Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 311 (2003) 553–560
BBRC
www.elsevier.com/locate/ybbrc

Tissue- and developmental stage-specific activation of α5 and α6(IV) collagen expression in the upper gastrointestinal tract of transgenic mice

Christine Herzog,[1] Liyan Zhuang,[1] Lucian Gorgan, Yoav Segal,* and Jing Zhou

Renal Division, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

Received 26 September 2003

Abstract

Little is known about mechanisms regulating gene expression for the α chains of basement membrane type IV collagen, arranged head-to-head in transcription units *COL4A1–COL4A2*, *COL4A3–COL4A4*, and *COL4A5–COL4A6*, and implicated broadly in genetic diseases. To investigate these mechanisms, we generated transgenic mouse lines bearing 5′-flanking sequences of *COL4A5* and *COL4A6*, cloned upstream of a *lacZ* reporter gene. A 3.8-kb fragment upstream of *COL4A6* directs reporter gene expression in the esophagus, stomach, and duodenum, whereas a 13.8-kb fragment directs expression in the esophagus only. A 10.6-kb fragment upstream of *COL4A5* directs expression in the esophagus. Coupled with evidence of long-range conservation between human and mouse non-coding sequences, described herein, our findings provide the first indication that highly specialized patterns characteristic of *COL4A5–COL4A6* expression in vivo arise from effects of distributed *cis*-acting regulatory elements on a bidirectional proximal promoter, itself transcriptionally competent.

Keywords: Collagen type IV; Gene expression regulation; Mice, transgenic; Esophagus; Gastric mucosa; Intestinal mucosa; Nephritis, hereditary

Basement membranes (BMs) are thin sheet-like extracellular matrices that compartmentalize and organize multicellular structures, and confer important signals for tissue differentiation, maintenance, and remodeling [1]. The integrity of BMs is based on the formation of a stable network of type IV collagen molecules, enabling aggregation with laminins and heparan sulfate proteoglycans, achieved by crosslinking via nidogen [2]. Six α chains, α1(IV)–α6(IV), contribute to type IV collagen triple-helical subunits. The functional importance of individual α(IV) chains is illustrated in inherited and acquired diseases, including Alport syndrome, diffuse leiomyomatosis, nail-patella syndrome, and Goodpasture syndrome, as recently reviewed [3].

Tissue distributions of the α chains underlie the major manifestations of type IV collagen disorders. The "major" α1(IV) and α2(IV) chains are widely distributed. The "minor" chains, α3(IV)–α6(IV), have restricted distributions. The α3(IV) and α4(IV) chains are consistently co-localized [4,5]. The α5(IV) chain is co-localized with the α3(IV) and α4(IV) chains within BMs of kidney, lung, eye, ear, brain, and neuromuscular junction, and is also present in BMs of skin and gastrointestinal tract [5,6]. The α6(IV) chain is co-localized with the α5(IV) chain in numerous BM structures, with the renal glomerular BM providing an exception [6]. Developmental regulation of α(IV) chain expression has been described for several tissues. A mid-gestational switch from major to minor α(IV) chain predominance is a defining feature of renal glomerular BM development [5]. Stage-specific expression patterns, important in cellular proliferation, migration, and differentiation in the developing gastrointestinal tract [7], have also been described, for esophagus [8,9] and small intestine [10].

* Corresponding author. Present address: Division of Renal Diseases and Hypertension, University of Minnesota, Box 736 UMHC, 420 Delaware Street SE, Minneapolis, MN 55455, USA. Fax: 1-612-626-3840.
E-mail address: ysegal@tc.umn.edu (Y. Segal).
[1] These two authors equally contributed to this work.

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2003.09.233

Virtually nothing is known about regulatory mechanisms accounting for type IV collagen distribution. The genes encoding the α(IV) chains are arranged in bidirectional transcription units not found for other collagen genes. The *COL4A1* and *COL4A2* genes on human chromosome 13 are arranged head-to-head and connected by a short common promoter of only 127 bp in length [11]. A similar arrangement has been found on human chromosome X for the *COL4A5* and *COL4A6* genes [12], and on chromosome 2 for the *COL4A3* and *COL4A4* genes [13]. Although self-evident, the importance of transcriptional mechanisms is illustrated directly in nail-patella syndrome, a multisystem disorder in which genetic mutations of the LMX1B transcription factor disrupt activation of the *COL4A3–COL4A4* genes, contributing to characteristic basement membrane abnormalities [14]. Focusing on the proximal promoter region, we demonstrated that *COL4A5* and *COL4A6* possess separate promoters immediately upstream of their respective start sites, enabling fine regulation of transcriptional activity in vitro [15]. However, these findings could not explain broad aspects of tissue-specific regulation.

To define the framework for transcriptional regulation of the α5(IV) and α6(IV) chains, as a basis for tissue- and developmental stage-specificities, we generated transgenic mice expressing a *lacZ* reporter gene, under the control of *COL4A5* and *COL4A6* regulatory elements.

Methods

Plasmid construction. For plasmid construction we used the pLacF vector, kindly provided by Dr. Jacques Peschon. In this vector, the *Escherichia coli* β-galactosidase (*lacZ*) gene is cloned upstream of the mouse protamine gene, which supplies an intron and a polyadenylation signal [16]. A unique restriction site (*Not*I) was introduced downstream of the mouse protamine gene by standard recombinant techniques. Inserts are specified in Results.

