

Mental Retardation and Abnormal Skeletal Development (Dyggve-Melchior-Clausen Dysplasia) Due to Mutations in a Novel, Evolutionarily Conserved Gene

Daniel H. Cohn,^{1,3,5} Nadia Ehtesham,¹ Deborah Krakow,^{1,2,6} Sheila Unger,⁷ Alan Shanske,⁸ Kent Reinker,⁹ Berkley R. Powell,¹⁰ and David L. Rimoin^{1,3,4,5}

¹Medical Genetics-Birth Defects Center, Ahmanson Department of Pediatrics, and ²Department of Obstetrics and Gynecology, Cedars-Sinai Research Institute, Los Angeles; Departments of ³Human Genetics, ⁴Medicine, ⁵Pediatrics, and ⁶Obstetrics and Gynecology, Geffen School of Medicine, University of California Los Angeles, Los Angeles; ⁷Medical Genetics/Metabolism, Division of Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto; ⁸Center for Congenital Disorders, Children's Hospital at Montefiore, Albert Einstein College of Medicine, Bronx; ⁹University of Texas Health Sciences Center, San Antonio; and ¹⁰Children's Hospital Central California, Madera and Department of Pediatrics, UCSF School of Medicine, Fresno, CA

Dyggve-Melchior-Clausen dysplasia (DMC) and Smith-McCort dysplasia (SMC) are similar, rare autosomal recessive osteochondrodysplasias. The radiographic features and cartilage histology in DMC and SMC are identical. However, patients with DMC exhibit significant developmental delay and mental retardation, the major features that distinguish the two conditions. Linkage studies localized the SMC and DMC disease genes to chromosome 18q12-21.1, providing evidence suggesting that they are allelic disorders. Sequence analysis of the coding exons of the *FLJ90130* gene, a highly evolutionarily conserved gene within the recombination interval defined in the linkage study, identified mutations in SMC and DMC patients. The affected individuals in two consanguineous DMC families were homozygous for a stop codon mutation and a frameshift mutation, respectively, demonstrating that DMC represents the *FLJ90130*-null phenotype. The data confirm the hypothesis that SMC and DMC are allelic disorders and identify a gene necessary for normal skeletal development and brain function.

Introduction

Dyggve-Melchior-Clausen dysplasia (DMC [MIM #223800]) and Smith-McCort dysplasia (SMC [MIM #607326]) are autosomal recessive osteochondrodysplasia phenotypes (Smith and McCort 1958; Dyggve et al. 1962; Kaufman et al. 1971; Spranger et al. 1976) that are classified together in the current nosology and classification of constitutional disorders of bone (International Nomenclature Group on Constitutional Disorders of Bone, in press). The radiographic features of both conditions include epiphyseal and metaphyseal irregularity, platyspondyly with a double-humped appearance of the vertebrae, and a characteristic lacy iliac crest. Short stature is progressive and complications include varus and valgus deformity at the knees, scoliosis, and limitation in joint movement. Both disorders exhibit abnormal cartilage histology, with the growth plate showing disordered chondrocyte columns that include degenerating cells and rough endoplasmic reticulum inclusions

(Horton et al. 1982; Engfeldt et al. 1983; Nakamura et al. 1997).

Linkage studies in SMC defined a 10.7-cM recombination interval at 18q12 that is likely to contain the disease gene (Ehtesham et al. 2002), and a similar interval was identified in DMC (Ehtesham et al. 2002; Thauvin-Robinet et al. 2002). Sequence analysis of the coding exons of *FLJ90130*, a highly evolutionarily conserved gene located within the interval defined in the linkage studies, identified mutations in both DMC and SMC families. These data demonstrate that DMC and SMC are allelic disorders and identify a gene necessary for normal skeletal development and brain function.

Subjects and Methods

Subjects

All families were ascertained by the International Skeletal Dysplasia Registry and were collected after informed consent was obtained under an approved institutional review board protocol. Affected individuals presented with characteristic clinical and radiographic (fig. 1) findings. This included marked disproportionate short stature with a short barrel-shaped chest, kyphoscoliosis, and joint limitation. The Dyggve-Melchior-Clausen dysplasia patients showed developmental delay and mental retardation, whereas the Smith-McCort dysplasia patients

Received October 17, 2002; accepted for publication November 11, 2002; electronically published December 16, 2002.

Address for correspondence and reprints: Dr. Daniel H. Cohn, Ph.D., Medical Genetics, SSB-3 Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048. E-mail: dan.cohn@cshs.org

© 2003 by The American Society of Human Genetics. All rights reserved. 0002-9297/2003/7202-0023\$15.00



Figure 1 *Left panel*, Lateral view of the spine, showing platyspondyly and the characteristic double-humped appearance of the vertebral bodies with anterior beaking. *Right panel*, Antero-posterior view of the hips, showing short, widened ilia with a lacy appearance of the iliac crests. The acetabulae are hypoplastic and irregular. The proximal femoral necks are short, with irregular metaphyses and epiphyses.

were of normal intelligence. A more detailed clinical and radiographic description of the Smith-McCort dysplasia families is presented in an article by Burns et al. (in press).

