

# Automated genomic sequence analysis of the three collagen VI genes: applications to Ullrich congenital muscular dystrophy and Bethlem myopathy

A K Lampe, D M Dunn, A C von Niederhausern, C Hamil, A Aoyagi, S H Laval, S K Marie, M-L Chu, K Swoboda, F Muntoni, C G Bonnemann, K M Flanigan, K M D Bushby, R B Weiss

*J Med Genet* 2005;42:108–120. doi: 10.1136/jmg.2004.023754

See end of article for authors' affiliations

Correspondence to: Professor Kate Bushby, Institute of Human Genetics, University of Newcastle upon Tyne, International Centre for Life, Central Parkway, Newcastle upon Tyne, NE1 3BZ, UK; Kate.Bushby@newcastle.ac.uk

Revised version received 21 August 2004  
Accepted for publication 23 August 2004

**Introduction:** Mutations in the genes encoding collagen VI (*COL6A1*, *COL6A2*, and *COL6A3*) cause Bethlem myopathy (BM) and Ullrich congenital muscular dystrophy (UCMD). BM is a relatively mild dominantly inherited disorder with proximal weakness and distal joint contractures. UCMD is an autosomal recessive condition causing severe muscle weakness with proximal joint contractures and distal hyperlaxity.

**Methods:** We developed a method for rapid direct sequence analysis of all 107 coding exons of the *COL6* genes using single condition amplification/internal primer (SCAIP) sequencing. We have sequenced all three *COL6* genes from genomic DNA in 79 patients with UCMD or BM.

**Results:** We found putative mutations in one of the *COL6* genes in 62% of patients. This more than doubles the number of identified *COL6* mutations. Most of these changes are consistent with straightforward autosomal dominant or recessive inheritance. However, some patients showed changes in more than one of the *COL6* genes, and our results suggest that some UCMD patients may have dominantly acting mutations rather than recessive disease.

**Discussion:** Our findings may explain some or all of the cases of UCMD that are unlinked to the *COL6* loci under a recessive model. The large number of single nucleotide polymorphisms which we generated in the course of this work may be of importance in determining the major phenotypic variability seen in this group of disorders.

Extracellular matrix molecules are critical for skeletal muscle stability, regeneration, and muscle cell matrix adhesion.<sup>1–3</sup> Collagen VI is a ubiquitous extracellular matrix protein<sup>4</sup> which forms a microfibrillar network in close association with the basement membrane around muscle cells and which is capable of interacting with several other matrix constituents.<sup>5–7</sup> Collagen VI is composed of three different peptide chains:  $\alpha 1(VI)$  and  $\alpha 2(VI)$ , both 140 kDa in size, and  $\alpha 3(VI)$  which is much larger at 260–300 kDa.<sup>8</sup> The  $\alpha 1(VI)$  and  $\alpha 2(VI)$  chains are encoded by two genes (*COL6A1* and *COL6A2*, respectively) situated in a head-to-tail organisation on chromosome 21q22.3<sup>9</sup> (NT\_011515) and separated by 150 kb of genomic DNA, whereas *COL6A3*, the gene for the  $\alpha 3(VI)$  chain, maps to chromosome 2q37<sup>10</sup> (NT\_005120). All three chains contain a central short triple helical domain of 335–336 amino acids with repeating Gly-Xaa-Yaa sequences flanked by large N- and C-terminal globular domains consisting of motifs of 200 amino acids each homologous to von Willebrand factor (vWF) type A domains.<sup>11</sup> In total, the three *COL6* genes consist of 107 coding exons spread over 150 000 bp of genomic DNA. *COL6A2* and *COL6A3* undergo extensive alternative splicing.<sup>12–14</sup>

The assembly of collagen VI is a complex multistep process. Association of the three subunits,  $\alpha 1(VI)$ ,  $\alpha 2(VI)$ , and  $\alpha 3(VI)$ , to form a triple helical monomer is followed by staggered assembly into disulfide bonded antiparallel dimers, which then align to form tetramers, also stabilised by disulfide bonds. Outside the cell, tetramers, the secreted form of collagen VI, associate end-to-end to form the characteristic beaded microfibrils.<sup>8 15 16</sup>

Mutations in the genes encoding collagen VI cause Bethlem myopathy (MIM 158810) and Ullrich congenital

muscular dystrophy (MIM 254090), two increasingly recognised conditions which were previously believed to be completely separate entities. Bethlem myopathy (BM) is a dominantly inherited disorder which usually follows a relatively mild course and is characterised by proximal muscle weakness and joint contractures mainly involving the elbows, ankles, and fingers.<sup>17 18</sup> Where contractures are prominent, this disorder may resemble an Emery-Dreifuss muscular dystrophy.<sup>19</sup> In other patients the contractures may be subtle, leading to confusion in diagnosis with cases of limb-girdle muscular dystrophy.<sup>20</sup> By contrast, Ullrich congenital muscular dystrophy (UCMD) is an autosomal recessive condition causing severe muscle weakness of early onset such that children may never walk independently, with proximal joint contractures and striking hyperelasticity of distal joints, and early respiratory failure.<sup>21–23</sup>

To date, 11 different dominant mutations in the *COL6* genes have been published in BM patients. Truncating and splicing mutations in *COL6A1* affecting protein folding and secretion as well as a missense mutation in the *COL6A3* N-terminal  $\alpha 3(VI)$  domain are thought to cause BM via functional haploinsufficiency.<sup>24–28</sup> In other cases single amino acid substitutions disrupting the Gly-Xaa-Yaa motif of the highly conserved triple helical domain of any of the three collagen VI chains<sup>20 29 30</sup> appear to interfere with extracellular microfibril formation. Ten recessive truncating mutations in *COL6A2* and *COL6A3* have been described in eight families with UCMD, eight in *COL6A2*,<sup>31–34</sup> and two in *COL6A3*.<sup>35</sup> The

**Abbreviations:** BM, Bethlem myopathy; PTC, premature termination codon; SCAIP, single condition amplification/internal primer; UCMD, Ullrich congenital muscular dystrophy; vWF, von Willebrand factor

mild phenotype of one of the latter, where a homozygous nonsense mutation leads to a stop codon in exon 5 of *COL6A3*, can possibly be explained by alternative splicing. In addition, a homozygous splice site mutation causing an in-frame deletion of 17 amino acids of the triple helical region of the  $\alpha 3$ (VI) chain has been associated with a comparatively mild phenotype.<sup>35</sup> Recently, there has also been one report of an apparently dominant de novo heterozygous in-frame deletion of exons 9 and 10 of the  $\alpha 1$ (VI) chain in a patient with a classical UCMD phenotype.<sup>36</sup>

In the patients with UCMD and *COL6* mutations reported so far, immunohistochemical analysis of muscle biopsy samples shows absent or severely reduced collagen VI expression especially in the basement membrane, while in most patients with BM collagen VI immunolabelling is normal (though a secondary reduction in laminin  $\beta 1$  has been reported in some patients). A few patients have also been recently described where sarcolemmal collagen VI expression is abnormal but mutations in the *COL6* genes have been excluded by RT-PCR and cDNA sequencing.<sup>32</sup> Because of the complexities of genetic counselling caused by the possibility of dominant as well as recessive mutations in the collagen VI genes, and the possibility that collagen VI abnormalities may also occur as a secondary phenomenon, accurate molecular diagnosis becomes very important. As there are currently no other techniques available for routine molecular diagnosis in this group of diseases, a rate limiting step in diagnosis is the ability to offer mutation detection.

We have developed a method for rapid, robust, and economical direct sequence analysis of all 107 exons of the *COL6* genes using single condition amplification/internal primer (SCAIP) sequencing. SCAIP relies on amplification of a large number of exons under a single set of PCR temperatures. Sufficient sequencing specificity is obtained by uniform use of a second, internal set of sequencing primers without optimisation of individual amplicon conditions.

SCAIP was developed for direct sequence analysis of the dystrophin gene<sup>37</sup> and avoids intermediate screening steps such as single-stranded conformation polymorphism (SSCP) or denaturing high-performance liquid chromatography (DHPLC). It is now used for the routine molecular diagnosis of dystrophinopathies. This is the first study to perform genomic sequence analysis of all three *COL6* genes. The three *COL6* genes were sequenced from genomic DNA in 79 patients with UCMD or BM, achieving over 97% coverage. Potential mutations were found in 62% of patients, thereby more than doubling the number of previously identified *COL6* mutations. Most of these putative mutations are consistent with straightforward autosomal dominant or recessive inheritance, though our results suggest that a proportion of UCMD patients may have dominant mutations rather than recessive disease. In addition, we have discovered in several UCMD patients examples of putative recessive mutations in each of two different *COL6A* genes, potentially indicating a novel mode of disease causation or modification in this group.

## METHODS

### Patients

We ascertained 79 patients from four muscle centres with clinical phenotypes suggestive of either UCMD (see Pepe *et al*<sup>19</sup> for diagnostic criteria) or BM (see Pepe *et al*<sup>38</sup> for diagnostic criteria) and also included patients with overlapping/intermediate phenotypes in the BM cohort who had features of both connective tissue as well as skeletal muscle involvement. Of the patients, 26 were UCMD patients and 53 BM, of which nine were sub-characterised as "mild BM" and nine as "severe BM". Their genomic DNA was extracted from peripheral lymphocytes or dermal fibroblast cultures by

standard methods. The genetic studies were performed in accordance with the ethical procedures of the participating institutions. The mutations in individuals #71 (harbouring a heterozygous splice donor change at *COL6A1* intron 14+1) and #72 (carrying a heterozygous internal *COL6A1* gene deletion which removes exons 9 and 10) have previously been published<sup>36</sup> and were used as positive controls.