Generation of transgenic mice. Before injection, inserts were released from the plasmid by *Kpn*I and *Not*I digestion, and purified using the QIAquick Gel Extraction Kit (Qiagen). DNA was eluted in injection buffer (10 mM Tris, 0.25 μM EDTA) to a concentration of 2 ng/μL and syringe-filtered. The DNA was microinjected into pronuclei of fertilized eggs, which were then transferred to pseudopregnant FVB foster mothers.

Founder tail DNA samples were prepared using the QIAamp Tissue Kit (Qiagen). Genomic DNA was examined for transgene integration by PCR and Southern blot analysis. For PCR, two *lacZ*-directed oligonucleotides (*lacZ*-for 5′-aag ctg gct gga gtg cga tct-3′ and *lacZ*-rev 5′-ggc ata acc acc acc acg ctc atc-3′) generated a 647-bp product. For conventional Southern blot analysis, *Eco*RI-digested DNA was hybridized against a radiolabeled 0.9-kb *lacZ* fragment.

Positive founders were bred to non-transgenic FVB mice to establish lines. β-Galactosidase activity was assessed on whole embryos and individual organs from post-natal animals by standard techniques [17].

Immunochemistry. Rabbit polyclonal antiserum to bacterial β-galactosidase was purchased from Clontech. Goat antiserum to human α1/α2(IV) collagen, also reactive with mouse, was purchased from Southern Biotechnology Associates. Rabbit polyclonal antiserum to mouse α5(IV) collagen was a generous gift of Drs. Joshua Sanes and Jeffrey Miner. The rabbit polyclonal antibody to mouse α6(IV) collagen was raised against a synthetic peptide of a non-consensus amino acid sequence at the COOH-terminal region of the mouse α6(IV) NCI domain (VEERGOFREEPVSE). The antibody was purified on a peptide affinity column, using the AminoLink Plus Immobilization Kit (Pierce). The purified antibody stained kidney tissues in accord with previously described patterns [6] and was blocked by pre-incubation with excess peptide (not shown). The FITC-conjugated secondary antibody to rabbit IgG was purchased from Cappel and the Texas Red-conjugated secondary antibody to goat IgG was from Vector Laboratories.

Individual organs were removed from animals and embedded in O.C.T. Compound (Sakura). Tissues were sectioned at 5 μm in a cryostat, mounted on pretreated slides, and stored at −80 °C. Sections were post-fixed in 95% ETOH for 5 min at 4 °C, rinsed with PBS, and denatured in 6 M urea/0.1 M glycine HCl, pH 3.5, for 1 h at 4 °C. For dual labeling of α1/α2(IV) and α6(IV), primary antibodies were diluted 1:50 in PBS containing 1% BSA and 0.1% Triton X-100, and incubated for 1 h at room temperature. Secondary antibody to rabbit IgG was diluted 1:600 in PBS containing 5% dried milk and incubated for 30 min at room temperature. Secondary antibody to goat IgG was diluted 1:100 in PBS containing 5% dried milk and incubated for 30 min at room temperature. For dual labeling of α1/α2(IV) and α5(IV), the α1/α2(IV) was diluted 1:50, and the α5(IV) antibody 1:100, in PBS containing 5% dried milk. All sections were mounted in VectaShield (Vector Labs) and examined with a Zeiss Axiovert microscope. Pictures were taken with a spot-camera (Diagnostic Instruments). Control experiments were performed omitting the primary antibody and using PBS containing 5% nonfat dried milk only.

Human–mouse sequence comparison. Human and mouse genomic sequences were obtained from GenBank accessions. Long-range sequence comparison was carried out using Dialign [18]. Graphic representation of the comparison was carried out using SynPlot [19]. Repetitive elements were identified using RepeatMasker (ftp.genome.washington.edu/cgi-bin/RepeatMasker).

Results

Generation of transgenic mice

The activities of three reporter gene constructs were investigated in transgenic mice (Fig. 1A). Fragments from human *COL4A5* and *COL4A6* were assembled in *lacZ* transgenic constructs, designated relative to translation start sites at position +1, and demarcated on the basis of GenBank Accession No. AL034369.1. These were: a 3.8-kb *Xba*I–*Xba*I fragment from positions −3984 to −156 of *COL4A6*, exon 1B (−3.8α6); a 13.8-kb *Kpn*I–*Xba*I fragment from positions −14,003 to −156 of *COL4A6*, exon 1B (−13.8α6); and a 10.6-kb *Bam*HI–*Hin*dIII fragment from positions −10,719 to 143 of *COL4A5* (−10.6α5). A translation start site in human *COL4A6*, exon 1A is not conserved in mouse [20]. For simplicity, we focused on exon 1B.

PCR and Southern blot analysis were used to identify founder mice and determine approximate transgene copy numbers (Fig. 1B; Table 1). Offspring from five transgenic line carrying the −13.8α6 fragment, three lines carrying the −3.8α6 fragment, and one line carrying