Mutation Analysis

Genomic DNA fragments containing the coding sequence and flanking splice-junction consensus sequences of each exon were amplified by PCR. Primer sequences are given in table 1. Amplified fragments were subjected to sequence analysis using the ABI Prism Big Dye version 3 sequencing kit (Applied Biosystems), and the products were separated by electrophoresis using an ABI model 377 automated DNA sequencer. Sequence data were processed using ABI software and analyzed using Sequencher.

RT-PCR

Total RNA from frozen cartilage was prepared by Trizol extraction. Reverse transcription used 1 μ g of the total RNA, 100 ng random hexamers, Superscript reverse transcriptase and buffers from a kit under conditions recommended by the manufacturer (Invitrogen). Amplification used primers derived from exon 7 (5'-TGTTTTCCTTTCCTGCCAAC-3') and exon 9 (5'-AGAGGGGAAGAAAGCTCTGG-3'). The cDNA fragments were separated by agarose gel electrophoresis.

Results and Discussion

Sequence analysis was carried out for genes in the DMC/SMC linkage interval (Ehtesham et al. 2002; Thauvin-Robinet et al. 2002) in affected individuals from three families with DMC and two families with SMC. There were no mutations identified in either the coding exons or flanking splice-junction consensus sequences of the *MADH2* (Ehtesham et al. 2002), *MADH4*, *MADH7*, and *TPMT* genes, so structural mutations in these genes were excluded as the cause of SMC and DMC.

Mutations were identified in the *FLJ90130* gene in all five families, and these were the only families studied. *FLJ90130* consists of 17 exons distributed over >400 kb of genomic DNA. Features of the gene structure are given in table 2. The initiating methionine codon was identified in exon 2. Although the initiation codon was in a poor Kozak consensus sequence (GAAGAUGG), there was an in-frame stop codon 6 bp to the 5' side, supporting the identity of the AUG as the initiator. The stop codon was in exon 17 and was followed by a polyadenylation signal (AATAAA) 228 bp into the 3' UTR. In the cDNA, the poly A followed 18 bp after the polyadenylation signal. The cDNA encodes a 669-amino acid protein of unknown function, with an implied molecular weight of 75,918 kDa. To identify the mutations, primers were derived from the genomic sequence for amplification and sequence analysis of the coding exons (2–17) and flanking splice junctions.

Table 1**Amplification of the *FLJ90130* Exons**

Exon	PCR Product Size (bp)	Primer Sequences
2	415	5'-GCCCCATATTTCTGTAACTTTG-3' 5'-AATCAGCATACCACACAAGTTGA-3'
3	326	5'-TTCCTACTGTGCACCATATTTTTC-3' 5'-AGCATCTGTGAATGCCACAA-3'
4	279	5'-AGCAATTTTAATATGACATTTGGTTTT-3' 5'-CCACAGTCCATTTCCATCAA-3'
5	310	5'-CGGTTGTCAAATCAGGAAGAA-3' 5'-TCAAAATCATACTTTTATTCCCTGAAA-3'
6	238	5'-GGAATGAAAATGTAACCTTCTGAACAA-3' 5'-GGACCCACAAACTTCTCAACA-3'
7	336	5'-AGTGTCTTAATCTTCCAAAGTGACTG-3' 5'-TAGAATGCGGTTCCCTATGC-3'
8	394	5'-TGTGAGTTGTGCCAATAGCC-3' 5'-TGAACCAAGCTGCACAAATTC-3'
9	283	5'-ATCTCATTGGTGGCAGCAC-3' 5'-AACAAGCAATACTCTCCCATT-3'
10	351	5'-AATGGGAATTGAGTTGTACACATTAC-3' 5'-GCTGTGCCATGTACCCTAAAA-3'
11	328	5'-AAGCAAAGAGCACAAAATGAAA-3' 5'-GGTTCTACAGGAACATCTACCA-3'
12	342	5'-CCTTCAGAGCAAAGTCATCGT-3' 5'-CAGTGACCGGTATTTTCTACCA-3'
13	245	5'-CAGTGACAACTGAGTTTACATTTC-3' 5'-ATAGCTCAGCCAGAGGCAAA-3'
14	240	5'-CCACAGAAGACAGAAAAGAAAACA-3' 5'-TGGGATTACAGGCAGGAGTT-3'
15	369	5'-CTTCATGATCTTTCCAATGCTG-3' 5'-CAAGGCTTATAAACATTCTCTGCT-3'
16	281	5'-CATGTGAGAAAAGGCGGCTAC-3' 5'-AGCAGCATGATTCTGGATAGG-3'
17	533	5'-ACCACCTGCAAACTCTACC-3' 5'-GCAACAGGATCTTGGGAAAAG-3'

The affected individual in DMC family R94-187, a consanguineous family originating in the Dominican Republic, was homozygous for a point mutation in exon 2 (C→G) that predicted a Y16X change (fig. 2a). Both parents were carriers of the mutation, and the unaffected child was homozygous for the normal sequence, compatible with the chromosome 18 haplotypes inherited from her parents (Ehtesham et al. 2002). The very early truncation of the protein implied by the mutation strongly suggests that the mutation is a null. Thus, DMC appears to be the phenotype resulting from absence of the *FLJ90130* gene product.

