The expanded collagen VI family: new chains and new questions

Jamie Fitzgerald[1,2], **Paul Holden**[1], and **Uwe Hansen**[3]

[1]Department of Orthopaedics and Rehabilitation, Oregon Health and Science University, Portland, OR 97239, USA

[2]Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, OR 97239, USA

[3]Institute for Physiological Chemistry and Pathobiochemistry, University Hospital of Muenster, 48129 Muenster, Germany

Abstract

Collagen VI is a component of the extracellular matrix of almost all connective tissues, including cartilage, bone, tendon, muscles and cornea, where it forms abundant and structurally unique microfibrils organized into different suprastructural assemblies. The precise role of collagen VI is not clearly defined although it is most abundant in the interstitial matrix of tissues and often found in close association with basement membranes. Three genetically distinct collagen VI chains, α1(VI), α2(VI) and α3(VI), encoded by the *COL6A1*, *COL6A2* and *COL6A3* genes, were first described more than 20 years ago. Their molecular assembly and role in congenital muscular dystrophy has been broadly characterized. In 2008, three additional collagen VI genes arrayed in tandem at a single gene locus on chromosome 3q in humans, and chromosome 9 in mice, were described. Following the naming scheme for collagens the new genes were designated *COL6A4*, *COL6A5* and *COL6A6* encoding the α4(VI), α5(VI) and α6(VI) chains, respectively. This review will focus on the current state of knowledge of the three new chains.

Keywords

Collagen VI; collagen assembly; extracellular matrix; skeletal muscle collagen

---

The collagens are the most abundant family of extracellular matrix (ECM) components. They are extremely diverse in structure and function, with at least 28 subtypes known to exist (collagens I to XXVIII), which are assembled from 46 distinct gene products (1–3). Most of these subtypes can be further sorted into sub-families on the basis of function or domain organization. These include the fibrillar collagens, beaded filament collagens, anchoring fibril collagens, the FACIT collagens, transmembrane collagens and the network-forming collagens.

One collagen subtype that is present in virtually all connective tissues is collagen VI. For more than 20 years three collagen VI chains (α1(VI), α2(VI) and α3(VI)) and the genes that

---

Correspondence: Jamie Fitzgerald, fitzgerj@ohsu.edu.

encode them (*COL6A1*, *COL6A2* and *COL6A3*) were known to exist (4–6). However, the recent discovery of three additional collagen VI genes at a single genomic locus on chromosome 3q in humans and chromosome 9 in mice expands the collagen VI family to six members (7,8). These new genes named *COL6A4*, *COL6A5* and *COL6A6*, encode the α4(VI), α5(VI) and α6(VI) chains, respectively.

Chromosomal organization

The collagen VI genes are present at three distinct chromosomal loci in mammalian genomes. *COL6A1* and *COL6A2* are arrayed in tandem (chr21 in humans and chr10 in mice) and *COL6A3* (on chr2 in humans and chr1 in mice). The chromosomal location of the three new collagen VI genes is conserved in all mammals with the exception of some primate species. In the vast majority of primate genomes the three new collagen VI genes are located in tandem in the same orientation and order (5′ to 3′): *Col6a4>Col6a5>Col6a6* (Figure 1). However, close inspection of the human collagen VI gene cluster on 3q24 revealed that only a remnant of *COL6A4* representing the 3′ end of the gene is present at the locus (7–9). The *COL6A4* exons on 3q encode most of the triple helix and all the C-terminal exons. Homology searching demonstrated that the 5′ half of *COL6A4*, together with other genes upstream of *COL6A4*, are located on the p arm of chromosome 3 and facing the opposite direction. This suggests that a pericentric inversion occurred sometime in the evolutionary history of humans effectively disrupting the *COL6A4* gene locus (7–9). Interrogation of genomes of other primate species revealed that two ape species (orangutan and gibbon), old world monkeys (rhesus and baboon), new world monkeys (squirrel monkey and marmoset) and two prosimian species (lemur and bushbaby) contain intact *Col6a4–Col6a5–Col6a6* gene clusters. In contrast to this and mirroring the human genome the chimpanzee and gorilla genomes contain a disrupted *Col6a4* gene locus. The pericentric inversion must have occurred following the orangutan split from the main primate lineage but before the subsequent hominidae radiation, currently estimated to be 8–16 MYA (10). Consequently, the separated halves of human *COL6A4* are assumed to be pseudogenes and have been named *COL6A4P1* for the 5′ end at 3p and *COL6A4P2* for the 3′ end at 3q. Inspection of other vertebrate genomes, including placental mammals, marsupials and reptiles, indicate that all vertebrates examined so far have an intact *Col6a4*, *Col6a5* and *Col6a6* gene locus confirming that this gene arrangement is the ancestral state (7).

Interestingly, a small number of ESTs representing the split halves of *COL6A4* are present in the EST database suggesting that the region is transcriptionally active. It is unclear whether these have functional significance or whether they represent background or stochastic transcription. The growing appreciation of the role of non-coding RNAs in mRNA stability and regulation means that this transcriptional activity cannot be ignored. This is especially relevant because transcripts originating from the *COL6A4P1* and *COL6A4P2* pseudogenes will have high homology with the other collagen VI paralogues and, potentially, be in a position to regulate expression of these genes via RNA-mediated mechanisms (11,12).

