its developmental expression pattern. During early embryonic stages, *col5a2* expression is low and diffusely distributed in the peritoneal membranes, intestinal and craniofacial mesenchymes. At later stages, *col5a2* gene expression is higher and more restricted to the primary ossified regions, perichondrium, joints, tendon, atrioventicular valve of the heart and selected portions of the head. Thus, this analysis provided further indirect evidence for the postulated cooperativity of types I and V collagens in the fibrillogenesis of non-cartilaginous matrices. In order to directly confirm this relationship, we targeted the *col5a2* exon coding for the N-telopeptide by homologous recombination in mouse embryonic stem cells. Homozygous mice present a severe kyphotic phenotype, substantial skin fragility, and a defective corneal stroma. These abnormalities were correlated to ultrastructural deficiencies in type I collagen fibrillogenesis. We found that type I collagen fibrils of homozygous animals are thicker and more disorganized than those of wild type littermates. Altogether, the data provide proof that type V collagen plays a critical role in regulating the growth and tri-dimensional organization of type I collagen fibrils.

To my parents Ioanni and Maria,

Στους γονείς μου Ιωάννη και Μαρία,

Χωρίς την απέραντη αγάπη τους αυτές οι λέξεις δεν θα ήταν δυνατό να γραφτούν.

ACKNOWLEDGEMENTS

I hope that all students in their graduate education will be as fortunate as I have been. For starters, a project that worked...! (...more or less). This would have been impossible to accomplish without the vision, continuous support (at all levels), and encouragement of my mentor Francesco Ramirez. A "Supreme Educator", you will always remain "The Boss", and my intellectual father (= "Godfather", see under the Greek religion chapter...!). Thank you for always being there, and for showing me how things should be done.

The unusual period of time I spent away from New York in collaborating laboratories, has been as important in contributing to my development as a scientist and as a person, as to the time I spent here at Mt. Sinai. Hence, my extensive acknowledgement list! My sincere Thank you, to:

Dr. Michael Solursh for allowing me to be in his lab in Iowa city, and to Dr. Hiroaki Suzuki for guiding me to the world of *in situ*'s.

Dr. Rudolf Jaenisch for allowing me to spent a considerable period of time in his lab in Boston, and for being patient with my defense date changes. I believe it was worth the waiting! Drs. Xin Liu and Hong Wu, for introducing me to the stem cell world; I admire your accomplishments, you will always be my dearest friends. The people in the Jaenisch lab for their hospitality.

Dr. Paul Bornstein while in his sabbatical at the Whitehead, for the lifts to the retreat and the Gordon conference, and the extra-dimension he added to the way I saw things in the extracellular matrix and life in general.

Dr. David Birk for allowing me to spent an incredibly educational week in his lab at Tufts. I will return...!

Back at Mt. Sinai ...

The wonderful people in the Ramirez lab for making it soooo much easier...

My Dean, Dr. Terry Krulwich for believing in me and for being there.

Dr. Yiannis Ioannou for sharing ideas and opinions with me, and for keeping me sane the last five years. Φίλε παντοτινέ σ' ευχαριστώ!

Dr. Tom Einhorn for guiding me with his clinical eye to the right direction.

Dr. Michael Klein for helping me with the bone "stuff".

Kathy Pelton-Henrion and Vladimir Protopopov for the "gorgeous" electron micrographs.

Dr. Sandy Masur for correcting my English and for helping me with the cell "stuff"; forever friends and collaborators. Dr. Stefan Antohi, also for help with the cell "stuff".

Dr. Anne Wang for being a true friend from the veeeery beginning, and for introducing me to the American culture. I owe you!

...Finally, my sister Angeliki for putting up with our parents and looking after them the time I was away. I do not think I will ever be able to "pay" you back for all you have given me. Σ' αγαπώ.

This work would not have been possible without the love, patience and continuous support of my parents. Their contribution shall remain unparallel. This thesis is dedicated to them. They have set "The Standard". It will be a very difficult task to follow. I thank them for showing me the way and the way.

=

"Η δουλειά αυτών των πέντε (είκοσι εννέα!) χρόνων δεν θα ήταν εφικτή χωρίς την αγάπη, στοργή και υπομονή σας. Η συνεισφορά σας θα παραμείνει ανεπανάλειπτη. Γι' αυτό και αφιερώνω τη διατριβή μου σε εσάς. Έχετε θέσει "Το Επίπεδο". Το παραδειγμά σας είναι δύσκολο να ακολουθηθεί. Σας ευχαριστώ που μου δείξατε το δρόμο και τον τρόπο. Σας λατρεύω και σας αγαπώ."

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

The extracellular matrix (ECM) has originally been viewed as an inert scaffolding whose main function is to maintain the integrity of the developing and mature organ systems and tissues. Recent evidence has shown that the ECM is also a dynamic milieu in which cells become organized, exchange signals, and differentiate (Hay, 1991). Examination of these processes has led to the identification of several new ECM components, matrix receptors, and cell-matrix interactions. The ECM is composed of multidomain macromolecules linked together by covalent and noncovalent interactions to form highly complex networks. There are two major types of matrices. One is the *interstitium*, which is synthesized by mesenchymal cells and forms the stroma of organs. The other is the *basement membrane*, which is in turn produced by endothelial and epithelial cells. These two matrices are comprised of the four major classes of the ECM's macromolecules:

i) collagens, ii) elastic fibers iii) nectins, and iv) proteoglycans.

Collagens

The collagens represent a family of structurally related but genetically distinct proteins (Bornstein and Sage, 1980; Miller, 1988). They all contain one or more domain consisting of three α-chains twisted together into a right-handed helix. The collagens are the most abundant class of vertebrate proteins, constituting almost 30% of the total body protein content. Their involvement in a variety of developmental programs such as cell adhesion and cell movement, and in several physiological processes, including tissue remodeling and wound healing, has been extensively documented (Bornstein and Sage, 1980). To date, nineteen collagen types with more than thirty genes specifying their subunits (α-chains) have been described (for recent review see Kivirikko, 1993). The

collagen heterogeneity is further amplified by the existence of alternative molecular species for the same collagen type, as well as by the ability of these molecules to form heterotypic trimers.

Table 1. Collagen types (adapted from Kivirikko, 1993).

Type	Constituent chains	Gene	Chromosomal localization	Distribution
I	α1(I)	COL1A1	17q21.3-q22	Ubiquitous
	α2(I)	COL1A2	7q21.3-q22	
II	α1(II)	COL2A1	12q13-q14	Cartilage, Vitreous humor
III	α1(III)	COL3A1	2q24.3-q31	Like type I
IV	α1(IV)	COL4A1	13q34	Basement membranes
	α2(IV)	COL4A2	13q34	
	α3(IV)	COL4A3	2q35-q37	
	α4(IV)	COL4A4	2q35-q37	
	α5(IV)	COL4A5	Xq22	
V	α1(V)	COL5A1	9q34.2-34.3	Interstitial
	α2(V)	COL5A2	2q24.3-q31	
	α3(V)	COL5A3		
VI	α1(VI)	COL6A1	21q22.3	Soft tissues
	α2(VI)	COL6A2	21q22.3	
	α3(VI)	COL6A3	2q37	
VII	α1(VII)	COL7A1	3p21	Anchoring fibrils
VIII	α1(VIII)	COL8A1	3q12-q13.1	Endothelium, Mesenchyme
	α2(VIII)	COL8A2	1p32.3-p34.3	
IX	α1(IX)	COL9A1	6q12-14	Like type II
	α2(IX)	COL9A2		
	α3(IX)	COL9A3		
X	α1(X)	COL10A1	6q21-q22	Hypertrophic cartilage
XI	α1(XI)	COL11A1	1p21	Like type II
	α2(XI)	COL11A2	6p21.2	
	α3(XI)	COL11A3	12q13-q14	
XII	α1(XII)	COL12A1	6	Many tissues
XIII	α1(XIII)	COL13A1	10q22	Many tissues
XIV	α1(XIV)	COL14A1		Skin, Tendon
XV	α1(XV)	COL15A1	9q21-p22	Many tissues
XVI	α1(XVI)	COL16A1	1p34-p35	Fibroblasts, Keratinocytes
XVII	α1(XVII)	COL17A1	10q24.3	Skin hemidesmosomes
XVIII	α1(XVIII)	COL18A1		Liver, Kidney, Placenta
Y	α1(Y)	D6S228E	6q12-q14	Rhabdomyosarcoma cells

In broad morphologic terms, the collagen types are segregated according to whether they form fibrillar or non-fibrillar aggregates. Based on their relative representation, the fibrillar collagens are further subdivided into "major" (collagens I, II and III) and "minor" types (collagens V and XI). The basic structure of a fibrillar collagen consists of three α-

chains made of uninterrupted collagenous domains. In contrast, the subunits of the non-fibrillar collagens contain interrupted collagenous domains and do not form quarter-staggered arrays. Examples include the extensively studied collagen types IV (an integral component of basement membrane) and VII (an integral component of anchoring fibrils), and the more recently described fibril-associated collagens with interrupted triple helices (FACIT's, types IX, XII, XIV, Y) (van der Rest and Garrone, 1991; Yoshioka *et al.*, 1992).

Human Fibrillar Collagen Proteins and Genes

The five members of the fibrillar collagen group share numerous features. Their similarities include the structure of their precursor subunits (procollagens), the biosynthetic steps giving rise to the trimers, and the extracellular aggregates that they form (fibrils). Furthermore, the similarities extend to the organization of the corresponding genes.

The fibrillar collagens are distributed in a wide variety of tissues (Table 1). Type I collagen is the major non-mineralized component of bone and is distributed in virtually all connective tissues (Bornstein and Sage, 1980; Nimni and Harkness, 1988). Type III collagen has a distribution pattern similar to that of type I, although at lower levels, but with relatively higher representation in the distensible tissues such as the internal organs and skin (Bornstein and Sage, 1980; Nimni and Harkness, 1988). Based on its pattern of expression , type II collagen was originally regarded as a cartilaginous molecule. Recent data has however shown a non-cartilaginous distribution of this collagen type (Cheah et al., 1991b; Swiderski and Solursh, 1992). The significantly lower levels of types V and XI have limited the characterization of their structural-functional features. This has been recently overcome by cloning experiments and immunohistochemical studies suggesting a wider spectrum of expression than before thought (Burgeson, 1988). These findings are discussed in more detail later in this chapter. Following is a brief synopsis of fibrillar collagen structure and metabolism, and their contribution to disease.

i) Protein structure and Fibrillogenesis:

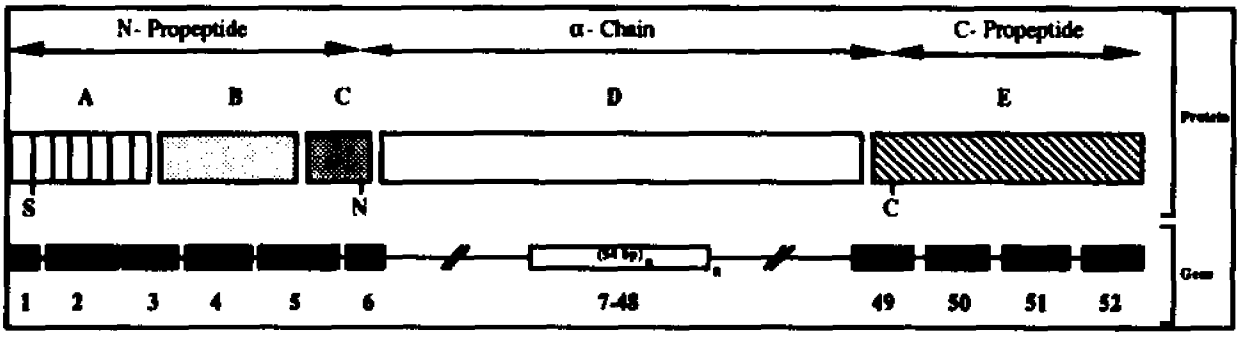


Figure 1a.

Schematic representation of the distribution of exon sequences coding for the major domains of the fibrillar procollagen.
(A) N-terminal cysteine-rich, globular domain containing the signal peptidase cleavage site (S);
(B) N-terminal triple helical domain;
(C) Short non-helical segment containing the N-proteinase cleavage site (N);
(D) Central α-helical domain that participates in triple helical assembly; and
(E) C-propeptide containing the C-proteinase cleavage site (C).

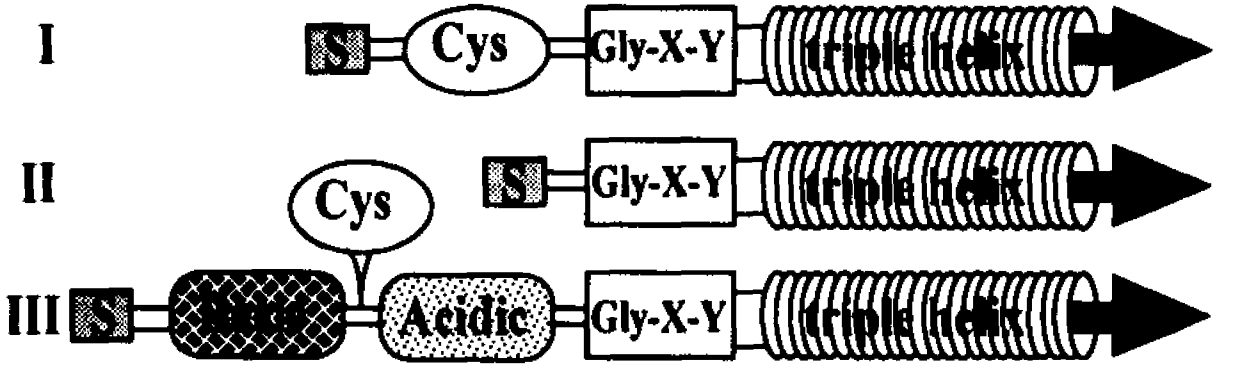


Figure 1b.

Schematic representation of the 3 different groups of N-propeptides of the fibrillar collagens.
(S) signal peptide, (Cys) cysteine-rich globular subdomain, (Basic) basic residue-rich globular subdomain, (Acidic) acidic residue-rich globular subdomain, (Gly-X-Y) small triple helical subdomain, (triple helix) long triple helical domain.

A typical fibrillar procollagen molecule consists of three major domains: a long triple helix and two shorter globular terminal projections [N(amino)- and C(carboxy)-propeptides, Fig. 1a). In contrast to the long triple helical domains and the C-propeptides, the N-propeptides differ greatly in composition and size and appear to conform to one of three basic architectures (Fig. 1b) (Timpl and Glanville, 1981; Su *et al.*, 1989; Yoshioka

and Ramirez, 1990; Greenspan *et al.*, 1991). The first architecture, exhibited by pro-α2(I), pro-α1(III) and pro-α2(V), displays the longitudinal arrangement of a cysteine-rich globular domain, a short collagenous sequence, and a shorter non-helical segment (Fig. 1b I). The second, which is characteristic of pro-α1(I) and the adult form of pro-α1(II), lacks in its entirety the cysteine-rich globular domain and as a consequence, the signal peptide is directly connected to the short collagenous sequence (Fig. 1b II). The third, is the longest N-propeptide described as yet, and is exhibited by pro-α1(V) and pro-α1(XI) collagen chains. It displays a very long globular domain followed by an interrupted collagenous domain. Additionally, the globular domain is bisected by a cluster of cysteines which separate a somewhat basic upstream subdomain from a downstream subdomain rich in tyrosines and acidic residues (Fig. 1b III) (Yoshioka and Ramirez, 1990; Greenspan *et al.*, 1991).

For the proper folding of the chains into the left-handed helical conformation, a glycine is required to be present at every third amino acid (aa) residue position. This requirement therefore results in the highly repeated (up to 1,000 aa) and uninterrupted Gly-X-Y motif of the α-chains. The X and the Y are frequently a proline and a hydroxyproline or a hydroxylysine, respectively. The hydroxylated residues have been shown to provide thermal stability and rigidity to the collagen trimer (Berg *et al.*, 1973). Concurrent with its translation, a pre-procollagen molecule translocates into the rough endoplasmic reticulum and begins to undergo post-translational modifications. These include the hydroxylation of proline and lysine residues, and N-linked glycosylation of the C-propeptide (Bailey *et al.*, 1974; Olsen *et al.*, 1977).

The formation of the triple-helix is initiated by disulfide bonding of interchain and intrachain cysteine residues at the C-propeptide. Triple-helical propagation nucleates in a progressively N-terminal direction (Bornstein and Sage, 1980). Parallel to this process, but limited by the rate of the triple helical assembly, is the glycosylation of the hydroxyproline

and hydroxylysine residues on the unassembled portions of the α-chains (Bailey *et al.*, 1974). These post-translational modifications stop concomitantly with the formation of the triple-helix.

Upon secretion into the extracellular matrix, procollagen trimers reach the mature collagen conformation by the action of specific endopeptidases that remove the globular N- and C-terminal domains. With the loss of these domains, the processed collagen homo/hetero-trimers self-assemble laterally in quarter staggered arrays to form the fibrils. Increasing evidence suggests that partially processed procollagen molecules may play an important role in fibrillogenesis by regulating both the growth and the diameter of the fibrils (Keene *et al.*, 1987; Fleischmajer *et al.*, 1988; Fleischmajer *et al.*, 1990; Romanic *et al.*, 1991). These partially processed molecules are usually referred to as pN-collagens (containing the N-terminal propeptides but not the C-terminal propeptides) and pC-collagens (containing the C-terminal propeptides but not the N-terminal propeptides). pN-type III molecules were the first to be experimentally implicated in determining the diameter of type I collagen fibrils (Miyahara *et al.*, 1984; Fleischmajer *et al.*, 1985). Most type I collagen rich tissues also contain type III collagen of which a considerable fraction is present in the partially processed monomer pN-type III (Fessler *et al.*, 1981; Timpl and Glanville, 1981). Subsequent studies involving several tissues and different developmental stages, showed that the ratio of collagen III and pN-type III to collagen I is inversely proportional to the diameter of collagen fibrils (Fleischmajer *et al.*, 1988). Romanic *et al.* (1991) more recently, demonstrated that pN-type III forms true copolymers with type I collagen. They in fact showed that pN-type III presence inhibits the rate at which type I assembled into fibrils and also decreases the amount of type I incorporated into the fibrils. The results were consistent with a model in which pN-collagen III regulates the diameter of type I fibrils by coating the surface of the fibrils, thus allowing tip growth but not lateral growth.

The spontaneous assembly of collagen into fibrils is dependent on the specific distribution and interaction of the charged and polar amino acids in the X- and Y- positions of the triple-helix. Finally, the last extracellular modification includes selective oxidative deamination of lysyl and hydroxy-lysyl residues by lysyl oxidase to produce aldehydes which participate in intermolecular crosslinks (via Schiff's bases) (Tanzer, 1973). This last step provides tensile strength to the collagen fibers.

ii) Gene structure:

Fibrillar collagens are encoded by distinct genes with typically more than 50 exons (Lee *et al.*, 1991a). Interspersed throughout the genome, their sizes range from 18kb to more than 60kb. The high structural homology of the fibrillar collagen genes supports the hypothesis that they may have evolved from a common ancestral gene.

The typical structure of a fibrillar collagen gene includes several conserved features. The last four exons encode the C-propeptide domain, and thus the following structural elements: the C-proteinase cleavage site, 7-8 similarly spaced cysteine residues and the N-linked glycosylation site. Exons usually 54bp in size, or differing in multiples of three (45, 99, 108, 162bp), code for the 333 Gly-X-Y repeats of the central helical domain. Because of the structural heterogeneity of the N-propeptides, the number and size of the exons coding for this part of the protein differ greatly from one gene to another (Lee *et al.*, 1991a). The only conserved feature of this coding domain is the exon which encodes the sequence that crosses from the N-propeptide to the helical domain, and includes the N-proteinase cleavage and telopeptide cross-linking sites (exon no. 6). It is also worth noting that the great majority of the exons are "in-frame", hence deletions of exons are expected not to disrupt mRNA translation.

iii) Pathologies:

It is now well established that mutations in the fibrillar collagen genes are the cause of many heritable disorders. They include osteogenesis imperfecta (OI), several chondro-dysplasias, and two forms of the Ehlers-Danlos syndrome (EDS). The genetic studies have provided the current theories regarding fibrillogenesis and structure-function relationships. Fibrillar collagen disorders are inherited in an autosomal dominant manner; most phenotypes result from point mutations that produce substitutions for glycine residues within the triple-helical domain of the chains. A relationship between the phenotypic effect of the mutations and their location has emerged. In general, substitutions near the C-terminal end of the triple-helix tend to be more severe, often resulting in lethality, while those occurring toward the N-terminal end give rise to milder phenotypes. There is also some evidence suggesting that the nature of the amino acid substituting for the obligatory glycine residue might also contribute to determining the phenotypic severity. It is likely that additional factors such as local chain structure, molecular interactions and domain functions modify the effect of specific substitutions.

Point mutations that disrupt the Gly-X-Y motif, as well as large and small deletions have been shown to have the following effects:

i) they increase the post-translational modification of the chains N-terminally to the site of the mutation,

ii) they decrease the thermal stability of the abnormal trimers,

iii) they decrease the efficiency of procollagen trimer secretion.

Collagen stability is thought to result largely from hydrogen bonding involving hydroxyproline residues and thus charge interactions. Small and large deletions disrupt the normal charge interactions, even when the triplet motif is maintained. It has been suggested that these disruptions are propagated along the full length of the structure (Kuivaniemi *et al.*, 1991). This effect may account for the increased modification, and the delay in secretion (because of the existence of an 'unfolded' structure N-terminal to the mutation

site). Abnormal molecules are generally secreted less efficiently than normal, but once secreted they become incorporated into the matrix and disrupt the assembly of fibrils. The number of abnormal molecules required to produce abnormal fibrils may be small because of the constraints on molecular dimensions needed to pack molecules into fibrils.

It must be emphasized that mutations rendering transcription inactive, and therefore abolishing peptide production, have been rarely documented. EDS VII, is a condition characterized by extreme joint laxity which often leads to congenital dislocation of the hips and other large joints, by skin with limited bruisability and minimal hyperextensibility, and by normal bone density (Beighton, 1993). The biochemical abnormality is impaired cleavage of the N-terminal propeptides of type I collagen. Wirtz *et al.* (1990) demonstrated that the cleaved N-propeptide of the normal chain is retained in noncovalent association with the mutant pN-type I chain in native mutant collagen molecules, both *in vivo* and *in vitro*. In subtypes A and B, mutations in both the pro-α1(I) and pro-α2(I) chains respectively, affect one of the nucleotides of the 5' splice donor site of intron 6, and cause exon 6 skipping (Weil *et al.*, 1988; Weil *et al.*, 1989; Weil *et al.*, 1990). Exon 6 contains the sequences encoding the N-proteinase cleavage site and a precursor lysine cross-linking site. Thus, loss of exon 6 ultimately prevents N-terminal processing.

Chiodo *et al.* (1992) described an EDS VIIB patient with a partial loss of the pro-α2(I) chain's exon 6, eliminating the N-proteinase cleavage site. In contrast to previously reported cases of EDS-VIIB, the lysine residue of the N-telopeptide, and with it one of the intramolecular cross-linkages, were not lost. More recently, an EDS VIIC patient has been described (Smith *et al.*, 1992). This subtype is believed to be caused by an N-proteinase deficiency, leading to a dermatosparactic-like condition (extreme skin fragility). Ultrastructurally, EDS VIIC is characterized by the "cauliflower" shape of fibrils (Smith *et al.*, 1992; Wertelecki *et al.*, 1992). It should be noted that, aside from EDS VII, no other human condition associated with impaired N-propeptide removal has been described for the other fibrillar collagen types.

The Minor Fibrillar Collagens

Because of their low representation in the extracellular matrix, the fibrillar collagen types V and XI are often referred to as "minor" when compared to the more abundant "major" types I, II and III.

Type V:

Type V collagen is widely distributed in the interstitial matrix of several tissues (Fessler and Fessler, 1987; van der Rest and Garrone, 1991). Its composition varies depending on the tissue. In the skin, bone, and cornea, type V collagen is composed predominantly of [$\alpha1(V)_2\alpha2(V)$] heterotrimers; other forms include the heterotrimer [$\alpha1(V)\alpha2(V)\alpha3(V)$], present in the placenta and synovial membranes, and the homotrimer [$\alpha1(V)_3$], present in hamster lung cultures. The presence of an additional pro-$\alpha1'$(V) chain in chicken tendon, has also been reported (Fessler and Fessler, 1987). It is thought that pro-$\alpha1$(V) and pro-$\alpha1'$(V) chains may be the products of different genes, or that they may arise from alternative splicing of transcripts of the same gene (Eyre and Wu, 1987; van der Rest and Garrone, 1991).

Type V collagen is found at its highest concentration in the corneal stroma where its relative representation to the rest of the collagen fibrils is 15-20%, as compared to less than 5% in other tissues (Tseng *et al.*, 1982; McLaughlin *et al.*, 1989). In the chick cornea, Birk *et al.*, (1988) were the first to show that type V collagen localizes within collagen type I fibrils. They could not detect the type V epitopes without disrupting the normal fibrillar structure with agents such as guanidine-HCl, or by the enzymatic digestion of type I. Thus, they suggested that a heterotypic fiber may be formed between the two, with type V present in the core of the above fibril (Fig. 2) (Linsenmayer *et al.*, 1993).

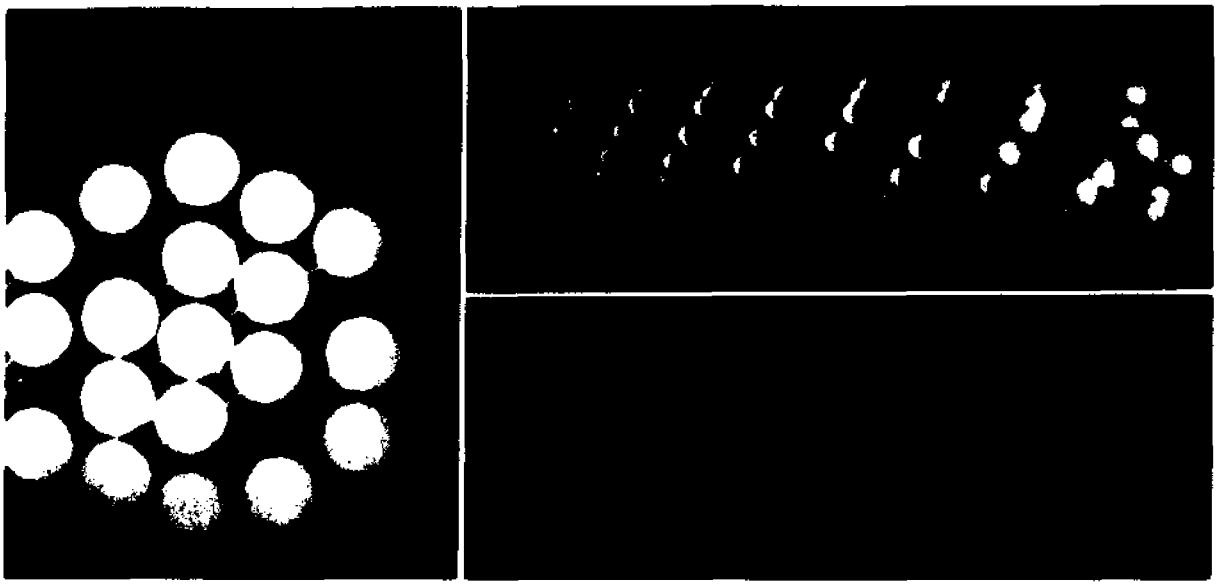


Figure 2.
Schematic showing the proposed type I type V heterotypic fibril in three different views. Type I collagen in white. Type V collagen molecules in red (dark) [reprinted with permission from Linsenmayer *et al.*, 1993].

In vitro co-polymerization experiments showed that increased amounts of type V progressively decrease the diameter of the heterotypic fiber. Since mature corneal stroma contains mostly fibrils of small (25nm) and uniform diameter thought to be required for transparency (Cox *et al.*, 1970), it was logical to suggest that the type V presence plays a major role in the regulation of the heterotypic fibril diameter (Birk *et al.*, 1988).

Type XI:

Type XI collagen is mostly found in hyaline cartilage (Eyre and Wu, 1987). Its co-expression with type II collagen resembles that of types V and I collagen in non-cartilaginous matrices. Like type V, type XI collagen could not be detected in cartilaginous fibers without first shearing or enzymatically digesting the heterotypic fiber containing types II and XI (Mendler *et al.*, 1989). Based on this finding, it was therefore proposed that type XI may play a role in the regulation of the type II containing cartilage fibrils.

Type XI collagen is composed of three distinct polypeptide subunits: α1(XI), α2(XI), and α3(XI) (Morris and Bachinger, 1987). Together they form a single molecule of the [α1(XI)α2(XI)α3(XI)] configuration. While the α1(XI) and α2(XI) chains are products of two different genes, the α3(XI) chain is believed to be a post-translational variant of the major collagen of cartilage α1(II). Interestingly, high level of structural homology is also observed along the entire length of the α1(XI) and α1(V) collagen chains (Greenspan *et al.*, 1991).

Niyibizi and Eyre, (1989) were the first to report that α1(XI) collagen is present in bone tissue in a 1:1:1 ratio with α1(V) and α2(V) collagen. This finding was recently confirmed by other investigators showing co-expression and concerted modulation of the α2(V) and α1(XI) collagens. These two peptides were shown to be present in a variety of tissues, including bovine vascular smooth muscle cells (Brown *et al.*, 1991), vitreous humor (Mayne *et al.*, 1993), and a human rhabdomyosarcoma cell line (Yoshioka and Ramirez, 1990; Kleman *et al.*, 1992). Thus, there is increasing evidence that types V and XI collagens may form heterotypic trimers.

In conclusion, emerging evidence indicates a potentially important role of the minor collagen types in regulating ECM architecture. This in turn implies possible contribution of the abnormal forms of these fibrillar collagen types to connective tissue pathology.

While there has been substantial progress in our understanding of the biological complexity of the ECM, numerous questions regarding the assembly and maintenance of the three-dimensional network remain unanswered. Studies of the collagenopathies have given us insights on how mutations alter the different biosynthetic levels leading to fibrillogenesis. One would therefore expect that we might be close to understanding the relationship between a molecular defect and the altered ultrastructure of the fibril. Unfortunately this is not the case. The primary defect appears to have little predictable relationship with the ultrastructural change. One possibility is that the interactions with other ECM members are probably as important as the individual collagen fibers in contributing to the connective tissue architecture. Although understanding the molecular basis of a disease is important, it represents only the first and simplest step toward uncovering the full spectrum of its pathology. In this respect, characterization of naturally occurring mutations has expanded our understanding of collagen pathophysiology only to a limited extent. An alternative and complementary strategy is to use the genetic approach to analyze logically predicted structural-functional relationships. Such an experimental system exists, and has proven to be a formidable tool in our quest for answers to these questions, both in the case of collagen and of other gene products. The model we are referring to is the transgenic mouse.