In family R02-106, which was affected with DMC, the two affected individuals were homozygous for deletion of the last nucleotide of exon 8, predicting a frameshift after codon 254 and a stop 9 codons further downstream (fig. 2b). Cultured cells from the family were not available, so it could not be determined if the mRNA derived from the mutant allele was stable to

nonsense-mediated decay (Maquat 2000) or if the predicted truncated protein was synthesized and stable. However, the marked truncation of the protein implied by the mutation suggests that a truncated protein would likely not be functional. Homozygosity for the mutation was expected because the parents were first cousins and the affected individuals were homozygous for markers throughout the interval defined in the linkage studies (Ehtesham et al. 2002). Both parents were carriers of the mutation, and the unaffected child was also heterozygous for the pathologic change.

In DMC family R96-219, the affected child was a compound heterozygote. The first mutation was a transversion in exon 5 (T→A) that predicted a Y132X change (fig. 2c), presumably a null allele. This mutation was inherited from the mother. The second mutation, which was inherited from the father, was a point mutation that predicted a N469Y missense change (fig. 2c). Because we are unsure of the biological function of the protein and the roles of specific amino acid residues in its activity, we cannot be absolutely sure if this change is a pathological mutation. However, no other sequence changes among the *FLJ90130* exons were identified. Furthermore, BLAST

Table 2**Structure of the *FLJ90130* Gene**

Exon/Intron Number	Exon Size (bp) ^a	Amino Acid Residues ^b	Splice Donor	Splice Acceptor	Intron Size (kb) ^c
1			GTGGGC	TAG	29.3
2	193	1–47	GTAAGA	CAG	38.6
3	53	48–64	GTAAGT	CAG	11.8
4	94	65–96	GTAAGC	TAG	1.0
5	133	97–140	GTAAGT	CAG	15.3
6	73	141–165	GTAAGT	TAG	29.3
7	126	166–207	GTATGT	TAG	1.7
8	143	208–254	GTAAGT	CAG	44.6
9	183	255–315	GTTTGT	CAG	4.3
10	179	316–375	GTAAGT	TAG	9.7
11	126	376–417	GTAAGT	CAG	13.6
12	114	418–455	GTAAGT	CAG	1.3
13	95	456–487	GTAAT	CAG	93.3
14	103	488–521	GTAAGT	AAG	44.7
15	183	522–582	GTGAGT	CAG	21.2
16	114	583–620	GTAAGA	CAG	53.2
17		621–669			

^a The size of exon 1 is undetermined, as the 5' nucleotide of the gene has not been defined. Similarly, the 3' nucleotide of the transcript is undetermined, so the size of exon 17 is uncertain.

^b Where a codon is split between two exons, the amino acid residue is included in the exon with two nucleotides of the codon.

^c The genome assembly for the region of chromosome 18 containing the *FLJ90130* gene has not been finalized, so intron sizes are approximate.