### SCAIP sequencing

The genomic organisation of the *COL6* genes was assembled from contigs downloaded from the Nov2002 assembly of the UCSC Human Genome Browser web page<sup>39</sup> (<http://genome.cse.ucsc.edu/>; see also the International Human Genome Sequencing Consortium 2001<sup>40</sup>). Assembly and exon-intron annotation was performed using task specific Perl scripts. Our assembly reveals that the collagen VI regions on chromosome 21q22.3 and 2q37 are currently contiguous and gap free for the *COL6A1* transcript (accession number NM\_001848), the *COL6A2* isoforms 2c2 (accession number NM\_001849), 2c2a (accession number NM\_058174), and 2c2a' (NM\_058175) as well as the full length isoform of *COL6A3* (accession number NM\_004369) and its four variant transcripts (NM\_057164, NM\_057165, NM\_057166, and NM\_057167). Primer systems for PCR were designed to span each coding exon including splice variants. Each amplicon was designed for an optimal size range of 1.2–1.4 kb with the exon centred within the amplicon to maintain uniform conditions. This resulted in 103 amplicons of nearly uniform size with the uniformity allowing prediction of likely amplification conditions using a single set of PCR temperatures. Aliquots (15 pmol) of each primer were placed into individual wells of a 96 well plate, evaporated to dryness, and stored at  $-20^{\circ}\text{C}$  until use. For PCR amplification, 25  $\mu\text{l}$  of a master PCR mixture were aliquoted into each well of a thawed plate so that each well contained 100 ng of genomic DNA, 200  $\mu\text{M}$  dNTPs, 15 pmol of each PCR primer, 1 $\times$ reaction buffer, 2 mM  $\text{Mg}_2\text{SO}_4$ , 1 M GC-RICH resolution solution (Roche, Indianapolis, IN, USA), and 0.5 U Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA, USA). The PCR conditions were 2 min at  $94^{\circ}\text{C}$ ; 10 cycles of  $94^{\circ}\text{C}$  for 10 s,  $55^{\circ}\text{C}$  for 30 s, and  $68^{\circ}\text{C}$  for 2 min; 20 cycles of  $94^{\circ}\text{C}$  for 10 s,  $55^{\circ}\text{C}$  for 30 s, and  $68^{\circ}\text{C}$  for 2 min, elongated by 20 s/cycle. The PCR products were then bound to a 96 well filter plate with a 1.0  $\mu\text{m}$  glass fibre type B filter (Millipore, Billerica, MA, USA) in the presence of 5 M guanidine HCl/1 M potassium acetate solution and washed with 80% ethanol followed by elution with warm nanopure  $\text{H}_2\text{O}$ .

Internal sequencing primers were designed and stored in 384 well plates so that both PCR set up and sequence reaction set up could be done with multichannel pipettors and pipetting robots. For sequencing 6  $\mu\text{l}$  of PCR product and 10 pmol of sequencing primer were evaporated to dryness in a speed vacuum system and then rehydrated with 0.5  $\mu\text{l}$  ABI Prism BigDye terminators (version 3.0) and 0.75  $\mu\text{l}$  5 $\times$ buffer mix, and made up to 5  $\mu\text{l}$  with nanopure  $\text{H}_2\text{O}$  before the plates were heat sealed with foil. The sequencing conditions were 30 s at  $96^{\circ}\text{C}$ , 46 cycles of  $96^{\circ}\text{C}$  for 10 s,  $50^{\circ}\text{C}$  for 5 s, and  $60^{\circ}\text{C}$  for 4 min following which the sequencing reactions were precipitated with 20  $\mu\text{l}$  of 62.5% ethanol/1 M potassium acetate (pH 4.5) and centrifuged at 4000 rpm at  $4^{\circ}\text{C}$  for 45 min. Following resuspension in 15  $\mu\text{l}$  HiDi formamide (ABI; Applied Biosystems, Foster City, CA, USA) the samples were electrophoresed on an ABI 3700 DNA analyser prepared with POP-5 capillary gel matrix.

The oligonucleotide sequences of primers used are available in supplementary table A (only available online at <http://jmg.bmjournals.com/supplemental/>).

### Sequence analysis

Following initial data processing on the ABI 3700 instruments, the base calls were analysed with the Phred programme<sup>41</sup> which adds a quantitative base quality value and provides a probability of the correctness of a base call. The quality values represent the log of the probability that the base call is correct, such that a Phred value of 20 corresponds to a 99% probability of accuracy and a Phred value of 30 corresponds to a 99.9% probability that a base call is accurate. Using the Phrap programme<sup>42</sup> the read sequence was assembled with consensus sequence and potential mutations were identified using the Consed programme.<sup>43</sup> The read assembly was done on a per-PCR fragment basis and a single Phrap assembly consisted of the consensus genomic sequence and all sequence reads relating to the PCR fragment. The read sequence and Phred quality values were compared with the assembled consensus sequence using Cross\_match,<sup>42</sup> and all discrepancies were tagged and ranked according to Phred quality of the base. Potential base discrepancies were catalogued using Perl scripts and their trace files underwent human review. The two control samples were analysed in a blinded fashion.

### Significance of sequence changes

We analysed all three genes in 158 chromosomes, thereby allowing internal controls for the possibility that sequence changes might represent polymorphisms. We adopted a conservative approach to the designation of sequence changes as possibly pathogenic. Sequences showing variation from the published data were assigned as likely to be polymorphisms if they were at intron positions  $\pm 4$  or more nucleotides away from the splice site or synonymous exonic changes. In addition, exonic variants leading to non-synonymous amino acid changes were designated as polymorphisms if they had been reported previously, were present as an additional change in patients with already defined mutations, were present in patients known to be unlinked to the particular gene in question, or the minor allele had been annotated as the consensus in the GenBank reference sequence. Conversely, sequence variations were designated as putative mutations if they were rare or unique exonic changes leading to non-synonymous amino acid changes, splice acceptor or splice donor changes at position  $\pm 1-3$  or changes of the last nucleotide in the exon, or exonic insertions or deletions leading to shift of the reading frame and a consequent premature termination codon, or involving a splice site.

Sequence changes assigned as putative mutations were resequenced from a second PCR product to confirm the reliability of the original finding.

### RESULTS

Details of all the putative mutations and polymorphisms found are stratified by gene in tables 1–3. The patient specific putative mutations are documented in table 4 and in fig 1.

#### BM

Our BM cohort of 53 patients included three affected siblings (two of whom had milder symptoms than the third), as well as four parent-offspring pairs. Counting each family only once, we found 23 different putative mutations in 26 of 47 patients with a clinical diagnosis of BM. Seven are in *COL6A1* (one found in two unrelated individuals and two each found in three unrelated individuals), six in *COL6A2* (one recurring in two unrelated individuals), and 10 in *COL6A3*. Of these, five are predicted to alter splicing (two in *COL6A1*, two in *COL6A2*, and one in *COL6A3*), one in *COL6A2* is predicted to lead to frameshift and a consequent premature termination codon, and 17 are predicted to cause non-synonymous amino

acid substitutions (five in *COL6A1*, three in *COL6A2*, and nine in *COL6A3*) and are thus putative missense mutations.

One of the putative missense mutations, a G to A change in exon 11 of *COL6A3*, which is predicted to change a glycine to glutamic acid in the N-terminal domain of the  $\alpha 3(VI)$  chain, was previously reported by Pan *et al.*<sup>25</sup> Another four are predicted to cause single amino acid substitutions disrupting the Gly-Xaa-Yaa motif of the highly conserved triple helical domain of any of the three *COL6* genes.

The three affected siblings, #3, #4, and #9 in family  $\epsilon$ , share a heterozygous G to A change in exon 26 of *COL6A2* which is predicted to change a glycine to serine in the C1-terminal domain of the  $\alpha 2(VI)$  chain. Two separate putative missense mutations, one in *COL6A1* and one in *COL6A3*, are present in both individuals (#28 and #29) of parent-offspring pair  $\beta$ . The heterozygous C to T change in exon 9 of *COL6A1* is predicted to change a proline to leucine in the triple helical domain of the  $\alpha 1(VI)$  chain and the heterozygous G to A change in exon 15 of *COL6A3* is predicted to change a glycine to aspartic acid right at the beginning of the triple helical domain of the  $\alpha 3(VI)$  chain. Two unrelated individuals (#43 and #74) both carry one putative missense mutation in *COL6A1* alongside a putative missense mutation in *COL6A3*. We also found a splice donor mutation in *COL6A1* intron 14+1 (first described by Lamande *et al.*<sup>44</sup>) in our parent-offspring pair  $\gamma$  and two other unrelated individuals, one of the latter being positive control individual #71 who has previously been reported on by Pan *et al.*<sup>36</sup>

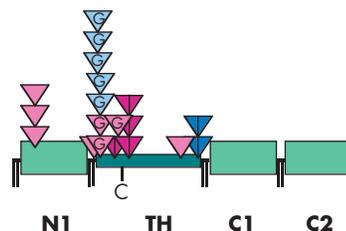
#### UCMD

In the 26 patients with a UCMD phenotype, a total of 21 different putative mutations were found in 20 unrelated patients. Four are in *COL6A1* (two of which are present in two unrelated individuals), 11 are in *COL6A2*, and six in *COL6A3* (one of which is present in three unrelated individuals). Of these, seven are predicted to alter splicing (one in *COL6A1*, three in *COL6A2*, and three in *COL6A3*), two are predicted to lead to premature termination codons in *COL6A2*, and 12 are predicted to cause non-synonymous amino acid substitutions (three in *COL6A1*, six in *COL6A2*, and three in *COL6A3*) and are thus putative missense mutations. Single amino acid substitutions disrupting the triple helical Gly-Xaa-Yaa motif are predicted to be caused by five of these changes. Within the UCMD cohort, five patients are homozygous for putative mutations and two patients have two different changes in the same gene, whereas in another 10 patients only one putative mutation was identified (five of these are predicted to result in a splicing error and five are putative missense mutations). Finally, two patients are compound heterozygotes with one putative mutation each in *COL6A1* and *COL6A2*.