Domain organization

The domain structures of the six collagen VI chains (mouse Col6a4 and human COL6A1, 2, 3, 5 and 6) are presented in Figure 2. It is clear from the domain architecture and phylogenetic analysis that the new chains are most similar to the α3(VI) chain (7,8). For example, the α3(VI) chain and the new chains all have large globular N-termini comprising multiple von Willebrand factor A-like (VWA) domains. α3(VI) has ten VWA modules (denoted N1 to N10) and α4(VI), α5(VI) and α6(VI) chain each have seven VWA domains (N1 to N7), compared to just a single VWA domain in α1(VI) and α2(VI). The function of these N-terminal domains are poorly understood although the N-terminal globular domains of α3(VI) appear to function in the extracellular space, probably as potential binding sites for other ECM molecules. Studies on recombinantly produced α3(VI) suggest that portions of the N-terminus project away from the microfibril and are well-positioned to interact with other ECM components (13). Data from experiments in transfected SaOs-2 cells indicate that the N5 domain of α3(VI) plays an active role in microfibril formation in the extracellular space (14). It remains to be seen whether the N-globular domains of the new chains function in a similar manner.

Each of the new chains has a 336 amino acid Gly-X-Y repeat triple helix and a globular C-terminus composed of several different types of protein modules. All six collagen VI chains have a single Cys residue in the triple helical domain that is important for stabilizing higher-order assembly structures (15–17). In the α3, α4, α5 and α6 chains, this Cys is located 50 amino acids from the start of the triple helix (7,8). In contrast, the Cys residues in α1(VI) and α2(VI) are in a different relative position along the triple helix at amino acid 89. These similarities between the α3, α4, α5 and α6 chains suggests an assembly model for collagen VI where the α4, α5 and α6 can substitute for α3 in assembling with α1 and α2 chains (7,8) although this has yet to be directly demonstrated.

The C-terminal globular region of the α chains of collagen VI is the most variable by domain organization. All six chains have a pair of VWA domains immediately following the triple helix (C1 and C2), but then diverge substantially after that. The (mouse) α4 chain continues with a 100 amino acid sequence lacking homology to any reported domains (C3) and a partial Kunitz protease inhibitor domain. Interestingly, the α3(VI) chain contains an intact Kunitz inhibitor domain at the C-terminus and recent evidence suggests a role for this domain in tumor biology (18,19).

The α5 chain has the largest C-terminus of the three new chains and continues after C2 with a unique sequence of approximately 130 amino acids followed by another VWA domain and a second 130 amino acid stretch (C3, C4 and C5). The α6 chain continues after the two VWA domains with a unique stretch of 100 amino acids (C3). Interestingly, the α4-C3, α5-C3, α5-C5 and α6-C3 do not show homology with any other reported domain, but demonstrate a short sequence of homology to each other centered around a conserved Cys residue (see (8), Figure 3C). The function of this stretch that is unique to the collagen VI chains is currently unknown, but it may be important for collagen VI chain selection or assembly processes that are known to involve C-terminal elements (20,21).

COL6A4 gene and α4(VI) protein

The *col6a4* gene is the most centromeric of the collagen VI gene cluster on mouse chromosome 9. It comprises 38 exons with a predicted MW of 251 kDa. *Col6a4* expression appears to be developmentally regulated. mRNA for *Col6a4* is found in multiple newborn tissues including intestine, sternum and lung (22), although expression in adult tissues is restricted to ovary and uterus (7,8). Consistent with RT-PCR data in newborn mice, immunohistochemical staining for α4(VI) protein in E14.5 embryos demonstrates high expression levels in stomach, bronchi, midgut and ovary.

In 2007, it was reported that a common variant in the human *DVWA* gene is associated with knee osteoarthritis (OA) in Japanese and Chinese populations (23). The *DVWA* gene, located on the p arm of chromosome 3, is the 5′ part of the split *COL6A4* gene (9,24). However, the association of *DVWA* with knee OA was not found in a Korean population (25) nor in two studies of European OA individuals (26,27). These contrary findings could indicate differences in OA selection criteria or control selection, or simply reflect ethnic differences in OA susceptibility genes. However, arguing against a role for *DVWA* in cartilage biology is its lack of expression in human and mouse articular cartilage although as mentioned before, RNA-mediated mechanisms could not completely be excluded (9,24).

COL6A5 gene and α5(VI) protein

In mouse, *Col6a5* transcripts are present in virtually all tissues examined (7,8). At the protein level α5(VI) is detected in skeletal muscle and diaphragm, bronchi, kidney, ovary, testis and blood vessels (22). The expression of *COL6A5* mRNA in humans is much more restricted with transcripts present only in skin, lung, testis, colon and small intestine (7,28). The reason for more restricted expression of *COL6A5* in humans compared to mouse is not clear but may be related to the chromosome rearrangement that disrupted *COL6A4*. The break, which occurred approximately 130 kb upstream of the first protein coding exon of *COL6A5*, may have disrupted upstream enhancer elements important for *COL6A5* expression or conversely, introduced transcriptional repressor elements that repress transcription.

In 2007 Soderhall et al. (28) reported that variants of *COL6A5* were associated with atopic dermatitis using a family-based association analysis. This paper was published prior to the initial description of the 3p collagen VI cluster first by Gara et al. (8) and then later by Fitzgerald et al. (7) in 2008. Soderhall et al. (28) failed to recognize that *COL6A5* is a paralogue of the collagen VI genes and named the gene *COL29A1*, the next collagen gene name available. Those working in the field refer to this gene as *COL6A5* in keeping with the accepted collagen nomenclature (3,7,8). A role for α5(VI) in a skin disorder is consistent with its localization in skin (29). However, two independent studies failed to find an association between *COL6A5* SNPs and atopic dermatitis or eczema (30). In addition, an association between *COL6A5* SNPs and other autoimmune diseases including inflammatory bowel disease, Crohn's disease and ulcerative colitis was also not found (31). However, these are relatively small studies and the large datasets now being generated by exome and whole genome sequencing studies may shed light on the link between the *COL6A5* locus and inflammatory disease in humans.