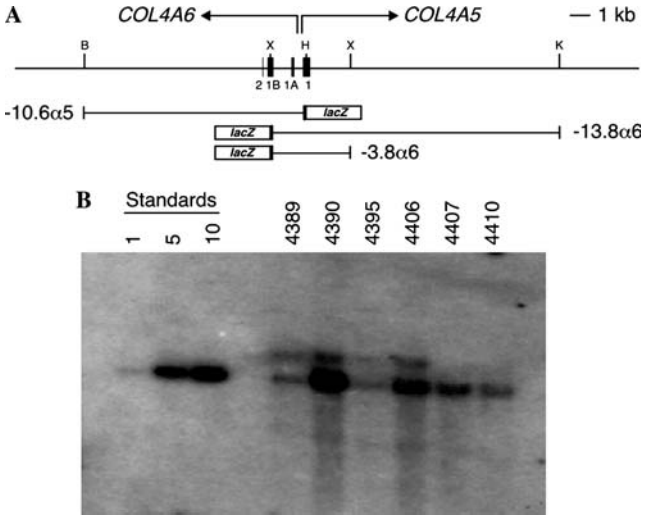

Fig. 1. (A) Schematic representation of DNA constructs used to generate transgenic mice. Cloning sites are indicated as B (*Bam*HI), K (*Kpn*I), H (*Hin*dIII), and X (*Xba*I). See text for details. (B) Determination of copy numbers for founder mice bearing the −13.8α6 fragment.

the −10.6α5 fragment were analyzed. Histochemical staining for β-galactosidase activity was carried out in tissues from whole-mounted embryos at 12.5 (E12.5) and 14.5 (E14.5) days of gestation, infants at 1 (P1), 3 (P3), and 7 (P7) days of age, weanlings, and adults at 6–12 weeks of age.

An esophagus-specific DNA element in the intergenic region of COL4A5 and COL4A6

β-Galactosidase activity was detected uniformly in superficial epithelial cells of the distal esophagus, in adult mice from one of one −10.6α5 line, five of five −13.8α6 lines, and three of three −3.8α6 lines (Fig. 2). Activity was not detected in proximal esophagus or muscle cell layers, or in any esophageal structures in sections from wild-type FVB littermates. Activity was also detected in stomach and duodenum, in mice carrying the −3.8α6 fragment (Fig. 2; see below). Other organs, including ileum, colon, skin, kidney, lung, liver, and skeletal muscle, examined in sections from transgenic mice, were negative, as summarized in Table 1.

To compare distributions of reporter gene activity with those of the endogenous proteins, immunofluorescence staining with α1/α2(IV), α5(IV), and α6(IV) antibodies was conducted. Double staining with α1/α2 chains showed that the endogenous expression of α5(IV) and α6(IV) is located, as expected, in epithelial BM of the esophagus, with weaker expression in surrounding muscle BM (Fig. 2).

An inhibitory element in the −13.8 to −3.8 kb region of the COL4A6 gene

Patterns of reporter gene activity differed between lines carrying the −3.8α6 and −13.8α6 fragments. Three of three lines carrying the −3.8α6 fragment showed β-galactosidase activity in the esophageal epithelium, in base regions of gastric glands within the pylorus, and in crypt bases of the duodenum (Fig. 2). Adult mice carrying the −13.8α6 fragment showed specific β-galactosidase activity in esophageal epithelium only. Thus, a negative element between positions −13.8 and −3.8 kb is implicated in *COL4A6* repression, in stomach and duodenum. β-Galactosidase activity was not detected in other organs, for any of the lines (Table 1). Immunofluorescence studies with α1/α2(IV), α5(IV), and α6(IV) antibodies demonstrated the presence of endogenous α5(IV) and α6(IV) chains in BMs of the stomach and duodenum (Fig. 2).

Developmental stage-specific expression of the lacZ transgene in esophagus

Whole-mounted embryos bearing the −10.6α5 and −13.8α6 fragments were studied for β-galactosidase activity. No activity was detected in any organs of E12.5 and E14.5 embryos. Activity was absent in organs of P1

Table 1
Copy numbers and β-galactosidase expression patterns in transgenic mouse lines

Fragment	Line	Copy number	Adult tissues									
			Esophagus	Stomach	Duodenum	Ileum	Colon	Skin	Kidney	Lung	Liver	Skeletal muscle
−3.8α6	2946	5	+++	++	+++	–	–	–	–	–	–	–
	2945	1	++	+	++	–	–	–	–	–	–	–
	2939	1	++	+	++	–	–	–	–	–	–	–
−13.8α6	4839	2	++	–	–	–	–	–	–	–	–	–
	4390	30	++++	–	–	–	–	–	–	–	–	–
	4410	2	++	–	–	–	–	–	–	–	–	–
	4406	5	+++	–	–	–	–	–	–	–	–	–
	4407	4	+++	–	–	–	–	–	–	–	–	–
−10.6α5	0907	1	++	–	–	–	–	–	–	–	–	–

β-Galactosidase activity was graded on an arbitrary scale from no (–) to intense (++++) staining.