The Mouse as an Experimental System

In the last ten years the mouse has revolutionized the field of mammalian genetics. It has proven to be the experimental system of choice for several reasons. Much more complex than the organisms commonly used to study genetic mechanisms and phylogenetically closer to humans, the mouse possesses genes and networks of regulatory interactions very similar to their human equivalents (Capecchi, 1989b). It is therefore likely that mutations in the mouse genes may yield phenotypes analogous to the human ones. Although the existing collection of mouse mutations is remarkably extensive and varied,

serendipity was the major force in exposing the visible variation in morphology and/or behavior. With the advent of recombinant DNA technology, three main methods for the modification of the mouse genome were developed:

i) Use of retroviruses,

ii) Pronuclear micro-injection of DNA, and

iii) Gene targeting by homologous recombination in embryonic stem (ES) cells.

To most of us, integration, expression and inheritance of foreign genetic material within the genome of an animal seems now a fairly straightforward process. Its successful implementation did however take a substantial amount of effort on the part of a small group of investigators. Jaenisch and Mintz, (1974) were the first to document transformation of mouse embryos in their blastocyst stage by the introduction of SV40 DNA. Successful integration of foreign DNA into the embryonic genome, and subsequent germ-line transmission in a Mendelian fashion was accomplished soon thereafter using the Moloney murine leukemia retrovirus (Mo-MuLV) (Jaenisch, 1976). The term "transgenic" was first used by Gordon *et al.* (1980) to describe the successful integration of foreign DNA into the mouse genome by pronuclear microinjection of the fertilized mouse oocytes (Gordon and Ruddle, 1981). The functional expression of integrated transgenes in mice followed soon thereafter (Brinster *et al.*, 1981; Constantini and Lacy, 1981). This approach has proven to be the method of choice for scientists interested in studying gene expression within the context of the whole animal. Cancer biology, immunology, embryology and developmental biology, as well as gene regulation are some of the several areas that have blossomed through the use of transgenic mouse technology. Collagen related research proved no exception.

The first animal model for a collagenopathy came with the use of a retroviral vector. The so-called Mov-13 mouse was derived experimentally by the random integration of Mo-MuLV into the mouse germ line, and resulted in the sudden death of embryos between days

11 and 14 *post coitum (p.c.)* (Schnieke *et al.*, 1983). The provirus integrated in the first intron of *col1a1* and did apparently render the gene transcriptionally inactive. Heterozygous Mov-13 mice though produce 50% of pro-α1(I) and of mature type I collagen, and display a phenotype that recapitulates the mild dominant form (type I) of human OI (Bonadio *et al.*, 1990).

The second transgene was generated by expressing a *col1a1* mutation (G859S) on the normal genetic background of transgenic mice (Stacey *et al.*, 1988). The transgenic mice displayed the clinical and biochemical features of human perinatal lethal OI type II. The result confirmed previous predictions based on the human studies which showed the "antimorphic" effect of collagen mutations (Muller, 1932; Herskowitz, 1987).

More recent examples of transgenic mice expressing mutant collagens on normal background include those of α1(II) collagen with a phenotype similar to severe human chondrodysplasia (Garofalo *et al.*, 1991; Vandenberg *et al.*, 1991; Metsäranta *et al.*, 1992); of a partially deleted α1(IX) chain with osteoarthrosis and mild chondrodysplasia (Nakata *et al.*, 1993); and of a truncated α1(X) with spondylometaphyseal dysplasia (Jacenko *et al.*, 1993).

Despite the wealth of new information, random integration of the exogenous DNA in the mouse genome has two major disadvantages: first, the integration site has profound effects on the expression of the introduced gene (position effect, Nandi *et al.*, 1988); and second, the introduced gene can disrupt the function of an endogenous gene (insertional mutagenesis, Gossler *et al.*, 1989). Although the later has been useful in identifying important genes, the former is more of a problem particularly in cases in which gene dosage plays an important role in determining clinical severity. To overcome these and related problems, several investigators have spent considerable effort perfecting the gene targeting technology. As a result, it is now possible to potentially generate transgenic mice that have any desired genotype.

Evans and Kaufman, (1981) and Martin, (1981) were the first to show the ability of mouse ES cells derived from the inner cell mass (ICM) of mouse blastocysts, to adapt and grow in culture. Under appropriate conditions they could be maintained in continuous culture retaining their pluripotentiality. When introduced back into the blastocyst, they were capable of continuing their developmental program after re-implantation into the uterus of a pseudopregnant female. Shortly thereafter, Thomas and Capecchi, (1987) demonstrated targeting of an exogenous DNA segment to a specific chromosomal site. This was achieved by taking advantage of the somatic cells' intrinsic ability (enzymatic machinery) to mediate recombination between homologous sequences. The combination of both ES cell and homologous recombination technology provided the opportunity to develop the currently used approach of gene targeting (Fig. 3).

Generating specific lines of mutant mice requires the use of standard recombinant DNA technology to introduce the desired mutation into a cloned DNA fragment. The mutation is then transferred via homologous recombination to the genome of a population of ES cells. In order to identify them, the cells are usually from a pigmented mouse strain (i.e. 129/Sv-agouti color). The targeted locus contains one or more genes which confer to the cells novel properties that can be selected for, thus the correct recombinatorial event can be identified. After selection and screening for the correctly targeted sequences, ES cells carrying the mutation are microinjected into blastocysts from an unpigmented mouse strain (i.e. BALB/C-white color). Injected blastocysts are then re-implantated into the uterus of a surrogate mother. The resulting chimeras are composed of cells from both ES and host-blastocyst genotypes, and are readily identifiable by their variegated coats. Since the pigment is a dominant trait, the chimeras are then bred with unpigmented mice and germ-line transmission is therefore monitored by the progeny color. The heterozygous, pigmented mice are sequentially interbred to generate animals homozygous for the mutation.

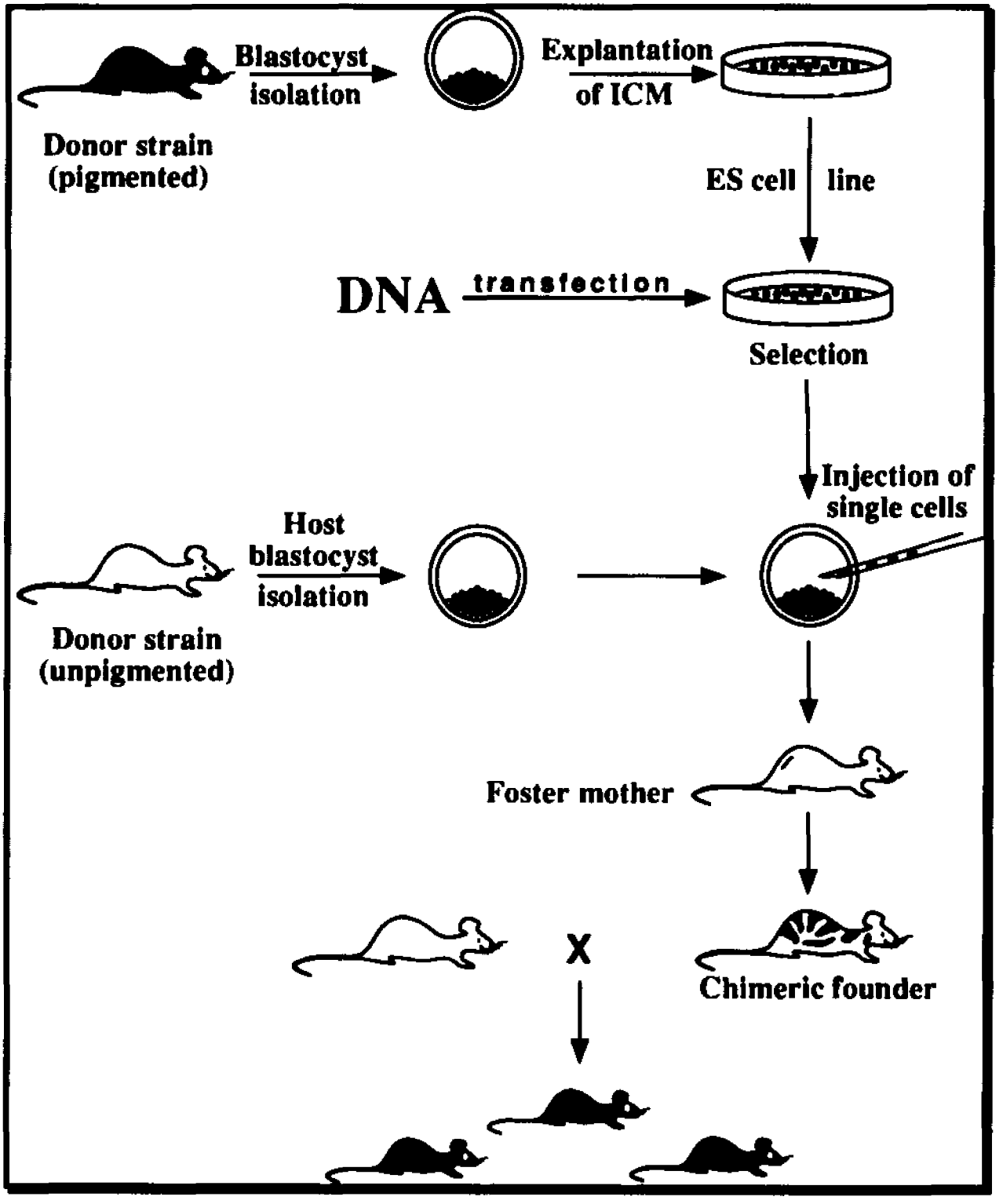


Figure 3.

Schematic representation of the gene targeting strategy (ICM: Inner Cell Mass).

It must be emphasized that gene targeting can be utilized not only to generate non-functional protein molecules (null alleles), but also to modify any structure of the gene that may affect its function. Examples include the gene's transcriptional pattern, its mRNA or protein maturation pattern, or the ability of the gene's protein product to interact with other gene products (Capecchi, 1989a). There are currently three methods for introducing the desired mutations into the ES cell genome. The one most widely used was developed by Mansour *et al.* (1988), and involves a replacement type targeting vector and positive-negative selection.

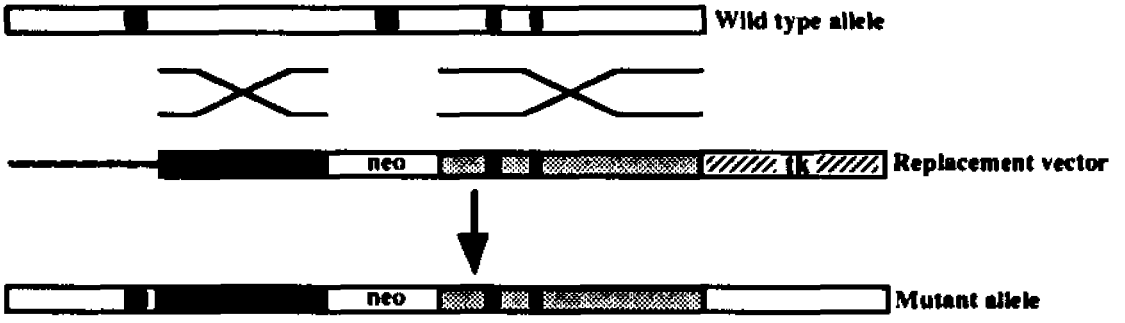


Figure 4.
Schematic representation of gene targeting using a typical replacement vector. Solid black boxes indicate exons.

An ideal targeting vector contains a total of ~5-7kb of identical sequences flanking the gene to be targeted (Thomas *et al.*, 1992). This is the segment that undergoes homologous recombination and is therefore called the homology region (Fig. 4). In the non-homologous portion, the vector contains the gene whose expression can be selected for in a positive manner; usually a gene that confers resistance to an antibiotic (e.g. the widely used neomycin conferring resistance to G418). A negatively selectable gene, usually the herpes simplex virus (HSV) thymidine kinase (tk), is incorporated at one end of the homologous arm of the targeting vector. The construct linearized outside the region of homology is introduced into the ES cells (most commonly by electroporation). Since homologous recombination has to be mediated by the homology region of the targeting vector, such an event will result in retention of the positive selectable gene and the loss of

the negative selectable gene. In contrast, non-homologous recombination is mediated by the free ends (Roth and Wilson, 1988), and cells in which such an event has occurred will have a functional HSV-tk. These cells can be selected against using the nucleotide analogue gancyclovir, or the less toxic FIAU [1-(2-deoxy-2-fluoro-β-D-arabinofuranosil)-5-iodouracil], since expression of the HSV-tk gene results in the conversion of this analogue into a compound lethal to cells (Borelli *et al.*, 1988). Cells which survive both G418 and FIAU selection are therefore expected to have a substantial enrichment for homologous events. Recombinatorial events are confirmed using the polymorphisms generated in the ES genome, and are detected by Southern hybridizations to probes from both outside and inside the regions of homology, and in conjunction with DNA amplification using the PCR (polymerase chain reaction) technique.

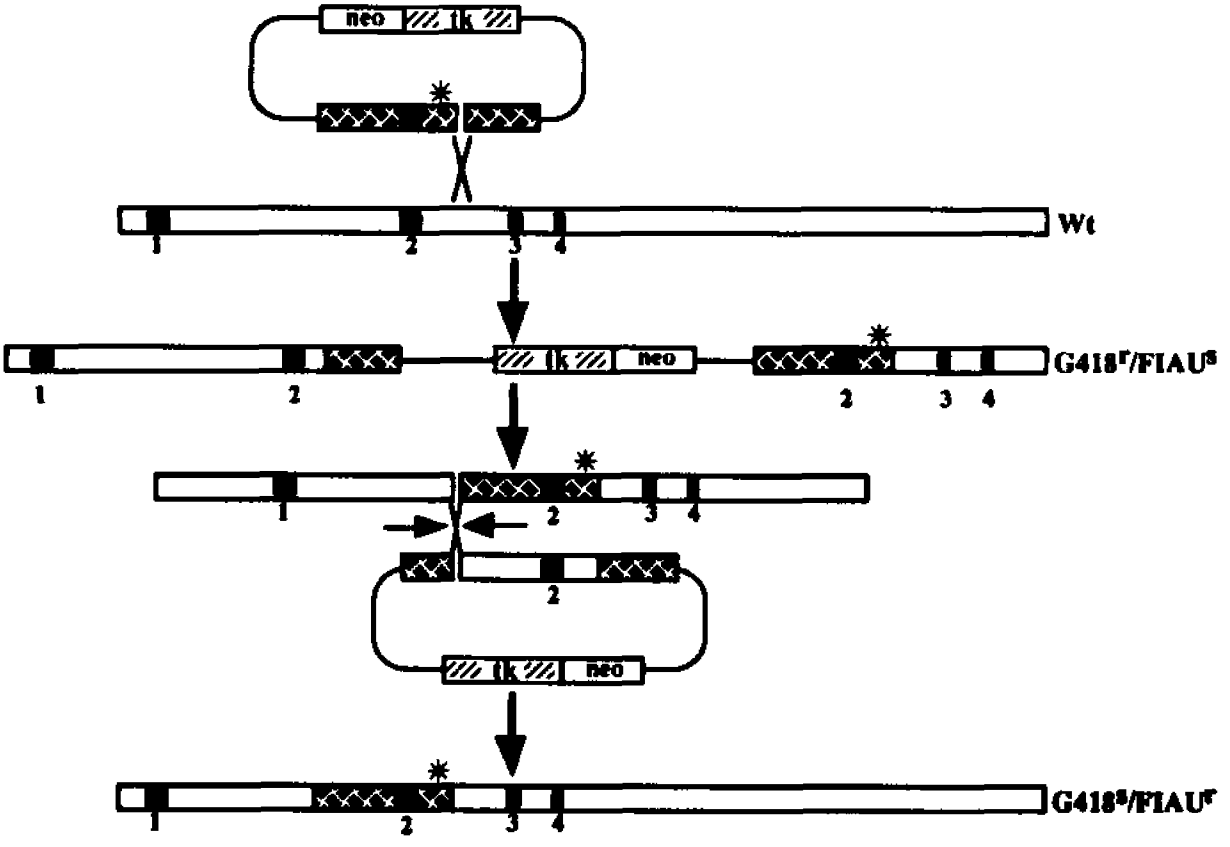


Figure 5.
Schematic representation of the gene targeting strategy using the *hit-and-run* approach. An arbitrary number is assigned to exons in order to facilitate orientation.

The second method, also known as the "hit-and-run" approach, was first described by Hasty *et al.* (1991). It is a two-step recombinatorial procedure that makes use of a sequence insertion vector to generate a subtle mutation (Fig. 5). In this case, the two selection genes are placed within the plasmid backbone and outside the region of sequence homology. The vector DNA, linearized within the region of sequence homology, is introduced into ES cells. Colonies in which the vector has integrated into the ES genome by single reciprocal recombination are resistant to G418 and sensitive to FIAU. Acquisition of these traits is the result of the inclusion of both selection genes within the duplicated arbitrary sequence of interest. Following DNA analysis, positive ES clones are therefore scored for revertants. Because of the sequence duplication, and as a result of reciprocal intrachromosomal recombination (reversion), a fraction of the ES clones are expected to become resistant to FIAU and sensitive to G418. These colonies should therefore contain the mutated sequence, a conclusion which is again confirmed by the usual means of DNA analysis.

The third method, is also a two-step recombinatorial procedure used to introduce subtle mutations, and may be described as a "double-hit" (Wu and Jaenisch, unpublished).

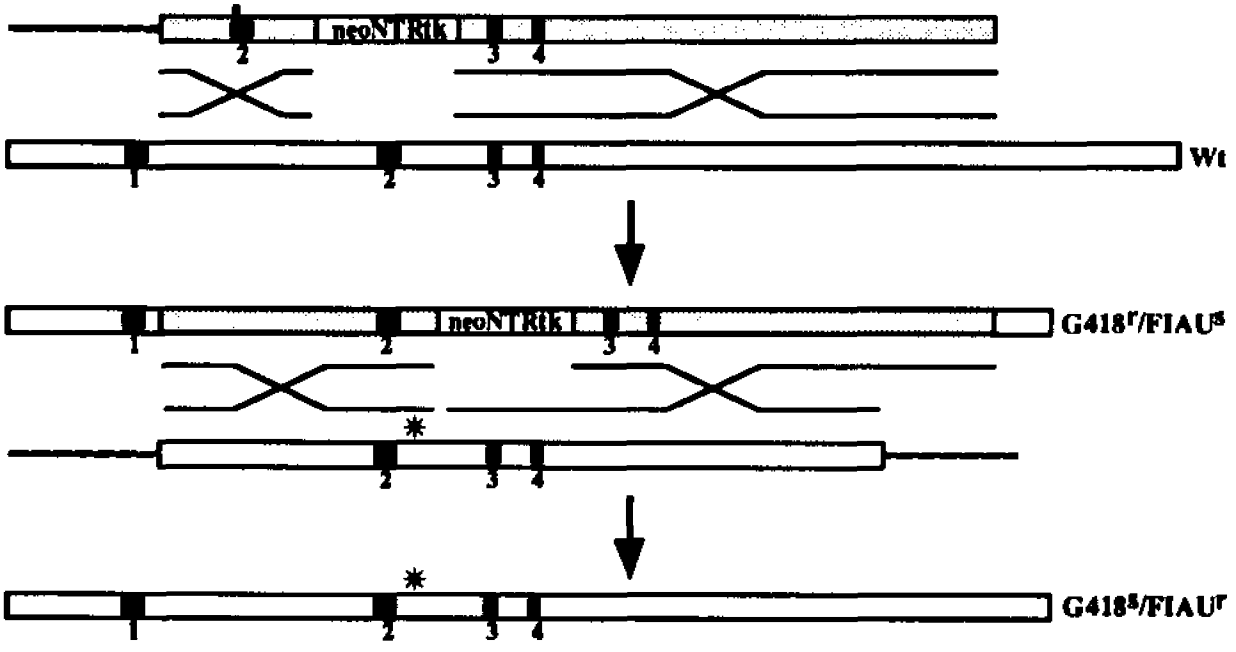


Figure 6.
Schematic representation of the gene targeting strategy using the *double-hit* approach.

Both steps require a replacement type vector (Fig.6). The first one is used to insert both the positive and negative selection markers (neo and HSV-tk), which are placed next to each other in the targeting region. After positive selection in G418, clones with the correctly integrated sequences are identified by the usual means, and are subjected to a second electroporation with another vector containing the desired mutation. ES cells are subjected to FIAU selection. Surviving clusters should only contain the desired mutation in their genome.

Several investigators have now reported successful applications of the above mentioned technology (Travis, 1992). The animals generated are proving to be invaluable tools for understanding the bases of diseases; they are also providing systems for exploration of new therapeutic protocols, including genetic intervention.

Specific Aims and Significance

It was the objective of this thesis project to explore some of the structural-functional properties of type V collagen by using the powerful technology of gene targeting in ES cells. Toward this end, the following experiments were performed:

1. Isolation of cDNA clones and characterization of the primary structure of the murine pro-α2(V) gene.
2. Establishment of a detailed temporal and spatial pattern of its expression during embryogenesis.
3. Use of gene targeting technology to generate mice harboring a single exon deletion (mutation pNA25) in the pro-α2(V) gene (*col5a2*).
4. Analysis of the phenotypic, histologic and ultrastructural consequences of the pNA25 mutation in transgenic mice.

The long-term goal of the project was to assign a biological function to type V collagen and, indirectly, to extend the collagenopathy taxonomy by using the transgenic mouse as a model.

Materials and Methods

Isolation and Characterization of Murine Pro-α2(V) cDNA clones:

Screening of an NIH 3T3 library in λZAP-II (Stratagene, La Jolla CA) was performed using the Eco RV (600-bp) fragment of clone OK25 as a hybridization probe. This fragment codes for most of the C-propeptide of the human pro-α2(V) collagen cDNA (Weil *et al.*, 1987).

Hybridization for the library screening with the DNA probe at ~2-3x10^8cpm/μg was in 50% formamide, 3x SSC, 5x Denhardt's, 100μg/ml salmon sperm DNA.

20x SSC	175.3g of NaCl and 88.2g of sodium citrate were dissolved in 800ml of H_2O. The pH was adjusted to 7.0, and the volume to 1 lt.
10% SDS	100g of sodium dodecyl sulphate were dissolved in 900ml of H_2O, at 68°C to help dissolution. The volume was adjusted to 1 lt.
50x Denhardt's	5g of Ficoll, 5g of polyvinylpyrrolidone and 5g of bovine serum albumin (fraction V) were dissolved in 500ml of H_2O, filtered through a Nalgene 0.45μm filter, and stored in aliquots at -20°C.

Filters were washed in 2x SSC, 0.1% SDS at room temperature (r.t.); 3x for 10' each, in 1x SSC, 0.1% SDS at 65°C; 1x for 10', in 0.1x SSC, 0.1% SDS at 65°C; 1x for 10'. Conditions for the isolation and purification of positive clones were by standard protocols (Sambrook *et al.*, 1989). Sequencing was performed in both orientations mainly on Pst I and Pst I-Xho I subclones (see Results Fig. 8), with a limited use of oligonucleotides, using the method by (Sanger *et al.*, 1977) , and reagents from the Sequenase™kit (USB).

Isolation and Characterization of Murine Pro-α1(II) cDNA clones:

Type II cDNA clones were obtained from a day 12 *p.c.* mouse limb library in λZAP-II (generous gift of Dr. Babu J. Padanilam, U. of Iowa), using oligonucleotide 3'-GGGCTTGGGACTTTGTTGTGTTAGG-5' from the mouse pro-α1(II) 3'-untranslated region as a probe (Cheah *et al.*, 1991b).

Hybridization with the oligonucleotide probe at ~2-3x10^8cpm/μg was in 6x SSC, 1x Denhardt's, 0.05% sodium pyrophosphate, 0.1% SDS, 100μg/ml salmon sperm DNA. The washes in 6x SSC; 2x for 10', 0.05% sodium pyrophosphate at r.t.; 1x for 10', in 6x SSC, 0.05% sodium pyrophosphate at Tm-10°C; 1x for 10'. Conditions for the isolation and purification of positive clones, and sequencing were essentially the same as described above for the type V.

Temporal and Spatial specific expression:

RNA Isolation:

Total RNA was isolated from BALB/C mouse embryos obtained from pregnant females at appropriate gestation times using the guanidinium isothiocyanate/CsCl method (Sambrook *et al.*, 1989). Briefly, embryos were pooled together and homogenized in 5 volumes of 4M guanidinium isothiocyanate, 5mM sodium citrate (pH 7.0), 0.1M β-mercaptoethanol, 0.5% sarcosyl.

After adding 1g of CsCl to 2.5ml of homogenate, this was layered onto a 1.2ml cushion of 5.7M CsCl in 0.1M EDTA (disodium ethylene diamine tetraacetate. $2H_2O$, pH 7.5) in an ultracentrifuge tube. This was followed by centrifugation at 35,000 rpm for 12 hours at 20°C.

The supernatant was discarded and the RNA pellet was disolved in 10mM Tris-HCl (pH 7.4), 5mM EDTA, 1% SDS. After extraction with a 4:1 mixture of chloroform and 1-butanol, the aqueous phase was transferred to a fresh tube. The organic phase was then re-extracted with an equal volume of 10mM Tris-HCl (pH 7.4), 5mM EDTA, 1% SDS. The two aqueous phases were combined, and 0.1 volumes of 3M sodium acetate (pH 5.2) and 2.2 volumes of ethanol were added. It was stored at -20°C overnight.

The RNA was recovered by centrifugation at 35,000 rpm for 12 hours at 20°C. The pellet was dissolved in 1ml of diethylpyrocarbonate (DEPC) treated, sterile distilled water (SDW) and re-precipitated with ethanol. The RNA was stored in 70% ethanol at -70°C.

Temporal pattern of expression:

Stage-specific RNA samples were reverse transcribed using a commercial kit (Stratagene, La Jolla CA). Briefly, 3µl of random primers (100ng/µl) were added to 10µg of total RNA (dissolved in 32µl of DEPC-treated sterile distilled water). The mixture was heated at 65°C for 5', and was cooled slowly (~10' at r.t.) to allow annealing of the random primers to the RNA. The following reagents were then added in order: 5µl of 10x first strand buffer, 5µl of 0.1M dithiothreitol (DTT), 1µl of RNase Block, 2µl of 25mM dNTP's (deoxyribonucleotide triphosphates), and 1µl of Mo-MuLV reverse transcriptase (20U/µl). All of the components were mixed gently, and incubated at 37°C for 1hour (hr). The completed reaction was placed on ice. Products were amplified by the polymerase chain reaction (PCR) as described by Perkin Elmer-Cetus (Norwalk CT) with the GeneAmp® DNA amplification reagent kit, using the following collagen-specific oligonucleotide primers:

Forward direction:

5'-TATTGGATTGATCCTAACCAGGG-3'[pro-α2(V)],

5'-GAACGGTCCACGATTGCATG-3'[pro-α2(I)],

5'-CACACTGGTAAGTGGGGCAAGACCG-3'[pro-α1(II)];

Reverse direction:

5'-GTGGTCAGGCACTTCAGATC-3' [pro-α2(V)],

5'-GGCATGTTGCTAGGCACGAAG-3' [pro-α2(I)],

5'-GGATTGTGTTGTTTCAGGGTTCGGG-3' [pro-α1(II)].

Sequences of oligonucleotide primers specific for *col1a2* (the pro-α2(I) gene) and *col2a1* (the pro-α1(II) gene) were derived from Rossi and de Crombrugghe (1987) and Cheah *et al.* (1991b), respectively. The location of the *col5a2* forward primer is between nucleotides 3,925 and 3,947 of the coding sequence (Fig. 9), while the sequence complementary to the reverse primer is in bold in Fig. 10. In all cases, primers correspond

to sequences within non-contiguous exons in order to monitor for the possible presence of contaminating genomic DNA.

RNA from the tails of adult mice was obtained using identical procedures. Reverse transcribed PCR (RT-PCR) was also performed in an identical manner. PCR products were subcloned into the Sma I site of Bluescript SK II after they were filled in using the Klenow (large fragment of the T4 DNA polymerase) and dNTPs, and then subjected to kinase treatment in the presence of ATP to increase the efficiency of the ligation. Total RNA dot blots were performed according to standard protocols (Sambrook *et al.*, 1989).

Primer extensions:

Primer extension assays were performed essentially as described in (Sambrook *et al.*, 1989) with some modifications. Briefly, the following reaction mixture was prepared: in 20 µl total volume 10 µg of total RNA, 10^5 cpm ^{32}P labeled oligonucleotide primer, 4µl of 5x hybridization buffer {200mM PIPES (piperazine-N, N' bis[2-ethane-sulfonic acid]) pH 6.4, 5mM EDTA pH 8.0, 2M NaCl}, 10 µl of formamide, were added together.

The reaction(s) was covered with light mineral oil (Sigma) and incubated at 55°C overnight. The following day it was phenol/chloroform extracted and ethanol precipitated. The primer:RNA hybrids were dissolved in 20 µl of reverse transcriptase buffer (50mM Tris.Cl pH 7.6, 60mM KCl, 10mM $MgCl_2$, 1mM of each dNTP, 1mM DTT, 1U/µl placental RNAase inhibitor). 50U of Mo-muLV reverse transcriptase (Stratagene, La Jolla CA) were added, and the reaction mix was incubated for 2 hr at 37°C. After phenol/chloroform extraction and ethanol precipitation, the product was resuspended in 6µl of TE (10mM Tris.Cl pH 8.0, 1mM EDTA pH 8.0). 4µl of sequencing loading buffer were added. The mixture was denatured at 95°C for 3', chilled on ice, and run on a 5% polyacrylamide sequencing gel to determine its exact size.

Spatial pattern of expression:

In situ hybridizations were performed following the protocol described by Angerer *et al.* (1987) with some modifications. All of this work was done in Dr. Michael Solursh's laboratory at the University of Iowa, and under the supervision of Dr. Hiroaki Suzuki.