COL6A6 gene and α6(VI) protein

The *COL6A6* gene is the third gene in the chromosome 3q collagen VI gene cluster. *COL6A6* mRNA is present in a wide range of fetal and adult tissues in mouse and human (7,8). Several recent papers have described the expression of the three new chains in mouse (22), and human skin and muscle (29,32). In skeletal muscle, a restricted and differential distribution of the α6(VI) chain was found, which is presumably connected with specific functions of the new chain in specialized ECM structures. In human skin the α6(VI) chain was localized around the vessels of the papillary and reticular dermis, with a weaker and discontinuous labeling below the dermal-epidermal junction zone whereas the α3(VI) chain is broadly distributed in the papillary and reticular dermis.

Role in muscle disease

Nearly 100 mutations have been described in the *COL6A1*, *COL6A2* and *COL6A3* genes that cause two types of congenital muscular dystrophies, Bethlem myopathy (BM) and Ulrich congenital muscular dystrophy (UCMD) (33–41). One study reports a homozygous mutation in *COL6A2* in a family with myosclerosis (42). However, mutations cannot be found in approximately 40% of BM and UCMD patients despite extensive mutation screening (37,41) suggesting that mutations in other genes besides *COL6A1*, *2* and *3* are involved in the molecular pathology of BM and UCMD. The discovery of two genes *COL6A5* and *COL6A6* raises the possibility that these genes are involved in BM and UCMD and can account for the "missing" mutations. In mouse skeletal muscle α5(VI) chains are primarily expressed in the perimysium and α6(VI) in the perimysium and endomysium, with both molecules showing partial overlap with α3(VI) (22). In human skeletal muscle the staining pattern is similar. Like the α3 chain, α6 is expressed in both the perimysium and endomysium with a discontinuous pattern around muscle fibers (32). While α3 is clearly part of the basement membrane, α6 appears to be a component of the interstitial tissue between adjacent fibers and absent from the basement membrane though there are some areas of overlapping expression. Since α1, α2 and α3 chains were found in basement membrane structures in mouse skeletal muscle by electron microscopy, the authors speculate that α1α2α3 heterotrimers represent the basement membrane form of collagen VI and that this form of collagen VI plays a critical role in basement membrane integrity (32). The α5(VI) chain in contrast, is present at the myotendinous junction and co-localizes with laminin 211, a basement membrane marker. α6(VI) was upregulated in skeletal muscle of patients with Duchenne muscular dystrophy (*DMD*) gene mutations and associated with increased fibrosis (32).

These observations suggest that *COL6A5* and *COL6A6* variants may be involved in neuromuscular disease. No mutations in *COL6A5* and *COL6A6* have been reported to date despite their inclusion in sequencing and CGH screening strategies (43). In addition, we have screened for mutations in patients presenting with "classic" BM and UCMD phenotypes and failed to find a mutation in *COL6A6* (unpublished data) The failure to find *COL6A6* mutations in particular is puzzling and suggests that *COL6A6* is not involved in the same congenital muscular dystrophies as *COL6A1*, *COL6A2* and *COL6A3*. Instead, α6(VI) may have a different role in muscle compared to α1, α2 and α3, perhaps related to

its expression in interstitial tissue instead of closely associated with the basement membrane. Other muscular dystrophies should be screened outside of the BM-UCMD spectrum.