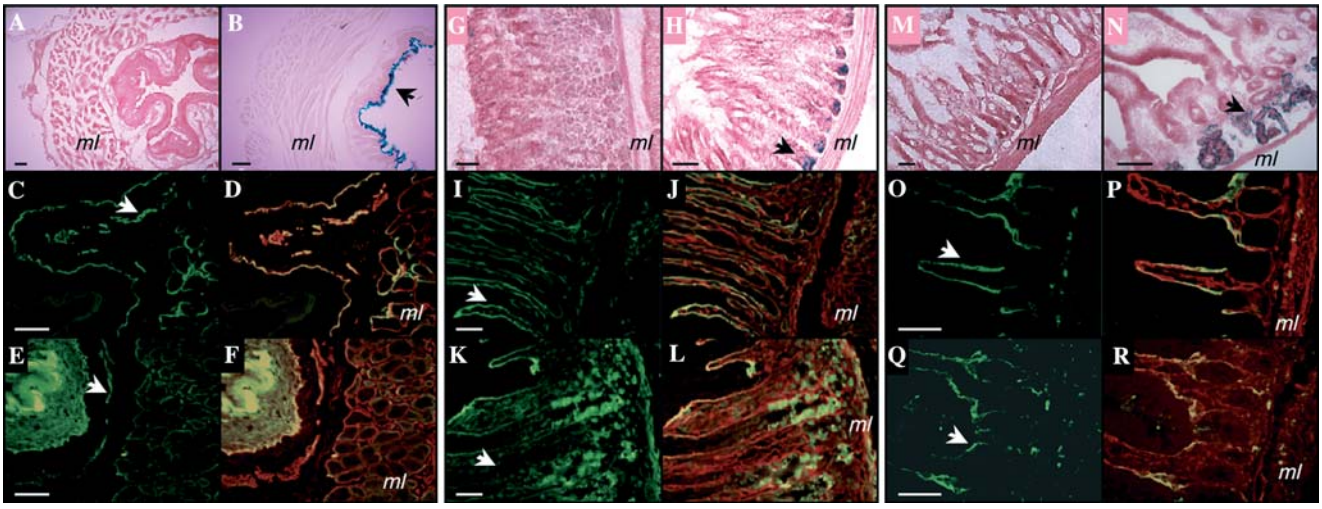

Fig. 2. Correlation of *lacZ* reporter and endogenous collagen IV expression. Data are shown for adult esophagus (A–F), stomach (G–L), and duodenum (M–R). Histochemical staining for β-galactosidase activity is shown for control (A, G, and M) and −3.8α6 transgenic (B, H, and N) mice. Positive staining is indicated (arrow). In dual immunofluorescence studies of adult mice, endogenous α6(IV) (C, D, I, J, O, and P) or α5(IV) (E, F, K, L, Q, and R) staining is green and α1, α2(IV) (D, F, J, L, P, and R) staining is red. In esophagus, the α5(IV) and 6(IV) chains are present in epithelial BM and show weak and patchy staining patterns in BM of the muscular layers (ml). α6(IV) antibody staining of the keratin layer is non-specific. In stomach, α5(IV) and α6(IV) staining is strongest in epithelial BM of the gastric pits, close to the mucosal surface. Some weak staining can also be seen in BM of the base regions. In duodenum, α5(IV) and α6(IV) staining is strongest in BM of the mid-villus. Some weak staining can also be seen in BM of the villus tip. The α1, α2(IV) chains are present in epithelial and mesenchymal BM, including those within muscular layer (ml), of all tissues. Bars = 50 μm.

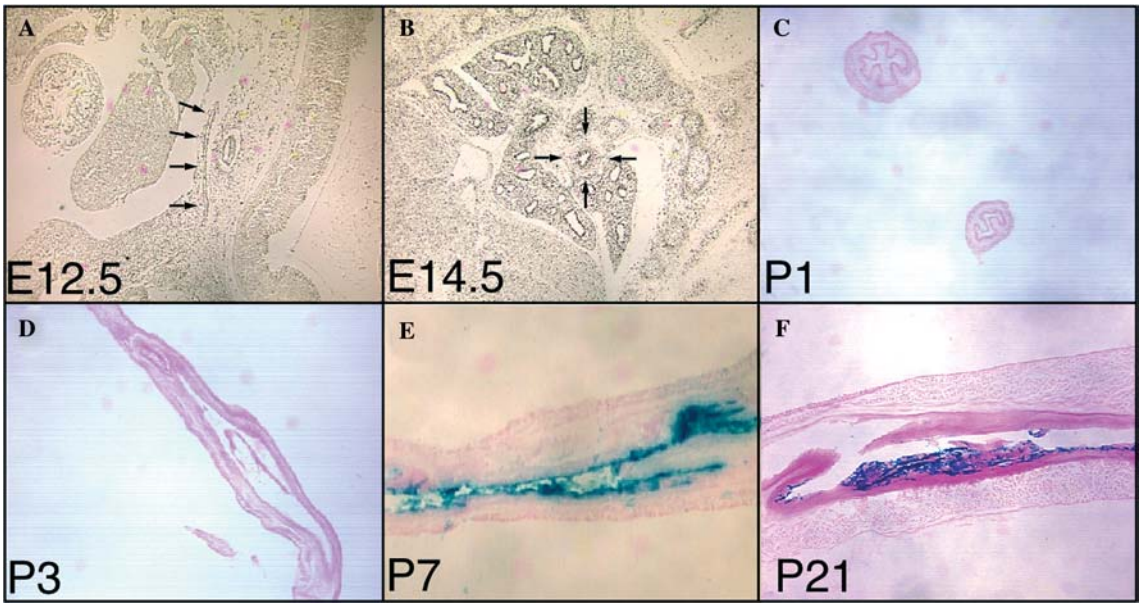

Fig. 3. Developmental *lacZ* expression in esophagus. Histochemical staining for β-galactosidase activity is shown for −13.8α6 transgenic mice, at low-power magnification. Staining for E12.5 (A) and E14.5 (B) embryos, and P1 (C), P3 (D), P7 (E), and P21 (F) post-natal animals reveals activity beginning at the P7 stage. Arrows indicate esophagus.

and P3 mice, and detectable in the esophageal epithelium of P7 mice carrying the −10.6α5 or −13.8α6 fragments (Fig. 3). β-Galactosidase activity was evident in superficial epithelial cells at day P28, by which time the mouse esophagus is fully keratinized. Importantly, the starting point of expression is the same for both constructs (not shown). Wild-type mice at the same stages displayed no detectable β-galactosidase activity.