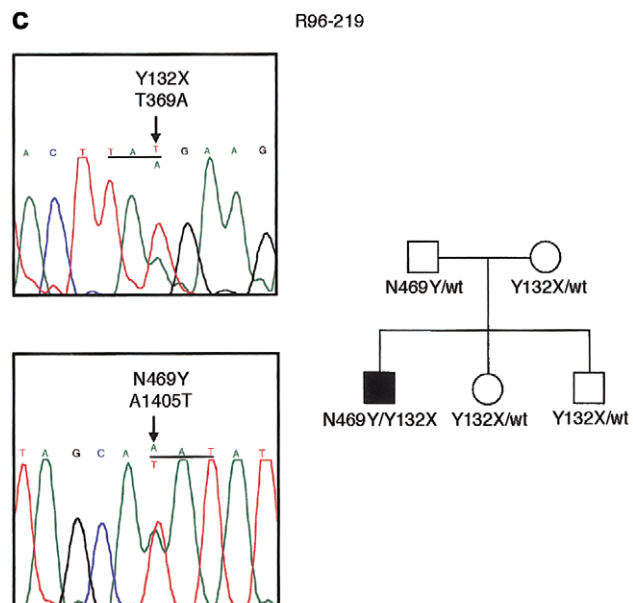
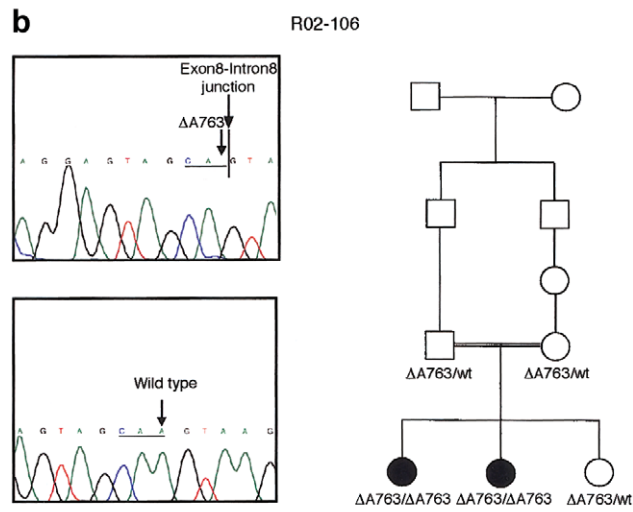
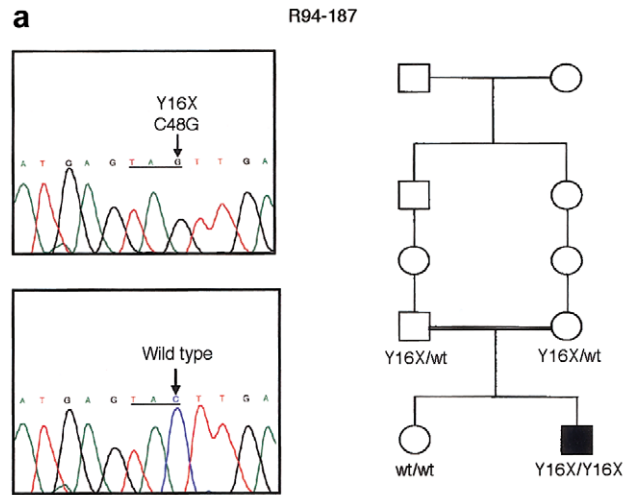


Table 3**Comparison of the FLJ90130 Protein Sequence and Homologous Proteins from other Species**

Species	Identity (%)	Similarity (%)
<i>M. musculus</i>	94	96
<i>C. elegans</i>	37	55
<i>D. melanogaster</i>	42	61
<i>A. gambiae</i>	42	59
<i>A. thaliana</i>	31	53

analysis of the predicted FLJ90130 protein sequence identified homologous proteins in other species (fig. 3; table 3), including *Mus musculus*, *Drosophila melanogaster*, *Anopheles gambiae*, *Caenorhabditis elegans*, and *Arabidopsis thaliana*, and amino acid residue N469 was conserved among all of them. Residue 469 is also within a region of the protein with very high cross-species conservation. The evolutionary conservation, from mammals to plants, suggests that residue 469 is important and therefore that the N469Y substitution could affect the function of the encoded protein. Because the patient has the more severe DMC phenotype, it is likely that the mutation has a markedly detrimental effect on protein function.

In SMC families R94-126 and R98-116 (Ehtesham et al. 2002; Burns et al., in press), both of which originated in Guam, the affected individuals were heterozygous for a mutation in the splice acceptor (TAG→TGG) immediately preceding exon 8, predicting skipping of exon 8 and a shift in the reading frame (fig. 4a). Skipping of the exon would lead to a protein with an altered sequence after amino acid residue 207 and a stop 62 codons later. The sequence change was inherited from the fathers in both families. Affected individuals in both families were also heterozygous for a point mutation (G→A) that predicted an E87K substitution, an acidic-to-basic charge change that was inherited from their respective mothers (fig. 4a). Since the mothers in both families share a chromosome 18 haplotype inherited by their affected offspring, it is likely that this is an ancestral mutation. Although we did not have access to a population of ethnically matched individuals with which to determine if the sequence change is a patho-

logical mutation or a polymorphism, like residue N469 described above, residue E87 is evolutionarily conserved among homologous proteins in other species (fig. 3).

To determine if the splice-junction mutation identified in SMC led to exon skipping, total RNA from cartilage of patient II-5 in family R94-126 and patient III-2 in family R98-116 was used for amplification by RT-PCR of a cDNA fragment spanning the exon. Compared with control RNA, a normal-sized cDNA fragment and a cDNA fragment of a size compatible with exon skipping was amplified from each patient sample (fig. 4b). Sequence analysis of the smaller amplified fragment confirmed that the exon 8 sequences were missing from the fragment, demonstrating that the mutation led to exon skipping. The abundance of the shorter amplified fragment was very much reduced relative to the normal-sized product. To determine if this resulted from exon skipping in only a proportion of transcripts derived from the abnormal allele, a cDNA fragment spanning both mutations was amplified, and gel electrophoresis was used to separate the fragment of normal size from the fragment with the skipped exon. Sequence analysis in the region of the E87K mutation showed both the mutant and normal sequences among the normal-sized fragments, indicating that these fragments were derived from both alleles (data not shown). Sequence analysis across the skipped exon region showed that the normal-sized fragments contained only the normal exon 8 sequence. These data demonstrated that exon skipping was leaky and that some normal protein is likely to be produced from the exon skipping allele.