We also documented 145 polymorphisms (34 in *COL6A1*, 49 in *COL6A2*, and 62 in *COL6A3*). Of these, 68 are intronic (24 in *COL6A1*, 30 in *COL6A2*, and 14 in *COL6A3*), 50 are exonic but predicted to be synonymous (seven in *COL6A1*, 11 in *COL6A2*, and 32 in *COL6A3*), and 27 are exonic and predicted to cause a non-synonymous amino acid substitution (three in *COL6A1*, eight in *COL6A2*, and 16 in *COL6A3*). Of the latter, nine had been reported previously,<sup>36</sup> four were present in a patient known to be unlinked to the particular gene in question, three were registered in the NCBI dbSNP database, three were designated as polymorphisms following personal communication with Shireen Lamande (Cell and Matrix Biology Research Unit, University of Melbourne, Australia, 2004) regarding unpublished polymorphism data, six were present as additional changes in patients with an already defined mutation, one was present in only one of two affected first degree relatives in a pedigree with autosomal dominant inheritance, and for one the minor allele had been

**UCMD**

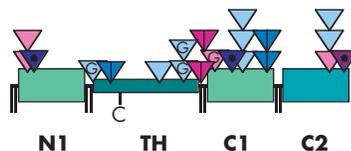
- ▼ Splice site change / exonic deletion
- ◄ Missense change affecting glycine
- ▼ Insertion /deletion /duplication causing frameshift and PTC



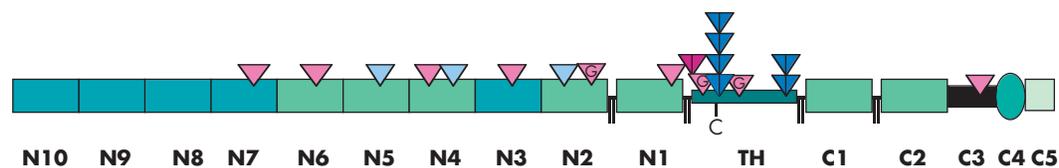
$\alpha 1(\text{VI})$  ~140 kDa  
COL6A1 – 21q22.3

**BM**

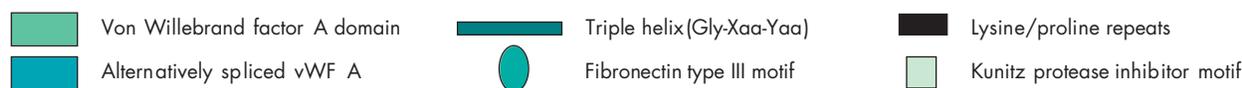
- ▼ Splice site change / exonic deletion
- ◄ Missense change affecting glycine
- ▼ Insertion causing frameshift and PTC



$\alpha 2(\text{VI})$  ~140 kDa  
COL6A2 – 21q22.3



$\alpha 3(\text{VI})$  ~300 kDa  
COL6A3 – 2q37



**Figure 1** Localisation of the identified putative mutations. Schematic diagram of the domain structure of the protein chains encoded by COL6A1-3. All three chains contain a triple helical domain of 335–336 amino acids flanked by large N- and C-terminal globular domains made up of vWF A-like motifs. Short cysteine rich connecting peptides separate N2, N1, triple helix, C1, and C2. The triple helical domains contain a single cysteine residue (depicted as “C”) which is important for dimer assembly. The localisation of the identified putative mutations is shown stratified by clinical phenotype.

annotated as the consensus in the GenBank reference sequence.

In addition, our genomic sequencing data revealed that the COL6A1 GenBank transcript NM\_001848 harbours two inaccuracies. At position c.1182\_1185 there are only three G present instead of the four found in the GenBank sequence. Similarly, at position c.1192\_1193 there are three G instead of the two found in GenBank. These results alter the predicted amino acid sequence: assuming an exon 17 splice acceptor sequence of CAG, we predict that the amino acids number 396 and 397 within the triple helix of the  $\alpha 1(\text{VI})$  chain would be glutamine and proline instead of proline and alanine within the Gly-Xaa-Yaa motif.

**DISCUSSION**

SCAIP has previously been demonstrated to be efficient in detecting mutations in the X linked gene dystrophin.<sup>37</sup> We have now applied it to the three COL6 genes and shown that it is a reliable method in these autosomal genes. We have sequenced all three COL6 genes from genomic DNA in a cohort of 79 patients clinically ascertained as either UCMD or BM with more than 97% coverage of the 107 coding exons for all patients. We found putative mutations in one of the COL6 genes in more than 62% of patients (66% of those clinically classified as BM, 56% of severe BM patients, and 79% of UCMD patients). The comparatively low detection rate in patients with mild BM (22%) and patients with UCMD and

short stature (0%) may be related to overlap of their phenotype with other forms of muscular dystrophy.

**Sequence changes detected**

The conservative approach we adopted to the designation of sequence changes as possibly pathogenic was guided by the criteria suggested by Cotton and Scriver.<sup>45</sup> In the clinical context a “mutation” is generally regarded as a genetic variant which has a harmful effect and causes disease. Our project focussed on genomic sequencing and did not extend to segregation or expression studies, work which will be necessary to confirm the pathogenicity of sequence variants we have detected but which was outside the scope of our project. We thus designated variants as putative mutations by using criteria of mutation type.

**Variants likely to affect splicing**

We found nine different sequence variations at intron position  $\pm 1-3$  from the splice site as well as one complex insertion/deletion involving a splice donor site and one deletion involving a splice acceptor site. These change the highly conserved “sequence logos” of the spliceosome described by Stephens and Schneider.<sup>46</sup> We also included a homozygous change of the last exonic nucleotide adjacent to the COL6A3 exon 27 splice donor site in a patient from a consanguineous relationship known to link to COL6A3 as a putative splice mutation.<sup>47</sup> In order to confirm the pathogenicity of all putative splice mutations it will be necessary to perform RT PCR on RNA from dermal

**Table 1** Putative mutations and probable polymorphisms found in *COL6A1*

<i>COL6A1</i>	Location	NM_001848 nucleotide change	Predicted change	Frequency	Previously reported	
Putative mutations	Exon 3	c.347	G→A	Ser116Asn	Heterozygous change in 3/54 individuals	
	Exon 9	c.821	C→T	Pro274Leu	Heterozygous change in 2/77 individuals*	
	Exon 9	c.841	G→A	Gly281Arg	Heterozygous change in 1/77 individuals	
	Exon 9	c.850	G→A	Gly284Arg	Heterozygous change in 2/77 individuals	
	Exon 10	c.868	G→A	Gly290Arg	Heterozygous change in 4/77 individuals	
	Intron 13	c.1003-3	C→A	Splice acceptor CAG to AAG	Heterozygous change in 1/78 individuals	
	Exon 14	c.1022	G→T	Gly341Val	Heterozygous change in 1/78 individuals	‡
	Intron 14	c.1056+1	G→A	Splice donor GTA to ATA	Heterozygous change in 4/78 individuals*	§
	Exon 26	c.1712	A→C	Lys571Thr	Heterozygous change in 1/78 individuals	
	Intron 27	c.1776+1	G→A	Splice donor GTA to ATA	Homozygous change in 1/76 individuals	
Probable polymorphisms	Exon 1	c.66	G→A	Synonymous	1/92 chromosomes	
	Intron 3	c.428+14	G→A	Glu22Glu	19/105 chromosomes	**
	Intron 3	c.429-19	G→A		42/154 chromosomes	
	Intron 4	c.588+13	A→C		35/154 chromosomes	
	Intron 8	c.805-36	G→A		2/154 chromosomes	
	Intron 9	c.859-19	G→A		23/152 chromosomes	**
	Intron 10	c.903+14	C→A		42/154 chromosomes	
	Exon 15	c.1095	C→T	Synonymous Gly365Gly	75/156 chromosomes	**
	Intron 19	c.1335+27	C→A		17/156 chromosomes	**
	Intron 20	c.1398+28	G→A		1/156 chromosomes	
	Intron 20	c.1399-32	C→T		75/158 chromosomes	**
	Intron 21	c.1461+18	A→C		62/158 chromosomes	
	Intron 21	c.1462-36	A→G		43/156 chromosomes	**
	Intron 22	c.1524+16	G→A		1/158 chromosomes	
	Intron 22	c.1524+19	G→C		17/158 chromosomes	**
	Intron 22	c.1524+27	A→G		1/158 chromosomes	
	Intron 25	c.1674-16	C→T		1/154 chromosomes	
	Intron 28	c.1814-6	C→G		3/146 chromosomes	
	Intron 29	c.1823-31	C→T		5/156 chromosomes	
	Exon 30	c.1950	G→A	Synonymous Leu650Leu	2/156 chromosomes	
	Intron 30	c.1956+15	C→T		76/156 chromosomes	
	Intron 30	c.1957-11	C→T		44/152 chromosomes	
	Intron 30	c.1957-5	C→T		3/152 chromosomes	
	Intron 31	c.2067-10	T→C		1/154 chromosomes	
	Exon 32	c.2130	G→A	Synonymous Thr710Thr	1/154 chromosomes	
	Intron 32	c.2251-29	T→C		61/148 chromosomes	
	Intron 33	c.2434+15	G→A		11/140 chromosomes	**
	Intron 33	c.2434+20	G→A		46/148 chromosomes	
	Exon 35	c.2549	G→A	Non-synonymous Arg850His	37/108 chromosomes	**
	Exon 35	c.2622	G→A	Synonymous Ala874Ala	1/114 chromosomes	
Exon 35	c.2643	C→T	Non-synonymous Thr881Met	1/112 chromosomes†		
Exon 35	c.2667	G→A	Synonymous Ala889Ala	39/110 chromosomes	**	
Exon 35	c.2669	C→T	Non-synonymous Ser890Leu	32/104 chromosomes	¶	
Exon 35	c.2796	C→T	Synonymous Ser932Ser	40/112 chromosomes	**	

\*Two affected first degree relatives; †present in only one of two affected first degree relatives in a pedigree of autosomal dominant inheritance; ‡reported by Scacheri *et al*<sup>20</sup> as Gly341Asp; §first reported by Lamande *et al*<sup>44</sup> and then by Pan *et al*<sup>36</sup> who provided patient #71 as a positive control; ¶reported by Pan *et al*<sup>36</sup> as Ser889Leu; \*\*registered in the NCBI dbSNP database.

fibroblasts of these patients. Seven of the splice site changes were in UCMD and five in BM patients.