Embryos were staged and fixed by immersion in Bouin's fixative{140ml picric acid, 50ml 37-40% formaldehyde, 10ml glacial acetic acid} for 2 to 7 days at 4°C. After dehydration in an ethanol series{50%, 70%, 80%, 90%, 95%, 100% ethanol each for 20', Histosol or Xylene 2x for 20', all at r.t.}, they were embedded in parafin{In a 1:1 Xylene/Paraplast Plus mix 3x for 20' in a 60°C oven}. 8 µm, serial, sagital sections of day 12 *p.c.* whole embryos, and separate head, body and forelimb sagital sections of day 16.5 *p.c.* embryos were placed on slides. Slides were previously rinsed briefly in DEPC-H_2O, 0.1 HCl 2x, each for 30', 70% ethanol 2x, each for 30', 95% ethanol briefly, air-dried, submerged in 2% (v/v)3-aminopropyltriethoxysilane (Aldrich) in acetone for 30", rinsed briefly in acetone and in DEPC-H_2O, and airdried (L. M. Angerer personal communication).

The sections were left for two days at 40°C to dry, and then baked at 60°C overnight to ensure their adherence on the slides (Sandberg and Vuorio, 1987). They were then rehydrated{ in Xylenes 2x, each for 15', 100% ethanol 2x, each for 5', 95%, 90%, 70%, 50%, 30%, dH_2O, $T_{100}E_{50}$ pH 8.0, each for 5' at r.t..

Proteins were digested by submersion in $T_{100}E_{50}$ pH 8.0 containing Proteinase-K (1 mg/ml) for 30' at 37° C, washed in the $T_{100}E_{50}$ buffer at 37° C for 30', and again for another 5' at r.t.. with a final wash in 0.1M glycine (mw=75.07, therefore 2.25g/300ml), 0.2M Tris-HCl pH 7.4 for 10' at r.t.. Subsequently, sections were acetylated {in 0.1M Triethanolamine-HCl pH 8.0 (19.97ml of triethanolamine/1480ml of dH_2O and adjusted pH to 8.0 with conc. HCl) for 5' at r.t., in 0.25% (v/v) acetic anhydride for 10' at r.t. It should be noted that 0.75ml of acetic anhydride were placed first into the jar, followed by the 300ml of the 0.1M Triethanolamine-HCl, and the mixture was immediately and

vigorously agitated. Finally, they were rinsed with 2xSSC for 5' at r.t., dehydrated in {30%, 50%, 70%, 90%, 95%} ethanol each for 5', and air-dried.

Radiolabeled ^{35}S antisense RNA probes were prepared using the T3 and T7 promoters of Bluescript II-SK vector (Stratagene, La Jolla CA) containing the type V and type II specific sequences. For the *col5a2* the sequence was the 5'-untranslated - Pst I, 702 bp cDNA fragment (Fig. 8), proven to be highly conserved between species (Fig. 17), and therefore specific for this transcript. For the *col2a1* gene, the Sty I - 3'-untranslated 445 bp fragment of the cDNA (sequence not shown). DNA dependent RNA polymerases were purchased from Promega. The transcripts were treated with DNase I, phenol-chloroform extracted and ethanol precipitated. They were then hydrolysed to an average size of ~150 bp (Angerer *et al.*, 1987) and were used at the 3.6×10^6 cpm/50 µl specific concentration in the hybridization buffer {0.3M NaCl, 20mM sodium acetate, 1 mM EDTA, 100 mM DTT, 50% formamide, 10% dextran sulphate, 1x Denhardt's, 250 µg/ml yeast tRNA}. The hybridization solution containing the antisense probe was then applied to the sections on the slides, siliconized coverslips were placed on top, and hybridized overnight at 50°C in a formamide saturated environment.

Slides were then washed 3x for a total of 2 hrs in 5xSSC, 10 mM DTT at 50°C, and 1x for 30' in 50% formamide, 2x SSC, 10mM DTT at 60° C. Subsequently they were rinsed in STE at 37°C for 10', and treated with 20 mg/ml RNase A, 1U/ml RNase T1 in STE at 37°C for 30'. Final washes were in STE at 37°C for 30', 2x SSC at 37°C for 15', 0.1x SSC at 50°C for 15', and 0.1x SSC at r.t. for 5', with all washes containing 10mM β-mercaptoethanol. Sections were again dehydrated in an ethanol series {30%, 50%, 70%, 90%, 95%}, all inclusive of 0.3M ammonium acetate, each for 5'.

For autoradiography, sections were dipped in Ilford K.5D emulsion and exposed for 7 days at 4°C; developed in Kodak D-19 for 3' at 10°C, fixed in Kodak fixer for >20' at 10°C, and lightly stained with Hematoxylin for 10-15' on ice. They were again

dehydrated through an ethanol series (xylenes last), air-dried, and siliconized coverslips were permanently fixed with Permount®.

Sections were viewed in an Olympus microscope under dark-field illumination, and pictures were taken with a Nikon camera and a 400 ISO black & white film by Kodak.

Isolation and Characterization of Murine Clones for the 5' of the Pro-α2(V) gene:

A genomic library was prepared from a size-fractionated, partially digested with Sau 3A, genomic DNA. The DNA was obtained from the liver of an 8-week old, 129/Sv genotyped, female mouse (a generous gift of Sharon Boast and Dr. Steven Goff of Columbia University) using standard protocols (Sambrook *et al.*, 1989). Fragment sizes of 9-23 kb were cloned into the Bam HI digested λ-DASH phage vector (Stratagene, La Jolla CA). Packaging was accomplished using the Gigapack-Gold packaging extract also from Stratagene. The primary titre was in the region of $1x10^6$ pfu's (plaque forming units). The library was amplified once before any subsequent screening. Conditions for screening were identical to the ones described earlier.

Culture of ES Cells and Generation of Germ-line Chimeras:

All of this work was performed in Dr. Rudolf Jaenisch's laboratory at the Whitehead Institute of Biomedical Research, M.I.T., and under the supervision of Dr. Xin Liu. The J1 ES cell line was maintained essentially as described before (Li *et al.*, 1992). Some modifications were employed after the electroporation.

Both embryos and ES cells were cultured in ES cell culture medium which was prepared from Hepes-buffered (20 mM, pH 7.3) Dulbecco's modified Eagles medium (D-MEM, high glucose, from Gibco-BRL) supplemented with 15% heat inactivated Fetal Calf Serum (FCS) (HyClone), 0.1 mM non-essential amino acids (100x stock from Gibco-BRL), 0.1 mM β-mercaptoethanol (4μl / 500ml medium), and antibiotics (penicillin,

streptomycin). Plates for ES cell culture were coated with 0.2% gelatin (Sigma) in H_2O. The gelatin solution was autoclaved and stored at 4°C.

The trypsin/EDTA solution used, consisted of 0.25% trypsin (from a 2.5% mycoplasma screened trypsin stock, Gibco-BRL) and 1 mM EDTA in Hepes saline, that was filtered and stored at -20°C. To avoid contamination, the trypsin/EDTA solution prepared this way was used only for cell culture involving ES cells, embryonic fibroblasts (EF), and embryos.

Hepes-buffered saline consisted of 121mM NaCl, 5.4 mM KCl, 0.44 mM KH_2PO_4, 0.3 mM Na_2HP0_4, 5.56 mM glucose, 20 mM Hepes (pH 7.3) and phenol red. The osmolarity was adjusted to 290 mmol/Kg. Phosphate-buffered saline (PBS) was sometimes used as a replacement (Robertson, 1987).

Freezing solution (1x) consisted of 10% DMSO, 10% FCS in Hepes-buffered (pH 7.3) D-MEM. The medium used for blastocyst-injection was prepared from Hepes-buffered (pH 7.3) D-MEM supplemented with 10% FCS. Sodium bicarbonate was not added since it would turn the medium basic upon air exposure. The osmolarity of the medium was adjusted to 290 mmol/Kg with 5M NaCl or H_20, since osmolarity which deviates too far from the standard could affect the viability of the ES cells being injected.

Embryonic fibroblast (EF) cells were cultured in Hepes-buffered (pH 7.3) D-MEM supplemented with 10% fetal calf serum and antibiotics. G418 resistant EF cells were isolated from mouse embryos at day 14 of gestation. The embryos were generated from a homozygous x homozygous cross of the viable β2-microglobulin null mutant mouse (Zijlstra *et al.*, 1989). They were isolated and washed once in Hepes saline. Individual embryos were dissected removing the head and soft tissues (e.g. liver, heart and other viscera). The carcasses were washed twice in Hepes saline. They were then minced into fine pieces in a small volume of trypsin/EDTA solution (just covering all the embryos). More trypsin/EDTA solution (2 ml trypsin solution/10 embryos) was added and mixed well with the embryonic tissues. The mixture was incubated at 37°C for 30'. After the addition

of 10 ml D-MEM/10% FCS, the digested tissues were transferred into a 50 ml tube. Tissues were dissociated by vigorous pipetting. Large pieces of tissue debris were allowed to settle down and the supernatant was transferred into a clean tube. More trypsin/EDTA solution was added. After the addition of another 10 ml of medium, tissues were further dissociated. This step was repeated three or more times. All the supernatants were combined, and the cell suspensions plated out into 175 cm^2 flasks (about 1 embryo/flask). Cells were trypsinized and frozen in two tubes/embryo after reaching confluency. For the preparation of mitotically inactive feeder cells, a tube of primary EF cells was passaged up to 4 times, and the cells were subjected to γ-irradiation (2000-3000 rads). Aliquots of mitotically inactivated EF cells were frozen at -70°C for several months or in liquid nitrogen for years.

ES cell culture:

ES cells were routinely cultured in ES medium in the absence of Leukemia Inhibitory Factor (LIF-differentiation inhibitor) on a feeder layer of mitotically inactivated EF cells. Normally, $1.5x10^6$ ES cells were seeded in a 25 cm^2 tissue culture flask with the medium being changed daily. Cells were split 2-3 days after seeding, usually when the flask was about 80% confluent. Good (but not over-) trypsinization of cells to achieve a single cell suspension, and seeding of cells at appropriate densities were critical. Cells were washed twice with Hepes saline and trypsinized with 0.5 ml trypsin/EDTA at 37°C for 4'. They were then detached off the plate and mixed thoroughly by agitation. 5 ml of medium were added, and cells were pipetted several times against the flask walls to achieve complete dissociation. ES cells were frozen at a density of about $4-5x10^6$ cells/ml at 0.5 ml/cryotube aliquots. Frozen cells were stored at -70°C overnight and transferred into liquid nitrogen the next day. Frozen cells were quickly thawed in a 37°C water bath, pelleted in 5 ml of ES medium to remove DMSO, and plated in 25 cm^2 flask with inactivated EF feeder cells.

Electroporation:

Rapidly growing ES cells were trypsinized, counted, washed and resuspended at $2x10^7$ cells/0.8ml in the electroporation buffer containing 20 mM Hepes (pH 7.0), 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 6 mM glucose, and 0.1 mM β-mercaptoethanol (Thomas and Capecchi, 1987). Linearized DNA was added to the cell suspension at 25 μg/ml. Electroporation was carried out with a Bio-Rad Gene Pulser. Cells were exposed to a single pulse at either 800V, 3μF or 400V, 25μF. The cuvette was left for 10' at room temperature, and the cells were plated at $2\text{-}4x10^6$ / 100mm plate. Plates were gelatinized and covered with γ-irradiated EF ($G418^r$) cells. Selection started 12-18 hours later, in LIF supplemented (500 U/ml) ES medium containing G418(350 μg/ml dry powder)-1 control plate, G418 (Gibco-BRL) and FIAU (Eli-Lily) -9 plates.

Picking, Freezing and Thawing colonies:

Colonies were picked 7-12 days after selection. The medium was replaced with Hepes buffered saline before the picking. Each colony was dissociated in a 20μl drop of trypsin/EDTA in a V-bottomed 96-well plate, and dissociated while pipetting up-and-down 20x. 75μl/colony of ES medium/LIF/G418/FIAU were added, and mixed while pipetting up-and-down 5x. 30μl of the above mixture containing a portion of the ES cell cluster destined to be used as stock, were transferred to a single well of a flat-bottomed 96-well plate with EF cells and LIF/G418/FIAU. The remaining 60μl were transferred to a single well of a 24-well plate, continued selection with LIF/G418/FIAU supplemented medium, and used 7-8 days later (when confluent) for DNA extractions. When the cells in the 96-well plate reached over 80% confluency, usually after 2-3 days, they were washed twice with Hepes buffered saline and trypsinized (30μl of trypsin/EDTA per well, pipetting up-and-down 5x). Freezing solution was added (150μl of ES medium/LIF/G418/FIAU/well), and the plates were frozen and stored at -70°C. After positive clones were identified, the plates were thawed out with each positive clone being washed once with 5 ml ES medium

and transferred into one well of a 24-well plate with feeder cells/LIF/G418/FIAU. Cells were passaged sequentially, upon confluency, to a 6-well plate, and to a 75cm^2 flask. Cells were trypsinized and frozen to 9x 0.5ml aliquots.

DNA extraction and characterization:

When the 24-well plates reached confluency, the method of Laird *et al.* (1991) was used to extract DNA. Briefly, the cells were washed once with Hepes, and 0.5ml of lysis buffer/well were added (Lysis Buffer: 100mM Tris-HCl pH 8.5, 5mM EDTA, 0.2% SDS, 200mM NaCl, 100µg/ml Proteinase K). Incubated overnight at 37°C. The plates containing the lysates were then transferred to a swirling table and agitated for ~30' at r.t.. One volume of isopropanol was then added to each well and agitation continued until precipitation was complete (~3hrs.). DNA from the individual ES clones was picked with a yellow tip and transferred to labeled Eppendorf tubes. Each DNA sample was then washed once with 70% ethanol, air-dried, resuspended in 50µl of TE (10mM Tris.Cl, 1mM EDTA) pH 8.0, and incubated overnight at 55°C to help dissolve it. DNAs were stored at 4°C. Correctly targeted clones were identified by Southern hybridizations. In a total reaction volume of 50µl, 10µg of DNA (~10µl) were digested in the presence of BSA (1µg/ml) with the appropriate restriction endonuclease (20U/µg of DNA). Reactions were incubated overnight at 37°C in a H_2O saturated incubator. 0.7% ethidium bromide gels were run at 150 V/cm for ~3hrs, and then transferred according to standard protocols (Sambrook *et al.*, 1989) to a nylon membrane (Hybond N$^+$, Amersham U.K.). Pre-hybridization was as short as 5'. Hybridization was standardly performed overnight. Both pre-hybridization and hybridization were in 50% formamide, 5x SSC, 1x Denhardt's, 0.01M Na_2HPO_4 pH 6.7, 1% SDS, 5% dextran sulphate, 100µg/ml salmon sperm DNA. The washes were in 2x SSC, 0.1% SDS at r.t.; 3x for 10', in 1x SSC, 0.1% SDS at 65°C; 1x for 10' (all in 500 ml final volume).

Production of Germ-line Chimeras:

Blastocysts were isolated from pregnant BALB/C or C57BL/6 mice of day 3.5 *p.c.* by flushing of the uterine horns with the injection medium. Blastocysts from C57BL/6 mice were observed to have large and fully expanded blastocoels which can be used for injection right away. However, blastocysts from BALB/C mice were often not well expanded, and about half the embryos were at late morula stage. Usually, morulae and blastocysts with small cavities were cultured in ES medium for 1-3 hrs. Expanded blastocysts were then chosen for injection. Positive ES cells (where homologous recombination for the chosen genetic locus had been identified) were harvested by trypsinization 1 day after seeding (one tube of the nine, in a $25cm^2$ flask), washed once in the injection medium, and resuspended in the same medium at 1-2 x 10^6 cells/ml. Injection was carried out in the injection medium or ES medium covered with dimethylpolysiloxane (Sigma). About 15-20 ES cells were injected into each blastocyst. After injection, embryos were cultured in the ES medium for 1 hr, and 6-10 embryos (usually including 2 non-injected blastocysts as carriers) were re-implanted into one uterine horn of each pseudopregnant foster mother (F1 of DBA x C57BL/6 or CBA x C57BL/6 were preferred and mice weighing 26-32 grams were found to perform best). Pregnant foster mothers were individually caged 14 days after injection without disturbance for a week (this measure was found to effectively reduce the incidence of offspring being eaten by the mother). Chimeras were identified by eye pigmentation (for BALB/C embryo-derived chimeras) or coat color. At 5-6 weeks of age, male chimeras were caged with BALB/C or C57BL/6 females to test germ-line transmission of the ES genome.

Germ-line carriers (heterozygous for the mutation introduced) were identified by Southern hybridizations in a similar manner to the one used to identify the correctly targeted ES cells. DNA was extracted from 1.5 cm mouse tails during weaning [3wks. *p.p. (post partum)*] using the Laird *et al.* protocol]. Tails were incubated overnight in lysis buffer, with shaking at 55°C. Hair was removed by centrifugation. 0.5ml of isopropanol were added. DNA was

precipitated for ~1hr. After washing once with 70% ethanol and air-drying, each sample was resuspended in 200µl of TE. To help dissolve the DNA samples were incubated overnight at 55°C. 25µl of each were used for digestions.

Histology:

Fixing of embryos, animals and tissues was performed as described earlier in the *in situ* hybridization section using Bouin's fixative, immediately after sacrifice. Staining of the paraffin sections was performed using the Mallory's tetrachrome staining method as described by Lufkin *et al.* (1992).

Mallory's Tetrachrome staining of paraffin sections:

Deparaffinization and rehydration of sections:

Slides were placed in a series of baths, each for 3' (changing the series every 150 slides). First in Americlear (a non-toxic substitute for the Xylenes) 2x, and then in ethanol 100% 2x, ethanol 90%, and ethanol 70%.

Staining of sections:

Slides were placed in Groat's hematoxylin for 3-5', and running tap water for at least 5'. They were then rinsed in ddH_2O for a few seconds, followed by Acid Fuschin for 3', Aniline Blue and Orange-G for 1hr, with no H_2O rinse in between.

Rinsing of sections:

Sections were rinsed in four baths of 95% ethanol at 3'/bath, with the last bath remaining clear, followed by a series of baths, each for 3', in ethanol 100% 2x and Americlear 2x. If the dehydration was poorly performed, the Americlear became cloudy and had to be repeated. The slides were drained, then edges were wiped without complete drying, and were coverslip mounted using a drop of Permount®.

Stain Preparation:

Reagents:

Hematoxylin:	Mallinckrodt CAS 517-28-2
Phosphotungstic Acid:	Mallinckrodt; CAS 12067-99-1
Orange G:	Fisher O-267; Colour index 16230
Aniline Blue (Ingrain Blue):	Fisher A-967; CAS 28631-66-5
Acid Fuschin:	Fisher F-97
Ferric Ammonium Sulphate:	Sigma F-3629; CAS 7783-88-7

Groat's Hematoxylin:

Solution 1:	10g ferric ammonium sulphate $\{FeNH_4(SO_4)_2 \cdot 12H_2O\}$, and 8ml of sulphuric acid were added to 500ml sterile H_2O.
Solution 2:	5g Hematoxylin were added to 500ml of 95% ethanol.

Solutions 1 & 2 were mixed thoroughly. This mix was incubated at least for 2 months at r.t. prior to use in order to "ripen". It was not used prior to this incubation. Once "ripened" it was used within 6 months. The baths were changed every 1000 slides.

Acid Fuschin:

For 500ml of solution 2.5g of Acid Fuschin and ddH_2O to 500ml were combined. This stain was used within 6 weeks.

Aniline Blue & Orange-G:

For 500ml of solution 2.5g of Aniline Blue, 10g of Orange-G, 5g of phosphotungstic acid and ddH_2O to 500ml were combined. This stain was used within 3-4 months. The baths were changed every 700 slides or so. Solutions were thoroughly mixed before each use.

Decalcification for Histology:

Embryos older than E16.5 were decalcified prior to sectioning. For *in situ* hybridization, this was achieved by adding EDTA to 5.5% in the 4% paraformaldehyde-PBS fix solution. However, for standard histology, the following protocol was used. All procedures required 50ml per embryo.

Embryos were fixed in Bouin's for at least 1 week at 4°C. They were then transferred to the decalcifying solution (40ml of conc. HCl, 30ml of glacial acetic acid,

100ml of chloroform, 100ml of sterile distilled H_2O and 730ml of 100% ethanol) which was changed at least 4 times during a 48 hr period. They were then transferred to a 95% ethanol solution which was again changed 4 times during a 24 hr period. This step (24hr incubation, 4 washes) was repeated in 100% ethanol, and finally for Americlear. The final incubation was in paraffin at 65°C for 24hr. Blocks (using Peel-A-Way disposable embedding molds from Polysciences Inc., Warrington, PA) were removed and stored at r.t. indefinitely.

Electron Microscopy:

Tissue processing for transmission electron microscopy, viewing and photography were performed with the help of Kathy Pelton-Henrion and Vladimir Protopopov, at Mount Sinai's Cell Biology and Anatomy departmental EM core. In addition, sectioning and microscopy were also performed by Dr. David Birk at Tufts University, Boston MA.

Briefly, organs and tissues of interest were fixed immediately after sacrifice in 2.5% glutaraldehyde, 2% paraformaldehyde, in PBS overnight at 4°C. Post-fixation was in 1% OsO_4 in 0.1M cacodylate buffer (pH 7.4) for 1 hr. Bone-containing tissues were decalcified in 4.5% EDTA in PBS for at least 3 weeks at 4°C. After post-fixation (and decalcification) in osmium the tissues were washed with cacodylate buffer, and dehydrated through a 30 and 50% ethanol series. At this stage the tissues were transferred to 2% uranyl acetate in 50% ethanol and *en bloc* stained with filtered ethanolic uranyl acetate (to enhance the staining of collagen fibrils) for 1 hr at 4°C. Dehydration was continued through a 70, 85, 95, and 100% ethanol series, propylene oxide, and the sections were embedded in Epon-Araldite. The embedded sections were mounted on epoxy cylinders and polymerized. Sections were cut with a pale gold interference color, picked onto mesh grids, stained with 2% aqueous uranyl acetate followed by 1% phosphotungstic acid pH 3.2. Sections were examined and photographed using a Hitachi H-7000 transmission electron microscope.

X-Ray Radiography:

High density X-rays were taken using a Xerox Micro-50 at 25keV with 30" exposure and Agfa-Gevaert Osray-C film.

Mineralization studies:

These studies were performed with the help of Dr. Michael Klein at the Dept. of Pathology. Littermates at 6 weeks of age, of all three groups of animals, were injected *i.p.* (*intra peritoneal*) with 1mg/10g body weight of tetracycline, followed with an identical dose 10 days later (Dr. T. Einhorn, personal communication). Animals were sacrificed 2 days after the second injection, their femurs were cleaned of soft tissues, and fixed in PBS buffered formaldehyde (pH 7.1) . Careful washing 5 days later removed all fixative, and were then embedded in methylmethacrylate. Blocks were trimmed, and sagital sections were cut at 6μm using a Reichert Polycut S microtome. They were then examined unstained under fluorescent light using a Zeiss Axiolphot microscope (Glorieux *et al.*, 1991).

Biomechanical testing:

These experiments were performed in Dr. Harold Alexander's laboratory at the Hospital for Joint Diseases, and with the help of Mr. Roberto Casar P.E.

Biomechanical studies were performed using an Instron-1321 servo-hydraulic test system. Briefly, skin from sacrificed animals was obtained as intact as possible. It was laid flat on paper that was saturated with PBS, and was stored at -20°C. Age matched specimens from wild type, heterozygous and homozygous animals were collected. An identical "dumbbell" shape section (see figure 7), was cut from identical locations on the back of all specimens, in similar orientations. This shape was chosen, in order to easier monitor the failure-breakage-points.

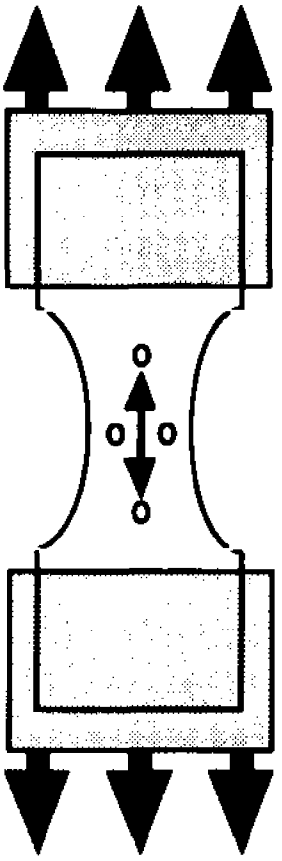


Figure 7.
Diagram showing the "dumbbell" shape of the skin specimen between clamps. Concentration of forces occurred in the middle of the specimen (indicated by the solid arrows). Black circles were used to monitor displacement.

Once clamped, the skin specimens were extended under linearly increasing force until failure. A video camera was used to monitor the displacement of the black circles and the force applied. Images were retrieved, digitized and measurements were made regarding the above parameters.

Establishment of Cell lines and Antibody staining:

Dermal fibroblasts:

Large skin pieces were removed from freshly sacrificed animals and washed briefly in absolute ethanol. These were then minced to finer samples, placed in 100mm tissue culture dishes and were just covered with enough D-MEM F12 medium (Gibco-BRL), 15% FCS, antibiotics and Fungizone (Gibco-BRL) so that they would not float. If contamination occurred, infected pieces would be removed and the remaining would be transferred to

individual wells of a 24-well plate. After 3-5 days when dense regions of cells were observed on the plate, tissue remnants were removed and fresh medium was added. When cells were >50% confluent, they were washed 2x with PBS, trypsinized and pooled together in a single dish (its size depending on total cell numbers). Cells were passaged not more than 8 times.

Corneal fibroblasts (keratocytes):

These cultures were established in Dr. Sandra Masur's laboratory by Dr. Stefan Antohi of the department of Ophthalmology. Eyes were obtained from freshly sacrificed animals. Corneas were removed and washed briefly in PBS. Epithelium side down, corneas were placed in contact with a drop of trypsin for ~2'. With the help of micro-forceps, the epithelium was removed by scraping. The corneas were turned over and Descemet's membranes were placed in contact with the trypsin for another 2', after which the endothelium was scraped away. Trypsinization was stopped by placing the corneas in D-MEM F12 medium supplemented with 15% FCS, antibiotics and fungizone. The corneal stroma was then cut in smaller pieces and placed in a 24-well plate in the above medium. After 3-4 weeks, corneal fibroblasts have migrated out of these fragments which were then discarded. The cells were used up to 8 passages.

Cytochemistry:

Both dermal and corneal fibroblasts destined for cytochemical studies were passaged on coverslips in a 24-well plate. They were then allowed to grow for another 24 hr. They were then rinsed once with BSS (Hank's solution) or PBS, and the medium replaced with one containing 1% FCS, and 50μg/ml Ascorbic acid (Sigma) to induce collagen production. Incubated overnight. Cultures were rinsed once with BSS and 1% FCS containing medium with Brefeldin (10μg/ml) was added 2.5-3.0 hours before fixation (in order to retain the normally secreted proteins in the endoplasmic reticulum for

visualization). The cells were rinsed 2x with PBS and fixed with 4% freshly prepared PBS-buffered paraformaldehyde pH 7.2, for 15' at room temperature. After rinsing 2x with PBS, autofluorescence was quenched for 5' in 50mM NH_4Cl in PBS. After another PBS wash, blocking with 3% normal goat serum (NGS) for 15' followed.

Cells were washed 2-3x with 1% PBSA (1% BSA in PBS) for 5' each. Coverslips were then removed and placed on parafilm. To demonstrate intracellular localization of both collagen types V and I in cultured fibroblasts, antibodies for each were applied sequentially. 50µl of the rabbit polyclonal anti-human type V collagen antibody [generous gift of Dr. D. Hartmann of the Institute Pasteur de Lyon](1:100 dilution in PBSAT[0.2% PBSA, 0.1% Triton]) was placed on each coverslip, at 4°C overnight. They were then rinsed 1x with 3% NGS/PBS, followed by 3-4 rinses with 1% PBSA, 5' each. The coverslips were then incubated with 50µl of biotinylated anti-rabbit (Sigma, at 1:100 dilution) and anti-mouse type I collagen antibodies (Chemicon, at 1:100 dilution) at r.t. in a humid chamber for 3 hours. They were rinsed again 1x with 3% NGS/PBS, followed by 3-4 rinses with 1% PBSA, 5' each.

Incubation with 50µl of streptavidin-FITC (fluoroscein isothiocyanate, Amersham U.K.) and anti-mouse TRITC (tetra-rhodamine isothiocyanate, Kirkagard & Perry) at 1:100 dilution in 0.2% PBSA, at r.t. in a humid chamber for 3 hours followed. They were finally rinsed 3x with PBS, covered with 50% glycerol and stored at 4°C. For viewing, coverslips were mounted on slides with antifade medium in 50% glycerol.
Viewing and photographing were performed using a Zeiss Axiomat microscope.

Collagen Purification from Supernatants of Cell Cultures:

Primary corneal and dermal fibroblasts were cultured to confluency using D-MEM containing 15% FCS. They were rinsed once with PBS, and D-MEM containing 1% FCS, ascorbic acid (50 µg/ml) and β-aminopropionitrile (64 µg/ml) was added, followed by 3 hr incubation at 37°C. Medium was removed, and D-MEM (which in most cases is proline-

free) containing 1% FCS and 2 μCi/ml ^{3}H-Proline (Amersham, U.K.) was added. Cells were incubated overnight at 37°C, and the supernatant was collected the following day on ice, in a tube containing protease inhibitors (0.2 mM PMSF (Phenyl Methane Sulfonyl Fluoride), 10 mM NEM (N-Ethyl Maleimide) and 2.5 mM EDTA). Dishes were rinsed once with PBS, and the collected liquid was combined with the supernatant. Tubes were centrifuged for 5 minutes at 1000 rpm to remove the debris of dead cells. Collagen type I at 2 mg/ml was added as a carrier to the supernatant. After gentle mixing it was transferred to a small beaker. The beaker was placed on ice, and ammonium sulfate (176 mg/ml) was added slowly, with constant stirring. The aqueous solution was transferred to a glass tube and collagen was allowed to precipitate at 4°C overnight. It was then centrifuged at 13,330 rpm (Sorvall SS34 rotor) for 30 minutes at 4°C. The pellet was resuspended in 0.5 M acetic acid and dialyzed against 0.5 M acetic acid to remove free radioactivity. The dialysis fluid was changed several times until the number of counts equaled background. Samples were lyophilized to concentrate counts. Immunoprecipitations were performed using standard protocols (Sambrook *et al.*, 1989).

Results

Primary Structure of the Murine Pro-α2(V) Collagen

Isolation and characterization of pro-α2(V) cDNA clones:

Initial screening of the NIH-3T3 cDNA library was performed under cross-hybridization conditions using a 600-bp EcoRV fragment of OK-25, a cDNA which codes for most of the C-terminal propeptide of the human pro-α2(V) chain (Weil *et al.*, 1987). The resulting positive clone, KA-273, extends from the 5' noncoding region of the message and includes most of the C-terminal propeptide coding sequences (Fig. 8). A subsequent screening of the NIH-3T3 library, using a 693-bp BamHI fragment located in the 3' foremost portion of clone KA-273, led to the identification of an overlapping cDNA, KA-311, whose sequence extends into the 3' noncoding segment of the pro-α2(V) mRNA (Fig. 8).