The *FLJ90130* cDNA predicts a 669-residue protein of unknown function. BLAST analysis of the protein sequence revealed highly homologous and thus likely orthologous proteins in many species (fig. 3; table 3), but in no species was a function ascribed to the protein. Although DMC and SMC result in defects in the brain and cartilage, database searching has identified *FLJ90130* transcripts derived from many tissues, suggesting that the *FLJ90130* gene is ubiquitously expressed. These data, however, do not allow assessment of the comparative level of expression among tissues. Molecular modeling using the Tmpred algorithm (Hofmann and Stoffel 1993) provided strong evidence that FLJ90130 is a transmembrane protein

Figure 2 Mutations in DMC. *a*, Chromatograms showing the mutation in family R94-187 and (*below*) a normal sequence. The sequence, with the mutant nucleotide and the corresponding normal nucleotide identified with an arrow, is shown above, and the reading frame is underlined. The nucleotide sequence is numbered from the beginning of the coding sequence. The pedigree, demonstrating segregation of the mutation in the family, is shown to the right. *b*, Chromatograms showing the mutation in family R02-106, together with a normal sequence. An arrow identifies the mutant nucleotide and the corresponding normal nucleotide, and the reading frame is underlined. A vertical line identifies the exon-intron junction. Segregation of the mutation in the family is shown on the pedigree to the right. *c*, Chromatograms showing the two mutations identified in family R96-219. The reading frame is underlined, and, in each panel, an arrow identifies the normal (*above*) and mutant (*below*) sequences. Segregation of the mutations in the family is shown on the pedigree to the right.

Hs	-----MGSNSSRIGDLPKNEYLKKLS	21
Mm	-----MGSNSSKISDLPKNEYLKKLS	21
Ce	-----MGSVITKETVLQENIYLKKLS	21
Dm	-----MGINVSRADLGSNQWLQRFV	21
Ag	VQRNTLYTVVKPILLNRAECLRVEEKFKYRSVQKFSTHRMGMNVSRQVDLTENEYLQRFV	60
At	-----MGGVPSVTPRKTTGGDDVVAEYLIATF	26
	: : *	
Hs	GTEISISEND-PFWNQLLSFSFPAPTSSSELKLLLEEATIS-VCRSLVKNNPRTGNLALIK	79
Mm	GPESISEND-PFWNQLFSSFPAPTSSSELKLLLEEATIS-VCKSLVENNPRGNLAALTK	79
Ce	GLEPVDYD-PFVNKLLSFLKFDDEDETRIALESALDDHLQCLMYNTQTGNFAAFIR	80
Dm	GRQHIAHDEAFWNGLLNYNIVLPENSQD-QLNLDLRLALCQSFIGNNLTGNFGSLVT	80
Ag	GKTHVPAydddfwNnflQYHINLPTNSEE-QLSLDSRLETLCQSFISHNLHTGNFGSLIN	119