Assembly of collagen VI molecules and higher order structures

The basic assembly steps in the formation of collagen VI molecules containing $\alpha 1$, $\alpha 2$ and $\alpha 3$ chains have been described from biochemical analyses of tissue extracts, cell transfection and human mutation studies. It is clear that these chains are present in collagen VI assemblages in a 1:1:1 ratio, with the assumption that $\alpha 1 \alpha 2 \alpha 3$ heterotrimers are the basic monomeric unit of collagen VI. As with all collagens, monomer assembly proceeds in a C- to N-terminal direction. Monomers then associate in an anti-parallel fashion to form dimers and dimers subsequently associate laterally to form tetramers. Dimer and tetramer assembly intermediates are stabilized by interchain disulfide bonds between Cys residues located in the triple helical domain of each chain. Each tetramer, which contains 12 distinct collagen VI chains, has a molecular weight in excess of 1 MDa. This large complex is assembled intracellularly so that only tetrameric collagen VI is secreted from the cells. Individual tetramers then associate end-to-end to form microfibrillar structures in the extracellular matrix by an unknown mechanism. Other than this basic description, few details about the collagen VI assembly process are known. The new chains add an additional layer of complexity to the process of assembly. We and others have speculated that the $\alpha 4$, $\alpha 5$ and $\alpha 6$ chains substitute for the $\alpha 3$ chain because of the overall similarity of these four chains (7,8) and their relative expression patterns in muscular dystrophy (29). For example, all collagen VI chains have a single triple helix Cys residue that participates in stabilizing assembly intermediates. In the $\alpha 3$, $\alpha 4$, $\alpha 5$ and $\alpha 6$ chains this Cys is in the same relative position 50 amino acids from the start of the Gly-XY repeat motif. The Cys residues in the $\alpha 1$ and $\alpha 2$ chains are further along the triple helix at position 89. The same relative positioning of triple helical Cys residues in $\alpha 3$, $\alpha 4$, $\alpha 5$ and $\alpha 6$ suggests that if the $\alpha 4$, $\alpha 5$ and $\alpha 6$ chains follow a similar assembly pathway with monomers, dimers and tetramers, then these new chains can substitute directly for $\alpha 3$(VI). In support of this hypothesis is the finding that in the *Col6a1*-null mice, no other collagen VI chains can be detected, including the new chains (8). This failure to identify the other collagen VI chains in mice deleted for *Col6a1* suggests that the $\alpha 1$ chain plays a key role in collagen VI assembly so that in its absence the other chains fail to assemble properly and presumably are degraded. While this seems reasonable, so far there is no direct biochemical evidence showing that the new chains assemble directly with the $\alpha 1$ chain. This is an important issue to resolve because, aside from the need to understand all possible assembly combinations, this information will be critical for interpreting mutations in the collagen VI-myopathies. In many tissues the protein expression of the $\alpha 5$(VI) and $\alpha 6$(VI) chains overlaps with that of the $\alpha 1$, $\alpha 2$ and $\alpha 3$ chains (11,29,32). However, in skeletal muscle it is clear that there are regions where $\alpha 5$ or $\alpha 6$ chains are expressed and $\alpha 1$, $\alpha 2$ and $\alpha 3$ chains are not expressed, such as at the neuromuscular junction. Also, in adult mouse skeletal muscle $\alpha 5$ and $\alpha 6$ chains are present in some regions that lack $\alpha 1$, $\alpha 2$ and $\alpha 3$ chains (see Figure 2F and I in Ref (22)) (11,32). These findings are potentially very significant because they imply that in some instances, $\alpha 6$ may not assemble with $\alpha 1$ and $\alpha 2$ chains, but actually be part of alternative collagen VI assemblies. If correct, this would force a re-evaluation of the collagen VI assembly paradigm.

Conclusions

The new collagen VI chains add new complexity to the collagen VI story as their roles in development, tissue homeostasis and the pathogenesis of inherited diseases are completely unknown. Moreover, the discovery of the new chains demonstrates that collagen VI is more than a ubiquitous component of connective tissues. Whereas the "classical" collagen VI chains are broadly distributed in almost all tissues it seems that the new α4, α5 and α6 chains are more restricted in their distribution, which implies potential new and specific functions for these chains.

Since the supramolecular aggregates of collagen VI are composite structures, a major challenge will be to describe how the new chains contribute to collagen VI supramolecular structures. One immediate question is whether the new chains substitute for α3(VI) in assembling with α1 and α2 to form α125 and α126 heterotrimers as has been suggested (7,8,29,32). While it has been assumed that this is likely to be the case, direct biochemical evidence is lacking. Also, in cells that synthesize more than one different collagen VI chain, are distinct pools of monomers and dimers produced to create a complex blend of more than one type of tetramer? In addition, since the final stage of microfibril assembly via end-to-end tetramer formation occurs in the extracellular space, do different cells produce different types of tetramers to create collagen VI "alloys" that vary by tissue type, or within tissues? For example, is the collagen VI composition at the myotendinous junction (32) different from collagen VI in the mid-tendon, and how do these differences relate to different functional properties? These questions will require sensitive investigative tools such as chain-specific antibodies and detailed *in vitro* experiments complemented by studies in tissues and animal models to answer. In addition, these findings will provide insights not only for this ubiquitous collagen subtype, but also for collagens in general.

Acknowledgments

References

1. Myllyharju J, Kivirikko KI. Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet. 2004; 20:33–43. [PubMed: 14698617]

2. Gordon MK, Hahn RA. Collagens. Cell Tissue Res. 2010; 339:247–257. [PubMed: 19693541]

3. Ricard-Blum S. The collagen family. Cold Spring Harb Perspect Biol. 2011; 3:a004978. [PubMed: 21421911]

4. Chu ML, Mann K, Deutzmann R, Pribula-Conway D, Hsu-Chen CC, Bernard MP, Timpl R. Characterization of three constituent chains of collagen type VI by peptide sequences and cDNA clones. Eur J Biochem. 1987; 168:309–317. [PubMed: 3665927]

5. Chu ML, Pan TC, Conway D, Kuo HJ, Glanville RW, Timpl R, Mann K, Deutzmann R. Sequence analysis of alpha 1(VI) and alpha 2(VI) chains of human type VI collagen reveals internal triplication of globular domains similar to the A domains of von Willebrand factor and two alpha 2(VI) chain variants that differ in the carboxy terminus. EMBO J. 1989; 8:1939–1946. [PubMed: 2551668]

6. Chu ML, Zhang RZ, Pan TC, Stokes D, Conway D, Kuo HJ, Glanville R, Mayer U, Mann K, Deutzmann R, Timpl R. Mosaic structure of globular domains in the human type VI collagen alpha 3 chain: similarity to von Willebrand factor, fibronectin, actin, salivary proteins and aprotinin type protease inhibitors. EMBO J. 1990; 9:385–393. [PubMed: 1689238]

7. Fitzgerald J, Rich C, Zhou FH, Hansen U. Three novel collagen VI chains, alpha4(VI), alpha5(VI), and alpha6(VI). J Biol Chem. 2008; 283:20170–20180. [PubMed: 18400749]