### Variants likely to cause functional haploinsufficiency

One exonic variant directly generates a premature termination codon in UCMD patient #78 and two duplications are predicted to change the reading frame creating downstream premature stop codons following six or nine non-synonymous amino acids, in BM patient #70 and UCMD patient #23, respectively. These changes would be expected to lead to premature mRNA decay and subsequent haploinsufficiency.<sup>46</sup>

### Likely pathogenic missense changes

For the sequence variations predicted to cause non-synonymous amino acid substitutions, the differentiation between putative missense mutation and probable polymorphism is more difficult. A polymorphic variation is without phenotypic effect and is formally defined as a change present at a frequency of more than 1% of alleles in a population. Given the absence of normal controls in our cohort, we did not deem it justified to assign every change occurring with a frequency of more than 1% as a probable polymorphism, especially because certain sequence variations conforming to a well-described pathogenic mechanism—the disruption of the Gly-Xaa-Yaa motif of the highly conserved triple helical

**Table 2** Putative mutations and probable polymorphisms found in *COL6A2*

<i>COL6A2</i>	Location	NM_001849 nucleotide change	Predicted change	Frequency	Previously reported		
Putative mutations	Exon 3	c.316	G→A	Glu106Lys	Heterozygous change in 2/76 individuals		
	Exon 3	c.688_689	dupAC	9 non-syn AA followed by PTC	Heterozygous change in 1/77 individuals		
	Exon 6	c.847	G→A	Gly283Arg	Heterozygous change in 1/78 individuals		
	Intron 9	c.955-2	A→G	Splice acceptor CAG to CGG	Heterozygous change in 1/78 individuals		
	Exon 18	c.1493	G→A	Arg498His	Heterozygous change in 1/78 individuals		
	Exon 20	c.1590	G→C	Gly531Arg	Homozygous change in 1/78 individuals		
	Intron 24	c.1817-3	C→G	Splice acceptor CAG to GAG	Heterozygous change in 1/78 individuals		
	Intron 25	c.1970-3	C→A	Splice acceptor CAG to AAG	Heterozygous change in 1/77 individuals		
	Exon 26	c.2098	G→A	Gly700Ser	Heterozygous change in 3/77 individuals*		
	Exon 26	c.2319	C→G	Tyr773PTC	Heterozygous change in 1/78 individuals		
	Exon 26	c.2329	T→C	Cys777Arg	Homozygous change in 1/78 individuals		
	Exon 26	c.2351	G→A	Arg784His	Heterozygous change in 1/78 individuals		
	Intron 26	c.2423-2	A→G	Splice acceptor CAG to CGG	Heterozygous change in 1/78 individuals		
	Exon 27	c.2455_2461+37delins33		Involves splice donor site	Heterozygous change in 1/78 individuals		
	Exon 28	c.2558	G→A	Arg853Gln	Heterozygous change in 1/75 individuals		
	Exon 28	c.2626	C→A	Arg876Ser	Homozygous change in 1/75 individuals		
	Exon 28	c.2663_2664	dupAG	6 non-syn AA followed by PTC	Heterozygous change in 1/75 individuals		
	Probable polymorphisms	Exon 3	c.228	C→T	Synonymous Tyr96Tyr	1/152 chromosomes	
		Exon 3	c.663	C→T	Synonymous Pro221Pro	17/154 chromosomes	
		Exon 3	c.679	G→A	Non-synonymous Asp227Asn	3/154 chromosomes†	
Intron 3		c.714+9	C→T		5/154 chromosomes		
Intron 4		c.735+34	C→T		1/156 chromosomes		
Intron 8		c.928-19	C→T		70/156 chromosomes	¶	
Intron 12		c.1116+22	C→T		1/156 chromosomes		
Intron 12		c.1116+32	G→A		72/156 chromosomes		
Exon 14		c.1196	A→G	Non-synonymous Asn399Ser	33/156 chromosomes	¶	
Intron 14		c.1269+32	C→T		1/156 chromosomes		
Intron 15		c.1332+26	G→A		30/156 chromosomes	¶	
Intron 15		c.1333-8	T→C		35/156 chromosomes		
Exon 17		c.1437	T→C	Synonymous Ala479Ala	1/156 chromosomes		
Exon 18		c.1466	G→A	Non-synonymous Arg489Gln	1/156 chromosomes	§	
Intron 18		c.1521+21	A→G		27/156 chromosomes		
Intron 18		c.1522-52	C→T		1/156 chromosomes		
Intron 18		c.1522-36	T→C		28/156 chromosomes		
Exon 19		c.1552	C→T	Non-synonymous Pro518Ser	3/156 chromosomes‡		
Intron 19		c.1573-32	C→T		17/156 chromosomes		
Intron 20		c.1609-10	C→T		27/156 chromosomes		
Intron 21		c.1671+10	G→A		30/156 chromosomes	¶	
Intron 21		c.1672-37	G→T		10/142 chromosomes		
Intron 21		c.1672-24	C→G		70/150 chromosomes	¶	
Intron 22		c.1734+35	A→G		16/148 chromosomes		
Intron 22		c.1735-30	A→G		16/150 chromosomes		
Intron 23		c.1770+4	G→A		7/150 chromosomes		
Intron 23		c.1771-34	G→A		1/156 chromosomes		
Intron 23		c.1771-25	G→A		31/156 chromosomes	¶	
Intron 24		c.1816+17	delG		6/156 chromosomes		
Intron 24		c.1817-33	C→T		7/154 chromosomes	¶	
Intron 24		c.1817-9	dupC		18/156 chromosomes		
Intron 25		c.1970-24	C→T		1/154 chromosomes		
Intron 25		c.1970-23	G→C		9/154 chromosomes		
Intron 25		c.1970-9	G→A		1/156 chromosomes		
Exon 26		c.2039	G→A	Non-synonymous Arg680His	72/154 chromosomes	§	
Exon 26		c.2094	G→A	Synonymous Ala698Ala	68/154 chromosomes		
Exon 26	c.2097	C→T	Synonymous Gly699Gly	68/154 chromosomes			
Exon 26	c.2163	G→A	Synonymous Gln721Gln	1/156 chromosomes			
Exon 26	c.2170	C→T	Non-synonymous Arg724Cys	1/156 chromosomes‡			
Exon 26	c.2184	G→A	Synonymous Val728Val	61/154 chromosomes	¶		
Exon 26	c.2411	T→G	Non-synonymous Val804Gly	1/156 chromosomes‡			
Intron 26	c.2423-35	C→A		19/156 chromosomes			

Table 2 Continued

COL6A2	Location	NM_001849 nucleotide change	Predicted change	Frequency	Previously reported
	Intron 26	c.2423-17 insGGCCCCGGCCCC			19/156 chromosomes
	Intron 27	c.2462-35	T→C	51/130 chromosomes	
	Exon 28	c.2697	G→T Synonymous Thr899Thr	4/150 chromosomes	
	Exon 28	c.2724	A→G Synonymous Thr908Thr	11/148 chromosomes	
	Exon 28	c.2769	C→T Synonymous His923His	1/152 chromosomes	
	Exon 28	c.2803	G→A Non-synonymous Gly935Arg	21/148 chromosomes	§
	Exon 28	c.2979	C→T Synonymous Arg993Arg	20/150 chromosomes	¶

\*Three affected first degree relatives; †annotated as A and Asn in GenBank reference sequence; ‡present as additional change in patient with proven mutation; §reported by Pan *et al*.<sup>26</sup>; ¶registered in NCBI dbSNP database.  
PTC, premature termination codon.

domain<sup>20 29 30</sup>—occurred with a frequency higher than that. For example, a heterozygous G to A change in exon 10 of *COL6A1*, predicted to change a glycine to arginine within the triple helical domain of the  $\alpha 1(VI)$  chain, occurred in 4/77 unrelated individuals, that is, with an allele frequency of 2.6%. Conversely, we noticed that some previously published polymorphisms, such as the G→A change in exon 18 of *COL6A2* thought to substitute a glutamine for arginine, occurred with extremely low frequency (in this case 1/156 chromosomes, that is, 0.6%) in our cohort.

We identified 55 different exonic sequence variations that are predicted to cause non-synonymous amino acid substitutions. Analysing their frequency, we noticed that all variants fell into one of two categories, occurring either with an allele frequency of more than 20% or with a frequency of less than 3%. The former, which had all been previously reported or were registered in the NCBI dbSNP database, were designated as probable polymorphisms. Within the latter group we assigned putative missense mutation status to all changes, unless they had been previously reported or registered, were present as an additional change in a patient with a proven mutation, were present in a patient known to be unlinked to the particular gene, or only found in one of two affected first degree relatives of a family with an autosomal dominant pedigree.