At	VGEKSFPLASDFWNKLELPLSSRWPSDR-----VQQACELFAQSNGYTRHLAKLLI	78
	*** ::. : .. .: . : . * ::. :	
Hs	VFLSRTEKELKLSAECQNHFIWQTHNALFIIICLLKVFICQMSEEEELQLHFTYE-----	133
Mm	VFLSRTELRLSAECQNHFIWQTHNALFIIICLLKVFIRELSEEEELQLHFTYE-----	133
Ce	LFLRRATELKTSEQCENKIYLWQTSNALLILRYIARFLTQRMTEKEFVRIFAKSHESTE-	139
Dm	VFLEKTESELLSLSDQESNMHVWQTFNALFIIRSLVKYINETGSEFQLLQHFAMPNAELL	140
Ag	VFLKVKSELLTSLDTESNVHIWHAFNALFIIRCLVKYMIETGSEYQLLQHFALPVQAAD	179
At	HLSWCLQELQLASDDQSSLYKKAVN-ATYISSVFLKYLIENGKSDSLQELHLSLDESEPV	137
	: * : : . . . * * : : : . . . :	
Hs	-EKSPGNYSSDSED-----LLEELLCCMLQITDIPLLDIT	168
Mm	-EKSPGSYSSDSED-----LLEELLCSLVQLITDTPLLDIT	168
Ce	TESSSTSSSSDEEDDDEQKSENENNGATKYIPIVFQNTAEFTYELVSIILINISVNETT	199
Dm	QAALQQQTPAESATIAMEA-TEQSAAAAPVIVDGSKFETFDALVNLIVVIPVKEFT	199
Ag	SDETPSDGTAAGTSTAIPIDMRTVEETKAAVAKLVDGSKFETFLEALVNIIVVIPVKEFT	239
At	PHGFVMDQDIQNFVMSVLSFIGSNEVSPNSYVLHQELLNFMVLVMTSQLLSGPSHGPTD	197
	: . : : .	
Hs	YEISVEAISTMVVFLSCQLFHKEVLRQSI SHKYLMRGCPCLPYTSKLVKTLLYNFIRQEKP	228
Mm	YEIAVEAISAMVVFLSCQLFHKEVLRQSI SHKYLMOGCPCLPYTSKLVKTLLYNFIRQEKP	228
Ce	LAIHVEAVRCLLTLSSQLYNESIVNTSIVFRFFIDGSCAKHAASLTKTLNLYLIHN--	257
Dm	YHLHLEAVNMLITLLSVHLFAQQPTDKSIVFRFTVFK--CQHANVMSALLHFVARIV--	254
Ag	YHLHLEAVNCMIILLVSLFSQQTMEKSTIYRTIYR--CQHANTLMSALLHFVSRMS--	294
At	ANPFIDAAMTQEKSLVSLVRRLLNYSRHRTPPNAKSYMYSQDGSQILERVGSAAAS	257
	::* * : . .: . . : : . :*	
Hs	PPPGAHVFPQQSDGGGLLYGPASGVATGLWTVFTLG-----GVGSKAAASPEL	276
Mm	PPPETHVFPQQSDGGGLLYGLASGVATGLWTVFTLG-----GAGSKAAASPEL	276
Ce	---SEYHMTVQKPQESIVFGLAS---SMWSMVQMA-----TGIDSABEEKKP	298
Dm	--EVPHTMFGSSSAGSFVFGIAE---SLLSIFTFR-KQPDILKAGQAAGGELSQRFR	307
Ag	--QAPQTMFGFGTGGSFVFGIAE---SLWSILTFARKQPDILTA-----HDLPTAFRE	342
At	LVLPLNLYLVSNSSGSKNP-LAECSLHVLILILINYH-----KSIMSDESMTDKS	305
	. . * . : : . .	
Hs	SSPLANQSLLLLLVLANLTDASDAPNPYRQAIMSFKNTQDSSPFPSSIPHAFQINFNSLY	336
Mm	TSPLANQSLLLLLVLVNLTDAPDIPNPYRQAVSSFKNTQDSSPFPSSIPHTFQINFNSLY	336
Ce	PLTLGNLSILLNLS-CHQPLNSSNPFKETTALFQNAQEVSTLPTQIVS-FKIDYNGLY	356
Dm	HYPLANQSLLLILILTNHCTAQD--NAYRTSLFSCADSKDS---PKQGAVSFQIDFSAVY	362
Ag	HYPLANQSLLLILILTNHHATKD--NPYRISLFGCSDSQDS---PKEDPATFKIDFSSLY	397
At	DDSATSESVSKVHVFSSDNTFSKALANARDVEFDRSDVEGN---AHPAGPHVRIPFASLF	362
	. . * : : . . : : : . . :* : : :	
Hs	TALCEQQTSDQATLLLYTLLHQNSNIRTYMLARTDMENLVLPFILEILYH--VEERNSHHV	394
Mm	TALCEQQTSDQATLLLYTLLHQNSNVRTYVLRARTDMENLVLPFILEILYH--VEERNSHHV	394