8. Gara SK, Grumati P, Urciuolo A, Bonaldo P, Kobbe B, Koch M, Paulsson M, Wagener R. Three novel collagen VI chains with high homology to the alpha3 chain. J Biol Chem. 2008; 283:10658–10670. [PubMed: 18276594]

9. Wagener R, Gara SK, Kobbe B, Paulsson M, Zaucke F. The knee osteoarthritis susceptibility locus DVWA on chromosome 3p24.3 is the 5′ part of the split COL6A4 gene. Matrix Biol. 2009; 28:307–310. [PubMed: 19486942]

10. Kehrer-Sawatzki H, Cooper DN. Understanding the recent evolution of the human genome: insights from human-chimpanzee genome comparisons. Hum Mutat. 2007; 28:99–130. [PubMed: 17024666]

11. Mello CC, Conte D Jr. Revealing the world of RNA interference. Nature. 2004; 431:338–342. [PubMed: 15372040]

12. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet. 2008; 9:102–114. [PubMed: 18197166]

13. Beecher N, Roseman AM, Jowitt TA, Berry R, Troilo H, Kammerer RA, Shuttleworth CA, Kielty CM, Baldock C. Collagen VI, conformation of A-domain arrays and microfibril architecture. J Biol Chem. 2011; 286:40266–40275. [PubMed: 21908605]

14. Fitzgerald J, Morgelin M, Selan C, Wiberg C, Keene DR, Lamande SR, Bateman JF. The N-terminal N5 subdomain of the alpha 3(VI) chain is important for collagen VI microfibril formation. J Biol Chem. 2001; 276:187–193. [PubMed: 11027693]

15. Furthmayr H, Wiedemann H, Timpl R, Odermatt E, Engel J. Electron-microscopical approach to a structural model of intima collagen. Biochem J. 1983; 211:303–311. [PubMed: 6307276]

16. Chu ML, Conway D, Pan TC, Baldwin C, Mann K, Deutzmann R, Timpl R. Amino acid sequence of the triple-helical domain of human collagen type VI. J Biol Chem. 1988; 263:18601–18606. [PubMed: 3198591]

17. Baker NL, Morgelin M, Peat R, Goemans N, North KN, Bateman JF, Lamande SR. Dominant collagen VI mutations are a common cause of Ullrich congenital muscular dystrophy. Hum Mol Genet. 2005; 14:279–293. [PubMed: 15563506]

18. Nanda A, Carson-Walter EB, Seaman S, Barber TD, Stampfl J, Singh S, Vogelstein B, Kinzler KW, St Croix B. TEM8 interacts with the cleaved C5 domain of collagen alpha 3(VI). Cancer Res. 2004; 64:817–820. [PubMed: 14871805]

19. Park J, Scherer PE. Adipocyte-derived endotrophin promotes malignant tumor progression. J Clin Invest. 2012; 122:4243–4256. [PubMed: 23041627]

20. Tooley LD, Zamurs LK, Beecher N, Baker NL, Peat RA, Adams NE, Bateman JF, North KN, Baldock C, Lamande SR. Collagen VI microfibril formation is abolished by an {alpha}2(VI) von Willebrand factor type A domain mutation in a patient with Ullrich congenital muscular dystrophy. J Biol Chem. 2010; 285:33567–33576. [PubMed: 20729548]

21. Ball SG, Baldock C, Kielty CM, Shuttleworth CA. The role of the C1 and C2 a-domains in type VI collagen assembly. J Biol Chem. 2001; 276:7422–7430. [PubMed: 11036066]

22. Gara SK, Grumati P, Squarzoni S, Sabatelli P, Urciuolo A, Bonaldo P, Paulsson M, Wagener R. Differential and restricted expression of novel collagen VI chains in mouse. Matrix Biol. 2011; 30:248–257. [PubMed: 21477648]

23. Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, Fujioka M, Sudo A, Uchida A, Yamamoto S, Ozaki K, Takigawa M, Tanaka T, Nakamura Y, Jiang Q, Ikegawa S. A functional polymorphism in the 5′ UTR of GDF5 is associated with susceptibility to osteoarthritis. Nat Genet. 2007; 39:529–533. [PubMed: 17384641]

24. Nakajima M, Miyamoto Y, Ikegawa S. Cloning and characterization of the osteoarthritis-associated gene DVWA. J Bone Miner Metab. 2011; 29:300–308. [PubMed: 21057832]

25. Lee SJ, Kim MJ, Kee SJ, Song SK, Kweon SS, Shin MH, Park DJ, Park YW, Lee SS, Kim TJ. Association study of the candidate gene for knee osteoarthritis in Koreans. Rheumatol Int. 2013; 33:783–786. [PubMed: 22083612]

26. Meulenbelt I, Chapman K, Dieguez-Gonzalez R, Shi D, Tsezou A, Dai J, Malizos KN, Kloppenburg M, Carr A, Nakajima M, van der Breggen R, Lakenberg N, Gomez-Reino JJ, Jiang Q, Ikegawa S, Gonzalez A, Loughlin J, Slagboom EP. Large replication study and meta-analyses of DVWA as an osteoarthritis susceptibility locus in European and Asian populations. Hum Mol Genet. 2009; 18:1518–1523. [PubMed: 19181678]

27. Valdes AM, Spector TD, Doherty S, Wheeler M, Hart DJ, Doherty M. Association of the DVWA and GDF5 polymorphisms with osteoarthritis in UK populations. Ann Rheum Dis. 2009; 68:1916–1920. [PubMed: 19054821]

28. Soderhall C, Marenholz I, Kerscher T, Ruschendorf F, Esparza-Gordillo J, Worm M, Gruber C, Mayr G, Albrecht M, Rohde K, Schulz H, Wahn U, Hubner N, Lee YA. Variants in a novel epidermal collagen gene *(COL29A1)* are associated with atopic dermatitis. PLoS Biol. 2007; 5:1952–1961.