Using these criteria, 28 of the 55 identified exonic sequence variations predicted to cause non-synonymous amino acid substitutions are considered to be putative missense mutations (11 found in UCMD patients, 16 in BM, and one in both UCMD and BM).

The pathogenic mechanism of single amino acid substitutions disrupting the Gly-Xaa-Yaa motif of the highly conserved triple helical domain of any of the three collagen VI chains has been previously described<sup>20 29 30</sup> and this is likely to apply to nine of our 28 changes which are distributed in equal numbers between BM and UCMD patients. One of these, a G to T change in exon 14 of *COL6A1* which is predicted to change a glycine to valine, corresponds to a glycine to aspartic acid change at this position and has previously been reported.<sup>20</sup> Another variant, a G to A change in exon 11 of *COL6A3* which is predicted to change a glycine to glutamic acid in the N-terminal domain of the  $\alpha 3(VI)$  chain, was previously reported by Pan *et al*.<sup>25</sup> Although other missense mutations not concerning glycines in the triple helical domain have been published,<sup>20</sup> for the remainder of our putative missense mutations, further expression studies as well as segregation analysis and testing of large numbers

of healthy controls will be necessary to confirm their pathogenicity or reassign them as polymorphisms.

### UCMD

Within the UCMD cohort, five patients were found to be homozygous for putative mutations and two patients have two different sequence changes in the same gene which would be in keeping with the autosomal recessive mode of inheritance originally described for this condition. However, in another 10 UCMD patients only one putative mutation was identified. While it is possible that these patients might carry a second mutation not detectable by our technique, such as a large exonic deletion, an alternative possibility is that the sequence variations identified in these patients might act in a dominant fashion, similar to the mutation reported by Pan *et al*.<sup>36</sup> If this is confirmed, dominant mutations may be a common finding in patients with features suggestive of UCMD. This could explain the apparent lack of linkage to *COL6A* genes in patients with a UCMD phenotype according to a recessive model. So could the finding of two patients with a clinical UCMD phenotype (#45 and #48) who are compound heterozygotes for putative mutations in *COL6A1* and *COL6A2*.<sup>48</sup> Both carry a sequence variation predicted to cause a glycine substitution within the triple helical domain of the  $\alpha 1(VI)$  chain alongside a separate unique putative missense mutation in *COL6A2*. These glycine substitutions may act in a dominant fashion or may be present together with an undetected large exonic deletion. However, given the potential pathogenicity associated with both variants, an important alternative possibility is inheritance of alterations in two different *COL6* genes causing disease. Family studies for inheritance of the changes and presence or absence of the phenotype in carriers as well as expression studies will be necessary to determine whether one or both changes for either of these patients are disease causing or disease modifying. Neither of these two families have a positive dominant family history of weakness.

### BM

Two separate putative missense mutations, one in *COL6A1* and one in *COL6A3*, are also present in one of our BM parent-offspring pairs. Given that it is predicted to change a glycine to aspartic acid right at the beginning of the triple helical domain of the  $\alpha 3(VI)$  chain, the heterozygous G to A change in exon 15 of *COL6A3* is likely to be disease causing. However, an additional heterozygous C to T change was also detected in this family and is predicted to change a proline to leucine in the triple helical domain of the  $\alpha 1(VI)$  chain. This change is

**Table 3** Putative mutations and probable polymorphisms found in *COL6A3*

<i>COL6A3</i>	Location	NM_004369 nucleotide change	Predicted change	Frequency	Previously reported		
Putative mutations	Exon 6	c.2030	G→A	Arg677His	Heterozygous change in 1/79 individuals		
	Exon 7	c.3040	A→G	Lys1014Glu	Heterozygous change in 1/79 individuals		
	Exon 8	c.3191	G→A	Arg1064Gln	Heterozygous change in 1/79 individuals		
	Exon 9	c.4156	G→A	Glu1386Lys	Heterozygous change in 1/79 individuals		
	Exon 9	c.4184	G→A	Arg1395Gln	Heterozygous change in 1/79 individuals		
	Exon 10	c.4399	A→G	Asn1467Asp	Heterozygous change in 1/79 individuals		
	Exon 11	c.5020	G→A	Asp1674Asn	Heterozygous change in 1/79 individuals		
	Exon 11	c.5036	G→A	Gly1679Glu	Heterozygous change in 1/79 individuals	§	
	Exon 14	c.5953	G→A	Val1985Met	Heterozygous change in 1/79 individuals		
	Exon 15	c.6140	G→A	Gly2047Asp	Heterozygous change in 2/79 individuals*		
	Intron 15	c.6156+1	G→A	Splice donor GTA to ATA	Heterozygous change in 1/79 individuals		
	Exon 16	c.6157-9_6177del30		Involves splice acceptor site	Heterozygous change in 1/78 individuals		
	Intron 16	c.6210+1	G→A	Splice donor GTA to ATA	Heterozygous change in 3/78 individuals		
	Exon 17	c.6239	G→A	Gly2080Asp	Heterozygous change in 1/79 individuals		
	Exon 27	c.6816	G→A	Nucleotide adjacent to splice donor	Homozygous change in 1/79 individuals		
	Probable polymorphisms	Exon 40	c.8822	C→T	Ala2941Val	Heterozygous change in 1/79 individuals	
		Exon 4	c.768	C→T	Synonymous Val256Val	3/158 chromosomes	
Exon 4		c.1182	C→T	Synonymous Thr394Thr	1/158 chromosomes		
Exon 4		c.1231	C→G	Non-synonymous Leu411Val	2/158 chromosomes†		
Exon 5		c.1389	C→T	Synonymous Ala463Ala	1/158 chromosomes		
Exon 5		c.1471	G→C	Non-synonymous Asp491His	1/158 chromosomes	¶	
Exon 5		c.1475	C→G	Non-synonymous Thr492Ser	1/158 chromosomes	¶	
Exon 6		c.2419	G→A	Non-synonymous Ala807Thr	2/158 chromosomes‡		
Exon 6		c.2463	T→C	Synonymous Ser821Ser	1/158 chromosomes		
Exon 6		c.2488	G→T	Non-synonymous Ala830Ser	2/158 chromosomes‡		
Exon 8		c.3087	C→T	Synonymous Asp1029Asp	2/158 chromosomes		
Exon 8		c.3129	C→T	Synonymous Gly1043Gly	40/158 chromosomes		
Exon 8		c.3262	A→C	Non-synonymous Lys1088Gln	2/158 chromosomes‡		
Exon 9		c.4107	C→T	Synonymous Ile1369Ile	1/158 chromosomes		
Intron 9		c.4285+17	G→A		2/156 chromosomes	‡‡	
Exon 10		c.4311	T→C	Synonymous Ile1437Ile	42/158 chromosomes	‡‡	
Exon 10		c.4533	G→T	Synonymous Gly1511Gly	34/158 chromosomes	‡‡	
Exon 10		c.4727	G→A	Non-synonymous Arg1576Gln	2/158 chromosomes	¶	
Exon 10		c.4895	G→A	Non-synonymous Arg1632Gln	2/158 chromosomes‡		
Exon 11		c.5059	C→T	Non-synonymous Pro1687Ser	1/158 chromosomes†		
Intron 14		c.6063+6	insT		32/156 chromosomes		
Intron 14		c.6003+100	insG		69/158 chromosomes		
Intron 15		c.6156+4	C→T		7/158 chromosomes		
Exon 20		c.6369	G→A	Synonymous Leu2123Leu	27/158 chromosomes	‡‡	
Exon 24		c.6618	C→T	Synonymous Pro2206Pro	1/158 chromosomes		
Exon 25		c.6653	C→T	Non-synonymous Pro2218Leu	13/158 chromosomes	**	
Exon 28		c.6855	G→C	Synonymous Gly2285Gly	77/158 chromosomes	‡‡	
Intron 28		c.6879+33	C→T		1/158 chromosomes		
Intron 29		c.6930+28	C→T		4/158 chromosomes		
Intron 29		c.6930+43	G→C		13/158 chromosomes	‡‡	
Exon 30	c.6945	C→T	Synonymous Phe2315Phe	13/158 chromosomes	‡‡		
Exon 31	c.6981	A→G	Synonymous Glu2327Glu	1/158 chromosomes			
Exon 32	c.7086	A→C	Synonymous Gly2362Gly	1/158 chromosomes			
Intron 32	c.7092+26	G→A		22/158 chromosomes	‡‡		
Intron 35	c.7175-32	G→A		7/156 chromosomes			