Ce	ERLCATAGQEPMLLLYMLLQANSQFRNYVLSRINLENLVVPVLRILHDGVATSNNSHHV	416
Dm	ETLCTIVTIDQATLLLYLLLHRNERFYRFVMMQQDLEQLVIPILQTLN--APDSNSHHI	420
Ag	NTLCRIVTIDQATLLLYLLLRYNRQRFQKYVMAQQNLQQLVIPILQTLN--APDSTSHHI	455
At	DTLGMFLADEGAVLLLYSLLQNSDFKEYVLRVTDLDTLLMPILETLN-ASKRTSSNQI	421
	* : . ***** ** * . . :: : : : : * : * : * . * : . * : : :	
Hs	YMALIILLILTEDDGFNRSIHEVILKNITWY-SERVLTEISLGSLLILVVIRTIQYNMTR	453
Mm	YMALIILLILTEDDGFNRSIHEVILRNITWY-SERVLTEISLGSLLILVVIRTIQYNMTR	453
Ce	YLALIVALILSEDDIFCKIIHETMIKDLGWLSDSDFSVREISLGGTLTLVLRVRAIQKNALK	476
Dm	YMSLIVLLILSEDEGFNKNVHTIMLKNITWY-TERTISEISLGGILILVIRTIQYNMMLK	479
Ag	YMSLIVLLILSEDDNFNKSVMHQIMLKNITWY-TERSISEISLGGLLILVVIRTIQYNMMLK	514
At	YMLLIVLLILSQDSSFNSSIHKMILPSVPWY-KEHLLHQTSLSGLMVIILIRTVQHNLSK	480
	* : ** : *** : * : * : * : : : * : . : : : * : * : * : * : * : * : :	
Hs	TRDKYLHTNCLAALANMSAQFRSLHQYAAQRIISLFSLLSKKHNVLEQA---TQSLRGS	510
Mm	TRDKYLHTNCLAALANMSAQFRSLHQYAAQRIISLFSLLSKKHNVLEQA---TQSLRGP	510
Ce	TRDRYLHTNCLAALANMSAFFKNLAPIVCQRLISLLDLLTKRHAKMVDHM---RVSSQND	533
Dm	MRDKYLHTNCLAALANMSGHFRALHPYVAQRLVSLFETLARKHTRLDAQL---KEPADSA	536
Ag	MRDKYLHTNCLAALANMSGQFQSLHPYVAQR-----	545
At	LRDVYLQTTCLATLANMAPHAHLSAYASQRLVSLFYMLSRKYNKLSDLTGDKLQSIKIN	540
	** * : * : * : * : * : * : : * : . . ** : : : :	
Hs	LS-SNDVPLPDYAQDLNVIEEVIRMMLEIINSCLTNSLHNNPNLVYALLYKRDLEQFRT	569
Mm	LS-SSDVPLPDYAQDLNVIEEVIRMMLEIINSCLTNSLHNNPNLVYALLYKRDLEQFRT	569
Ce	VA-DGQP--INFHDDITALEEGIRTLLEIINSALCGGLRHNSHLIYNLLYHRALFDAYVQ	590
Dm	VFVNVSTTPEDMLQDLVLEEVLRMVLEILNSCLTNQLVYCPNLVYTLTYKRSVFEGFRS	596
Ag	-----LLQDLVLEEVLRMVLEILNSCLSHQLVYCPNLVYTLTYKRNVFQFQFRS	594
At	LSGEDVGVSEDLAAELQIFTFDLRLVLDILNAILTYALPRNPEIVYAIMHRQEVFQPFKN	600
	. . : : : * : * : * : * : * : * : . . . : * : : : : * : : :	
Hs	HPSFQDIMQNIDLVISFFSSRLLQAG---AELSVERVLEI IKQGVVALPKDRLKFFPELK	626
Mm	HPSFQDIMQNIDLVISFFSSRLLQAG---AELSVERVLEI IKQGVVALPKDRLKFFPELK	626
Ce	HMPFQDLLVNIAAVISHFSSKVIHVP---AGDGST-MLQIIEKEANWPTDRLAKFFPELK	646
Dm	HHAFQDVIQNI DMVVGFFSSRLQRVQEQRGELGVNEVLEVISKQASQWSSDRLRKFDPDLK	656
Ag	HSAFQDIIQNI DMVVGFFSSRLQRVQEQRGELGVSEVLEVISKQASQWSSDRLRKFDPDLK	654
At	HPRFHELVENIYTVLDFFNMRMDSQRS-REWSVQKVLQFIINNCRSWRGEGMKMFTQLH	659
	* * : : : * : * : . . . * : : . . : : * : * : * : * : * : * : :	
Hs	FKYVEEQPEEFFIPYVWVSLVYNSAVGLYWNPDQIQLFMTDSD-----	669
Mm	FKYVEEQPEEFFIPYVWVSLVYNSAVGLYWNPDQIQLFAMDS-----	669
Ce	FRYVEDEYTVDFVPYVWVRLSVQHS-GIYFETSRIKIFNAQSI-----	689
Dm	FKYVEEDAPEEFFIPYVWVTLVCKYG-CVHFSSESISKSVTTDIAC-----	699
Ag	FKYVEEDAPEEFFIPYVWVTLVCKAG-CVHFSSETIKGVSAEIV-----	696
At	FSYEQESHPEEFFIPYVWQLAFSRC-GFGFNPDAINLFPVPHVPEKEIEDERGESEEGKA	718
	* * : : . . : * : * : * : * : * : * : . .	
Hs	-----	
Mm	-----	
Ce	-----	
Dm	-----	
Ag	-----	
At	KVQELNEQRIVFDP 732	