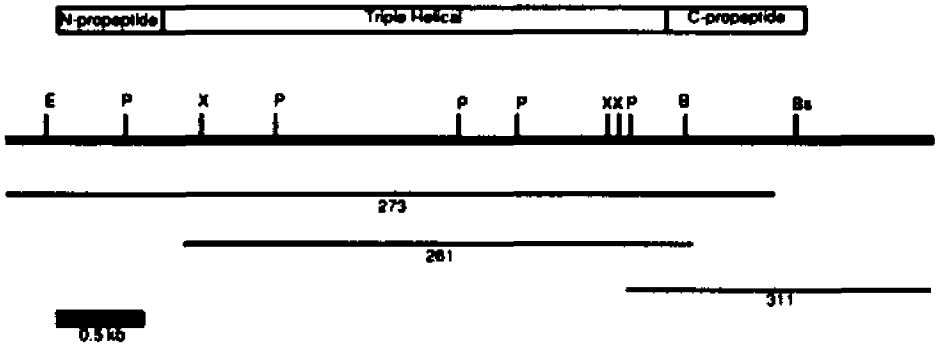
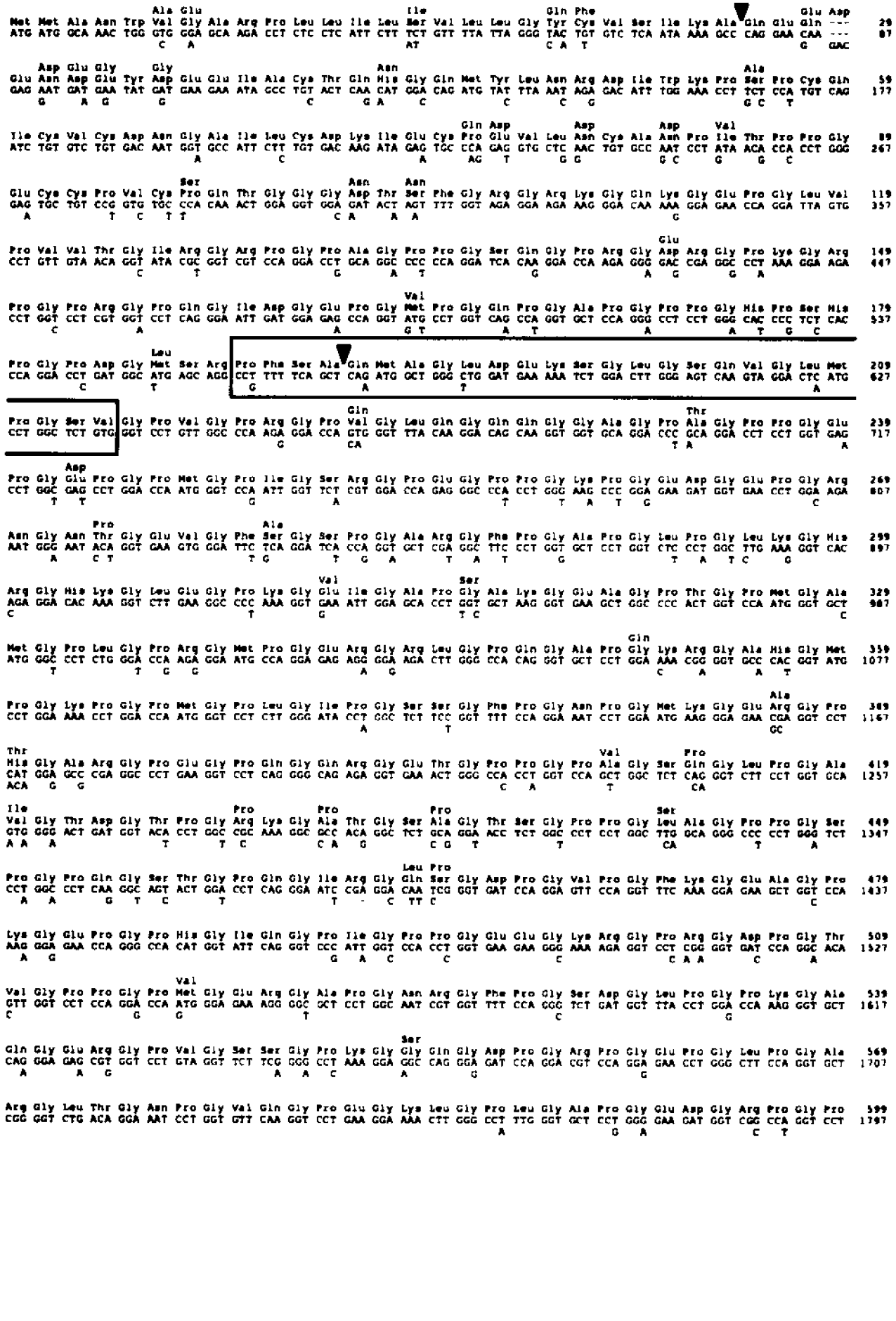
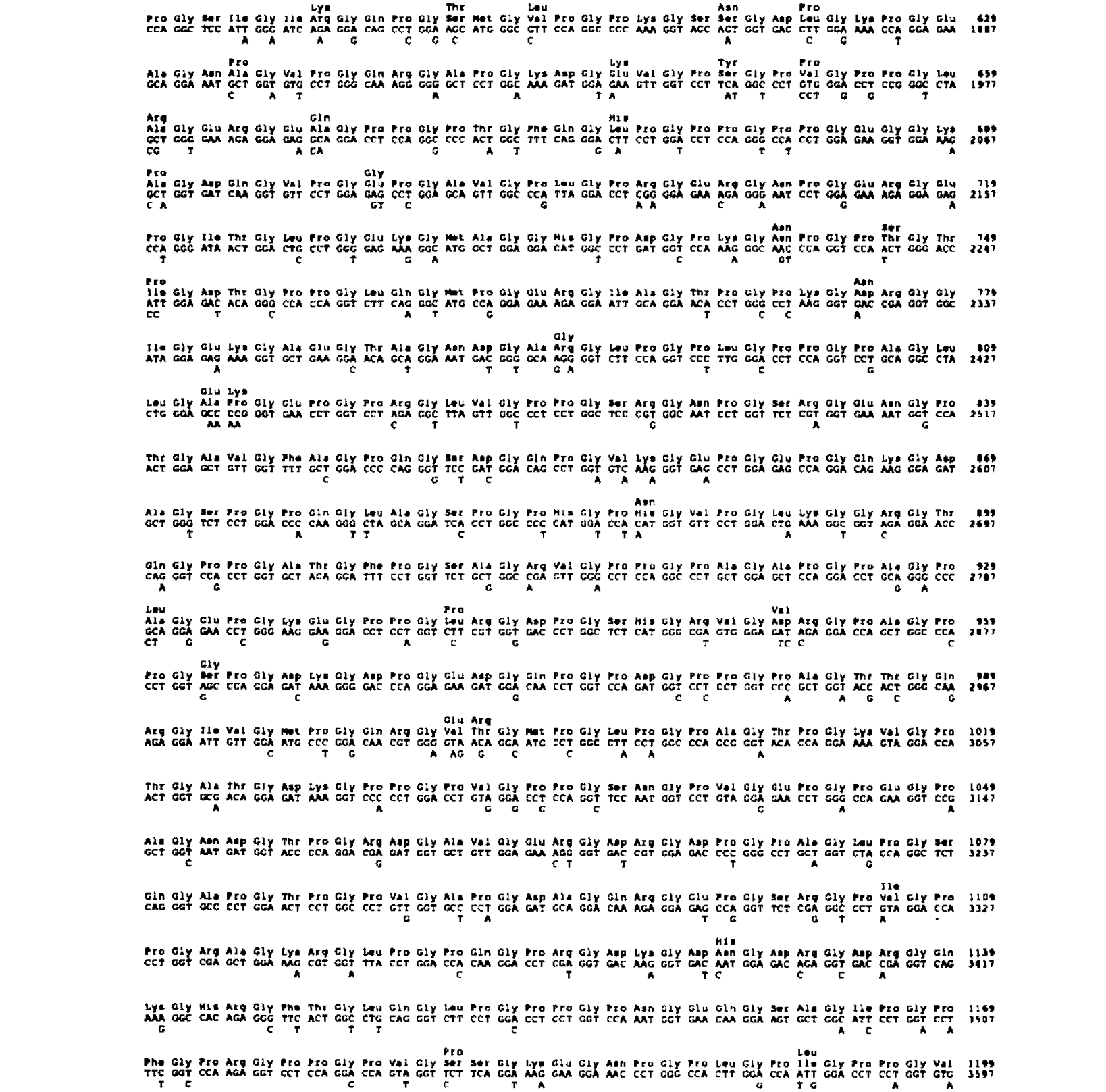


Figure 8.
A schematic showing the relative positions of the overlapping clones and the relative positions of E(Eco RI), P(Pst I), X(Xho I), B(Bam HI), and Bs(Bst XI).

The composition of the coding and noncoding sequences of *col5a2* are shown next in Figs. 9 and 10, respectively.

Note: The results described up to the gene targeting experiments (p.52) have been described in Andrikopoulos *et al.* (1992), with the exception of Fig. 12.





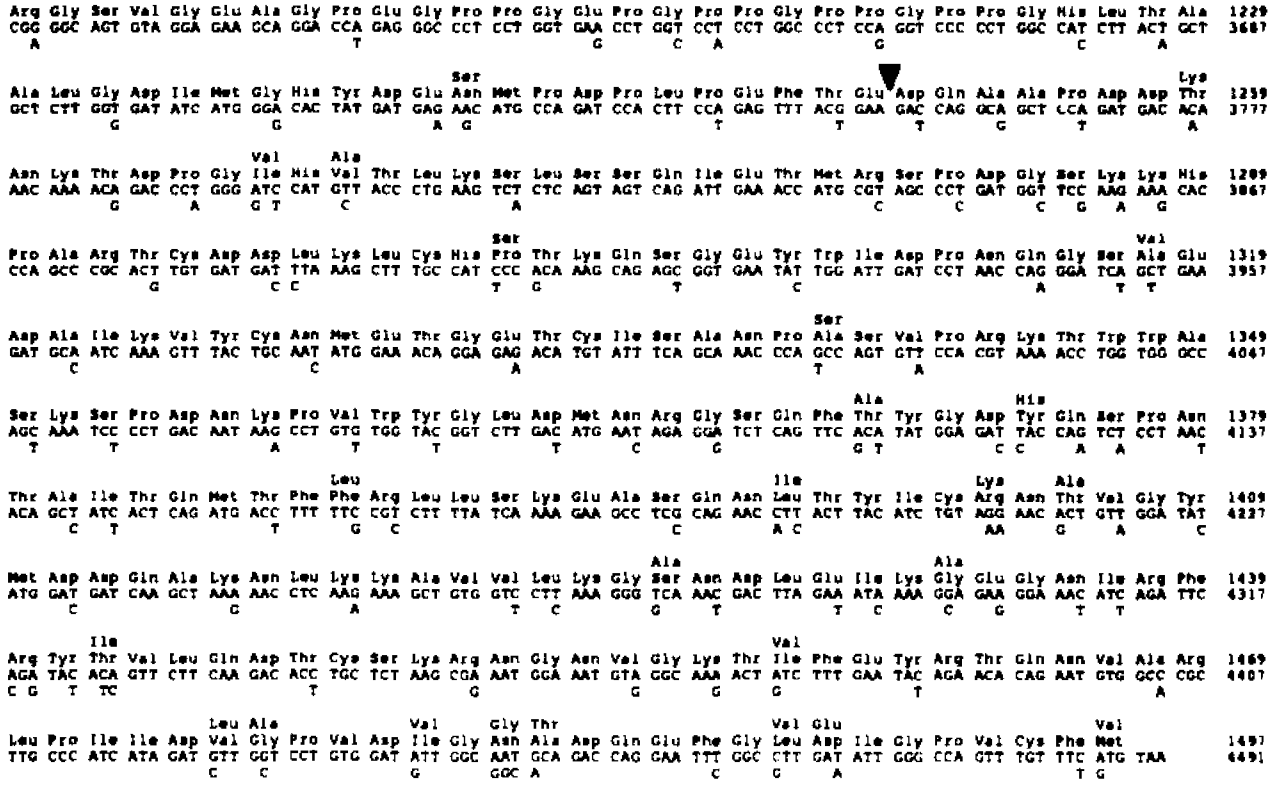


Figure 9.
Nucleotide and deduced amino acid sequences of mouse pro-α2(V) collagen.
Differences with the human sequences are shown above and below the mouse sequences. Putative cleavage sites for various endopeptidases are indicated by black triangles. Exon 6 is boxed. Nucleotide sequences have been deposited in the GenBank/EMBL Data Bank under the accession number LO2918.

A.

tcgcagaggcgcgcgctctgatttggttgtttaccatcaatcagacggttgcttggcagacactggatggttatgagcctgaacaagctg
aaaggggcaggaaaagaagtggaggcagcattcttcctattaaagctgcaccgcttgaaaaaagtttcgcagactgtgccggacctg
gtgctgaaacagactgaggcagcgcgggactggtgctgaaaaagagcccccagcggaattcatggtgctacagcgacctcagcatc
tactctccgaccccagcaggacccacatgtctcatctttcagac

B.

cagaagcccagacgtatcgacaatgagcactaccatcaatgaccaccacaagaactgtgactgtttaaagttcacctgagactcttgaag
taatggctggtcctgcatcagctttgtacatac**gatctgaagtgcctgaccac**cttatccttcagaatatttatttacttacaatcctcaa
gtttaattgacttaatttttcaatatggccgtttagattaaagcaaccaatgacaatgaccatctttacaaaaagtaaactgatcaaaaaatnaaa
tatgttttcatcaattcatttcactgtaatgaatacgttgcttagtatttatgagaaaattcttcttcctagcaggtagcttacatgatggggtatgt
aaagcctcagcatgtgttgattttgcttggctgggaat

Figure 10.
Nucleotide composition of 5' (panel A) and 3' (panel B) untranslated sequences of *col5a2*. The former extend from the ATG codon, while the latter begin immediately after the TAA codon. The sequence in bold corresponds to the oligonucleotide primer used in the RT-PCR amplification.

The deduced amino acid sequence of the murine chain exhibits a 93.5% identity to the human counterpart, with absolute conservation of several important structural elements. They include: the putative cleavage sites of the amino- and carboxy-proteinases; the lysine-mediated cross-linking sites; the N-linked glycosylation site of the C-terminal propeptide; and the inter- and intra-chain cysteinyl-mediated disulfide bonds of the globular domains (Weil *et al.*, 1987; Woodbury *et al.*, 1989). The major difference between the two mammalian polypeptides is the presence of an additional residue in the N-terminal propeptide of the human chain (Fig. 9).

Expression Patterns of the Pro-α2(V) collagen

Temporal pattern of expression:

To analyze *col5a2* expression during development, RT-PCR amplification was first employed on RNA samples from 9, 10, 11 and 13 days *p.c.* mouse fetuses. As controls, parallel RT-PCR amplifications were performed on the same RNA samples using primers specific for mouse pro-α2(I) and pro-α1(II) collagen sequences (Rossi and de Crombrugghe, 1987; Cheah *et al.*, 1991a). This revealed that the three collagen transcripts are readily detectable in 9 day *p.c.* mouse embryos (Fig. 11). Incidentally, the RT-PCR results are in agreement with the temporal data of types I and II collagen expression previously determined by *in situ* hybridizations (Cheah *et al.*, 1991b).

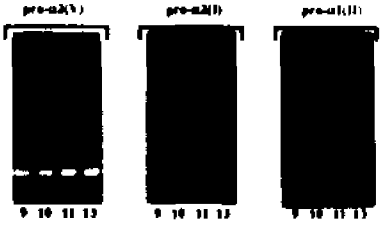


Figure 11.
Ethidium bromide stained gels showing the RT-PCR amplified products from days 9, 10, 11, 13 *p.c.* embryos using primers specific for pro-α2(V), pro-α2(I) and pro-α1(II) collagens. Sizes of PCR bands are 713-bp [pro-α2(V)], 167-bp [pro-α2(I)], and 172-bp [pro-α1(II)]. It should be noted that the intensity of the PCR products does not reflect the relative amount of each mRNA species but the difference in PCR amplification conditions.

Primer extension assays were also performed with total RNA from days 9 and 10 *p.c.* reconfirming the RT-PCR results, and also identifying the start site of transcription 84bp upstream from the ATG (Fig. 17).

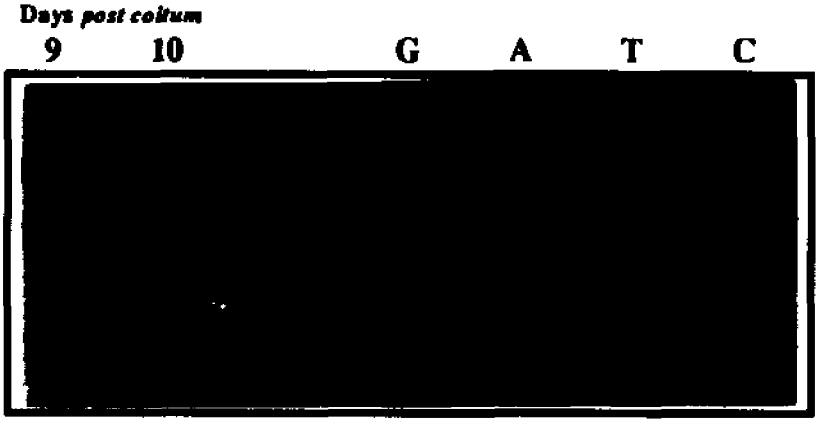


Figure 12.
Primer extension assays showing the presence of pro-α2(V) message for both days 9 and 10 *p.c.*

Spatial pattern of expression:

Based on these results, mouse embryos of days 12.5 and 16.5 *p.c.* were sagittally sectioned and utilized for *in situ* hybridization to examine the tissue distribution of the *col5a2* transcripts at these two developmental stages. As a positive control and useful point of reference, a *col2a1* antisense riboprobe was used in adjacent sagittal sections. The fragment used to generate a ^{35}S antisense RNA probe for the *col5a2* gene was a 5'-untranslated - Pst I, 702 bp fragment (Fig. 8), proven to be highly conserved between species (See Fig. 17 later in this chapter), and therefore specific for this transcript. The pro-α1(II) specific anti-sense RNA probe was generated using a Sty I - 3'-untranslated, 445 bp fragment was used in adjacent parasagittal sections.

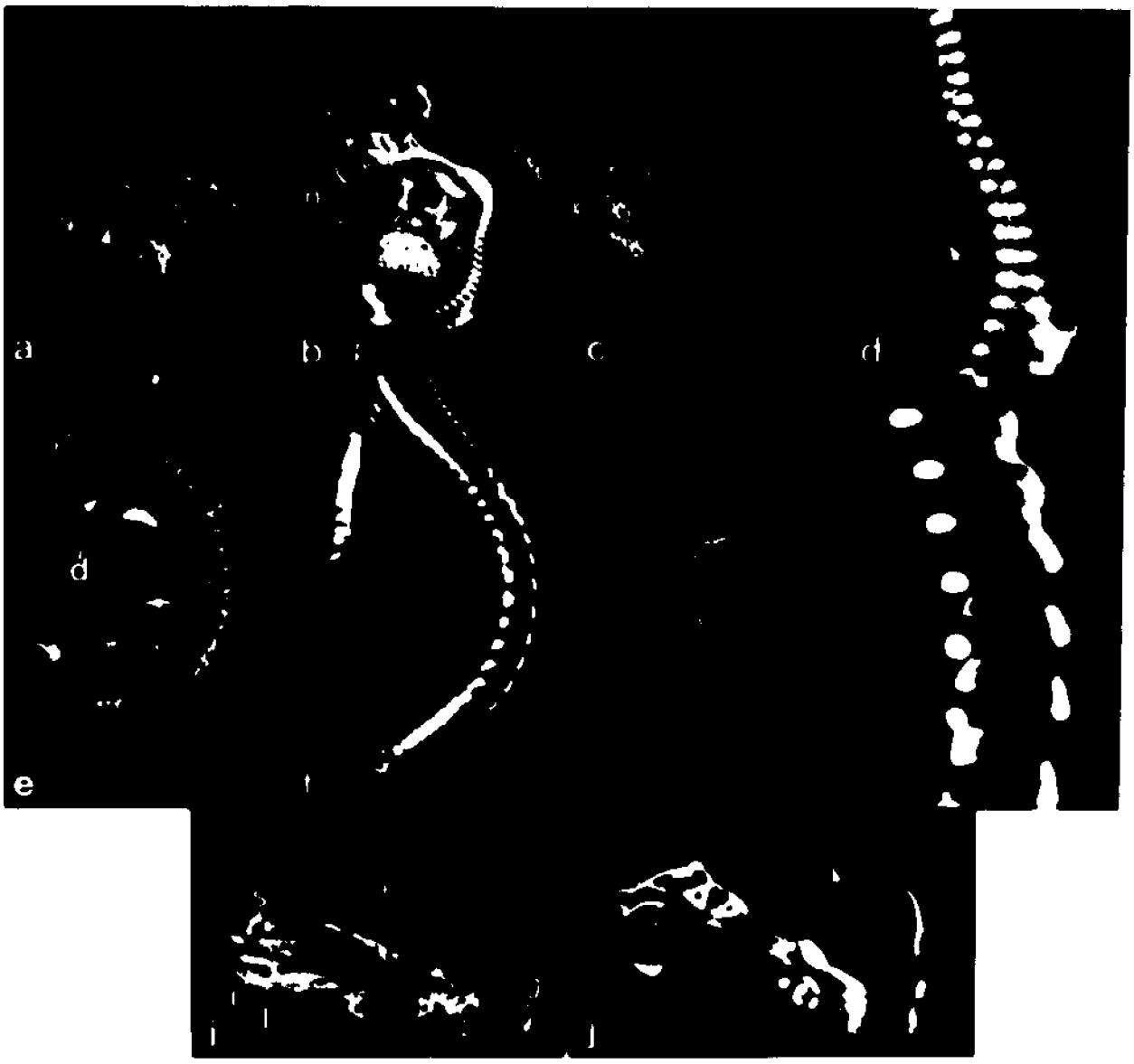

a
b
c
d
d
e
f
g
h
i
j

Figure 13.

Dark-field visualization of *col5a2* (a,c,e,g,i) and *col2a1* (b,d,f,h,j) messages in sagittal sections of days 12.5 (a-d) and 16.5 *p.c.* (e-j) mouse embryos. (a); *col5a2* is widely expressed in the mesenchymal tissues; high signal can be seen in the craniofacial mesenchyme and peritoneal tissue (arrow head). Bar = 1mm. (b); *col2a1* is heavily expressed in cartilaginous tissues, such as vertebral column, nasal septum (n), Meckel's cartilage (arrow head), the primordia of hyoid and thyroid cartilages (arrow) and pelvic girdle (p). Signal can be also seen in neuroepithelium of the roof and floor of the brain. (c); Enlargement of (a) showing a weak signal throughout the mesenchymal tissue. Bar = 500 μm. (d); Enlargement of (b) showing a weak signal associated with the lung epithelium (arrow) and a strong signal with the vertebrae. (e); *col5a2* is highly expressed in perichondrium, diaphragm (d) and intestinal mesenchyme (arrow). Transcripts are also detectable in the atrioventricular valve of the heart (arrow head). Note that the apparently high signal in the lung is an artefact. Bar= 1mm. (f); *col2a1* expression persists in the cartilaginous tissue of the vertebrae and sternum (s). (g); Enlargement of (e) showing a strong signal associated with the perichondrium and fascia (arrow) and a weak signal throughout mesenchymal tissue; in contrast there is no detectable expression in vertebral cartilage. Same magnification as in (c). (h); Enlargement of (f) showing *col2a1* expression restricted to the cartilaginous region of the vertebrae and its absence in the ossification centers and in other mesenchymal tissue, as well. (i); *col5a2* expression in the head. Signal can be seen in the mesenchymal tissue of the fronto-nasal region, palate, tongue, and mandibular arch particularly in the area around the *incisivus inf.* (i), lip furrow (l) and *incisivus sup.* (s). There is also some positive signal around the eye (e), otic capsule and other perichondrial tissue. Same magnifications as in (e). (j); In the head, *col2a1* expression is mostly restricted to cartilaginous tissue but a weak signal can still be noted in the neuroepithelium (arrow).

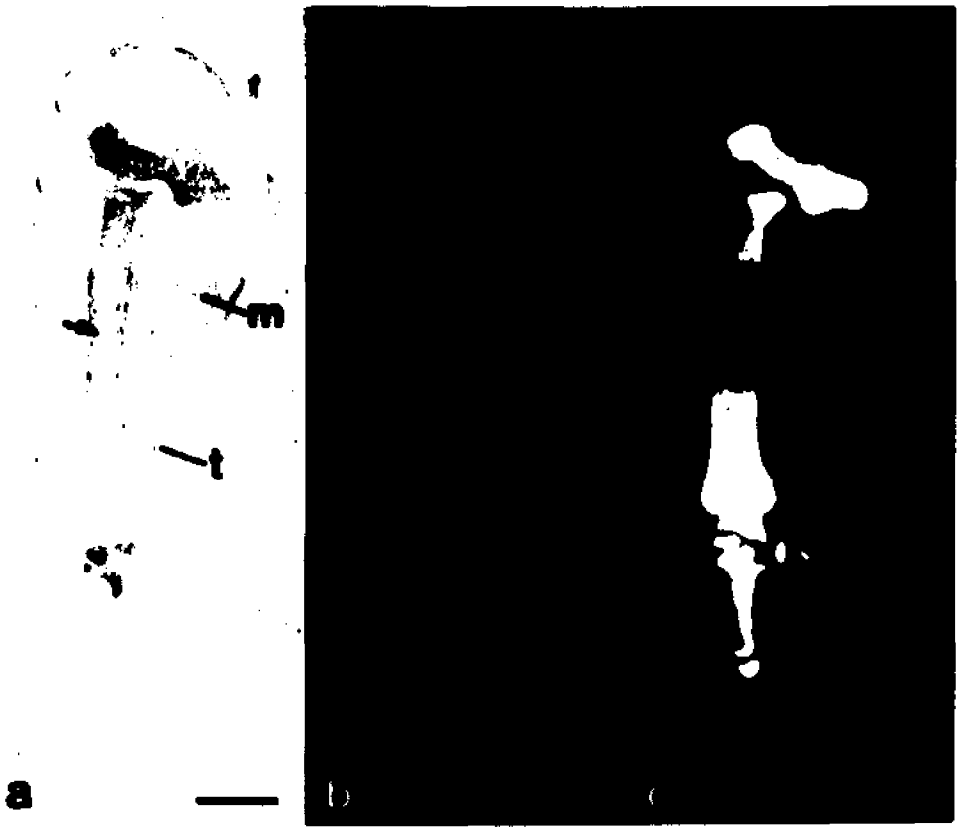


Figure 14.
Expression of *col5a2* (a,b) and *col2a1* (c) in sagittal sections of a day 16.5 *p.c.* mouse forelimb. *col5a2* transcripts are visible throughout the mesenchymal tissue except for cartilage. High signal can be seen in the perichondrium, periosteum, ossification center (arrow) and tendon (t). A much weaker signal is also noted in the muscle region (m). In contrast, *col2a1* expression is solely restricted to cartilage. Bar = 500 μm.

In day 12.5 *p.c.* embryos, a low and rather diffuse level of *col5a2* expression was observed in the peritoneal membranes, intestinal mesenchyme, and craniofacial mesenchyme (Fig. 13 panels a and b, c). In contrast, *col2a1* hybridization signals are substantially higher and confined to cartilaginous structures, such as the vertebral column, the nasal septum, Meckel's cartilage, and the primordia of hyoid and thyroid cartilage (Fig. 13, panels b and d) (Cheah *et al.*, 1991b).

In day 16.5 *p.c.* fetuses, levels of both transcripts are significantly augmented. Additionally, a distinct and more restricted pattern of *col5a2* expression begins to emerge (Fig. 13, panels e, g and i). At this stage of development most of the vertebral bodies

contain initial ossification centers (Theiler, 1989). As can be appreciated by comparing the relevant data of Fig. 13, as well as the sagittal sections of Fig. 14, *col5a2* transcripts are confined to primary ossified regions and perichondrium while *col2a1* expression is localized to the cartilaginous regions of the developing bones. Overall, the tissue specific expression of these collagen genes is mutually exclusive.

Additional pro-α2(V) expressing organs are the joints, tendon, the diaphragm, the peritoneal membrane surrounding the internal organs and the atrioventricular valve of the heart (Fig. 13, panels e and g and 14b). No *col5a2* hybridization is seen in the developing brain, where *col2a1* expression has been recently noted in the neuroepithelium and chondrocranium (Fig. 13, panels i and j, and Cheah *et al.*, 1991b). Other head structures showing high levels of *col5a2* transcripts are the mandibular arch, the tongue, the optic nerve region (probably perichondrial surrounding the sclera of the developing eye), the otic capsule and the calvarial mesenchyme (Fig. 13i).

Gene Targeting Experiments

Isolation and Characterization of Murine Pro-α2(V) 5'-genomic clones:

Absolute identity between the region of homology of the targeting vector and that of the chromosomal DNA of the stem cells, was emerging as an important factor for efficient gene-targeting. Since the ES cells were derived from a 129/Sv/ter mouse strain, a genomic library was prepared using isogenic DNA (Dr. Rudolf Jaenisch personal communication, and Deng and Capecchi, 1992). An example of an ethidium bromide stained gel with the different size fractions of the 129/Sv partially digested DNA is shown in Fig. 15.

Figure 15.
Ethidium bromide stained gel showing the different fractions obtained from the Sau3A partial digestion of the genomic DNA.

Using the same 5'-untranslated - Pst I, 702 bp cDNA fragment that was used for the *in situ* studies, 2 x 10^5 pfu's of the 129/Sv genomic library were screened. Two positives were identified, clones F1 and D2.6. In order to proceed with the gene targeting experiments, a detailed restriction enzyme mapping of the genomic clones was generated (Fig. 16). Clone F1 (~16.0 kb) included ~7.3 kb of the 5' region (upstream from the first exon), and up to ~8.3 kb into the first intron. Clone D2.6 (~14.0 kb) included ~4.5 kb of intron 3 proximal to exon 4/5, and up to ~1.0 kb of intron 8.