Human–mouse sequence comparison of proximal introns for COL4A5 and COL4A6

Our findings indicate that DNA fragments from *COL4A5–COL4A6*, overlapping across the proximal promoter region for this gene pair, are transcriptionally competent, directing reporter gene expression in the upper gastrointestinal tract, in a position-independent

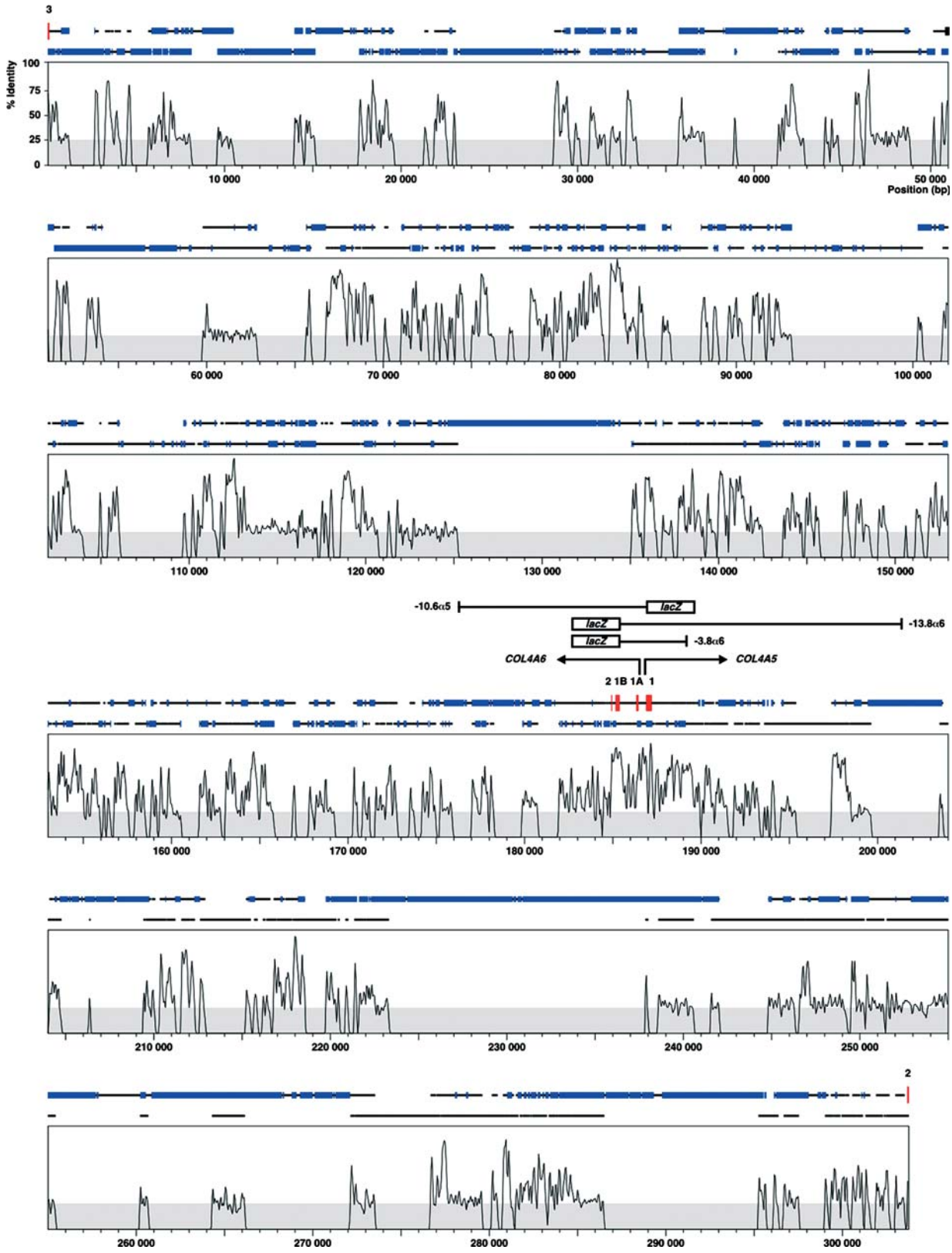

Fig. 4. Human–mouse sequence comparison. Genome sequences from exon 3 of *COL4A6* to exon 2 of *COL4A5* are represented for human (top line) and mouse (bottom line). Exons are depicted as red boxes. Interspersed repeats including LINE's and SINE's are depicted as blue boxes. Interruptions correspond to gaps in the sequence alignment. Percent nucleotide identity within fixed windows is plotted against nucleotide position within the aligned sequence (see [19] for details). Gray bars indicate 25% identity, characteristic of alignment for two random DNA sequences. Reporter gene constructs are shown schematically.

and bidirectional manner. Tissue-specific reporter gene expression is regulated by additional elements, evidenced by effects of the negative element in the −13.8 to −3.8 kb region of the *COL4A6* gene. Although >10 kb in size, the DNA fragments investigated here appear to lack *cis*-acting elements conferring tissue- or developmental stage-specificity for other important organs, including kidney, skin, and lower gastrointestinal tract.

To gain further insights into *COL4A5–COL4A6* regulation, we undertook a computational analysis. The *COL4A5–COL4A6* and remaining type IV collagen gene pairs contain expansive (>100 kb) proximal introns [21,22], raising the possibility that *cis*-acting elements distributed within these introns contribute to characteristic patterns of gene expression. Indeed, conserved non-coding sequences are implicated in gene regulation [23]. We therefore compared human and mouse sequences between exon 3 of *COL4A6* and exon 2 of *COL4A5* (Fig. 4).