Table 3 Continued

COL6A3	Location	NM_004369 nucleotide change	Predicted change	Frequency	Previously reported	
	Exon 36	c.7329	C→T	Synonymous Ala2443Ala	9/154 chromosomes	
	Exon 36	c.7509	G→A	Synonymous Arg2503Arg	3/154 chromosomes	
	Exon 36	c.7512	C→T	Synonymous Asn2504Asn	14/154 chromosomes	‡‡
	Exon 36	c.7596	G→A	Synonymous Lys2532Lys	17/156 chromosomes	‡‡
	Intron 36	c.7668+21	G→C		1/146 chromosomes	
	Intron 36	c.7669–35	C→T		15/154 chromosomes	
	Exon 38	c.7842	C→T	Synonymous Ser2614Ser	12/154 chromosomes	
	Exon 38	c.7929	G→A	Synonymous Ala2643Ala	55/154 chromosomes	
	Exon 38	c.7995	A→C	Synonymous Ala2665Ala	1/156 chromosomes	
	Exon 38	c.8145	A→G	Synonymous Leu2715Leu	1/156 chromosomes	
	Exon 38	c.8346	G→A	Synonymous Glu2782Glu	1/158 chromosomes	
	Exon 38	c.8451	A→G	Synonymous Pro2817Pro	12/156 chromosomes	
	Intron 38	c.8465–34	C→T		1/158 chromosomes	
	Exon 39	c.8491	G→C	Non-synonymous Asp2831His	15/156 chromosomes	††
	Exon 40	c.8780	T→C	Non-synonymous Met2927Thr	66/154 chromosomes	**
	Exon 40	c.8820	G→A	Synonymous Thr2940Thr	12/156 chromosomes	
	Exon 40	c.8868_8869	insTGC	insAla2957	139/146 chromosomes	**
	Exon 41	c.9031	G→C	Non-synonymous Ala3011Pro	42/156 chromosomes	‡‡
	Exon 41	c.9120	G→A	Synonymous Thr3040Thr	11/158 chromosomes	
	Exon 41	c.9125	G→A	Non-synonymous Arg3042His	1/156 chromosomes†	
	Exon 41	c.9126	C→T	Synonymous Arg3042Arg	8/158 chromosomes	
	Exon 41	c.9144	C→T	Synonymous Leu3048Leu	1/158 chromosomes	
	Exon 41	c.9203	C→T	Non-synonymous Thr3068Ile	47/156 chromosomes	**
	Exon 41	c.9210	C→T	Synonymous His3070His	13/158 chromosomes	
	Intron 42	c.9325+8	G→T		1/158 chromosomes	
	Exon 43	c.9408	T→C	Synonymous Cys3136Cys	4/158 chromosomes	
	Exon 43	c.9414	A→G	Synonymous Arg3138Arg	1/158 chromosomes	
	Intron 43	c.9491–26	C→T		51/158 chromosomes	

\*Two affected first degree relatives; †present as additional change in patient with proven mutation; ‡present in patient unlinked to COL6A3; §reported by Pan *et al*<sup>25</sup>; ¶personal communication by Shireen Lamande, Cell and Matrix Biology Research Unit, University of Melbourne, Australia, 2004; \*\*reported by Pan *et al*<sup>36</sup>; ††reported by Pan *et al*<sup>36</sup> as Asp2831Thr; ‡‡registered in NCBI dbSNP database.

not present in the other 77 individuals tested and could therefore be either be a rare polymorphism or convey additional pathogenicity. Similar reasoning can also be applied to the case of another unrelated BM patient, who carries a COL6A1 sequence variation predicted to change a glycine in the triple helical domain of the  $\alpha 1(VI)$  chain to arginine alongside a putative missense mutation in COL6A3, and a further BM patient where two putative missense mutations in two different COL6 genes have been identified. As discussed before, further expression studies as well as segregation analysis and testing of large numbers of healthy controls will be necessary to confirm their pathogenicity or possible role as modifiers or to reassign them as polymorphisms.

#### Recurrent putative mutations

Within our BM cohort of patients, a parent-offspring pair and two further unrelated individuals (one of the latter being

positive control individual #71 who has previously been reported<sup>36</sup>) carry an identical splice donor mutation in COL6A1 intron 14+1 which was first described as a BM mutation by Lamande *et al*.<sup>44</sup> A heterozygous exon 16 splice donor change was identified as the only putative mutation in three unrelated individuals of our UCMD cohort of patients.

#### Mutations found in both BM and UCMD

The putative missense mutation predicted to cause a substitution of glycine by arginine in the triple helical region of COL6A1 is present as a heterozygous change in two unrelated individuals assigned to the BM cohort. The same mutation was detected in two unrelated individuals assigned to the UCMD cohort. These two BM patients had a more severe phenotype than average, while the clinical phenotype of the two UCMD patients was comparatively mild. This underlines our impression that there is in fact a spectrum of collagen VI related disorders rather than two strictly separate

**Table 4** Putative COL6 mutations for each patient

Patient	Family*	Clinical phenotype	Putative mutations found
#14		Mild BM	
#16	α	Mild BM	
#17	α	Mild BM	
#24		Mild BM	
#2		Mild BM	
#76		Mild BM	
#10		Mild BM	
#8		Mild BM	Heterozygous COL6A1 exon3 Ser116Asn
#21		Mild BM	Heterozygous COL6A1 exon3 Ser116Asn
#51		BM	Heterozygous COL6A1 exon3 Ser116Asn
#25		BM	Heterozygous COL6A3 exon7 Lys1014Glu
#55		BM	Heterozygous COL6A3 exon9 Glu1386Lys
#30		BM	Heterozygous COL6A3 exon10 Asn1467Asp
#75		BM	Heterozygous COL6A3 exon40 Ala2941Val
#74		BM	Heterozygous COL6A1 exon26 Lys571Thr; heterozygous COL6A3 exon14 Val1985Met
#7		BM	Heterozygous COL6A2 exon3 Glu106Lys
#61		BM	Heterozygous COL6A2 exon3 Glu106Lys
#56		BM	Heterozygous COL6A2 exon28 Arg853Gln
#36		BM	Heterozygous COL6A1 exon14 Gly341Val
#19		BM	Heterozygous COL6A3 exon11 Gly1679Glu
#57		BM	
#58		BM	
#59		BM	
#62		BM	
#63		BM	
#28	β	BM	Heterozygous COL6A1 exon9 Pro274Leu; heterozygous COL6A3 exon15 Gly2047Asp
#29	β	BM	Heterozygous COL6A1 exon9 Pro274Leu; heterozygous COL6A3 exon15 Gly2047Asp
#52		BM	
#53		BM	
#54		BM	
#32		BM	Heterozygous COL6A1 intron13 splice acceptor change
#71		BM	Heterozygous COL6A1 intron14 splice donor change
#12	γ	BM	Heterozygous COL6A1 intron14 splice donor change
#13	γ	BM	Heterozygous COL6A1 intron14 splice donor change
#77		BM	Heterozygous COL6A1 intron14 splice donor change
#18		BM	Heterozygous COL6A2 intron24 splice acceptor change
#6		BM	Heterozygous COL6A2 intron25 splice acceptor change
#33	δ	BM	
#34	δ	BM	
#70		BM	Heterozygous COL6A2 exon28 insertion leading to frameshift and PTC
#20		BM	
#11		BM	
#3	ε	BM	Heterozygous COL6A2 exon26 Gly700Ser
#4	ε	BM	Heterozygous COL6A2 exon26 Gly700Ser
#9	ε	Severe BM	Heterozygous COL6A2 exon26 Gly700Ser
#31		Severe BM	
#26		Severe BM	
#64		Severe BM	Heterozygous COL6A3 intron15 splice donor change
#69		Severe BM	
#73		Severe BM	
#66		severe BM	Heterozygous COL6A3 exon17Gly2080Asp
#43		Severe BM	Heterozygous COL6A1 exon10 Gly290Arg; heterozygous COL6A3 exon6 Arg677His
#5		Severe BM	Heterozygous COL6A1 exon10 Gly290Arg
#47		UCMD	Heterozygous COL6A1 exon10 Gly290Arg
#48		UCMD	Heterozygous COL6A1 exon10 Gly290Arg; heterozygous COL6A2 exon26 Arg784His
#45		UCMD	Heterozygous COL6A1 exon9 Gly281Arg; heterozygous COL6A2 exon18 Arg498His
#46		UCMD	Heterozygous COL6A1 exon9 Gly284Arg
#41		UCMD	Heterozygous COL6A1 exon9 Gly284Arg
#49		UCMD	Heterozygous COL6A2 exon6 Gly283Arg
#79		UCMD	Heterozygous COL6A3 exon8 Arg1064Glu
#15		UCMD	
#72		UCMD	
#44		UCMD	Homozygous COL6A2 exon26 Cys777Arg
#38		UCMD	Homozygous COL6A2 exon20 Gly531Arg
#65		UCMD	Homozygous COL6A2 exon28 Arg876Ser; heterozygous COL6A3 exon9 Arg1395Gln; heterozygous COL6A3 exon11 Asp1674Asn
#42		UCMD	Homozygous COL6A1 intron27 splice donor change
#22		UCMD	Homozygous COL6A3 exon27 change of nucleotide adjacent to splice donor
#23		UCMD	Heterozygous COL6A2 exon3 insertion leading to frameshift and PTC; heterozygous COL6A2 exon27 complex insertion/deletion involving splice donor
#78		UCMD	Heterozygous COL6A2 exon26 Tyr 773 PTC; heterozygous COL6A2 intron26 splice acceptor change
#37		UCMD	
#50		UCMD	
#60		UCMD	
#67		UCMD	Heterozygous COL6A3 exon16 deletion involving splice acceptor
#35		UCMD	Heterozygous COL6A3 intron16 splice donor change
#39		UCMD	Heterozygous COL6A3 intron16 splice donor change
#40		UCMD	Heterozygous COL6A3 intron16 splice donor change
#68		UCMD	Heterozygous COL6A2 intron9 splice acceptor change
#27		UCMD+short stature	
#1		UCMD+short stature	

α, Parent-offspring pair; β, parent-offspring pair; γ, parent-offspring pair; δ, parent-offspring pair; ε, three affected siblings.

entities, but further work will be necessary to delineate the exact pathogenic mechanisms and correlate these to the clinical and pathological findings.

### Probable polymorphisms

Overall our results confirm that the three *COL6* genes are highly polymorphic. We identified a high number of intronic and exonic variants designated to be probable polymorphisms. Of the 77 non-pathogenic exonic sequence changes we have observed in the three *COL6* genes, 55 resulted in an amino acid substitution. In theory, it is possible that some of the 23 rare or unique exonic synonymous changes as well as some of the 24 rare intronic changes which we classified as probable polymorphisms might influence splicing. In patients with no immediately convincing pathogenic mutation, it will thus be important to perform RT-PCR studies even for these changes.