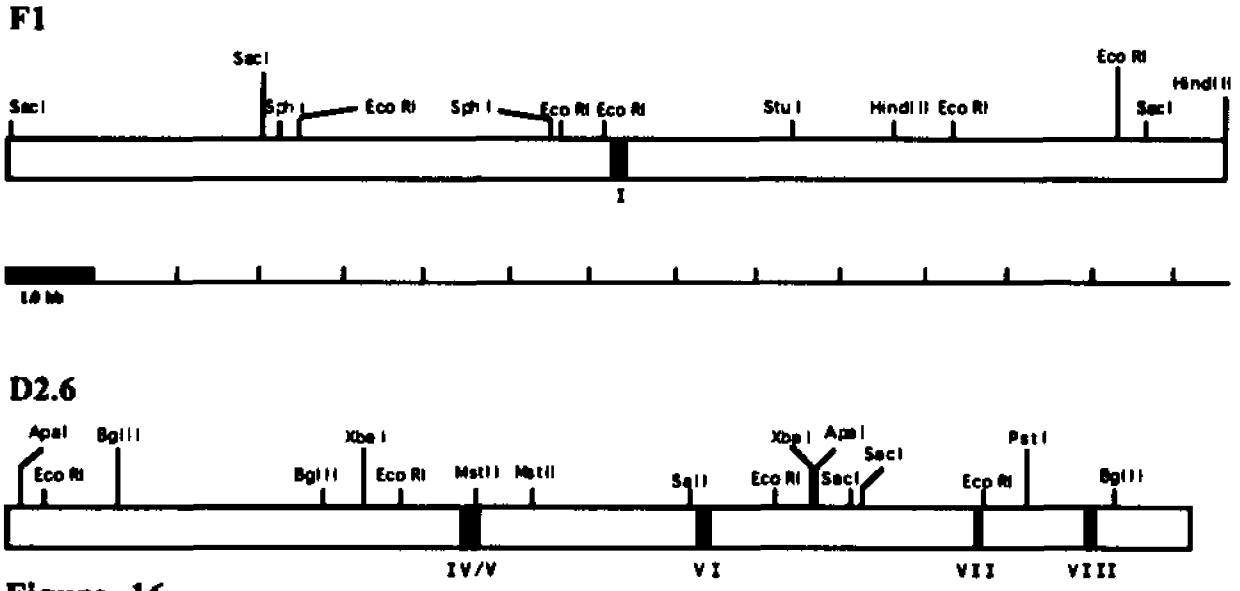


Figure 16.
Detailed restriction enzyme maps of the *col5a2* genomic clones F1 and D2.6. Exons and the relative distances between them were mapped by Southern blotting with cDNA probes, and by PCR.

The 2.0 kb Bgl II - Bgl II fragment of F1, containing exon 1 and 906 bp upstream of the ATG region, was subcloned into the Bluescript-SK vector (Stratagene, La Jolla CA) and sequenced. This revealed some interesting structural features. When the segment extending from nucleotide -150 (5' to the ATG) to nucleotide +160 was compared to its human counterpart , an overall sequence identity of 87% was noted (Weil *et al.*, 1987; Truter *et al.*, 1992) (Figs. 9 and 10). Such a high level of structural similarity of putative regulatory elements has been observed in previous interspecies comparisons of other fibrillar collagen genes, and is a nearly unique hallmark of this family of genes (Ramirez, 1989). More striking is the homology (95%) of the distinct regions A and B which have recently been shown to bind two, apparently novel transcription factors that seem to specify *col5a2* expression (Truter *et al.*, 1992). In addition, ubiquitous regulatory elements such as a TATA-like box and a binding site for activator protein-2 (AP-2), were also noted (Mitchell and Tjian, 1989) (Figs. 10 and 17). Upstream of position -152 a steady decrease in sequence homology was observed, and after nucleotide -300 the interspecies alignment became uninformative (data not shown).

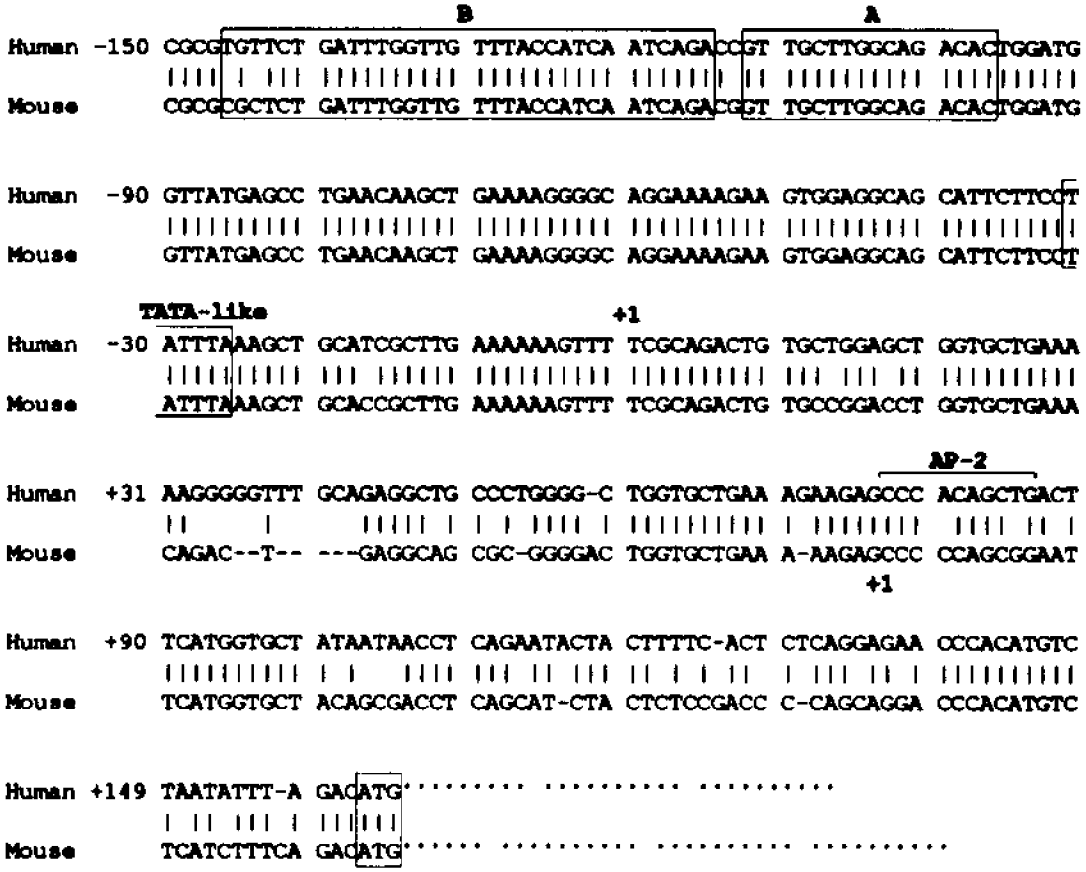


Figure 17.
Pairwise comparison of the 5' sequences of human and mouse *col5a2*. The start sites of translation are indicated as +1. The AP-2 consensus sequence, TATA-like motif and A and B *cis*-acting elements are also indicated.

Generation of the targeting construct:

In order to reproduce in mice a phenotype that would resemble as closely as possible one that has been described in the humans, the nature of the human mutations described in EDS VII was taken into consideration. The α2(V) chain has been shown to undergo proteolytic processing in the step-wise manner characteristic of the fibrillar collagens, both in the avian systems (Fessler and Fessler, 1987), and more recently in the A204 rhabdomyosarcoma cell line (Kleman *et al.*, 1992). By analogy, a similar structural abnormality was therefore introduced in the pro-α2(V) gene. The targeting vector resulting in the dominant/negative mutation was generated (Fig. 18).

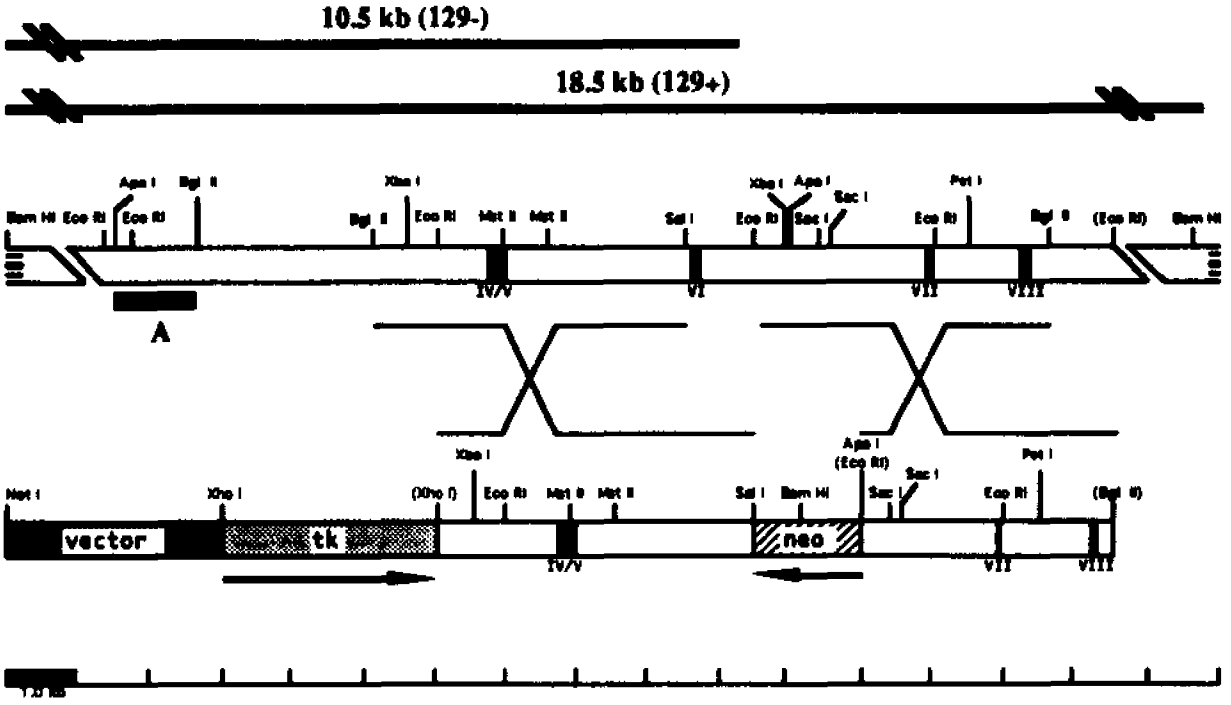


Figure 18.
Schematic representation of the targeting construct used to generate the pNa25 mutation. The bars above the constructs indicate the Bam HI expected sizes of the wild type (129+), and mutated alleles (129-), in the ES genome when the Apa I-Bgl II fragment (A) was used as a probe.

The 9.5 kb Bgl II-Bgl II fragment containing exons 4/5 to 8, was subcloned into the Bam HI site of Bluescript II-SK. All subsequent manipulations involved this genomic DNA segment of the pro-α2(V) gene. The 2.0 kb Sal I-Eco RI fragment containing exon 6 of the gene was replaced by the 1.8 kb neo cassette (conferring resistance to G418 for positive selection) (Tybulewicz *et al.*, 1991). The 3.0kb HSV-tk cassette (conferring sensitivity to FIAU for negative selection) was placed on the 5' end of the construct. The plasmid was linearized with Xho I before being introduced to ES cells by electroporation. The above construct was not expected to destroy the reading frame.

Identification of the correctly targeted clones was essential for the success of the targeting experiment. For this purpose, several fragments outside the homology region were used as hybridization probes to 129/Sv DNA digested with different restriction

endonucleases. This analysis was necessary in order to exclude regions containing repetitive elements. Since such sequences occur at high frequency in the mouse genome, background hybridization and uninformative bands could make identification of the correctly targeted clones impossible. The Apa I-Bgl II fragment 5' to the region of homology was shown to satisfy the above criterion by producing single hybridizing bands (Fig. 19).

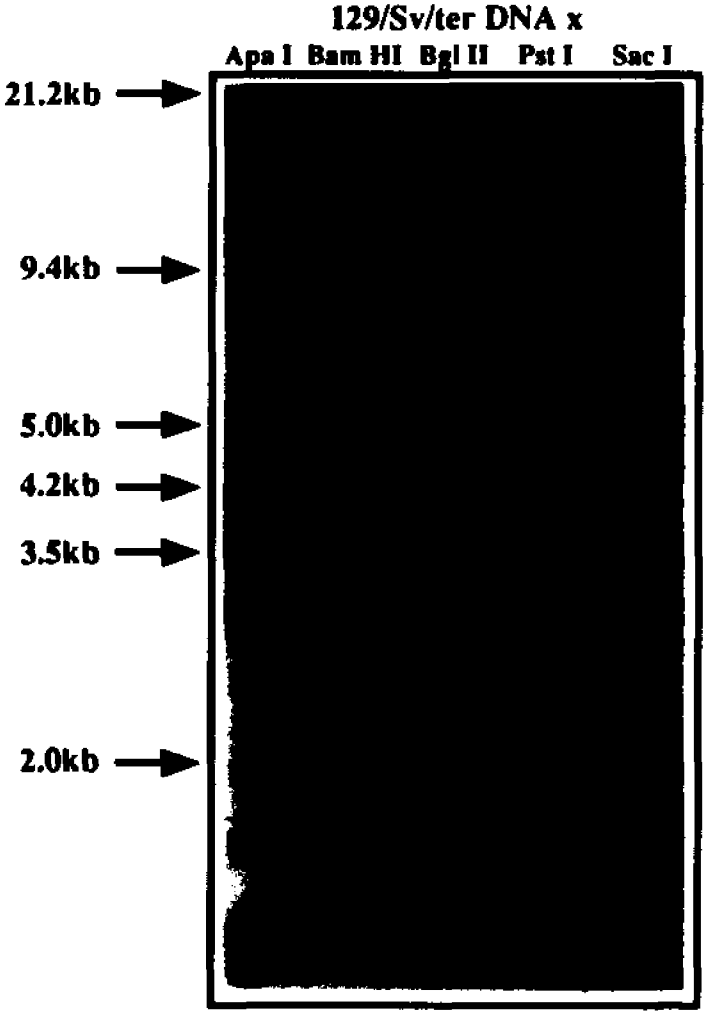


Figure 19.
A 129/Sv genomic DNA southern blot hybridization with the 5' Apa I-Bgl II fragment (A in Fig. 18) as a probe.

The Bam HI digest was chosen because probe A recognizes a ~19.0kb fragment which contains the 9.5kb Bgl II-Bgl II region of homology. Introduction of the neo

cassette also generates a Bam HI site. If homologous recombination occurs, a Bam HI digest and hybridization to probe A (of DNA from the correctly targeted clone) would identify two bands: one of ~19kb for the wild type allele, and one of ~10kb of the targeted allele.

One hundred and ninety two doubly selected ES cells were picked and processed. G418 and FIAU selected clones were on average half in numbers from the G418 only selected clones, thus indicating FIAU enrichment of ~2.0. Five of these lines were found to be correctly targeted (~1:40).

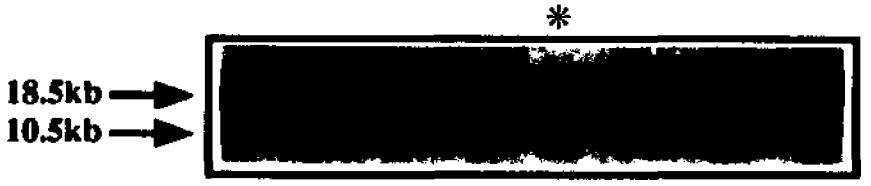


Figure 20.
Southern blot hybridization with an external probe, showing one ES clone (asterisk), where homologous recombination had occurred.

After expansion of these five ES cell lines, and prior to micro-injection, DNAs were isolated and tested again for homologous recombination events using the neo cassette as an internal probe (Fig. 21).

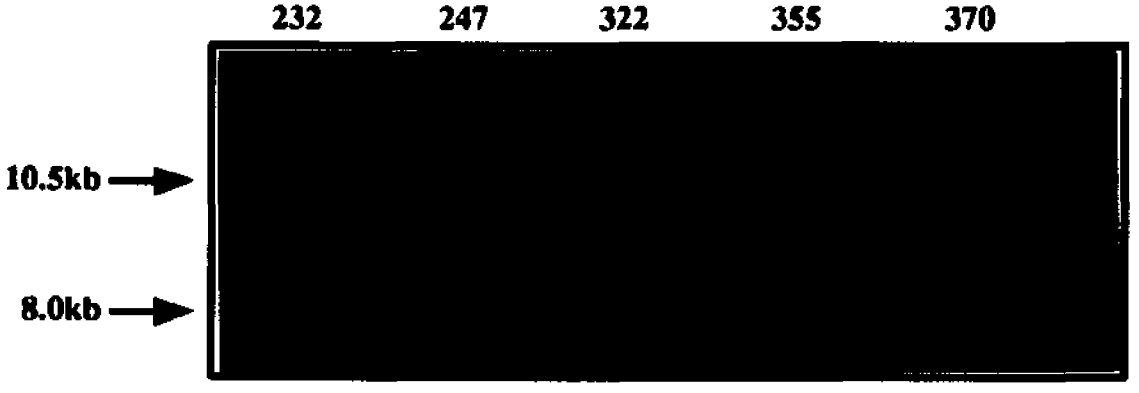


Figure 21.
Southern blot hybridization of the five ES clones with the neo cassette (internal probe).

Generation of chimeras:

Two of these lines, number 232 and 355, were injected into both C57BL/6 and BALB/C host blastocysts, which in turn were re-implanted into F1 CBA x C57BL/6 foster mothers. Line 232 produced low percentage chimeras (<10% as deduced by their coat colors). Line 355 on the other hand produced high percentage chimeric animals (>60%) on both backgrounds (Figs. 22 and 23).

Figure 22.
A nearly 100% chimeric animal (right), on the C57BL/6 background (left, control).

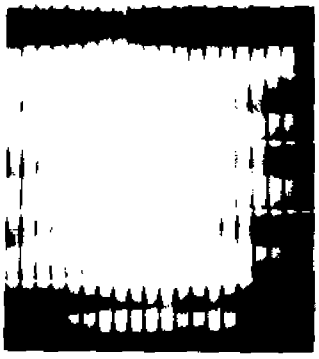
Figure 23.
An approximately 70% chimeric animal on the BALB/C background.

Table 2. Numbers of chimeric offspspring generated from the two different ES lines.

	C57BL/6		BALB/C		
ES Line	Male	Female	Male	Female	Total
232	-	-	1	2	3
355	3	-	4	4	11
Total	3	-	5	6	14

Chimeras in the C57BL/6 background were mated with C57BL/6 wild type animals only, and chimeras in the BALB/C background only with BALB/C wild types, in order to obtain germline transmission of the introduced mutation (and keep the backgrounds separate). Tail DNAs of the agouti colored offspring was processed as described in the previous chapter. Line 232 chimeras never produced any germline offspring during a 4 month period and were therefore sacrificed. On the contrary, all of line 355's chimeras gave germline transmission (Fig. 24).

A.



B.

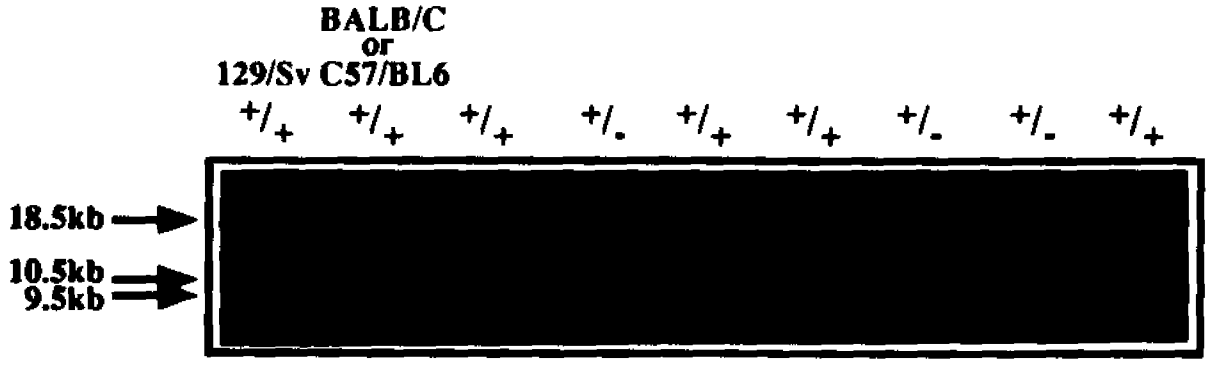


Figure 24.
Panel A shows the ethidium bromide stained gel of 96 Bam HI digested tail DNAs. Panel B shows a portion of its southern blot hybridized to the external probe A. The smaller hybridizing band (~9.5kb) is a result of a Bam HI polymorphism present in the C57BL/6 and BALB/C genomes.

The existence of three bands instead of the expected two was explained when C57BL/6 DNA was included as a control. All of the lanes have the smaller hybridizing band (~9.0kb) in common. This is the wild type band in the C57BL/6 and BALB/C backgrounds instead of the 18.5kb detected in the 129/Sv/ter ES cell genome. Chimeric offspring therefore have the wild type band, and either the mutated band (10.5kb, and are heterozygous for the mutation), or the 18.5kb 129/Sv wild type band. When the chimeras from line 355 were mated with wild type 129/Sv/ter animals, germline offspring gave the predictable hybridization pattern shown in the middle lane of figure 25.

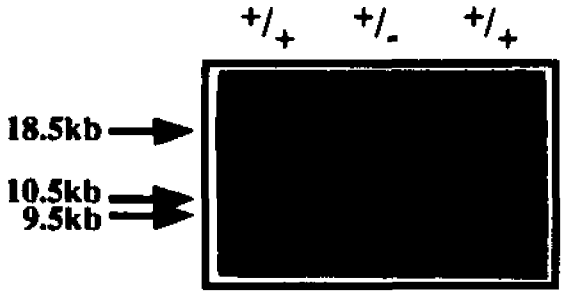


Figure 25.
Southern blot hybridization of DNA from chimeric offspring in the "inbred" (129/Sv/ter) background, with probe A.

Effect of the pNA25 mutation at the gene and protein levels:

In order to examine the effect of the mutation at both the gene and protein levels, several experiments were performed. Total RNA was isolated from the tails of all three groups of animals and RT-PCR was performed. Exon 4/5 and 7 specific oligonucleotides (spanning the deleted exon) were used for the amplification. The amplification product from the homozygous pNA25 mouse RNA was subsequently subcloned and sequenced.

Initial examination of the amplified fragments from the three animals revealed bands of the expected number, and migrations of the expected sizes, when run on a 2% agarose gel (Fig. 26A). The single wild type band (lane 1) was in the range of 283bp (size predicted by the sequence). The single homozygous band (lane 3) was in the order of

205bp, consistent with the deletion of the 26aa encoded by exon 6. The heterozygous sample contained both species (lane 2). Subsequent sequencing of the homozygous fragment revealed correct splicing of exons 4/5 to exon 7 (Fig. 26B, left), as compared to the wild type sequence (Fig.26B, right).

In order to exclude the possibility that the mutation decreases the transcriptional level of the target gene, RNA dot blot hybridizations were performed. Filters with identical amounts of total RNA were hybridized to a *col5a2* specific probe (Fig. 26C, left), and an actin specific probe as a control (Fig.26C, right). No significant changes are observed between the hybridization signal for the wild type (Fig. 26C, left, dot 1), heterozygous (dot 2), or homozygous (dot 3) mouse total RNA. Dot 4 contains 7.5µg of day 10 *p.c.* total mouse RNA (used as an additional positive control). The identical filter on the right (Fig.26C) is hybridized with the actin probe; the small variation in the signals reflects the different amounts of RNA loaded onto the filters, and of course applies to the filter on the left (*col5a2* probe).

Keratocyte lines were established and maintained from all three groups of animals (by Dr. S. Antohi, Dept. of Opthalmology). Both, the anti-V and anti-I collagen antibodies were used to detect their respective trimers (Fig.27) (by Dr. S. Masur, Dept. of Opthalmology). The monoclonal anti-I collagen antibody showed sharp, discrete localization of epitopes (Fig. 27A). Although the polyclonal anti-V antibody did not appear to stain as discretely in most cells, specificity for recognition was evident (Fig. 27B). In addition, staining of the wild type cells revealed that type V was recognized in cells that type I did not (data not shown). In all of the homozygous cells examined, type-V epitopes co-localized to the pockets where type-I epitopes were also recognized. Since the anti-V antibody recognized nascent trimers, binding to epitopes in the pNA25 keratocytes did not conclusively prove the existence of an α2(V) chain. Unfortunately, when ^{3}H-Proline labelled collagens were purified from supernatant of the primary cell cultures, immunoprecipitation for both types I and V was unsuccessful. This was probably due to

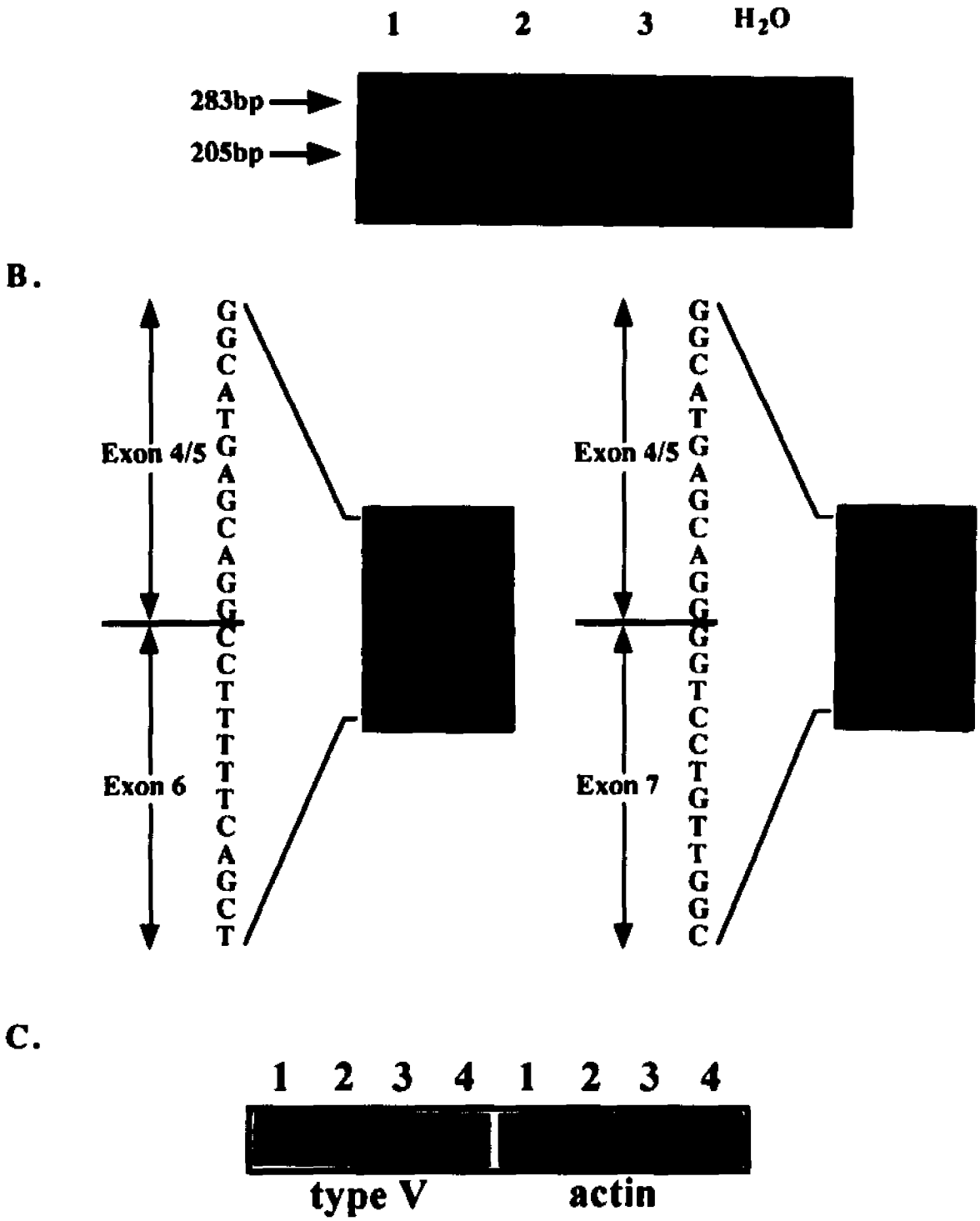


Figure 26.
Effect of the pNA25 mutation at the gene level.
Panel A shows an ethidium bromide stained gel of the RT-PCR products from the three groups of mice. Panel B shows the wild type (left) and homozygous (right) sequences, indicating the missing exon 6 in the pNA25 homozygous transcript. Panel C shows a total RNA dot blot from all three groups of mice hybridized to a *col5a2* specific probe (left), and an actin specific probe (right).

the low amounts of proteins present. Until larger amounts are purified, or an α2(V) specific antibody is available for testing, no direct evidence can confirm exclusive presence of pNA25 molecules in the homozygous mice.

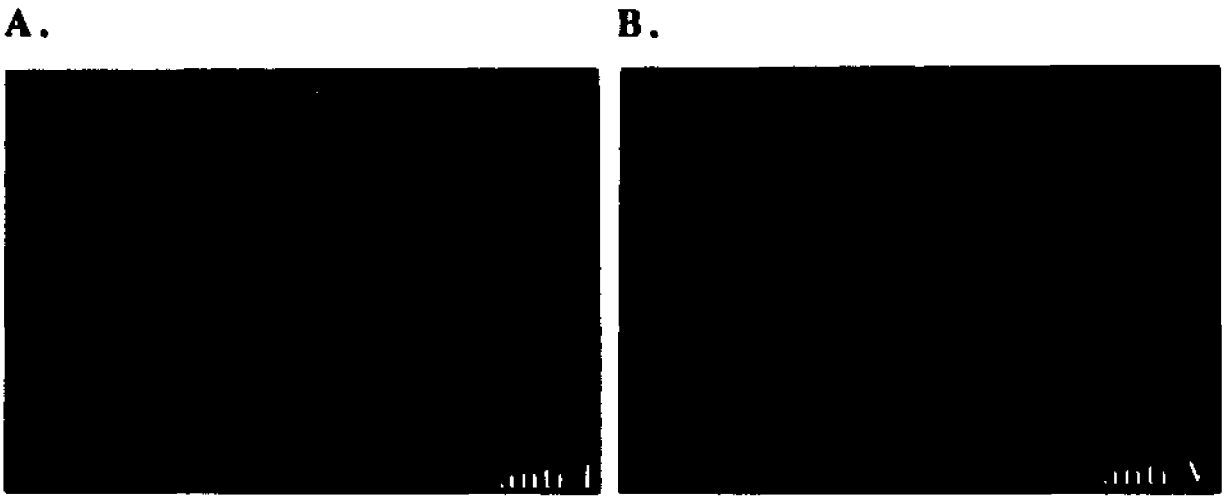


Figure 27.
Fluorescent micrographs of the same pNA25 homozygous keratocyte, stained with the anti-I collagen (panel A) and anti-V collagen (panel B) antibodies after treatment with Brefeldin.