29. Sabatelli P, Gara SK, Grumati P, Urciuolo A, Gualandi F, Curci R, Squarzoni S, Zamparelli A, Martoni E, Merlini L, Paulsson M, Bonaldo P, Wagener R. Expression of the collagen VI alpha5 and alpha6 chains in normal human skin and in skin of patients with collagen VI-related myopathies. J Invest Dermatol. 2011; 131:99–107. [PubMed: 20882040]

30. Harazin M, Parwez Q, Petrasch-Parwez E, Epplen JT, Arinir U, Hoffjan S, Stemmler S. Variation in the COL29A1 gene in German patients with atopic dermatitis, asthma and chronic obstructive pulmonary disease. J Dermatol. 2010; 37:740–742. [PubMed: 20649719]

31. Zucchelli M, Torkvist L, Bresso F, Halfvarson J, Soderhall C, Lee YA, Lofberg R, Kere J, D'Amato M. No association between the eczema genes COL29A1 and IL31 and inflammatory bowel disease. Inflamm Bowel Dis. 2009; 15:961–962. [PubMed: 18839421]

32. Sabatelli P, Gualandi F, Gara SK, Grumati P, Zamparelli A, Martoni E, Pellegrini C, Merlini L, Ferlini A, Bonaldo P, Maraldi NM, Paulsson M, Squarzoni S, Wagener R. Expression of collagen VI alpha5 and alpha6 chains in human muscle and in Duchenne muscular dystrophy-related muscle fibrosis. Matrix Biol. 2012; 31:187–196. [PubMed: 22226732]

33. Demir E, Ferreiro A, Sabatelli P, Allamand V, Makri S, Echenne B, Maraldi M, Merlini L, Topaloglu H, Guicheney P. Collagen VI status and clinical severity in Ullrich congenital muscular dystrophy: phenotype analysis of 11 families linked to the COL6 loci. Neuropediatrics. 2004; 35:103–112. [PubMed: 15127309]

34. Demir E, Sabatelli P, Allamand V, Ferreiro A, Moghadaszadeh B, Makrelouf M, Topaloglu H, Echenne B, Merlini L, Guicheney P. Mutations in COL6A3 cause severe and mild phenotypes of Ullrich congenital muscular dystrophy. Am J Hum Genet. 2002; 70:1446–1458. [PubMed: 11992252]

35. Giusti B, Lucarini L, Pietroni V, Lucioli S, Bandinelli B, Sabatelli P, Squarzoni S, Petrini S, Gartioux C, Talim B, Roelens F, Merlini L, Topaloglu H, Bertini E, Guicheney P, Pepe G. Dominant and recessive COL6A1 mutations in Ullrich scleroatonic muscular dystrophy. Ann Neurol. 2005; 58:400–410. [PubMed: 16130093]

36. Lucarini L, Giusti B, Zhang RZ, Pan TC, Jimenez-Mallebrera C, Mercuri E, Muntoni F, Pepe G, Chu ML. A homozygous COL6A2 intron mutation causes in-frame triple-helical deletion and nonsense-mediated mRNA decay in a patient with Ullrich congenital muscular dystrophy. Hum Genet. 2005; 117:460–466. [PubMed: 16075202]

37. Lucioli S, Giusti B, Mercuri E, Vanegas OC, Lucarini L, Pietroni V, Urtizberea A, Ben Yaou R, de Visser M, van der Kooi AJ, Bonnemann C, Iannaccone ST, Merlini L, Bushby K, Muntoni F, Bertini E, Chu ML, Pepe G. Detection of common and private mutations in the COL6A1 gene of patients with Bethlem myopathy. Neurology. 2005; 64:1931–1937. [PubMed: 15955946]

38. Lamande SR, Shields KA, Kornberg AJ, Shield LK, Bateman JF. Bethlem myopathy and engineered collagen VI triple helical deletions prevent intracellular multimer assembly and protein secretion. J Biol Chem. 1999; 274:21817–21822. [PubMed: 10419498]

39. Lampe AK, Bushby KM. Collagen VI related muscle disorders. J Med Genet. 2005; 42:673–685. [PubMed: 16141002]

40. Jimenez-Mallebrera C, Maioli MA, Kim J, Brown SC, Feng L, Lampe AK, Bushby K, Hicks D, Flanigan KM, Bonnemann C, Sewry CA, Muntoni F. A comparative analysis of collagen VI production in muscle, skin and fibroblasts from 14 Ullrich congenital muscular dystrophy patients with dominant and recessive COL6A mutations. Neuromuscul Disord. 2006; 16:571–582. [PubMed: 16935502]

41. Lampe AK, Dunn DM, von Niederhausern AC, Hamil C, Aoyagi A, Laval SH, Marie SK, Chu ML, Swoboda K, Muntoni F, Bonnemann CG, Flanigan KM, Bushby KM, Weiss RB. Automated genomic sequence analysis of the three collagen VI genes: applications to Ullrich congenital muscular dystrophy and Bethlem myopathy. J Med Genet. 2005; 42:108–120. [PubMed: 15689448]

42. Merlini L, Martoni E, Grumati P, Sabatelli P, Squarzoni S, Urciuolo A, Ferlini A, Gualandi F, Bonaldo P. Autosomal recessive myosclerosis myopathy is a collagen VI disorder. Neurology. 2008; 71:1245–1253. [PubMed: 18852439]