We identified numerous peaks of homology by a sequence alignment algorithm (see [18] for details), suggesting that conserved long-range as well as proximal non-coding sequences participate in transcriptional regulation. A total of 808 ungapped conserved non-coding sequences were identified. Among these, the most highly conserved is a single-copy sequence in intron 2 of *COL4A6*, 80 bp in length, 100% identical between the two species (GenBank Accession No. AL109943.18, nt 88,757–88,836) and containing consensus sites for several transcription factors. Within the region depicted in Fig. 4, human and mouse sequences differ largely in lineage-specific repeats. Coverage of conserved regions provided by the −10.6α5, −13.8α6, and −3.8α6 fragments is limited to those within and around the 5′-terminal and intergenic regions.

Discussion

Despite advantages in elucidating gene expression and function in vivo, transgenic reporter methodologies have been applied sparingly to basement membrane genes [24,25]. Little is known about molecular mechanisms regulating type IV collagen gene expression, despite obvious implications for organ involvement and potential therapy in human genetic diseases. We have used a transgenic system to identify *cis*-acting elements regulating the expression of *COL4A5* and *COL4A6* in vivo, the first such elements described for type IV collagen genes.

Each of three fragments from the human *COL4A5–COL4A6* region directed reporter gene expression to superficial epithelial cells of the distal esophagus, in several independently derived lines. Since these fragments overlap in a 1.6-kb region from *COL4A5* exon 1 to *COL4A6* exon 1B, the simplest explanation for our findings is that *cis*-acting elements, shared by *COL4A5* and *COL4A6*, direct expression in esophagus. Importantly, these elements function in the *COL4A5* or *COL4A6* orientation, thereby conferring a mechanism for co-expression of the genes.

Further experiments are required to identify individual regulatory elements, which seem to be esophagus-specific. Sequence analysis of the intergenic region between exon 1 of *COL4A5* and exon 1B of *COL4A6* revealed the presence of a CACCC box (not shown), reported to direct esophagus-specific endogenous and viral gene expression, upon activation by keratinocyte specific factor [26]. The region of overlap among our constructs also contains a positive regulatory element, shown previously to bind keratinocyte transcription factors and to activate bidirectional *COL4A5–COL4A6* transcription in vitro [15].

To compare endogenous tissue distribution with β-galactosidase activity, immunofluorescence studies with α1(IV)/α2(IV), α5(IV), and α6(IV) antibodies were performed. In esophagus, we saw strong epithelial BM and weaker muscular layer BM staining, consistent with previous reports [9]. Further immunofluorescence studies with α5(IV) and α6(IV) antibodies showed that both chains are already expressed by day P1 (not shown), also consistent with previously described developmental expression patterns [8]. By contrast, no β-galactosidase activity could be detected in esophagus until day P7. We conclude that the DNA fragments under investigation confer tissue-specificity for esophageal epithelium, but not embryonic or newborn stage-specificity. Moreover, with the observation of β-galactosidase activity in superficial rather than basal cells directly overlying BM, these fragments do not clearly confer specificity for differentiation state, as this term applies to progressively differentiating cells within stratified epithelia. Superficial cells of the esophageal epithelium undergo loss of the nucleus and cytoplasmic organelles, during their differentiation from deeper intermediate cells [27]. Thus, reporter gene activity, undetectable in control tissues, probably reflects transcription occurring in deeper cell layers. As expected, our findings indicate that developmental activation of *COL4A5–COL4A6* in mesenchyme-derived smooth muscle is programmed independently.

Mice carrying the −3.8α6 fragment showed β-galactosidase activity in esophagus. With this fragment, we also observed *lacZ* expression in base regions of gastric units within the pylorus, and in crypt bases of the duodenum. We infer that the region 13.8–3.8 kb upstream of *COL4A6* contains negative *cis*-acting elements, inhibiting expression in stomach and duodenum. Giving evidence of their differential activity along the proximal-to-distal axis, these negative elements may be responsive to regional signaling gradients, which direct late organ

morphogenesis and terminal differentiation in the upper gastrointestinal tract [28].

In the pylorus, the α5(IV) and α6(IV) antibodies showed the strongest staining in BM of the upper pit region of the gastric gland, whereas weaker staining was evident in the deeper base region. In duodenum, staining was strongest in the BM of villi. By comparison, β-galactosidase activity could only be observed in the base region of the gastric gland and the crypt base of the duodenum. It is possible to reconcile these observations if type IV collagen in epithelial BMs of the gastrointestinal tract is synthesized by fibroblasts [29], themselves not directed to express *lacZ* in our transgenic lines. In a recent study of type IV collagen synthesis in colon, Kirkland and Henderson [30] demonstrated the presence of type IV collagen mRNA in pericryptal fibroblasts, as expected, and in enteroendocrine cells, exclusively among colonic epithelial cells. They posed the interesting hypothesis that restricted expression of type IV collagen within heterogeneous gastrointestinal epithelia contributes to specialized pericellular matrices, for induction and maintenance of highly differentiated cellular phenotypes. Indeed, it is tempting to speculate that *lacZ* expression driven by the −3.8α6 fragment marks specific cell lineages and/or states of differentiation, within gastric and duodenal epithelia.