Altogether, the results of these experiments demonstrated correct targeting of exon 6 with no effect on the level of gene expression. As a result, homozygous animals are predicted to produce only shortened pro-α2(V) chains with an internal in-frame deletion of 26 amino acids. Based on structural homologies to other fibrillar procollagen chains, the deleted peptide should contain a potential N-proteinase cleavage site and one of the putative lysine cross-linking sites.

Phenotypic analyses:

Heterozygous mice appeared normal, and were therefore intercrossed to bring the mutation to homozygosity. Animals in the "outbred" backgrounds (C57BL/6 or BALB/C)) were used first, since 129/Sv/ters do not breed at such high frequencies. Litters from such crosses seemed normal at first. However, after the 6-7th day *p.p.* and as they grew older, animals of the same litter did not seem to be growing at the same rate. Upon weaning (3 weeks *p.p.*), the smaller animals were found to be between 1/2 to a 1/3 of the size and weight of their "normal" littermates (Fig. 28).

Figure 28.
A picture of three F1 littermates from a heterozygous x heterozygous cross.

Upon genotyping, the smaller mice were confirmed to be homozygous for the mutation, in that their DNA yielded only a 10.5kb hybridization band (Fig.29). This analysis revealed the mouse shown on the left in Fig. 28 is a heterozygous, while the other two are homozygous. Wild type and heterozygous littermates were indistinguishable by appearance. Interestingly, by the third or fourth month of age the homozygous mice had

grown to a size and weight similar to those of their littermates. Six weeks *p.p.* seemed the critical point of survival. A large number of smaller animals were cannibalized by their mothers before reaching this point. It is likely that these animals died before being cannibalized by their mothers. This was probably due to their overall weakness and poor posture which prevented them from competing for food with their much bigger and stronger littermates. The difference in size between the groups was obvious. The closed eyelids in the more severely retarded of the homozygotes were also noted. Inability to maintain posture, and problems with their ambulation were also part of the homozygous phenotype.

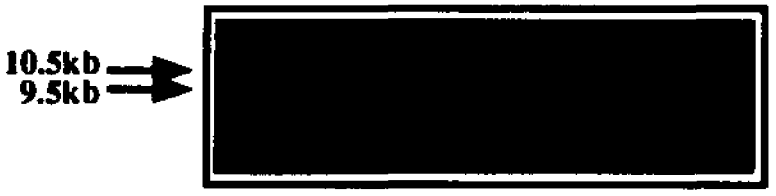


Figure 29.
An example of a Southern blot hybridization of tail DNA from F1s of the heterozygous x heterozygous cross.

Upon closer physical examination homozygotes seemed to have an unusually protuberant sternum. High penetrance X-raying was therefore performed. Severe kyphosis affecting the upper lumbar vertebral area was eminent (Fig. 30).

Defects in bone formation and mineralization were therefore suspected. To examine these possibilities, analysis of the mineralization fronts after intravital double tetracycline labeling was performed, using fluorescence microscopy. Although labels were clearly visible on most surfaces of the femurs examined, a representative region of the periosteal envelope was chosen for observation (Fig. 31). Consecutive labels were distinct and sharply linear, allowing for precise assessment of the distance between the two labels.

Unfortunately the heterozygous animal died 4 days after the first injection, and thus was discarded.

Nevertheless, the results clearly documented that the homozygous animal has normal bone formation and mineralization, as determined by histomorphometric analysis performed on von Kossa stained bone sections (data not shown). On the other hand, intravital double tetracycline labeling showed that the homozygous animals display a significantly slower mineralization rate, and thus a reduced rate of bone growth.

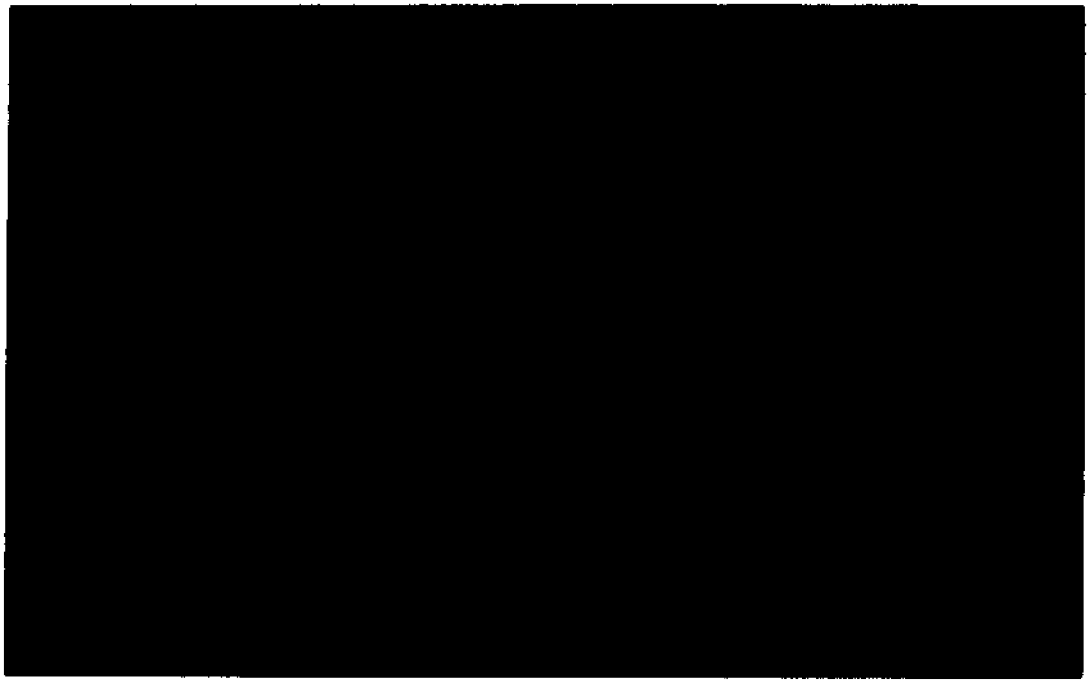


Figure 30.
Severe kyphosis, evident by the X-rays taken of the two 3-week old homozygous animals (middle and right), as compared to their heterozygote (left) and wild type (not shown) littermates.

A.

B.

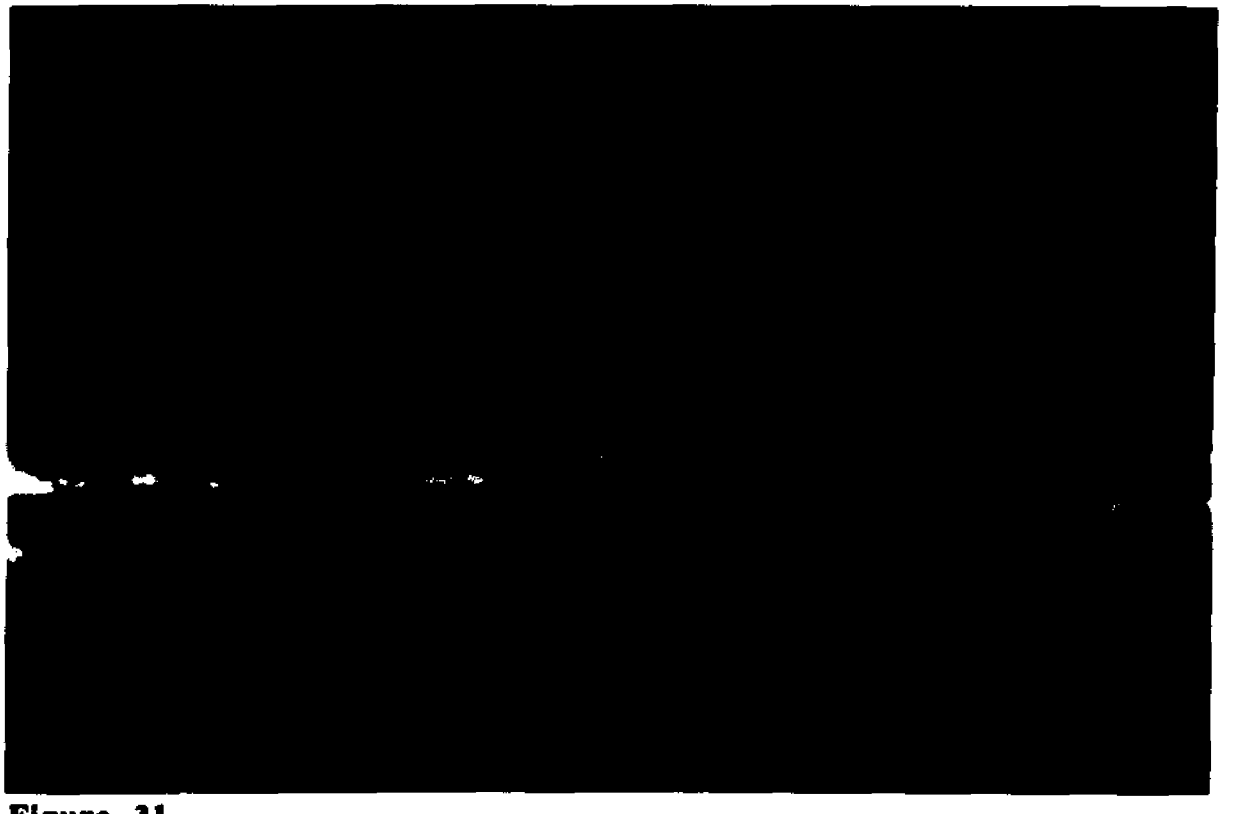


Figure 31.
Panel A shows a fluorescent micrograph of the two tetracycline labeled mineralization fronts in a wild type animal. Panel B those of a homozygous littermate (Both panels at 320x magnification).

Skin

i) Biomechanical properties:

Upon sacrifice of the animals, another aspect of the homozygous phenotype became apparent. Notable extreme skin fragility. An extensive analysis of this organ was therefore performed at multiple levels.

Although the number of animals available for testing was relatively small (two of each group: wild type, heterozygous, homozygous), the results were consistent within each group, and dramatically different between the homozygote group and the other two. The results are shown in table 3, and are represented in a graphical form in figure 32.

Table 3. Results from the extensiometry studies performed on the skins of wild type, heterozygous and homozygous mice.

Wild type		Heterozygous		Homozygous	
Force in N	*Displacement in mm*	*Force in N*	*Displacement in mm*	*Force in N*	*Displacement in mm*
0.00	0.00	0.00	0.00	0.00	0.00
0.47	0.30	0.98	0.62	1.14	0.31
1.39	1.47	2.29	2.48	2.12	1.24
3.40	2.05	2.95	3.72	3.26	3.72
4.95	2.34	3.93	4.66	3.43	5.90
6.34	3.22	4.26	4.97	3.55	8.07
7.73	3.51	5.90	5.19	3.75	12.72
9.42	4.68	7.54	5.87	3.95	14.27
11.59	4.97	8.85	6.29		
13.90	5.85	11.31	6.52		
15.45	6.72	13.44	6.63		
		14.10	6.83		
		18.36	6.93		
		20.00	7.14		

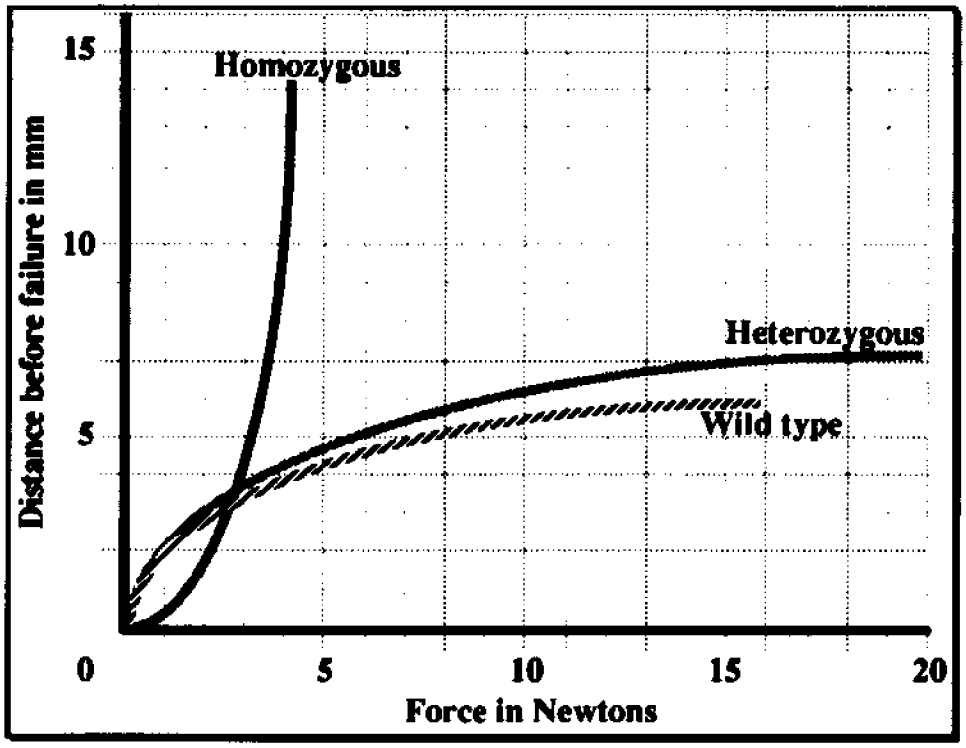


Figure 32.
Graphical representation of the load applied (N) vs. displacement (mm), of skin samples until breakage.

Although wild type and heterozygote animal skin samples failed with a force difference of 4.5N they both failed at a displacement of ~7mm (see overlapping stress curves and their slopes). The individual failure forces were at 15.45N for the wild type, and 20.00N for the heterozygous. In contrast, the different properties of the homozygote skin were immediately apparent. It failed under a force of only ~3.95N (1/4 to 1/5 that of the wild type and heterozygous groups). It also proved to be more elastic (stretchable) since it was observed to fail at a displacement of 14.27mm, nearly twice that of the wild type and heterozygous groups.

ii) Histological analysis:

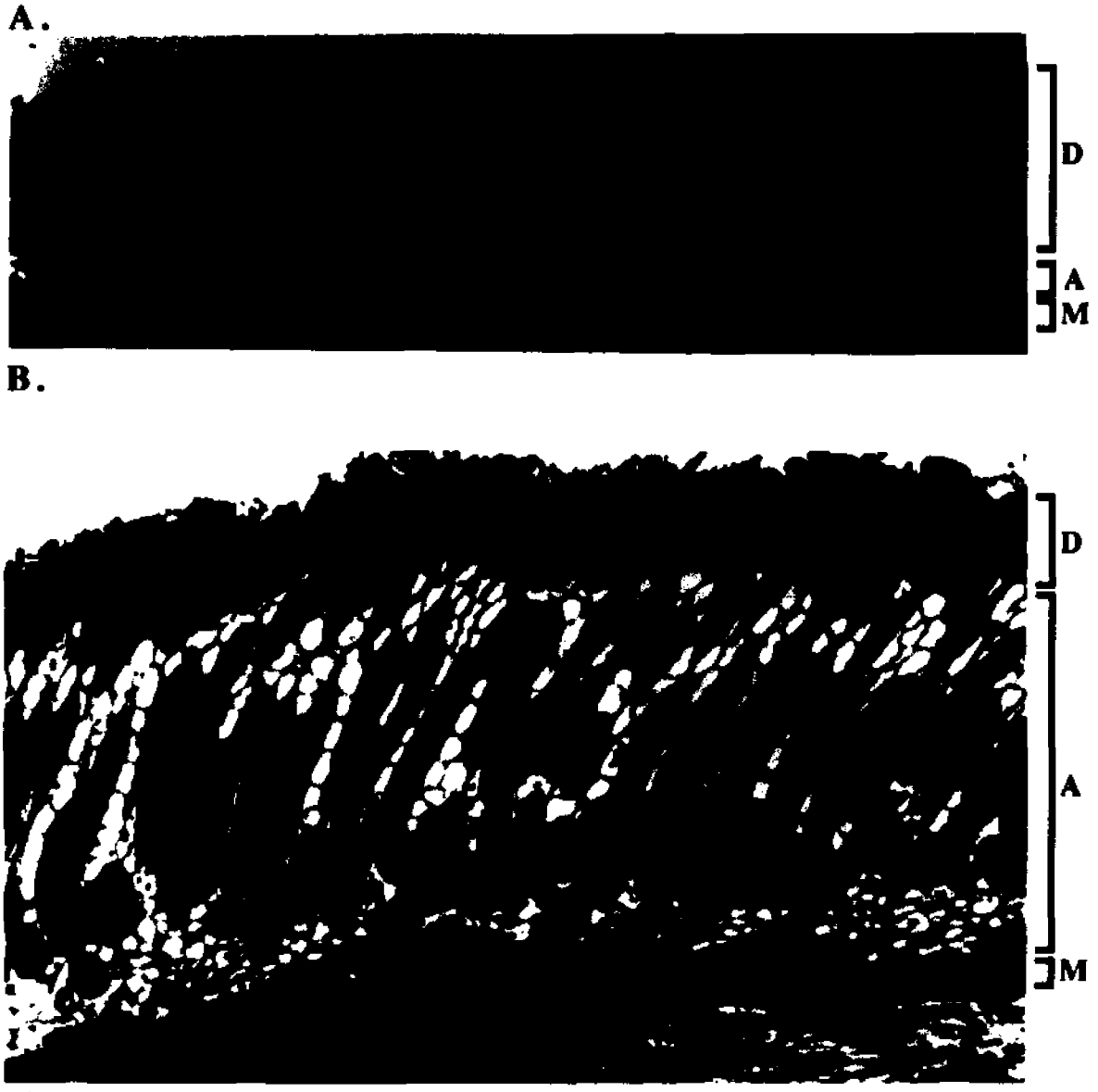


Figure 33.
Panel A shows a Mallory's tetrachrome stained skin section of a wild type mouse skin at 3.125x magnification. Panel B below shows a similar skin section of a homozygous pNA25 mouse littermate at an identical magnification. Dermal (D), adipose (A) and muscle (M) layers are indicated.

Examination of a section through the skin of a wild type animal shows the dense, dermal layer staining blue (Fig. 33A). This layer consists mostly of collagen type I fibrils, and less of type III and V fibrils. Immediately above it is the epidermis which consists of a layer of keratinocytes and the stratum corneum. The bulbs of the hair follicles (stained orange-red) are found embedded in the dermis. Immediately underneath the dermis is a fat layer composed of adipocytes followed by a layer of muscle (stained gray).

Examination of a similar section through the skin of a pNA25 homozygous animal revealed some striking differences (Fig. 33B). First, the dermal layer was approximately half in thickness of the wild type one. Second, the adipose layer is on the average 4-fold thicker than the control. Third, hair follicles seemed to be in their anagen (growth) phase, and mostly in the adipose layer; consequently the skin shows twice as thick as the one of the wild type. In contrast, the sections of the heterozygotes appeared to be identical to those of the wild type animals (data not shown).

iii) Ultrastructural analysis:

To examine further the skin abnormality, appropriate sections were subjected to electron microscopic analysis. Fig. 34 shows a skin section through a wild type animal's stratum corneum, keratinocyte epithelium, and apical part of the dermis. The longitudinal and cross- sections through the type I collagen fibrils reveal their characteristic architecture in tightly compacted bundles. In contrast, examination of a similar section through the skin of a pNA25 homozygous animal at an identical magnification showed a completely different organization (Fig. 35). The compactness and bundling of the collagen type I fibrils are no longer present, and in their place complete disarray and disorganization of the fibrillar network are evident. Furthermore, empty spaces are noted in the ECM probably filled with other structural components (such as proteoglycans).

When the fibrils were examined under higher magnification, no major differences were observed between the wild type and homozygote samples (Fig. 36). The periodic D-spacing at 67nm, characteristic of the fibrillar collagens was not affected. However, some variation between the individual fibril diameters was evident. Instead, there was no indication of the distinct "cauliflower" shape of the fibrils found in EDS VIIC (the dermatosparactic-like) phenotype. Once again, wild type and heterozygote animals showed no differences.

Figure 34.

Electron micrograph of a section through the stratum corneum (SC), epidermal (E) and dermal (D) layers of a wild type mouse skin at 9,500x magnification. Cross- and longitudinal- sections of the type I collagen fibrils, as well as their organization, are also depicted.

Figure 35.
Electron micrograph through a similar section in the skin of a homozygous pNA25 mouse littermate at the identical magnification of 9,500x. The disorganization of the type I collagen fibrils is striking.

A.

B.

Figure 36.
Higher magnification (27,000x) of the type I fibrils in wild type (panel A) and pNA25 homozygous (panel B) skin.

Eye

i) Histological analysis:

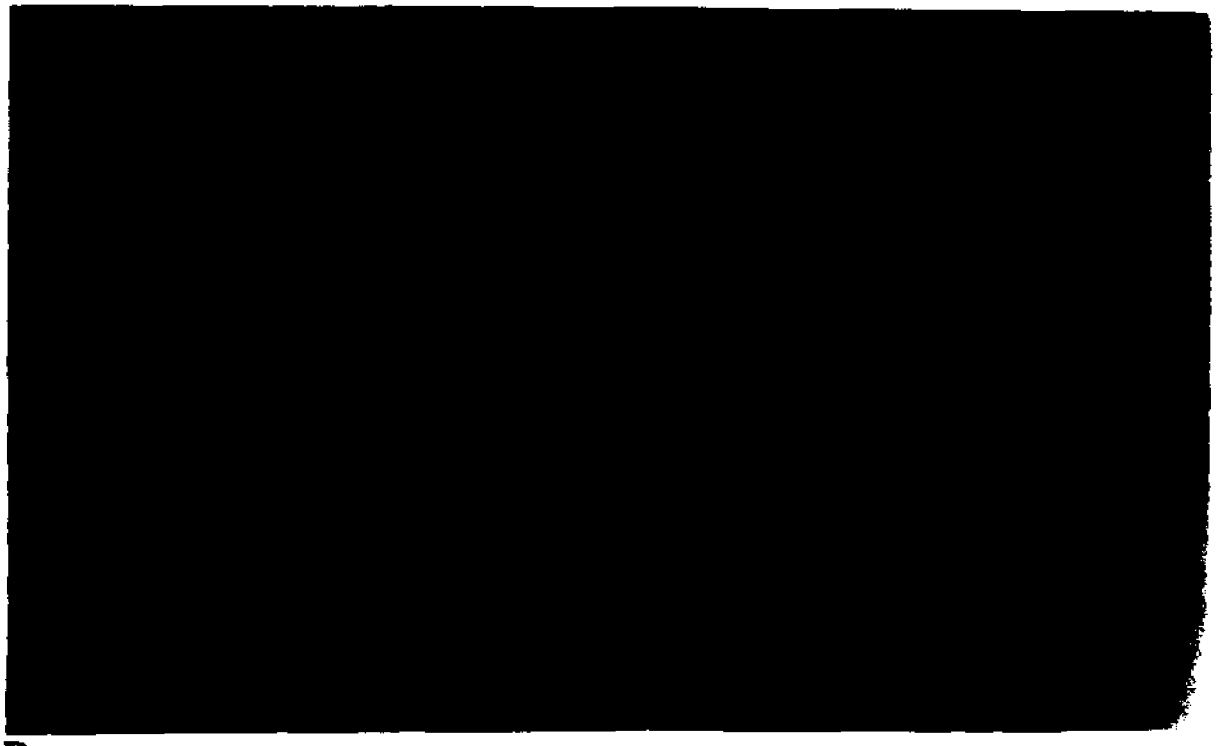
A section through the eye of a wild type animal showed the normal architecture of the cornea (Fig. 37A). The corneal epithelium is stained light brown-gray, and the corneal stroma, which is composed mainly of collagen type I and type V fibrils, is light blue. The lens (orange-red), the iris (black), and the retina behind the lens are also shown.

Examination of a similar section through the eye of a pNA25 homozygous animal revealed differences in the corneal architecture (Fig. 37B). Although the epithelium appears of the same thickness as the wild type, the corneal stroma is much thinner. Its collapsed architecture is more reminiscent of the sclera. The heterozygote eyes appeared identical to those of the wild type littermates. These differences were more accentuated upon electron microscopical examination.

ii) Ultrastructural analysis:

At lower magnifications (55,000x), the wild type and heterozygous corneal stromas appeared identical in structural organization (Fig. 38). Cross-section through the tissue revealed the regularity in the spacing, and three-dimensional arrangement of the fibrils. A similar section through the corneal stroma of a pNA25 homozygous animal showed a very different picture at identical magnification (Fig. 39). There is in fact a substructural collapse of the fibrillar network, the fibrils are highly compacted with no apparent orientation, and individual fibrils are somewhat thicker than the wild type ones. At higher magnification however, the differences in fibril diameter were obvious in the tissues of all three groups of animals (Fig. 40). Wild type fibrils are the narrowest and homozygous the widest, with heterozygous in between the two of them (Fig. 40). Assuming that the D-band spacing is in the order of 67nm, the diameters of the three groups of fibrils are about ~25nm, ~35nm and ~50nm for the wild type, heterozygous and homozygous mice, respectively.

A.



B.

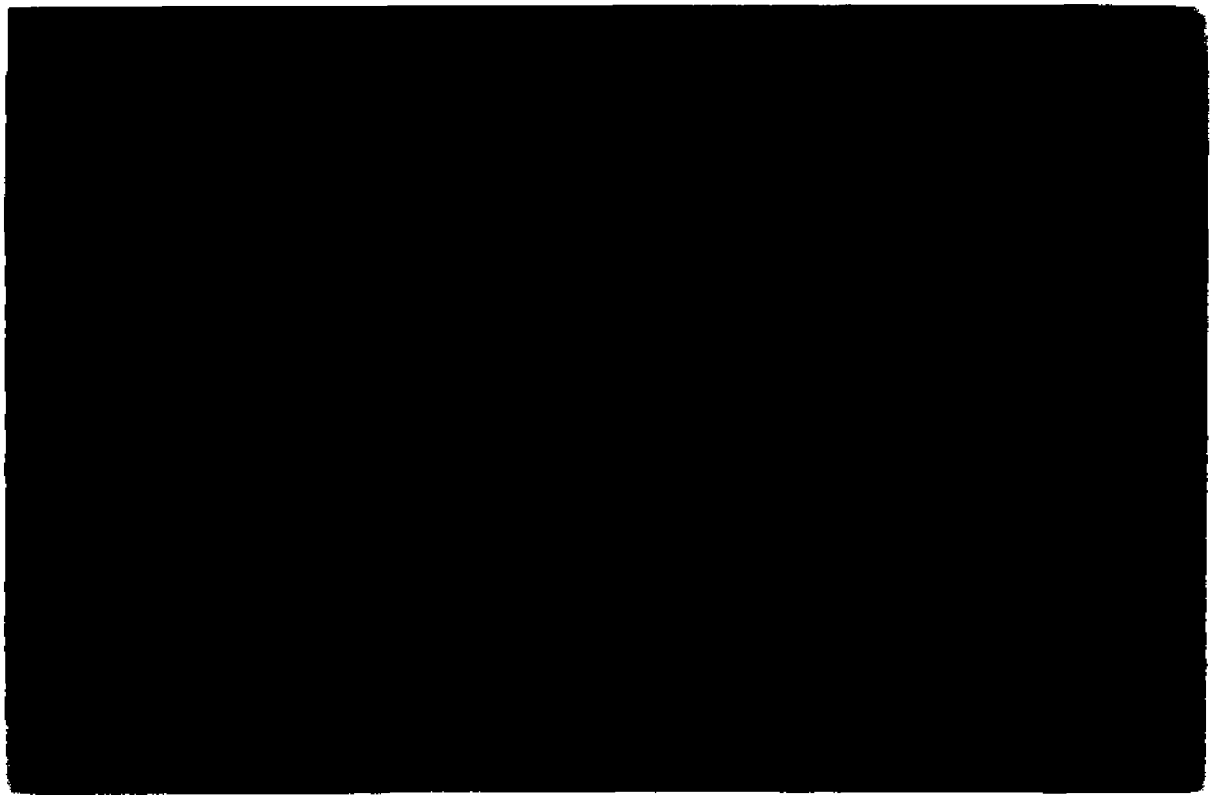


Figure 37.
Light micrographs of wild type (panel A) and pNA25 homozygous (panel B) eye sections.

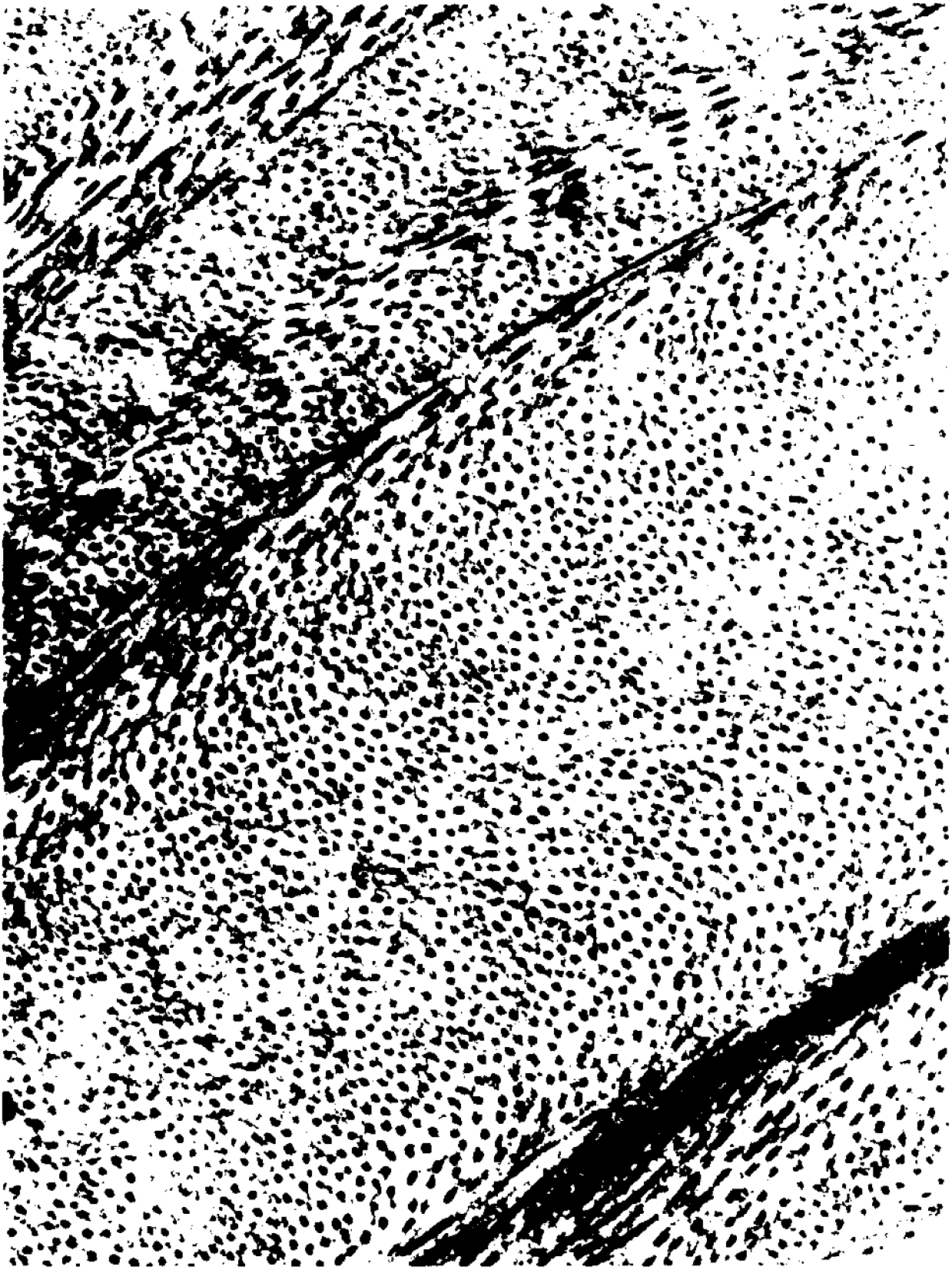


Figure 38.
Electron micrograph of a section through the corneal stroma of a wild type mouse eye at 55,000x magnification. Cross- and longitudinal sections of the type I collagen fibrils, as well as their regular spacing, are also depicted.