43. Bovolenta M, Neri M, Martoni E, Urciuolo A, Sabatelli P, Fabris M, Grumati P, Mercuri E, Bertini E, Merlini L, Bonaldo P, Ferlini A, Gualandi F. Identification of a deep intronic mutation in the COL6A2 gene by a novel custom oligonucleotide CGH array designed to explore allelic and genetic heterogeneity in collagen VI-related myopathies. BMC Med Genet. 2010; 11:44. [PubMed: 20302629]
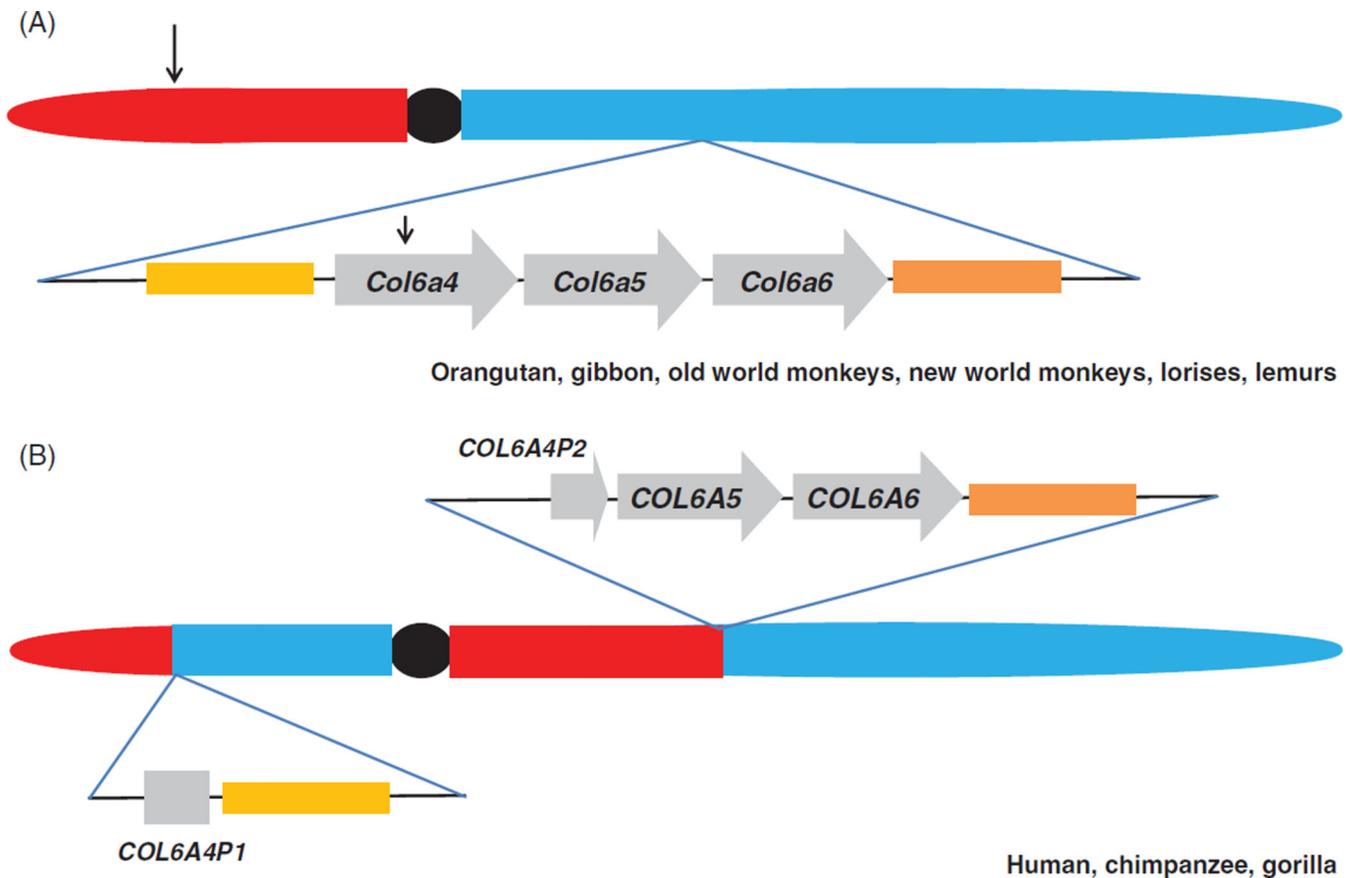
Figure 1.
The collagen VI gene locus is disrupted in a subset of primate species. (A) Schematic showing the organization of the collagen VI gene locus in several ape (orangutan, gibbon), old world monkey (rhesus monkey, baboon), new world monkey (squirrel monkey, marmoset), lemur (mouse lemur) and loris (bushbaby) genomes. The p arm (to the left of the centromere, in red), q arm (to the right of the centromere, in blue) and centromere (filled black circle) of the chromosome containing the collagen VI gene locus is shown. Below the chromosome, the *Col6a4*, *Col6a5* and *Col6a6* genes are represented by *arrowed boxes* to indicate the orientation of each gene. All genes flanking the collagen VI gene locus are conserved in all these primate species (orange and yellow boxes flanking collagen VI gene cluster). Note that this gene organization is the ancestral state because it is conserved in all vertebrate species that genome data is available for including mouse, rat, dog, platypus and lizard. (B) The same locus in the human, chimpanzee and gorilla genomes. In these species the locus has been disrupted by an evolutionary pericentric inversion with one breakpoint located within the *COL6A4* gene and the other breakpoint in the p arm of chromosome 3 (in humans) represented by arrows in (A). In human, chimpanzee and gorilla, genomic DNA telomeric to the inversion remains contiguous with *COL6A5* and *COL6A6*. The block of genomic DNA on the centromeric side of the break has been translocated to the p arm of chromosome 3. The disrupted halves of the *COL6A4* gene are presumably non-functional and have been designated *COL6A4P1* and *P2*.
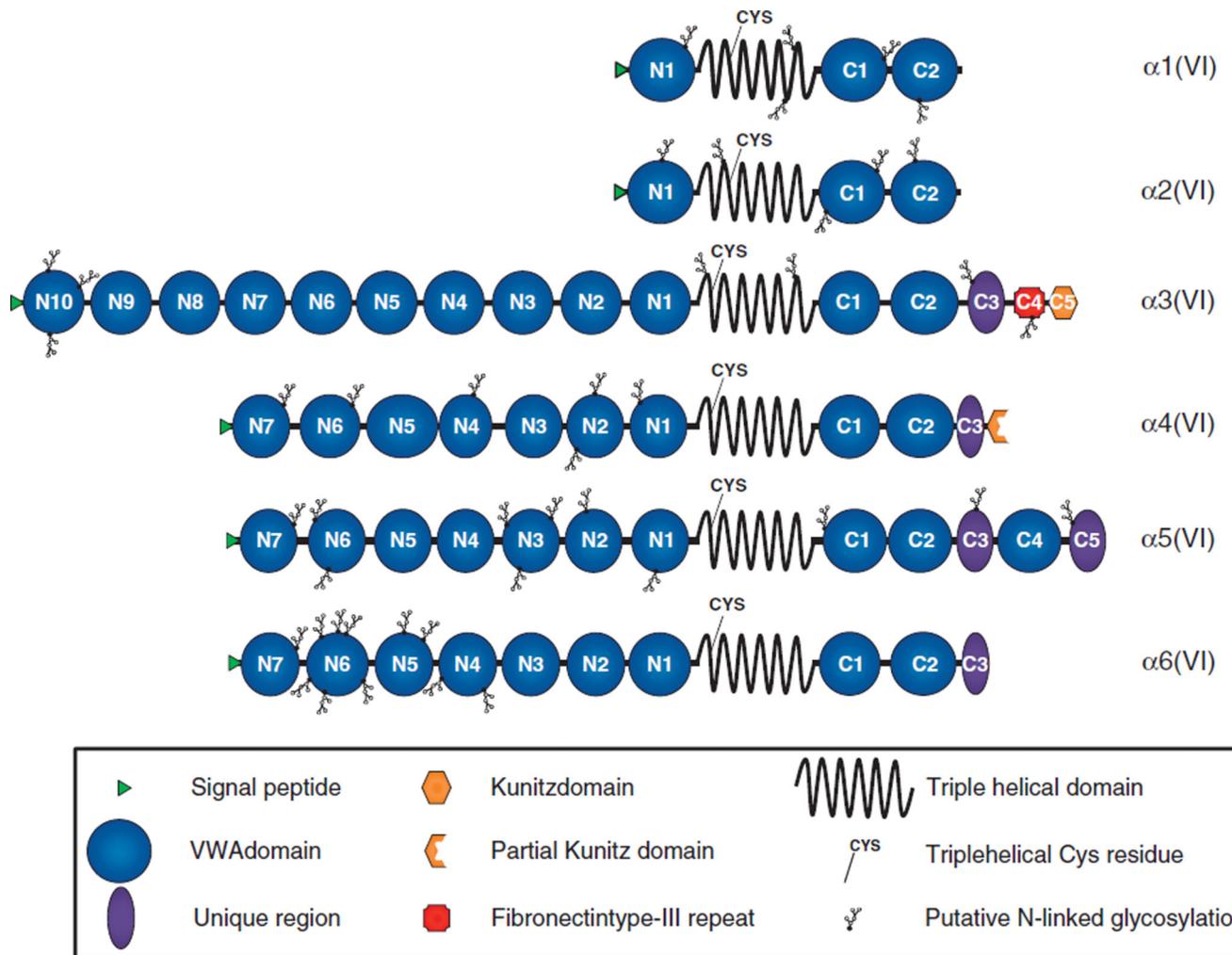
Figure 2.
Domain structures of the six collagen VI chains. Representation of the human α1(VI), α2(VI), α3(VI), α5(VI) and α6(VI), and mouse α4(VI) collagen chains. Each chain is drawn to the same relative scale to reflect the differences in size between the six chains. Multiple N- and C-terminal von Willebrand factor A-like (VWA) domains (blue circles) are numbered consistent with previous reports (8,15). The fibronectin type-III repeat unique to α3(VI) (red square labelled C4), Kunitz and partial Kunitz domains of α3(VI) and α4(VI) respectively (orange hexagon labelled C5 and partial hexagon) and a conserved region unique to the α3, α4, α5 and α6 chains (purple oval shapes) are shown. The position of the triple helical cysteine (Cys) residue important for stabilizing assembly intermediates is shown in all six chains. These are located at amino acid 89 from the start of the triple helix domain in α1(VI) and α2(VI) and closer to the start of the triple helix at amino acid 50 in the α4, α5 and α6(VI) chains. Putative N-linked glycosylation sites were determined using NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/).