### SUMMARY

The genetics of the collagen VI related muscular dystrophies/myopathies are complex. Collagen VI is composed of three different peptide chains which are encoded by three large genes and the assembly of collagen VI involves a number of different stages. Different mutations have been shown to have variable effects on protein assembly, secretion, and its ability to form a functioning extracellular network. The multitude of possible mutational mechanisms and modes of inheritance has important consequences for the genetic counselling of patients and their families.

Our results suggest that mutations in the *COL6* genes are likely to be found in the majority of patients with a clinical diagnosis of UCMD and BM, but that stringent criteria need to be applied to the designation of likely pathogenic changes in these genes because of their highly polymorphic nature. For many sequence changes demonstration of pathogenicity by segregation analysis or RT-PCR and cDNA sequencing will be necessary.

This study reports the first application of SCAIP methodology to an autosomal gene locus, and demonstrates that one of the limitations of this method for analysis of autosomal genes is the inability to detect reliably large exonic deletions, as illustrated by the fact that we were unable to detect the large heterozygous *COL6A1* gene deletion known to be present in positive control individual #72. However, the intronic sequences co-amplified as part of our PCR strategy provide a rich source of polymorphisms, which when heterozygously present can help to exclude the possibility of large deletions. In addition, although SCAIP analysis provides sequence information on exon/intron junctions and on the immediately adjacent intronic sequence, it cannot detect mutations located more distantly in the intron which cause altered splicing (such as those resulting in cryptic splice site activation and inclusion of pseudoexons into mRNA).

No one technique as yet stands out as definitive for diagnosis in BM and UCMD, but detection of pathogenic changes in one of the *COL6* genes remains the gold standard. We have demonstrated that SCAIP is a robust methodology by which this can be approached. An additional advantage of this methodology whereby all genes were screened in all patients is that we have been able to identify potentially novel modes of inheritance and that the full range of polymorphisms in the three genes can be appreciated, which may potentially be of importance for determining the major phenotypic variability seen in this group of disorders, especially as some patients showed changes in more than one of the *COL6* genes. The high frequency of polymorphic changes also has implications for further evaluation of primary sequence information. Reliable classification of sequence changes as polymorphisms will reduce the ambiguity of a

direct sequence approach as more sequence data accumulate. Details on detected sequence variations are available online at <http://www.genome.utah.edu/DMD/collagensnps/>.

### ACKNOWLEDGEMENTS

The authors thank the patients and their families for their participation in this study. The authors are grateful to K O'Brien and D Sudano for their technical assistance and thank UC Reed, M van der Knaap, RS Finkel, H Marks, and M Scavina for contributing patients to this study.

### ELECTRONIC-DATABASE INFORMATION



Accession numbers and URLs for data presented herein are as follows: GenBank, <http://www.ncbi.nlm.nih.gov/GenBank/> (for sequences of the *COL6A1* transcript (accession number NM\_001848), the *COL6A2* isoforms 2c2 (accession number NM\_001849), 2c2a (accession number NM\_058174), and 2c2a' (NM\_058175) as well as the full length isoform of *COL6A3* (accession number NM\_004369) and its four variant transcripts (NM\_057164, NM\_057165, NM\_057166, and NM\_057167)); Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for BM (MIM 158810) and UCMD (MIM 254090)); UCSC Genome Bioinformatics, <http://genome.cse.ucsc.edu/> (for contigs); and dbSNP, <http://www.ncbi.nlm.nih.gov/SNP/>. The supplementary table is available at <http://img.bmjournals.com/supplemental/> and details on detected sequence variations are available online at <http://www.genome.utah.edu/DMD/collagensnps/>.

### Authors' affiliations

**A K Lampe, S H Laval, K M D Bushby**, Institute of Human Genetics, University of Newcastle upon Tyne, Newcastle upon Tyne, UK

**D M Dunn, A C von Niederhausern, C Hamil, A Aoyagi, K Swoboda, K M Flanagan, R B Weiss**, Department of Human Genetics, University of Utah, Salt Lake City, UT, USA

**S K Marie**, Department of Medicine, University of São Paulo, São Paulo, Brazil

**M-L Chu**, Department of Dermatology and Cutaneous Biology, Jefferson Institute of Molecular Medicine, Thomas Jefferson University, Philadelphia, PA, USA

**F Muntoni**, Dubowitz Neuromuscular Centre, Department of Paediatrics and Neonatal Medicine, Imperial College, London, UK

**C G Bonnemann**, The Children's Hospital of Philadelphia, Philadelphia, PA, USA

This work was supported by MDC (Muscular Dystrophy Campaign) financial aid to the Newcastle Muscle Centre, Myoclueter and MDC grants to FM as well as NIH grant AR38912 to MLC. KMF is supported by NIH grants M01 RR00064-39 and R01 NS043264-03. CGB is a Pew Scholar in the Biomedical Sciences. He is also supported by an Ethel Brown Foerder award and by the Florence R.C. Murray Fellowship Program at the Children's Hospital of Philadelphia. AKL gratefully acknowledges funding by a Wellcome Entry Level Fellowship and subsequently by a Patrick Berthoud Fellowship. SL is funded by the MDC.

Conflict of interest: none declared.

### REFERENCES

- 1 Sewry CA, Muntoni F. Inherited disorders of the extracellular matrix. *Curr Opin Neurol* 1999;12:519-26.
- 2 Emery AE. The muscular dystrophies. *Lancet* 2002;359:687-95.
- 3 Helbling-Leclerc A, Zhang X, Topaloglu H, Cruaud C, Tesson F, Weissenbach J, Tome FM, Schwartz K, Fardeau M, Tryggvason K. Mutations in the laminin alpha 2-chain gene (LAMA2) cause merosin-deficient congenital muscular dystrophy. *Nat Genet* 1995;11:216-8.
- 4 von der Mark H, Aumailley M, Wick G, Fleischmajer R, Timpl R. Immunohistochemistry, genuine size and tissue localization of collagen VI. *Eur J Biochem* 1984;142:493-502.
- 5 Wiberg C, Klatt AR, Wagener R, Paulsson M, Bateman JF, Heinegard D, Morgelin M. Complexes of matrilin-1 and biglycan or decorin connect collagen VI microfibrils to both collagen II and aggrecan. *J Biol Chem* 2003;278:37698-704.