Figure 39.
Electron micrograph through a similar section in the corneal stroma of a homozygous pNA25 mouse littermate at 55,000x magnification. The disorganization of the fibrillar architecture is striking.

Figure 40.
Electron micrographs through similar sections in the corneal stroma of wild type (panel A), heterozygous (panel B), and homozygous pNA25 (panel C) mouse littermates at 120,000x magnification. The increasing fibril diameter is evident.

Bone

Light microscopic examination did not reveal any differences in the bone tissues of the three groups of animals, apart from "gross" size differences possibly due to the delay in growth associated with the homozygous animals. However, transmission electron micrographs, revealed differences in fibril organization between the wild type and homozygous tissues (Fig. 41). Cross sections through the wild type animal femurs showed that the fibrils are uniformly dispersed in a single orientation (Fig. 41A). In contrast, and only in the homozygous animals, a small percentage of the fibrils exhibits a different pattern. They are at a perpendicular orientation to the rest, an uncommon finding in healthy tissues (arrowed, Fig. 41B).

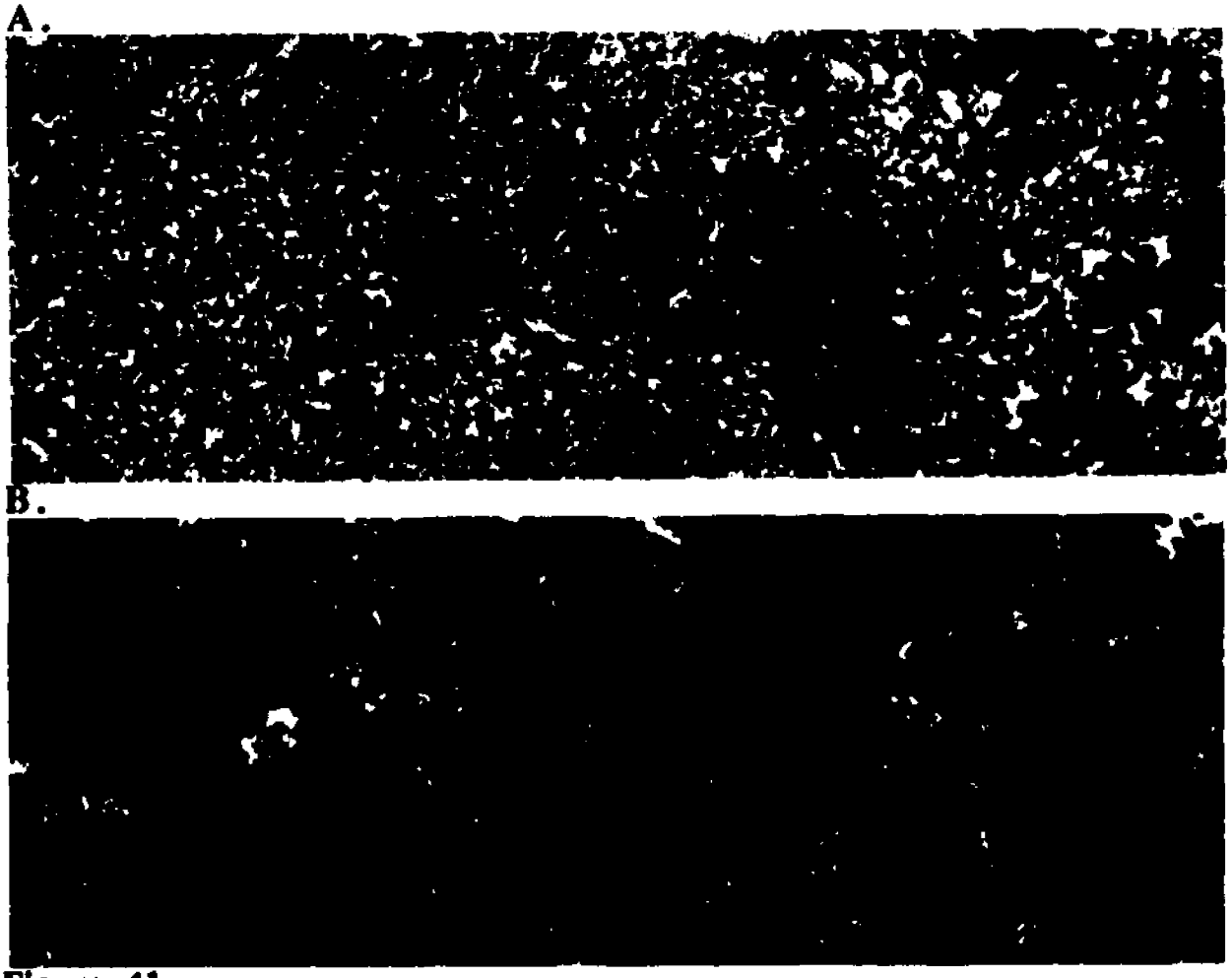


Figure 41.
Electron micrographs through cross sections of decalcified femur in wild type (panel A), and homozygous pNA25 (panel B) mouse littermates at 32,400x magnification. Regions of fibril disorganization in panel B are arrowed.

Perichondrium

Close examination of the wild type animal's perichondrium by electron microscopy demonstrated the integrity and compactness of the fibrillar network (Fig. 42). Once again, cross-section of the type I fibrils revealed their tight packing, with individual fibrils in uniform diameter (high magn. micrograph not shown). Additionally, fibroblasts in the vicinity had a "slender" shape, indicative of the network's architectural integrity around them.

A similar section through the perichondrium of a pNA25 homozygous animal was examined at the identical magnifications (Figs. 43 and 44). Great variation in the diameters of the individual fibrils was observed. There is an apparent looseness of the tissue, and the fibroblasts appear much wider and with enlarged cytoplasms. Moreover, the overall thickness of the tissue is twice that of the wild type littermates.

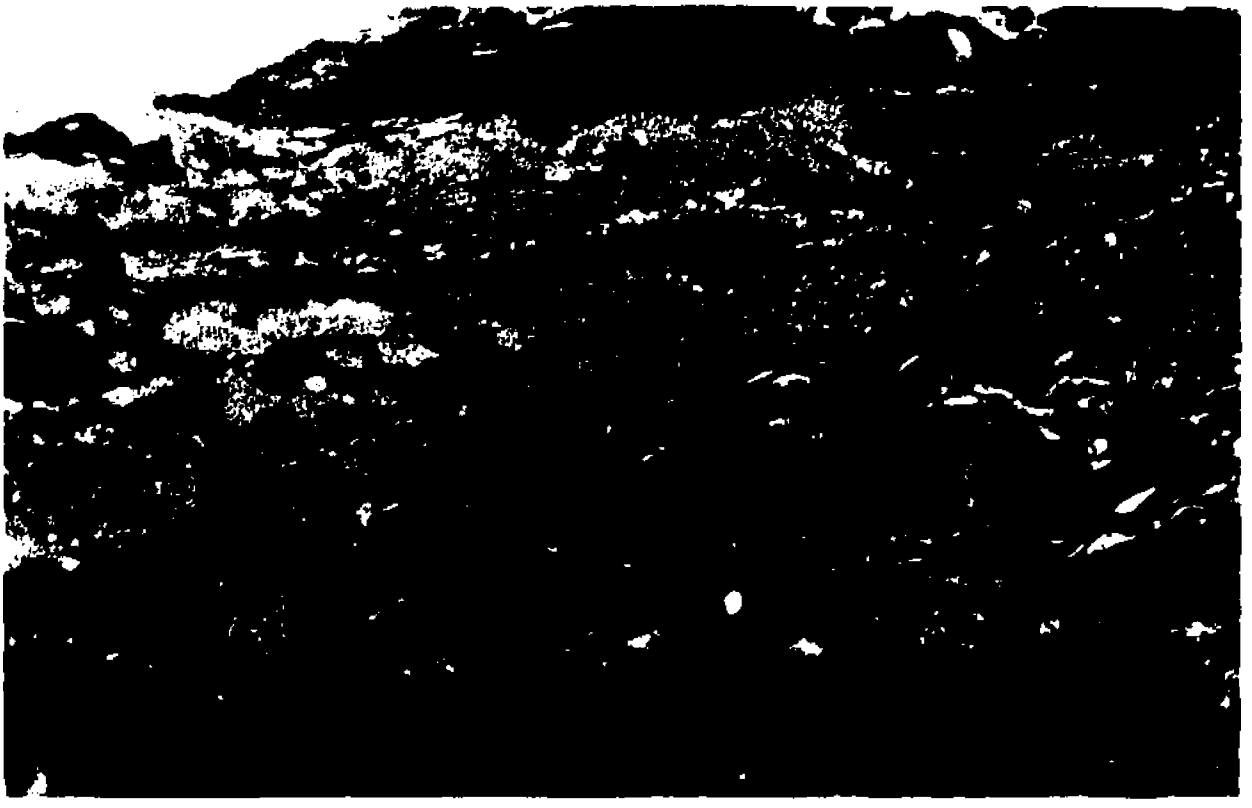


Figure 42.
Electron micrograph of a cross section through the perichondrium in the femur of a wild type mouse at 13,500x magnification.

Figure 43.
Electron micrograph of a cross section through the perichondrium in the femur of a homozygous pNA25 mouse, at 13,500x magnification. The overall increase in thickness as compared to that of the wild type (previous figure), is evident.

Figure 44.
Detail of the previous electron micrograph showing an enlarged pNA25 fibroblast and its surrounding fibrillar network, at 27,000x magnification.

Additional phenotypic observations:

The number of mice on the C57BL/6 and BALB/C genetic backgrounds increased with time and enabled us to make a series of additional observations. It should be noted that a statistically higher number of animals is required to verify these observations.

Table 4. Numbers of animals obtained.

a.

	Wild type	Heterozygous	Homozygous	Total
Observed	60	182	16	258
Male/Female	33/27	97/85	12/4	258

b.

	Litter	Wild type	Heterozygous	Homozygous	Total
Observed	3	6	12	9	27

c.

	Litter	Wild type	Heterozygous	Homozygous	Total
Observed	3	4	10	1	15

The very small number of homozygous animals (16 of a total 258), suggested a nearly 70% lethality rate when compared to the wild type (60/258) and heterozygous (182/258) number of animals (Table 4a). This data reflects the numbers of surviving animals up to weaning age (upon their genotyping). To determine the time of death, embryos were obtained by cesarean section at day 18.5 of gestation. Although the total number of offspring is statistically small, the ratios show a compliance with the expected mendelian ratios, and suggest that the animals are dying after birth (Table 4b). An interesting observation was made with the first offspring from the heterozygous crosses of mice in the 129/Sv/ter genetic background ("inbred"). There was only a single surviving homozygote from the three litters (Table 4c). This mouse began to show unhealthy signs, such as difficulty in breathing, 4 days before its death (at day 27 *p.p.*) while it was still under observation.

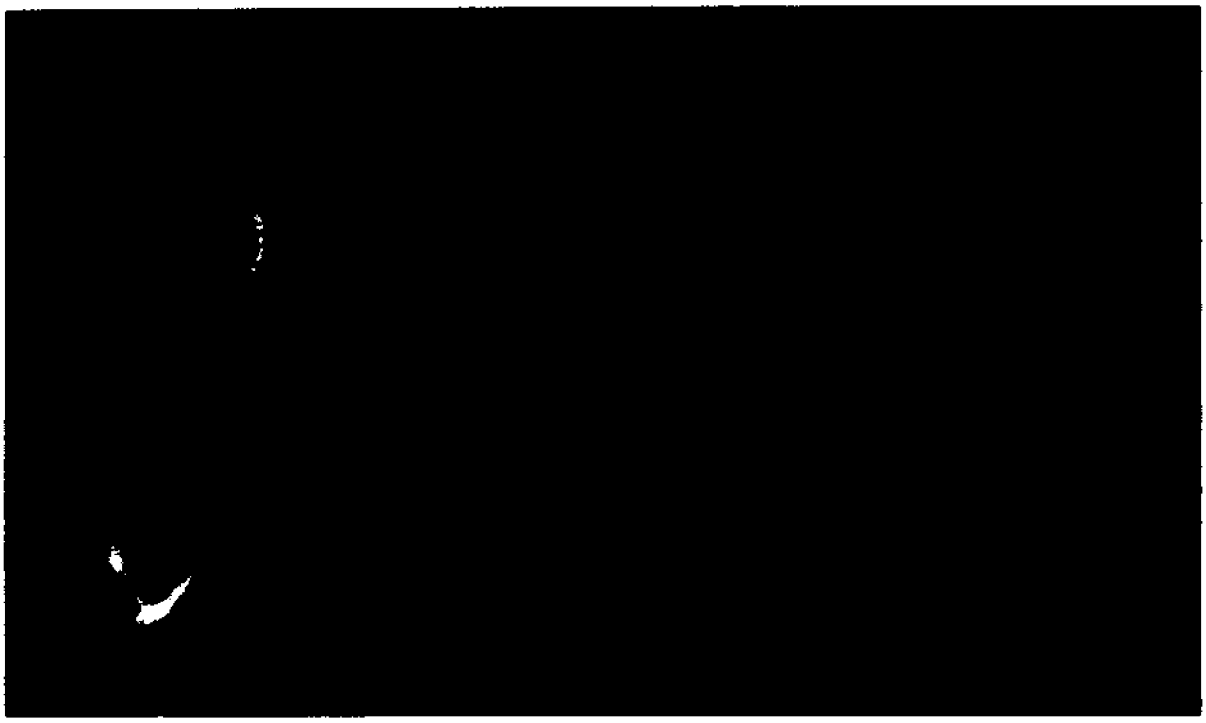


Figure 45.
X-ray of the heterozygous mother (left), a heterozygous offspring (middle) and the only homozygous pNA25 (right) mouse littermate in the 129/Sv/ter genetic background. The severe scolio-kyphosis is evident.

A radiogram taken after its death shows an extreme scolio-kyphotic phenotype (Fig. 45). In addition, both the small size and skin fragility were comparable to those previously noted in the "outbred" genetic backgrounds. The older homozygote is a male, now 8 months old, and has given offspring. Unfortunately, no homozygous female has survived. It will be of interest to obtain more homozygous females in order to observe their behavior during gestation (high expression of type V in the placenta).

Summary:

In conclusion, the study presented here reports on several experiments aimed at understanding the pathophysiology of type V collagen. The results of the *in situ* analyses were consistent with the notion that type V and type I collagens may interact in the formation of the complex fibrillar networks in a wide variety of non-cartilaginous matrices. The gene targeting data revealed that removal of a single exon of *col5a2* has a major impact

in organismal integrity and survival. Homozygous mice present a severe kyphotic phenotype, probably due to defective ligament attachment sites. In addition they display skin fragility conceivably caused by the inability of the type I fibrils to form properly packed bundles, a pre-requisite essential for the maintenance of the tissue's structural properties. Disorganized and thicker type I fibrils are probably responsible for the disorganization of the corneal stroma in the transgenic mice. Altogether, the data provided the first direct evidence for a role of type V collagen in regulating the tri-dimensional assembly of type I collagen fibrils.

Discussion

This report represents the first comprehensive study of the structure, function and expression of type V collagen in a mammalian system. The premises of this work was based on established structural similarities of the fibrillar collagen genes, and on functional inferences from the collagenopathies. Studies of types I-III collagen mutations have in fact revealed how deregulation of distinct steps in collagen biosynthesis (triple helical assembly, secretion and enzymatic processing) gives rise to very different phenotypes. We reasoned that these general principles could be applied to, and thus tested for fibrillar collagen types of unknown function. Hence, this work was undertaken to assign a role to type V collagen in the formation and maintenance of the tri-dimensional aggregates of the ECM.

Early *in vitro* work suggested that type V collagen can form pure fibrils by itself, or can direct the growth of type I and type III fibrils when associated to them (Adachi *et al.*, 1989). Subsequent studies in the chicken embryo showed that the striated collagen fibrils of the corneal stroma are heterotypic structures composed of both type V and type I collagen molecules (Birk *et al.*, 1990). Determination of this co-localization required pre-treatment of the tissue (enzymatic, mildly acidic, or by temperature) to partially disrupt fibril structure, and therefore "unmask" the type V epitopes. Later studies demonstrated the co-polymerization of the two in an *in vitro* self-assembly assay (Birk *et al.*, 1990). These heterotypic fibrils had a sub-structure similar to that observed *in vivo*. Increasing amounts of type V in the reaction mixture resulted in the progressive decrease of the mean fibril diameter. Parallel studies showed co-localization of collagen types I and III; types II and IX; and types II with XI (Keene *et al.*, 1987; Vaughan *et al.*, 1988; Mendler *et al.*, 1989). As a result, many investigators began to realize that heterotypic collagen assembly may

represent an additional regulatory level in fibrillogenesis (Keene *et al.*, 1987; Birk *et al.*, 1988; Fleischmajer *et al.*, 1990).

Very recently, Linsenmayer *et al.* (1993) proposed a model in order to explain how the physical relationships of the components of the heterotypic fibrils influence the ultrastructural assembly of these aggregates. They first demonstrated that type V and type I molecules are likely to be arranged parallel to one another, with the type V being present in the core of this fibril. Furthermore, upon rotary shadowing they noted the presence of a kink at the N-terminal end of type V molecules purified from the corneal fibrils. They further postulated that the flexibility for the kink formation might be provided by the partial processing of type V procollagen trimers.

Type V collagen maturation is unfortunately a long-standing and still unresolved controversy sustained by the low representation and compositional heterogeneity of the protein, and by differences of the *in vivo* and *in vitro* means of analysis. Although subsequent cloning data revealed the presence of canonical propeptide cleavage sites, early biochemical analyses had suggested that pro-α2(V) is partially processed (Fessler *et al.*, 1981; Myers *et al.*, 1985; Woodbury *et al.*, 1989). These studies however, disagreed on whether the mature α2(V) chain was in the pC or the pN form. Very recent evidence seems to support the latter hypothesis (Moradi-Améli *et al.*, 1994). However, work on a rhabdomyosarcoma line producing $[\alpha1(XI)]_2\alpha2(V)$ heterotrimers has shown that the pro-α2(V) chain undergoes full processing (Kleman *et al.*, 1992). Likewise, there has always been general agreement that the natural partner of the α2(V) collagen, α1(V), contains a partially processed N-propeptide (Kumamoto and Fessler, 1980; Fessler *et al.*, 1981). In this case, however, the controversy still centers around the size of the retained globular extension of α1(V) (Broek *et al.*, 1985; Fessler and Fessler, 1987). One cannot in fact apply the fibrillar processing rules to the pro-α1(V) chain, because of the unique structure of its N-propeptide (Greenspan *et al.*, 1991). Incidentally, the same structure characterizes the pro-α1(XI) chain, the molecule that can often substitute α1(V) as a partner of α2(V)

(Yoshioka and Ramirez, 1990). Interestingly, recent evidence indicates that a small amount of α1(V) chains is fully processed in bone and umbilical cord tissues. This fully processed form of type V collagen is believed to be made of three identical pro-α1(V) chains (Moradi-Améli *et al.*, 1994). It is therefore conceivable that different forms of the type V trimer and of the type V/XI cross-trimers determine the structure and the extent of subunit processing. This might in turn dictate the diversity of the regulatory signals imparted by these macromolecules on the growing fibrillar bundles of type I collagen. One such signal is the one that Linsenmayer *et al.* (1993) proposed to be mediated by the type V collagen kink. They argued that this structure may be the result of the full processing of the pro-α2(V) subunit, with the consequential projecting on the surface of the heterotypic fibrils of the partially processed pN-α1(V). They therefore concluded that the kink plays a function in regulating the diameter of the fibrils by sterically preventing the further addition of collagen molecules.

The present study presents supporting evidence for a role of type V collagen in regulating the tri-dimensional architecture of type I collagen fibers. First, the high degree of sequence conservation in the mouse and human exon 6-coded peptides implies a functional role for these sequences. Second, the superimposable pattern of type V and type I collagen gene expression during mouse embryogenesis infers that they cooperate in forming the fibrillar network of non-cartilaginous matrices. Third, the disruption of the tri-dimensional arrangement of type I fibrils and fibers in the pNA25 mice, correlates the integrity of the type V with the fibrillogenesis of type I collagen.

The pNA25 mutation depleted the *col5a2* gene of the exon coding for the putative N-proteinase cleavage sequence, and the N-telopeptide cross-linking site. The exon deletion was engineered by replacing exon 6 with the PGK promoter-driven neo cassette. The novelty of this approach initially raised concerns about the effect of positioning the PGK promoter within the body of the *col5a2* gene. Formally, we had in fact to prove that

the pNA25 allele is correctly spliced, and expressed at the same level as the wild type allele. Deviations from either of these two pre-requisites would in fact bias the interpretation of the transgenic data. Three lines of evidence suggest that the mutation had the expected effect. First, the RT-PCR test correlated the number and size of the *col5a2* transcripts with the genotype of the animals. Second, the sequencing of the PCR products demonstrated that all of the transcripts of the homozygous mice lack precisely the sequence corresponding to exon 6. Since removal of exon 6 does not change the translation frame of the mRNA, one expects the production of a single polypeptide species missing 26 internal residues. Finally, the dot-blot hybridizations excluded the possibility that replacement of exon 6 with the neo gene affects the level of expression of the mutated allele. Although we were not able to extend these data at the protein level, the results are nevertheless consistent with the notion that the pNA25 allele is a structural mutant and not a null allele.

As already mentioned, the transgenic experiments provided for the first time a direct demonstration that type V collagen regulates the three-dimensional organization of type I fibers. The disruption in the assembly of the fibrillar network has severe effects in several but not all of the tissues known to express types I and V collagen. Both the integrity and biomechanical properties of skin are changed, resulting in its extreme fragility. The mutation has also a curious and unexpected effect on the behavior of some of the cells of the outer dermal layer. We are obviously referring to the increased number of hair follicles, and their ectopic location in the adipose layer. The cause and significance of this finding is obscure. It is however tempting to speculate that this effect may reflect a change in cellular programming because of the "reading" of a different ECM by the hair-follicle cells. Alternatively, the phenomenon may simply be caused by the loss of pure type V fibrils at the basement membrane barrier. This loss could cause the hair-follicle cells to sink into the lower and nutritionally richer cellular stratum, which may in turn increase their proliferation rate.

Another organ affected by the pNA25 mutation is the eye. The normal corneal stroma is a highly organized connective tissue in which optical transparency depends on the rigid control of fibril diameter and the precise arrangement of the collagen fibrils (Cox *et al.*, 1970). The collapse of the stroma, as well as the significant changes in fibril diameters are evident in the homozygous mice. Slit-lamp examination of these mice confirmed changes in the physiological properties of the tissue, for it suggested visual problems closely resembling human myopia. In addition we noted a change in the surface tension of the cornea upon its removal. Unlike the cornea of the wild type mouse, the homozygous cornea becomes more elliptically shaped once outside of the eye. This anecdotal observation is consistent with the abnormal biomechanical properties of the corneal ECM and the animals' postulated myopia.

At first, we thought that the reduced size of the skeleton was caused by abnormal mineralization. This may have been the case since the homozygotes also exhibited severe kyphosis. However, more rigorous analyses excluded this possibility and suggested a delay in bone growth. Since one of the highest level of *col5a2* expression is in the perichondrium, it is conceivable to argue that disruption of fibrillogenesis at this site affects somehow the rate of long-bone growth. The effect is diametrically opposite to that caused by fibrillin mutations in the Marfan syndrome (Ramirez *et al.*, 1993). In this case, the matrix microfibrils of the perichondrium are structurally abnormal, and the rate of long bone growth is increased. The perichondrium is probably responsible for the severe kyphosis too, since it is there that the sustaining ligaments attach. We envision that the distortion of the vertebral column is not an intrinsic deficiency of the vertebral bodies, but the indirect effect of loss of tensile strength in the supporting ligaments. Incidentally, the distortion of the column and the consequential crashing of the abdominal cavity is one of the major causes of death in the transgenic animals.

Both in the case of the skin and cornea, one can offer further speculations based on the apparently earlier onset of type V collagen expression compared to type I collagen (day

4 of mouse embryogenesis vs. day 5 respectively; Adamson, 1982) . Accordingly, it is conceivable that type V collagen may form a primary three-dimensional matrix upon which type I fibrils are deposited, become oriented and polymerize (nucleation sites). Consequently, structural changes in the pro-α2(V) peptide may result in disruption of this highly organized meshwork. This can occur either by affecting intracellular type V assembly and its extracellular processing, or by interfering with type V collagen interactions with other ECM components. Although our work did not elucidate the primary consequence of the mutation, it provided the necessary reagents to study this important point in more detail. It also laid the ground to test several other structural-functional hypotheses by comparing the effects of different type V collagen mutations, or by cross-breeding different mouse mutants.

We should also note that several candidate tissues and organs did not show any noticeable abnormality. Based on the pattern of expression, we expected anomalies in the lung, kidney, heart and, most important of all, in the vascular system. There are three possible explanations for the absence of pathology in these and other organs. First, we may have not looked at enough animals, or late enough in adulthood. Second, additional other anomalies together with the severe kyphosis may be responsible for the apparent loss of peri-natal viability of the homozygotes. Third, *col5a2* expression in these organs and tissues may be so critical that other gene products, or alternative molecular forms of type V collagen fulfill its function when absent. Recently, Saga *et al.* (1992) generated a mouse completely lacking tenascin-C (TN-C). Despite the critical role played by this ECM component in neuronal development and morphogenesis, mice homozygous for the null TN-C alleles are normal. The authors argued that the functions of TN-C might be substituted by other proteins co-expressed in the same tissues. The same argument was used to explain the normal *in vitro* morphogenesis of organ explants from the type I collagen-deficient Mov-13 mouse, and more recently the normal development of a type X collagen-deficient mouse (Kratochwil *et al.*, 1986; Rosati *et al.*, 1994).

The last important point that still remains to be clarified is exactly how the pNA25 mutation affects type I fibrillogenesis. Is this caused by the lack of N-propeptide removal, by the loss of the cross-linking site, by combination of both, or simply by the deletion itself causing a change in the overall conformation of the trimer? The loss of the cross-linking site would be in agreement with the loosened packing of the pNA25 fibrils into fibers. However, work on the etiopathogenesis of the human EDS VII has clearly shown that loss of the corresponding site in type I collagen does not lead to a similar effect. All of the reported cases of EDS VII are caused by deletion of type I procollagen sequences encoded by exon 6, with one important exception. Chiodo *et al.* (1992) in fact described a splicing mutation in a proband with EDS VIIB which results in the partial deletion of exon 6 because of the activation of an internal cryptic splice site. This in turn causes the elimination of the N-proteinase site with the retention of the cross-linking site. Although this analogy supports the idea that the mouse phenotype is due to the lack of N-propeptide processing, the hypothesis contradicts recent evidence suggesting that the tissue form of α2(V) is pN-α2(V) (Moradi-Améli *et al.*, 1994).

To reconcile these two points we propose that pro-α2(V) is differentially processed in different tissues; we also argue that pro-α2(V) is fully processed in the tissues which are abnormal in the pNA25 mouse. Thus, we concur with Linsenmayer *et al.* (1993) in arguing that the type V kink is probably due to the full processing of the pro-α2(V) chain, and thus only consists of unprocessed pNα1(V) subunits. Collaborative work with Dr. David Birk is currently examining the pNA25 trimers by rotary shadowing electron microscopy. If our hypothesis is correct, we expect to see all of the rod-like shaped type V molecules to be longer than control because of the lack of the kink. This observation will provide the final proof of how the growth of the type I fibrils and tri-dimensional organization of the type I fibers are regulated during ECM formation.

The expression of the phenotype only in heterozygosity was somewhat of a surprise in view of the expected "dominant/negative" nature of the mutation. Instead, all of the chimeras (including the high percentage ones) reached adulthood, and most of the animals gave germline transmission in all three genetic backgrounds (C57BL/6, BALB/C and 129/Sv/ter). Heterozygous animals for the pNA25 mutation were in good health; they were therefore intercrossed to bring the mutation to homozygosity. At this point in time, we can only speculate on why the heterozygous animals do not display an intermediate pathology between the wild type and homozygous animals. As in the case of the Mov-13 mice, it is possible that we failed to detect abnormal manifestations in the heterozygote. In this respect an analogy could be drawn with the Mov-13 mice.

The Mov-13 strain of mice was generated by a proviral integration into the first intron of the α1(I) collagen gene. Mice homozygous for the null mutation die *in utero* because of abnormalities in their vascular system (Lohler *et al.*, 1984). In contrast, heterozygous Mov-13 mice are apparently asymptomatic. After careful examination these heterozygous mice were however found to exhibit a mild OI phenotype (OI type I) . They in fact present reduced amounts of collagen in non-mineralized tissues, profound and progressive hearing loss, and reduced biomechanical properties of long-bones (Bonadio *et al.*, 1990). Although seemingly normal, heterozygous pNA25 animals did show some subtle differences upon examination at the ultrastructural level. Most notably, they exhibited differences in the diameters of the corneal stroma fibrils. Thus, it is possible that additional abnormalities will become more apparent as the animal tissues are more carefully and more extensively examined at the ultrastructural and biomechanical levels.