The experiments showed that the constructs used in this study only directed expression in the upper gastrointestinal tract. Since the expression of *COL4A5* and *COL4A6* is not restricted to these tissues, and underlies manifestations of Alport kidney disease and other type IV collagen disorders, we conclude that additional regulatory elements must be present outside of the fragments under study. Enhancer sequences are usually located at considerable distance from the transcription start sites. It is also possible that negatively acting elements are present within the constructs we examined, which suppress the transgene expression in other tissues. To clarify these issues, a variety of more constructs need to be analyzed.

What, then, are the special considerations applying to studies of basement membrane collagen gene transcription, and how are these considerations illuminated by our findings? By comparison to "prototypical" cell-specific genes, e.g., those for globins, those for basement membrane collagens are transcribed in widely diverse cell types. To begin with, cells within stratified epithelia undergo progressive differentiation, and turn over on a time scale of days, whereas those in the kidney glomerulus are terminally differentiated and have limited regenerative capacity. Basement membrane components turn over on a time scale of weeks to months [31,32]. It is conceivable that the homeostatic needs of divergent cell types are met by transcription of the minor type IV collagen genes at very low rates and/or by consignment of transcription to dedicated cell subsets or discrete temporal windows. We have shown that overlapping DNA fragments encompassing the *COL4A5–COL4A6* proximal promoter region are transcriptionally competent, directing robust and bidirectional *lacZ* expression in several independent transgenic lines. An arrangement, conserved in evolution, in which expansive proximal introns of *COL4A5* and *COL4A6* are arrayed with *cis*-acting elements, would serve to provide diverse—and at the same time strict—regulation of bidirectional transcription. Indeed, conserved non-coding sequences within these introns can be used to guide further studies on *cis*-regulatory elements directing expression of *COL4A5* and *COL4A6* in other organs.

In summary, our experiments indicate that *cis*-acting elements located in the overlapping 5′-flanking regions of *COL4A5* and *COL4A6* direct position-independent and bidirectional tissue- and stage-specific expression, in esophagus from mice. Furthermore, our data imply that inhibitory elements are present in the region 13.8–3.8 kb upstream of the *COL4A6* gene, which suppress expression in stomach and duodenum. Elements identified thus far can be used to direct tissue-specific expression of target genes, for analysis of gene function or gene therapy. Further studies will identify remote elements important for regulation of expression in other organs.

Acknowledgments

This work was supported by National Institutes of Health grants to J. Zhou and Y. Segal. We thank Drs. Jacques Peschon, Joshua Sanes, and Jeffrey Miner for reagents, Stefan Kren for assistance with computational analysis, and Drs. Samuel Ho and Clifford Kashtan for critical reading of the manuscript.

References

[1] R. Timpl, Structure and biological activity of basement membrane proteins, Eur. J. Biochem. 180 (1989) 487–502.
[2] P.D. Yurchenco, J.J. O'Rear, Basal lamina assembly, Curr. Opin. Cell Biol. 6 (1994) 674–681.
[3] B.G. Hudson, K. Tryggvason, M. Sundaramoorthy, E.G. Neilson, Alport's syndrome, Goodpasture's syndrome, and type IV collagen, N. Engl. J. Med. 348 (2003) 2543–2556.
[4] R.J. Butkowski, J. Wieslander, M. Kleppel, A.F. Michael, A.J. Fish, Basement membrane collagen in the kidney: regional localization of novel chains related to collagen IV, Kidney Int. 35 (1989) 1195–1202.
[5] J.H. Miner, J.R. Sanes, Collagen IV α3, α4, and α5 chains in rodent basal laminae: sequence, distribution, association with laminins, and developmental switches, J. Cell Biol. 127 (1994) 879–891.
[6] B. Peissel, L. Geng, R. Kalluri, C. Kashtan, H.G. Rennke, G.R. Gallo, K. Yoshioka, M.J. Sun, B.G. Hudson, E.G. Neilson, et al., Comparative distribution of the α1(IV), α 5(IV), and α6(IV) collagen chains in normal human adult and fetal tissues and in kidneys from X-linked Alport syndrome patients, J. Clin. Invest. 96 (1995) 1948–1957.