- 6 **Kuo HJ**, Maslen CL, Keene DR, Glanville RW. Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen. *J Biol Chem* 1997;**272**:26522-9.
- 7 **Burg MA**, Tillet E, Timpl R, Stallcup WB. Binding of the NG2 proteoglycan to type VI collagen and other extracellular matrix molecules. *J Biol Chem* 1996;**271**:26110-6.
- 8 **Engvall E**, Hessel H, Klier G. Molecular assembly, secretion, and matrix deposition of type VI collagen. *J Cell Biol* 1986;**102**:703-10.
- 9 **Heiskanen M**, Saitta B, Palotie A, Chu ML. Head to tail organization of the human COL6A1 and COL6A2 genes by fiber-FISH. *Genomics* 1995;**29**:801-3.
- 10 **Weil D**, Mattei MG, Passage E, N'Guyen VC, Pribula-Conway D, Mann K, Deutzmann R, Timpl R, Chu ML. Cloning and chromosomal localization of human genes encoding the three chains of type VI collagen. *Am J Hum Genet* 1988;**42**:435-45.
- 11 **Chu ML**, Pan TC, Conway D, Saitta B, Stokes D, Kuo HJ, Glanville RW, Timpl R, Mann K, Deutzmann R. The structure of type VI collagen. *Ann N Y Acad Sci* 1990;**580**:55-63.
- 12 **Saitta B**, Stokes DG, Vissing H, Timpl R, Chu ML. Alternative splicing of the human alpha 2(VI) collagen gene generates multiple mRNA transcripts which predict three protein variants with distinct carboxyl termini. *J Biol Chem* 1990;**265**:6473-80.
- 13 **Stokes DG**, Saitta B, Timpl R, Chu ML. Human alpha 3(VI) collagen gene. Characterization of exons coding for the amino-terminal globular domain and alternative splicing in normal and tumor cells. *J Biol Chem* 1991;**266**:8626-33.
- 14 **Dziadek M**, Kazenwadel JS, Hendrey JA, Pan TC, Zhang RZ, Chu ML. Alternative splicing of transcripts for the alpha 3 chain of mouse collagen VI: identification of an abundant isoform lacking domains N7-N10 in mouse and human. *Matrix Biol* 2002;**21**:227-41.
- 15 **Lamande SR**, Sigalas E, Pan TC, Chu ML, Dziadek M, Timpl R, Bateman JF. The role of the alpha3(VI) chain in collagen VI assembly. Expression of an alpha3(VI) chain lacking N-terminal modules N10-N7 restores collagen VI assembly, secretion, and matrix deposition in an alpha3(VI)-deficient cell line. *J Biol Chem* 1998;**273**:7423-30.
- 16 **Furthmayr H**, Wiedemann H, Timpl R, Odermatt E, Engel J. Electron-microscopical approach to a structural model of intima collagen. *Biochem J* 1983;**211**:303-11.
- 17 **Bethlem J**, Wijngaarden GK. Benign myopathy, with autosomal dominant inheritance. A report on three pedigrees. *Brain* 1976;**99**:91-100.
- 18 **Jobsis GJ**, Boers JM, Barth PG, de Visser M. Bethlem myopathy: a slowly progressive congenital muscular dystrophy with contractures. *Brain* 1999;**122**:649-55.
- 19 **Pepe G**, Bertini E, Bonaldo P, Bushby K, Giusti B, de Visser M, Guicheney P, Lattanzi G, Merlini L, Muntioni F, Nishino I, Nonaka I, Yau RB, Sabatelli P, Sewry C, Topaloglu H, van der Kooi A. Bethlem myopathy (BETHLEM) and Ullrich scleroatonic muscular dystrophy: 100th ENMC international workshop, 23-24 November 2001, Naarden, The Netherlands. *Neuromuscul Disord* 2002;**12**:984-93.
- 20 **Scacheri PC**, Gillanders EM, Subramony SH, Vedanarayanan V, Crowe CA, Thakore N, Bingler M, Hoffman EP. Novel mutations in collagen VI genes: expansion of the Bethlem myopathy phenotype. *Neurology* 2002;**58**:593-602.
- 21 **Ullrich O**. Kongenitale, atonisch-sklerotische Muskeldystrophie, ein weiterer Typus der hereditären degenerativen Erkrankungen des neuromuskulären Systems. *Z Ges Neurol Psychiatr* 1930;**126**:171-201.
- 22 **Nonaka I**, Une Y, Ishihara T, Miyoshino S, Nakashima T, Sugita H. A clinical and histological study of Ullrich's disease (congenital atonic-sclerotic muscular dystrophy). *Neuropediatrics* 1981;**12**:197-208.
- 23 **Furukawa T**, Toyokura Y. Congenital, hypotonic-sclerotic muscular dystrophy. *J Med Genet* 1977;**14**:426-9.
- 24 **Pepe G**, Giusti B, Bertini E, Brunelli T, Saitta B, Comeglio P, Bolognese A, Merlini L, Federici G, Abbate R, Chu ML. A heterozygous splice site mutation in COL6A1 leading to an in-frame deletion of the alpha1(VI) collagen chain in an Italian family affected by Bethlem myopathy. *Biochem Biophys Res Commun* 1999;**258**:802-7.
- 25 **Pan TC**, Zhang RZ, Pericak-Vance MA, Tandan R, Fries T, Stajich JM, Viles K, Vance JM, Chu ML, Speer MC. Missense mutation in a von Willebrand factor type A domain of the alpha 3(VI) collagen gene (COL6A3) in a family with Bethlem myopathy. *Hum Mol Genet* 1998;**7**:807-12.
- 26 **Lamande SR**, Bateman JF, Hutchison W, McKinlay Gardner RJ, Bower SP, Byrne E, Dahl HH. Reduced collagen VI causes Bethlem myopathy: a heterozygous COL6A1 nonsense mutation results in mRNA decay and functional haploinsufficiency. *Hum Mol Genet* 1998;**7**:981-9.
- 27 **Sasaki T**, Hohenester E, Zhang RZ, Gotta S, Speer MC, Tandan R, Timpl R, Chu ML. A Bethlem myopathy Gly to Glu mutation in the von Willebrand factor A domain N20 of the collagen alpha3(VI) chain interferes with protein folding. *FASEB J* 2000;**14**:761-8.
- 28 **Vanegas OC**, Zhang RZ, Sabatelli P, Lattanzi G, Bencivenga P, Giusti B, Columbaro M, Chu ML, Merlini L, Pepe G. Novel COL6A1 splicing mutation in a family affected by mild Bethlem myopathy. *Muscle Nerve* 2002;**25**:513-9.
- 29 **Jobsis GJ**, Keizers H, Vreijling JP, de Visser M, Speer MC, Wolterman RA, Baas F, Bolhuis PA. Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nat Genet* 1996;**14**:113-5.
- 30 **Pepe G**, Bertini E, Giusti B, Brunelli T, Comeglio P, Saitta B, Merlini L, Chu ML, Federici G, Abbate R. A novel de novo mutation in the triple helix of the COL6A3 gene in a two-generation Italian family affected by Bethlem myopathy. A diagnostic approach in the mutations' screening of type VI collagen. *Neuromuscul Disord* 1999;**9**:264-71.
- 31 **Higuchi I**, Shiraishi T, Hashiguchi T, Suehara M, Niiyama T, Nakagawa M, Arimura K, Maruyama I, Osame M. Frameshift mutation in the collagen VI gene causes Ullrich's disease. *Ann Neurol* 2001;**50**:261-5.
- 32 **Ishikawa H**, Sugie K, Murayama K, Awaya A, Suzuki Y, Noguchi S, Hayashi YK, Nonaka I, Nishino I. Ullrich disease due to deficiency of collagen VI in the sarcolemma. *Neurology* 2004;**62**:620-3.
- 33 **Camacho VO**, Bertini E, Zhang RZ, Petriani S, Minosse C, Sabatelli P, Giusti B, Chu ML, Pepe G. Ullrich scleroatonic muscular dystrophy is caused by recessive mutations in collagen type VI. *Proc Natl Acad Sci U S A* 2001;**98**:7516-21.
- 34 **Ishikawa H**, Sugie K, Murayama K, Ito M, Minami N, Nishino I, Nonaka I. Ullrich disease: collagen VI deficiency: EM suggests a new basis for muscular weakness. *Neurology* 2002;**59**:920-3.
- 35 **Demir E**, Sabatelli P, Allamand V, Ferreira 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**:1446-58.
- 36 **Pan TC**, Zhang RZ, Sudano DG, Marie SK, Bonnemant CG, Chu ML. New molecular mechanism for Ullrich congenital muscular dystrophy: a heterozygous in-frame deletion in the COL6A1 gene causes a severe phenotype. *Am J Hum Genet* 2003;**73**:355-69.
- 37 **Flanigan KM**, von Niederhausern A, Dunn DM, Alder J, Mendell JR, Weiss RB. Rapid direct sequence analysis of the dystrophin gene. *Am J Hum Genet* 2003;**72**:931-9.
- 38 **Pepe G**, de Visser M, Bertini E, Bushby K, Vanegas OC, Chu ML, Lattanzi G, Merlini L, Muntioni F, Urtizberea A. Bethlem myopathy (BETHLEM) 86th ENMC international workshop, 10-11 November 2000, Naarden, The Netherlands. *Neuromuscul Disord* 2002;**12**:296-305.
- 39 **Kent WJ**, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res* 2002;**12**:996-1006.
- 40 **Lander ES**, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Showkhen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissole SL, Wendt MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madaen A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt L, McCombie WR, de la BM, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korfi I, Kulp D, Lancet D, Lowe TM, Mclysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, Szustakowski J, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860-921.
- 41 **Ewing B**, Hillier L, Wendt MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 1998;**8**:175-85.
- 42 **Ewing B**, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 1998;**8**:186-94.
- 43 **Gordon D**, Abajian C, Green P. Consed: a graphical tool for sequence finishing. *Genome Res* 1998;**8**:195-202.
- 44 **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**:21817-22.
- 45 **Cotton RG**, Scriver CR. Proof of "disease causing" mutation. *Hum Mutat* 1998;**12**:1-3.
- 46 **Stephens RM**, Schneider TD. Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. *J Mol Biol* 1992;**228**:1124-36.

47 **Dunn DM**, Ishigami T, Pankow J, von Niederhausern A, Alder J, Hunt SC, Leppert MF, Lalouel JM, Weiss RB. Common variant of human NEDD4L activates a cryptic splice site to form a frameshifted transcript. *J Hum Genet* 2002;**47**:665–76.

48 **Mercuri E**, Yuva Y, Brown SC, Brockington M, Kinali M, Jungbluth H, Feng L, Sewry CA, Muntioni F. Collagen VI involvement in Ullrich syndrome: a clinical, genetic, and immunohistochemical study. *Neurology* 2002;**58**:1354–9.

### Clinical Evidence—Call for contributors

*Clinical Evidence* is a regularly updated evidence-based journal available worldwide both as a paper version and on the internet. *Clinical Evidence* needs to recruit a number of new contributors. Contributors are healthcare professionals or epidemiologists with experience in evidence-based medicine and the ability to write in a concise and structured way.

#### Areas for which we are currently seeking authors:

- Child health: nocturnal enuresis
- Eye disorders: bacterial conjunctivitis
- Male health: prostate cancer (metastatic)
- Women's health: pre-menstrual syndrome; pyelonephritis in non-pregnant women

However, we are always looking for others, so do not let this list discourage you.

#### Being a contributor involves:

- Selecting from a validated, screened search (performed by in-house Information Specialists) epidemiologically sound studies for inclusion.
- Documenting your decisions about which studies to include on an inclusion and exclusion form, which we keep on file.
- Writing the text to a highly structured template (about 1500–3000 words), using evidence from the final studies chosen, within 8–10 weeks of receiving the literature search.
- Working with *Clinical Evidence* editors to ensure that the final text meets epidemiological and style standards.
- Updating the text every six months using any new, sound evidence that becomes available. The *Clinical Evidence* in-house team will conduct the searches for contributors; your task is simply to filter out high quality studies and incorporate them in the existing text.
- To expand the topic to include a new question about once every 12–18 months.

If you would like to become a contributor for *Clinical Evidence* or require more information about what this involves please send your contact details and a copy of your CV, clearly stating the clinical area you are interested in, to Klara Brunnhuber (kbrunnhuber@bmjgroup.com).

### Call for peer reviewers

*Clinical Evidence* also needs to recruit a number of new peer reviewers specifically with an interest in the clinical areas stated above, and also others related to general practice. Peer reviewers are healthcare professionals or epidemiologists with experience in evidence-based medicine. As a peer reviewer you would be asked for your views on the clinical relevance, validity, and accessibility of specific topics within the journal, and their usefulness to the intended audience (international generalists and healthcare professionals, possibly with limited statistical knowledge). Topics are usually 1500–3000 words in length and we would ask you to review between 2–5 topics per year. The peer review process takes place throughout the year, and our turnaround time for each review is ideally 10–14 days.

If you are interested in becoming a peer reviewer for *Clinical Evidence*, please complete the peer review questionnaire at [www.clinicalevidence.com](http://www.clinicalevidence.com) or contact Klara Brunnhuber (kbrunnhuber@bmjgroup.com).