Along these lines, one of our initial interests was to explore the possibility that the pNA25 phenotype has a human counterpart. Although the mutation was modelled on the EDS VII paradigm, we have yet to find a human correlate. In retrospect, however, we should have taken into consideration two major points before using this approach to search for a human type V collagenopathy. By analogy to EDS VII, a phenotype caused by this

molecular lesion is in fact expected to be very rare (Beighton, 1993). A more productive approach to identify human correlates would have been to disrupt the intracellular assembly of type V collagen, in a manner similar to the established paradigms of the types I-III collagenopathies (Lee *et al.*, 1991a). Nevertheless, the pNA25 phenotype does present some of the clinical hallmarks of the EDS spectrum of diseases. EDS in fact encompasses a variety of conditions manifesting hypermobility of the joints, malformations of the skeleton and hyperextensibility, thinness, and fragility of the skin. The second point we should have considered is that the human mutation may be lethal in heterozygosity. Mutations generated in mice often behave differently from the human counterparts. For example, Hirschsprung's disease, a common autosomal dominant condition (1 in 5,000 live births), causes intestinal obstruction in neonates and megacolon in infants and adults. It has been recently demonstrated that mutations in the *ret* proto-oncogene are responsible for this condition (Edery *et al.*, 1994; Romeo *et al.*, 1994). Transgenic mice carrying a null *ret* gene exhibit the Hirschsprung phenotype too, but with a number of important differences. They are more severely affected than Hirschsprung patients; they have no intestinal parasympathetic neurons; they have either abnormal or absent kidneys; and they die within 16-24 hr of birth. More interestingly, the transgenic mice with the *ret* mutation lack detectable abnormalities in heterozygosity (Schuchardt *et al.*, 1994).

A final issue to be considered is the phenotypic variability between homozygous animals of the same litter. This is an observation common to most transgenic experiments (Schuchardt *et al.*, 1994). The current hypothesis attributes this variation to epigenetic factors specific for the different genetic backgrounds involved. This is very much in accordance to the way human mutations are manifested in individuals of the same family. Examples of ECM-associated disorders exhibiting significant clinical variability within families include the Marfan syndrome, OI and the type II collagenopathies (Ramirez *et al.*, 1990; Lee *et al.*, 1991b; Ramirez *et al.*, 1992). On the other hand, the phenotypic variability of a type II collagen transgene in inbred mice was recently associated with

stochastic changes in the uterine micro-environment during embryogenesis (Helminen *et al.*, 1993). Generation of inbred pNA25 lines will enable us to test the validity of this hypothesis in a genetic situation that resembles more closely the natural one.

The successful elimination of the *col5a2* N-telopeptide by exon-targeting is likely to become a general strategy to modify functionally important regions in multidomain proteins. Based on the findings of the pNA25 mutation, it will be of great interest to generate additional mutations in the *col*a*25* gene. Two of such mutations are already underway. The first will replace a number of Gly-X-Y coding exons with the neo cassette. As already mentioned, this deletion is expected to affect trimer assembly and thermostability, secretion and fibrillogenesis. The Gly-X-Y mutation will probably be more severe than the pNA25, and also more likely to give a correlation with a human disorder.

We have different expectations for the null allele. With one single exception, there are no examples of human disorders associated with null collagen alleles. The exception is Stickler syndrome, a dominantly inherited arthroopthalmopathy, which has been shown to be caused by nonsense mutations in the *col2a1* gene (Ahmad *et al.*, 1991). The Stickler phenotype differs from the rest of the type II collagenopathies which are caused by structural mutations. Unlike such individuals, Stickler patients are tall rather than short, and have a marfanoid habitus. Alternatively, we may fail to obtain an abnormal phenotype. As already mentioned, there are no abnormalities in the transgenic mice carrying null TN-C and type X collagen alleles. The lack of an abnormal phenotype or a milder phenotype than expected are common findings in transgenes with null alleles of genes known to play regulatory roles in morphogenesis and cell differentiation (Donehower *et al.*, 1992; Lee *et al.*, 1992; Luetteke *et al.*, 1993; Mann *et al.*, 1993). This phenomenon has been interpreted as an indication of functional redundancy of these gene products. Functional redundancy can similarly be invoked to hypothesize that other proteins co-expressed in the same tissues substitute TN-C and type X collagen functions. We may therefore predict that one of the type XI subunits may substitute for the loss of type V. Although the null *col5a2* mutant

might be normal, it could eventually be crossed to mice with a null α1(V) allele, generating transgenic animals completely deficient in type V collagen. This will enable us to test whether or not there is functional redundancy of type V collagen.

In conclusion, this study demonstrated that the minorly expressed type V collagen plays an important role in the maintenance of the ECM architecture by regulating the growth and tri-dimensional organization of type I collagen fibrils. Although the exact nature of this role remains obscure, this work has generated the reagents which will enable others to dissect the metabolic consequences of the mutation. This will eventually assign a specific role to the N-telopeptide domain of pro-α2(V) collagen. More generally, the generation of this and other exon-deletion mutants has proven to be a powerfull tool in dissecting the complexity of the structural-functional interactions responsible for ECM integrity.

Bibliography

Adachi, E., Hayashi, T., and Hashimoto, P. H. (1989). Immunoelectron microscopical evidence that type V collagen is a fibrillar collagen: importance for an aggregating capability of the preparation for reconstituting banding fibrils. Matrix *9*, 232-7.

Adamson, E. D. (1982). The effect of collagens in cell division, cellular differentiation, and embryonic development. In Collagen In Health And Disease, J. B. Weiss, and Jayson, M. I. V., eds. (Edinburgh, UK: Churchill-Livingston), pp. 218-29.

Ahmad, N. N., Ala, K. L., Knowlton, R. G., Jimenez, S. A., Weaver, E. J., Maguire, J. I., Tasman, W., and Prockop, D. J. (1991). Stop codon in the procollagen II gene (COL2A1) in a family with the Stickler syndrome (arthro-ophthalmopathy). Proc Natl Acad Sci U S A *88*, 6624-7.

Andrikopoulos, K., Suzuki H. R., Solursh M., and Ramirez, F. Localization of pro-α2(V) collagen transcripts in the tissues of the developing mouse embryo. Devel Dyn *195*, 113-20 (1992).

Angerer, L. M., Stoler, M. H., and Angerer, R. C. (1987). *In situ* hybridization with RNA probes: An annotated recipe. In *In situ* hybridization: Applications to Neurobiology, K. L. Valentino, eds. (New York: Oxford Univ. Press), pp. 42-70.

Bailey, A. J., Robins, S. P., and Balian, G. (1974). Biological significance of the intermolecular crosslinks of collagen. Nature *251*, 105-9.

Beighton, P. (1993). The Ehlers-Danlos Syndromes. In McKusick's Heritable Disorders of The Connective Tissue, P. Beighton, eds. (St Louis, Missouri: Mosby), pp. 189-252.

Berg, R. A., Kishida, Y., Kobayashi, Y., Inouye, K., Tonelli, A. E., Sakakibara, S., and Prockop, D. J. (1973). A model for the triple-helical structure of (Pro-Hyp-Gly)10 involving a cis peptide bond and inter-chain hydrogen-bonding to the hydroxyl group of hydroxyproline. Biochim Biophys Acta *328*, 553-9.

Birk, D. E., Fitch, J. M., Babiarz, J. P., Doane, K. J., and Linsenmayer, T. F. (1990). Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. J Cell Sci *95*, 649-57.

Birk, D. E., Fitch, J. M., Babiarz, J. P., and Linsenmayer, T. F. (1988). Collagen type I and type V are present in the same fibril in the avian corneal stroma. J Cell Biol *106*, 999-1008.

Bonadio, J., Saunders, T. L., Tsai, E., Goldstein, S. A., Morris, W. J., Brinkley, L., Dolan, D. F., Altschuler, R. A., Hawkins, J. E. J., Bateman, J. F., Mascara, T., and Jaenisch, R. (1990). Transgenic mouse model of the mild dominant form of osteogenesis imperfecta. Proc Natl Acad Sci U S A *87*, 7145-9.

Borelli, E., Heyman, R., Hsi, M., and Evans, R. M. (1988). Targeting of an inducible toxic phenotype in animal cells. Proc Natl Acad Sci U S A *87*, :7572-6.

Bornstein, P., and Sage, H. (1980). Structurally distinct collagen types. Annu Rev Biochem *49*, 957-1003.

Brinster, R. L., Chen, H. Y., Trumbauer, M., Senear, A. W., Warren, R., and Palmiter, R. D. (1981). Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. Cell *27*, 223-31.

Broek, D. L., Madri, J., Eikenberry, E. F., and Brodsky, B. (1985). Characterization of the tissue form of type V collagen from chick bone. J Biol Chem *260*, 555-62.

Brown, K. E., Lawrence, R., and Sonenshein, G. E. (1991). Concerted modulation of α1(XI) and α2(V) collagen mRNAs in bovine vascular smooth muscle cells. J Biol Chem *266*, 23268-73.

Burgeson, R. E. (1988). New Collagens, new concepts. Annu Rev Cell Biol *4*, 551-77.

Capecchi, M. R. (1989a). Altering the genome by homologous recombination. Science *244*, 1288-92.

Capecchi, M. R. (1989b). The new mouse genetics: altering the genome by gene targeting. Trends Genet *5*, 70-6.

Cheah, K. S. E., Au, P. K. C., Lau, E. T., Little, P. F. R., and Stubbs, L. (1991a). The mouse COL2A1 gene is highly conserved and is linked to Int-1 on chromosome 15. Mamm Genome *1*, 171-83.

Cheah, K. S. E., Lau, E. T., Au, P. K. C., and Tam, P. P. L. (1991b). Expression of the mouse α1(II) collagen gene is not restricted to cartilage during development. Development *111*, 945-953.

Chiodo, A. A., Hockey, A., and Cole, W. G. (1992). A base substitution at the splice acceptor site of intron 5 of the COL1A2 gene activates a cryptic splice site within exon 6 and generates abnormal type I procollagen in a patient with Ehlers-Danlos syndrome type VII. J Biol Chem *267*, 6361-9.

Constantini, F., and Lacy, E. (1981). Introduction of a rabbit β-globin gene into the mouse germ line. Nature *294*, 92-4.

Cox, J. L., Farrell, R. A., Hart, R. W., and Langham, M. E. (1970). The transparency of the mammalian cornea. J Physiol *210*, 601-616.

Deng, C., and Capecchi, M. R. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol Cell Biol *12*, 3365-71.

Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., Montgomery, C. A. J., Butel, J. S., and Bradley, A. (1992). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. Nature *356*, 215-21.

Edery, P., Lyonnet, S., Mulligan, L. M., Pelet, A., Dow, E., Abel, L., Holder, S., Nihoul-Fekete, C., Ponder, B. A. J., and Munnich, A. (1994). Mutations of the *RET* proto-oncogene in Hirschsprung's disease. Nature *367*, 378-80.

Evans, M. J., and Kaufman, M. H. (1981). Establishment or culture of pluripotential cells from mouse embryos. Nature *292*, 154-6.

Eyre, D., and Wu, J. J. (1987). Type XI or 1α2α3α Collagen. In Structure and Function of Collagen Types, R. Mayne, and Burgeson, R. E., eds. (Orlando, Florida: Academic Press), pp. 261-80.

Fessler, J. H., and Fessler, L. I. (1987). Type V Collagen. In Structure and Function of Collagen Types, R. Mayne, and Burgeson, R. E., eds. (Orlando, Florida: Academic Press), pp. 81-103.

Fessler, L. I., Timpl, R., and Fessler, J. H. (1981). Assembly and processing of procollagen type III in chick embryo blood vessels. J Biol Chem *256*, 2531-7.

Fleischmajer, R., Perlish, J. S., Burgeson, R. E., Shaikh, B. F., and Timpl, R. (1990). Type I and type III collagen interactions during fibrillogenesis. Ann N Y Acad Sci *580*, 161-75.

Fleischmajer, R., Perlish, J. S., and Timpl, R. (1985). Collagen fibrillogenesis in human skin. Ann N Y Acad Sci *460*, 246-57.

Fleischmajer, R., Perlish, J. S., Timpl, R., and Olsen, B. R. (1988). Procollagen intermediates during tendon fibrillogenesis. J Histochem Cytochem *36*, 1425-32.

Garofalo, S., Vuorio, E., Metsaranta, M., Rosati, R., Toman, D., Vaughan, J., Lozano, G., Mayne, R., Ellard, J., Horton, W., and de Crombrugghe, B. (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 alpha 1-chain gene. Proc Natl Acad Sci U S A *88*, 9648-52.

Glorieux, F. H., Salle, B. L., Travers, R., and Audra, P. H. (1991). Dynamic histomorphometric evaluation of human fetal bone formation. Bone *12*, 377-81.

Gordon, J. W., and Ruddle, F. H. (1981). Integration and stable germ line transmission of genes injected into mouse pronuclei. Science *214*, 1244-6.

Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A., and Ruddle, F. H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. Proc Natl Acad Sci U S A *77*, 7380-4.

Gossler, A., Joyner, A. L., Rossant, J., and Skarnes, W. C. (1989). Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science *244*, 463-5.

Greenspan, D. S., Cheng, W., and Hoffman, G. G. (1991). The pro-alpha 1(V) collagen chain. Complete primary structure, distribution of expression, and comparison with the pro-alpha 1(XI) collagen chain. J Biol Chem *266*, 24727-33.

Hasty, P., Ramirez-Solis, R., R., K., and Bradley, A. . (1991). Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. Nature *350*, 243-6.

Hay, E. D. (1991), Cell Biology of Extracellular Matrix, 2nd edn. (New York: Plenum), pp. 468.

Helminen, H. J., Kiraly, K., Pelttari, A., Tammi, M. I., Vandenberg, P., Pereira, R., Dhulipala, R., Khillan, J. S., Ala, K. L., Hume, E. L., and Prockop, D. J. (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*, 582-95.

Herskowitz, I. (1987). Functional inactivation of genes by dominant negative mutations. Nature *329*, 219-22.

Jacenko, O., LuValle, P. A., and Olsen, B. R. (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 U S A *73*, 1260-4.

Jaenisch, R., and Mintz, B. (1974). Simian virus 40 DNA sequences in DNA of healthy adult mice derived from preimplantation blastocysts injected with viral DNA. Proc Natl Acad Sci U S A *71*, 1250-4.

Keene, D. R., Sakai, L. Y., Bachinger, H. P., and Burgeson, R. E. (1987). Type III collagen can be present on banded collagen fibrils regardless of fibril diameter. J Cell Biol *105*, 2393-402.

Kivirikko, K. I. (1993). Collagens and their abnormalities in a wide spectrum of diseases. Ann Med *25*, 113-26.

Kleman, J. P., Hartmann, D. J., Ramirez, F., and Van der Rest, M. (1992). The human rhabdomyosarcoma cell line A204 lays down a highly insoluble matrix composed mainly of alpha 1 type-XI and alpha 2 type-V collagen chains. Eur J Biochem *210*, 329-35.

Kratochwil, K., Dziadek, M., Lohler, J., Harbers, K., and Jaenisch, R. (1986). Normal epithelial branching morphogenesis in the absence of collagen I. Dev Biol *117*, 596-606.

Kuivaniemi, H., Tromp, G., and Prockop, D. J. (1991). Mutations in collagen genes: causes of rare and some common diseases in humans. FASEB J *5*, 2052-60.

Laird, P. W., Zijderveld, A., Linders, L., Rudnicki, M. A., Jaenisch, R., and Berns, A. (1991). Simplified mammalian DNA isolation procedure. Nucleic Acids Res *19*, 4293.

Lee, B., D' Alessio, M., and Ramirez, F. (1991a). Modifications in the Organization and Expression of Collagen Genes Associated with Skeletal Disorders. Critical Rev Euk Gene Expr *1*, 173-187.

Lee, B., Godfrey, M., Vitale, E., Hori, H., Mattei, M. G., Sarfarazi, M., Tsipouras, P., Ramirez, F., and Hollister, D. W. (1991b). Linkage of Marfan syndrome and a phenotypically related disorder to two different fibrillin genes. Nature *352*, 330-4.

Lee, E. Y., Chang, C. Y., Hu, N., Wang, Y. C., Lai, C. C., Herrup, K., Lee, W. H., and Bradley, A. (1992). Mice deficient for Rb are nonviable and show defects in neurogenesis and haematopoiesis. Nature *359*, 288-94.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-26.

Linsenmayer, T. F., Gibney, E., Igoe, F., Gordon, M. K., Fitch, J. M., Fessler, L. I., and Birk, D. E. (1993). Type V collagen: molecular structure and fibrillar organization of the chicken alpha 1(V) NH2-terminal domain, a putative regulator of corneal fibrillogenesis. J Cell Biol *121*, 1181-9.

Lohler, J., Timpl, R., and Jaenisch, R. (1984). Embryonic lethal mutation in mouse collagen I gene causes rupture of blood vessels and is associated with erythropoietic and mesenchymal cell death. Cell *38*, 597-607.

Luetteke, N. C., Qiu, T. H., Peiffer, R. L., Oliver, P., Smithies, O., and Lee, D. C. (1993). TGF alpha deficiency results in hair follicle and eye abnormalities in targeted and waved-1 mice. Cell *73*, 263-78.

Lufkin, T., Mark, M., Hart, C. P., LeMeur, M., and Chambon, P. (1992). Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its postral domain of expression. Cell *66*, 1105-19.

Mann, G. B., Fowler, K. J., Gabriel, A., Nice, E. C., Williams, R. L., and Dunn, A. R. (1993). Mice with a null mutation of the TGF alpha gene have abnormal skin architecture, wavy hair, and curly whiskers and often develop corneal inflammation. Cell *73*, 249-61.

Mansour, S. L., Thomas, K. R., and Capecchi, M. R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature *336*, 348-52.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-8.

Mayne, R., Brewton, R. G., M., M. P., and Baker, J. R. (1993). Isolation and characterization of the chains of type V / type XI collagen present in bovine vitreous. J Biol Chem *268*, 9381-6.

McLaughlin, J. S., Linsenmayer, T. F., and Birk, D. E. (1989). Type V collagen synthesis and deposition by chicken embryo corneal fibroblasts in vitro. J Cell Sci *94*, 371-9.

Mendler, M., Eich, B. S. G., Vaughan, L., Winterhalter, K. H., and Bruckner, P. (1989). Cartilage contains mixed fibrils of collagen types II, IX, and XI. J Cell Biol *108*, 191-7.

Metsäranta, M., Garofalo, S., Decker, G., Rintala, M., de Crombrugghe, B., and Vuorio, E. (1992). Chondrodysplasia in transgenic mice harboring a 15-amino acid deletion in the triple helical domain of pro alpha 1(II) collagen chain. J Cell Biol *118*, 203-12.

Miller, E. J., in Collagen: Biochemistry E. M. Nimni, Eds. (CRC Press, Boca Raton, Florida, 1988), pp. 139-156.

Mitchell, P. J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-8.

Miyahara, M., Hayashi, K., Berger, J., Tanzawa, K., Njieha, F. K., Trelstad, R. L., and Prockop, D. J. (1984). Formation of collagen fibrils by enzymic cleavage of precursors of type I collagen in vitro. J Biol Chem *259*, 9891-8.

Moradi-Améli, M., Rousseau, J.-C., Kleman, J.-P., Champliaud, M.-F., Boutillon, M.-M., Bernillon, J., and van der Rest, M. (1994). Diversity in the processing events at the N-terminus of collagen V. Submitted.

Morris, N. P., and Bachinger, H. P. (1987). Type XI collagen is a heterotrimer with the composition (1 alpha, 2 alpha, 3 alpha) retaining non-triple-helical domains. J Biol Chem *262*, 11345-50.

Muller, H. J., in Proceedings of the Sixth International Congress of Genetics D. Jones, Eds. Brooklyn Botanic Gardens, Menasha, Wisconsin (1932), pp. 213-5.

Nakata, K., Ono, K., Miyazaki, J., Olsen, B. R., Muragaki, Y., Adachi, E., Yamamura, K., and Kimura, T. (1993). Osteoarthritis associated with mild chondrodysplasia in transgenic mice expressing alpha 1(IX) collagen chains with a central deletion. Proc Natl Acad Sci U S A *90*, 2870-4.

Nandi, A. K., Roginski, R. S., Gregg, R. G., Smithies, O., and Skoultchi, A. I. (1988). Regulated expression of genes inserted at the human chromosome β-globin locus by homologous recombination. Proc Natl Acad Sci U S A *85*, 3845-9.

Nimni, E. M., and Harkness, R. D. (1988). Molecular structures and function of collagens. In Collagen: Biochemistry, E. M. Nimni, eds. (Boca Raton, Florida: CRC Press), pp. 1.

Niyibizi, C., and Eyre, D. R. (1989). Identification of the cartilage alpha 1(XI) chain in type V collagen from bovine bone. FEBS Lett *242*, 314-8.

Olsen, B. R., Guzman, N. A., Engel, J., Condit, C., and Aase, S. (1977). Purification and characterization of a peptide from the carboxy-terminal region of chick tendon procollagen type I. Biochemistry *16*, 3030-6.

Ramirez, F., in Collagen: Molecular Biology B. R. Olsen, and Nimni, E. M., Eds. (CRC Press, Boca Raton, Florida, 1989), pp. 21.

Ramirez, F., Boast, S., DAlessio, M., Lee, B., Prince, J., Su, M. W., Vissing, H., and Yoshioka, H. (1990). Fibrillar collagen genes. Structure and expression in normal and diseased states. Ann N Y Acad Sci *580*, 74-80.

Ramirez, F., Lee, B., and Vitale, E. (1992). Clinical and genetic associations in Marfan syndrome and related disorders. Mt Sinai J Med *59*, 350-6.

Ramirez, F., Pereira, L., Zhang, H., and Lee, B. (1993). The fibrillin-Marfan syndrome connection. Bioessays *15*, 589-94.

Robertson, E. J. (1987). Embryo-derived stem cell lines. In Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, E. J. Robertson, eds. (Oxford: IRL Press), pp. 71-112.

Romanic, A. M., Adachi, E., Kadler, K. E., Hojima, Y., and Prockop, D. J. (1991). Copolymerization of pNcollagen III and collagen I. pNcollagen III decreases the rate of incorporation of collagen I into fibrils, the amount of collagen I incorporated, and the diameter of the fibrils formed. J Biol Chem *266*, 12703-9.

Romeo, G., Ronchetto, P., Luo, Y., Barone, V., Serl, M., Ceccherini, I., Pasini, B., Bocclardi, R., Lerone, M., Käärlälnen, H., and Martucciello, G. (1994). Point mutations affecting the tyrosine kinase domain of the *RET* proto-oncogene in Hirschsprung's disease. Nature *367*, 377-8.

Rosati, R., Horan, G. S. B., Garofalo, S., Pinero, G. J., Vuorio, E., de Crombrugghe, B., and Behringer, R. R. (1994). Normal long bone growth and development in type X collagen-null mice. Nature Genetics in press.

Rossi, P., and de Crombrugghe, B. (1987). Identification of a cell-specific transcriptional enhancer in the first intron of the mouse alpha2 (type I) collagen gene. Proc Natl Acad Sci U S A *84*, 5590-4.

Roth, D., and Wilson, J. (1988). Illegitimate recombination in mammalian cells. In Genetic Recombination, R. Kucherlapati, and Smith, G. R., eds. (Washington, DC: ASM), pp. 621-54.

Saga, Y., Yagi, T., Ikawa, Y., Sakakura, T., and Aizawa, S. (1992). Mice develop normally without tenascin. Genes Dev *6*, 1821-31.

Sambrook, E., Fritsch, E. F., and Maniatis, T., (1989), Molecular Cloning: A Laboratory Manual, 2nd edn. (New York: Cold Spring Harbor), pp. 545.

Sandberg, M., and Vuorio, E. (1987). Localization of types I, II and III collagen mRNAs in developing human skeletal tissues by *in situ* hybridizations. J. Cell Biol. *104*, 1077-84.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A *74*, 5463-7.

Schnieke, A., Harbers, K., and Jaenisch, R. (1983). Embryonic lethal mutation in mice induced by retrovirus insertion into the alpha 1(I) collagen gene. Nature *304*, 315-20.

Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Constantini, F., and Pachnis, V. (1994). Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. Nature *367*, 380-3.

Smith, L. T., Wertelecki, W., Milstone, L. M., Petty, E. M., Seashore, M. R., Braverman, I. M., Jenkins, T. G., and Byers, P. H. (1992). Human dermatosparaxis: a form of Ehlers-Danlos syndrome that results from failure to remove the amino-terminal propeptide of type I procollagen. Am J Hum Genet *51*, 235-44.

Stacey, A., Bateman, J., Choi, T., Mascara, T., Cole, W., and Jaenisch, R. (1988). Perinatal lethal osteogenesis imperfecta in transgenic mice bearing an engineered mutant pro-alpha 1(I) collagen gene. Nature *332*, 131-6.

Su, M. W., Benson, C. V., Vissing, H., and Ramirez, F. (1989). Organization of the exons coding for pro alpha 1(II) collagen N-propeptide confirms a distinct evolutionary history of this domain of the fibrillar collagen genes. Genomics *4*, 438-41.

Swiderski, R. E., and Solursh, M. (1992). Localization of type II collagen, long form alpha 1(IX) collagen, and short form alpha 1(IX) collagen transcripts in the developing chick notochord and axial skeleton. Dev Dyn *194*, 118-27.

Tanzer, M. L. (1973). Cross linking of collagen. Science *180*, 561-6.

Theiler, K., (1989), The House Mouse: Atlas of Embryonic Development, edn. (New York: Springer-Verlag), pp. 152.

Thomas, K. R., and Capecchi, M. R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell *51*, 503-12.

Thomas, K. R., Deng, C., and Capecchi, M. R. (1992). High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. Mol Cell Biol *12*, 2919-23.

Timpl, R., and Glanville, R. W. (1981). The aminopropeptide of collagen. Clin Orthop Rel Res *158*, 224-42.

Travis, J. (1992). Scoring a Technical Knockout in Mice. Science *256*, 1392-4.

Truter, S., Di, L. M., Inagaki, Y., and Ramirez, F. (1992). Identification of an upstream regulatory region essential for cell type-specific transcription of the pro-alpha 2(V) collagen gene (COL5A2). J Biol Chem *267*, 25389-95.

Tseng, S. C. G., Smuckler, D., and Stern, R. (1982). Comparison of collagen types in adult and fetal bovine corneas. J Biol Chem *257*, 2627-33.

Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T., and Mulligan, R. C. (1991). Neonatal lethality and Lymphopenia in mice with a homozygous disruption of the c-abl protooncogene. Cell *65*, 1153-63.

van der Rest, M., and Garrone, R. (1991). Collagen family of proteins. FASEB J *5*, 2814-23.

Vandenberg, P., Khillan, J. S., Prockop, D. J., Helminen, H., Kontusaari, S., and Ala, K. L. (1991). Expression of a partially deleted gene of human type II procollagen (COL2A1) in transgenic mice produces a chondrodysplasia. Proc Natl Acad Sci U S A *88*, 7640-4.

Vaughan, L., Mendler, M., Huber, S., Bruckner, P., Winterhalter, K. H., Irwin, M. I., and Mayne, R. (1988). D-periodic distribution of collagen type IX along cartilage fibrils. J Cell Biol *106*, 991-7.

Weil, D., Bernard, M., Combates, N., Wirtz, M. K., Hollister, D. W., Steinmann, B., and Ramirez, F. (1988). Identification of a mutation that causes exon

skipping during collagen pre-mRNA splicing in an Ehlers-Danlos syndrome variant. J Biol Chem *263*, 8561-4.

Weil, D., Bernard, M., Gargano, S., and Ramirez, F. (1987). The pro alpha 2(V) collagen gene is evolutionarily related to the major fibrillar-forming collagens. Nucleic Acids Res *15*, 181-98.

Weil, D., DAlessio, M., Ramirez, F., de, W. W., Cole, W. G., Chan, D., and Bateman, J. F. (1989). A base substitution in the exon of a collagen gene causes alternative splicing and generates a structurally abnormal polypeptide in a patient with Ehlers-Danlos syndrome type VII. EMBO J *8*, 1705-10.

Weil, D., DAlessio, M., Ramirez, F., and Eyre, D. R. (1990). Structural and functional characterization of a splicing mutation in the pro-alpha 2(I) collagen gene of an Ehlers-Danlos type VII patient. J Biol Chem *265*, 16007-11.

Weil, D., DAlessio, M., Ramirez, F., Steinmann, B., Wirtz, M. K., Glanville, R. W., and Hollister, D. W. (1989). Temperature-dependent expression of a collagen splicing defect in the fibroblasts of a patient with Ehlers-Danlos syndrome type VII. J Biol Chem *264*, 16804-9.

Wertelecki, W., Smith, L. T., and Byers, P. (1992). Initial observations of human dermatosparaxis: Ehlers-Danlos syndrome type VIIC. J Pediatr *121*, 558-64.

Wirtz, M. K., Keene, D. R., Hori, H., Glanville, R. W., Steinmann, B., Rao, V. H., and Hollister, D. W. (1990). In vivo and in vitro noncovalent association of excised

alpha 1 (I) amino-terminal propeptides with mutant pN alpha 2(I) collagen chains in native mutant collagen in a case of Ehlers-Danlos syndrome, type VII. J Biol Chem *265*, 6312-7.

Woodbury, D., Benson, C. V., and Ramirez, F. (1989). Amino-terminal propeptide of human pro-alpha 2(V) collagen conforms to the structural criteria of a fibrillar procollagen molecule. J Biol Chem *264*, 2735-8.

Yoshioka, H., and Ramirez, F. (1990). Pro-alpha 1(XI) collagen. Structure of the amino-terminal propeptide and expression of the gene in tumor cell lines. J Biol Chem *265*, 6423-6.

Yoshioka, H., Zhang, H., Ramirez, F., Mattei, M. G., Moradi-Ameli, M., Van der Rest, M., and Gordon, M. K. (1992). Synteny between the loci for a novel FACIT-like collagen locus (D6S228E) and alpha 1 (IX) collagen (COL9A1) on 6q12-q14 in humans. Genomics *13*, 884-6.

Zijlstra, M., Li, E., Sajjadi, F., Subramani, S., and Jaenisch, R. (1989). Germ-line transmission of a disrupted beta 2-microglobulin gene produced by homologous recombination in embryonic stem cells. Nature *342*, 435-8.