[7] P. Simon-Assmann, M. Kedinger, A. De Arcangelis, V. Rousseau, P. Simo, Extracellular matrix components in intestinal development, Experientia 51 (1995) 883–900.
[8] L. Heidet, Y. Cai, Y. Sado, Y. Ninomiya, P. Thorner, L. Guicharnaud, E. Boye, V. Chauvet, L.C. Solal, A. Beziau, R.G. Torres, C. Antignac, M.C. Gubler, Diffuse leiomyomatosis associated with X-linked Alport syndrome: extracellular matrix study using immunohistochemistry and in situ hybridization, Lab. Invest. 76 (1997) 233–243.
[9] T. Seki, I. Naito, T. Oohashi, Y. Sado, Y. Ninomiya, Differential expression of type IV collagen isoforms, α5(IV) and α6(IV) chains, in basement membranes surrounding smooth muscle cells, Histochem. Cell Biol. 110 (1998) 359–366.
[10] A. Simoneau, F.E. Herring-Gillam, P.H. Vachon, N. Perreault, N. Basora, Y. Bouatrouss, L.P. Pageot, J. Zhou, J.F. Beaulieu, Identification, distribution, and tissular origin of the α5(IV) and α6(IV) collagen chains in the developing human intestine, Dev. Dyn. 212 (1998) 437–447.
[11] E. Poschl, R. Pollner, K. Kuhn, The genes for the α1(IV) and α2(IV) chains of human basement membrane collagen type IV are arranged head-to-head and separated by a bidirectional promoter of unique structure, EMBO J. 7 (1988) 2687–2695.
[12] J. Zhou, T. Mochizuki, H. Smeets, C. Antignac, P. Laurila, A. de Paepe, K. Tryggvason, S.T. Reeders, Deletion of the paired α5(IV) and α6(IV) collagen genes in inherited smooth muscle tumors, Science 261 (1993) 1167–1169.
[13] M. Mariyama, K. Zheng, T.L. Yang-Feng, S.T. Reeders, Colocalization of the genes for the α3(IV) and α4(IV) chains of type IV collagen to chromosome 2 bands q35–q37, Genomics 13 (1992) 809–813.
[14] R. Morello, G. Zhou, S.D. Dreyer, S.J. Harvey, Y. Ninomiya, P.S. Thorner, J.H. Miner, W. Cole, A. Winterpacht, B. Zabel, K.C. Oberg, B. Lee, Regulation of glomerular basement membrane collagen expression by LMX1B contributes to renal disease in nail patella syndrome, Nat. Genet. 27 (2001) 205–208.
[15] Y. Segal, L. Zhuang, E. Rondeau, J.D. Sraer, J. Zhou, Regulation of the paired type IV collagen genes *COL4A5* and *COL4A6*. Role of the proximal promoter region, J. Biol. Chem. 276 (2001) 11791–11797.
[16] E.H. Mercer, G.W. Hoyle, R.P. Kapur, R.L. Brinster, R.D. Palmiter, The dopamine beta-hydroxylase gene promoter directs expression of *E. coli lacZ* to sympathetic and other neurons in adult transgenic mice, Neuron 7 (1991) 703–716.
[17] B. Hogan, Manipulating the mouse embryo: a laboratory manual, second ed., Cold Spring Harbor Laboratory Press, Plainview, NY, 1994.
[18] B. Morgenstern, K. Frech, A. Dress, T. Werner, DIALIGN: finding local similarities by multiple sequence alignment, Bioinformatics 14 (1998) 290–294.
[19] B. Gottgens, J.G. Gilbert, L.M. Barton, D. Grafham, J. Rogers, D.R. Bentley, A.R. Green, Long-range comparison of human and mouse *SCL* loci: localized regions of sensitivity to restriction endonucleases correspond precisely with peaks of conserved noncoding sequences, Genome Res. 11 (2001) 87–97.
[20] K. Saito, I. Naito, T. Seki, T. Oohashi, E. Kimura, R. Momota, Y. Kishiro, Y. Sado, H. Yoshioka, Y. Ninomiya, Differential expression of mouse α5(IV) and α6(IV) collagen genes in epithelial basement membranes, J. Biochem. (Tokyo) 128 (2000) 427–434.
[21] J. Zhou, A. Leinonen, K. Tryggvason, Structure of the human type IV collagen *COL4A5* gene, J. Biol. Chem. 269 (1994) 6608–6614.
[22] X. Zhang, J. Zhou, S.T. Reeders, K. Tryggvason, Structure of the human type IV collagen *COL4A6* gene, which is mutated in Alport syndrome-associated leiomyomatosis, Genomics 33 (1996) 473–479.
[23] R.C. Hardison, Conserved noncoding sequences are reliable guides to regulatory elements, Trends Genet. 16 (2000) 369–372.
[24] P. Braghetta, C. Fabbro, S. Piccolo, D. Marvulli, P. Bonaldo, D. Volpin, G.M. Bressan, Distinct regions control transcriptional activation of the α1(VI) collagen promoter in different tissues of transgenic mice, J. Cell Biol. 135 (1996) 1163–1177.
[25] K.A. Sharif, C. Li, L.J. Gudas, *cis*-acting DNA regulatory elements, including the retinoic acid response element, are required for tissue specific laminin B1 promoter/*lacZ* expression in transgenic mice, Mech. Dev. 103 (2001) 13–25.
[26] H. Nakagawa, T. Inomoto, A.K. Rustgi, A CACCC box-like *cis*-regulatory element of the Epstein-Barr virus *ED-L2* promoter interacts with a novel transcriptional factor in tissue-specific squamous epithelia, J. Biol. Chem. 272 (1997) 16688–16699.
[27] P.F. Parakkal, An electron microscopic study of esophageal epithelium in the newborn and adult mouse, Am. J. Anat. 121 (1967) 175–195.
[28] A. Grapin-Botton, D.A. Melton, Endoderm development: from patterning to organogenesis, Trends Genet. 16 (2000) 124–130.
[29] P. Simon-Assmann, F. Bouziges, J.N. Freund, F. Perrin-Schmitt, M. Kedinger, Type IV collagen mRNA accumulates in the mesenchymal compartment at early stages of murine developing intestine, J. Cell Biol. 110 (1990) 849–857.
[30] S.C. Kirkland, K. Henderson, Collagen IV synthesis is restricted to the enteroendocrine pathway during multilineage differentiation of human colorectal epithelial stem cells, J. Cell Sci. 114 (2001) 2055–2064.
[31] R.G. Price, R.G. Spiro, Studies on the metabolism of the renal glomerular basement membrane. Turnover measurements in the rat with the use of radiolabeled amino acids, J. Biol. Chem. 252 (1977) 8597–8602.
[32] J.S. Trier, C.H. Allan, D.R. Abrahamson, S.J. Hagen, Epithelial basement membrane of mouse jejunum. Evidence for laminin turnover along the entire crypt-villus axis, J. Clin. Invest. 86 (1990) 87–95.