Figure 3 Comparison of the protein sequences of FLJ90130 and homologous proteins in other species. ClustalW was used for the alignment. The species from which each sequence is derived is shown on the left (*M. musculus* [Mm], *D. melanogaster* [Dm], *A. gambiae* [Ag], *C. elegans* [Ce], and *A. thaliana* [At]), and the amino acid residue numbers are shown on the right. Colors are used to indicate classes of amino acid residues, including hydrophobic and aromatic (green), acidic (blue), basic (dark red), and hydrophilic (green) residues. Symbols below the amino acid sequences indicate identity (*), presence of conserved substitutions (:), or presence of semiconserved substitutions (.). A gray background highlights the residues changed by the two identified missense mutations, E87 and N469.

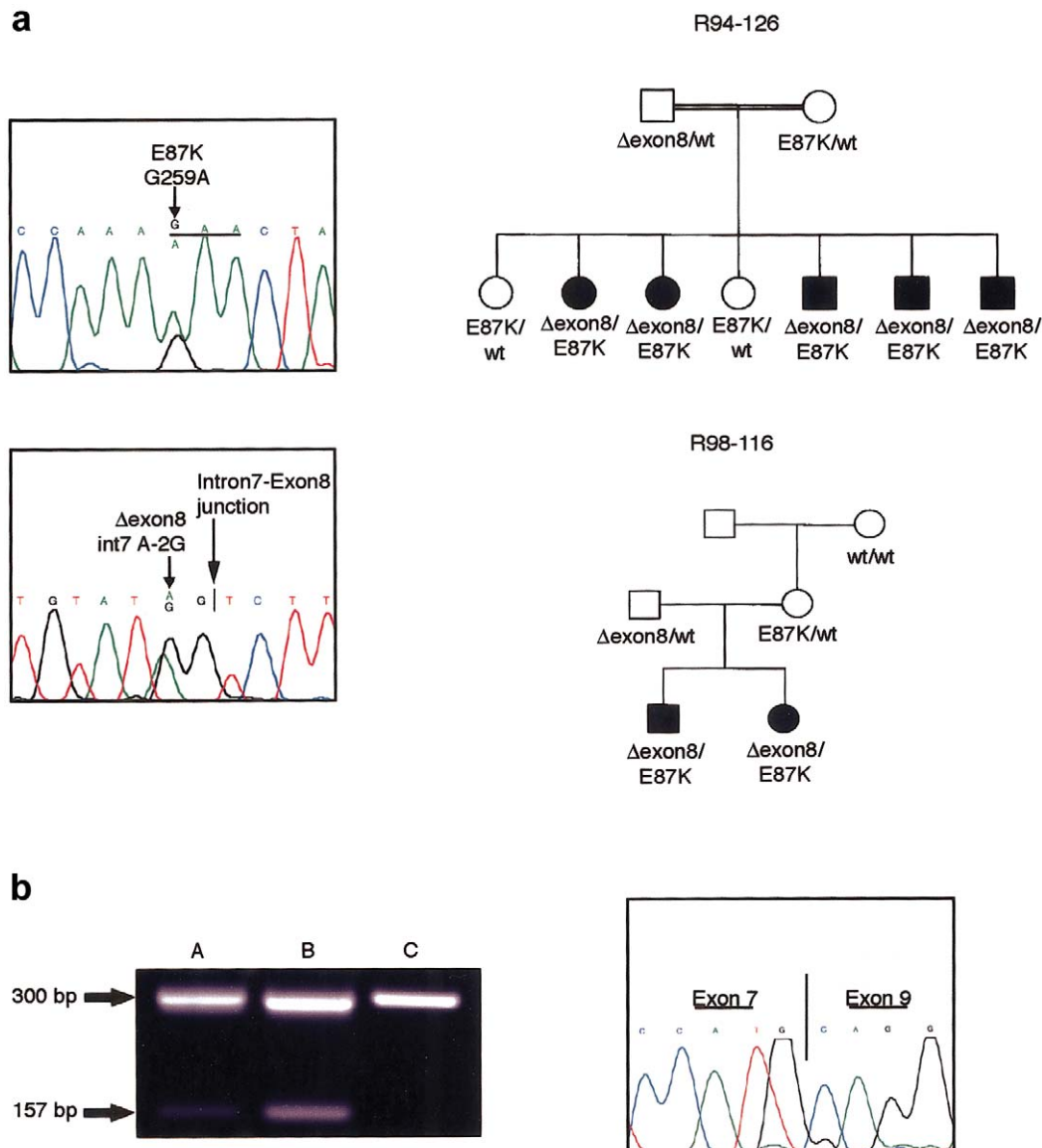


Figure 4 Mutations in SMC. *a*, Chromatograms showing the two mutations identified in families R94-126 and R98-116. For the E87K mutation, shown above, the reading frame is underlined, and the arrow identifies the normal (*above*) and mutant (*below*) sequences. For the exon-skipping mutation, the smaller arrow identifies the normal (*above*) and mutant (*below*) sequences. The intron-exon junction is identified with a vertical line and marked with the larger arrow. Segregation of the mutations in each family is shown on the pedigrees to the right. *b*, Skipping of exon 8 in cartilage RNA from an affected member of each family. On the left, an image of the agarose gel used to separate the amplified cDNA fragments from cartilage of patient II-5 in family R94-126 (“A”) and patient III-2 in family R98-116 (“B”), along with a fragment derived from normal fetal cartilage (“C”). Fragment sizes are indicated to the left. The chromatogram on the right shows the sequence across the exon 7/exon 9 boundary, which is identified by a vertical line, in the shorter cDNA fragment from patient II-5 in family R94-126.

with six transmembrane segments and a cytoplasmic amino terminus. The six predicted membrane-spanning regions, with the positions and/or consequences of the identified mutations noted, are diagrammed in figure 5. However, experimental verification will be required both to test the hypothesis that FLJ90130 is located in a membrane and to define the function of the protein in the normal and disease states. In particular, it is likely

that the null mutations lead to mRNA instability and no protein rather than a truncated protein.

Biochemical and morphologic data have suggested the hypothesis that DMC and SMC may result from a defect in proteoglycan metabolism (Engfeldt et al. 1983). The aggregated proteoglycan content of DMC iliac crest cartilage was reduced, and there was a concomitant increase in monomeric proteoglycan, suggest-

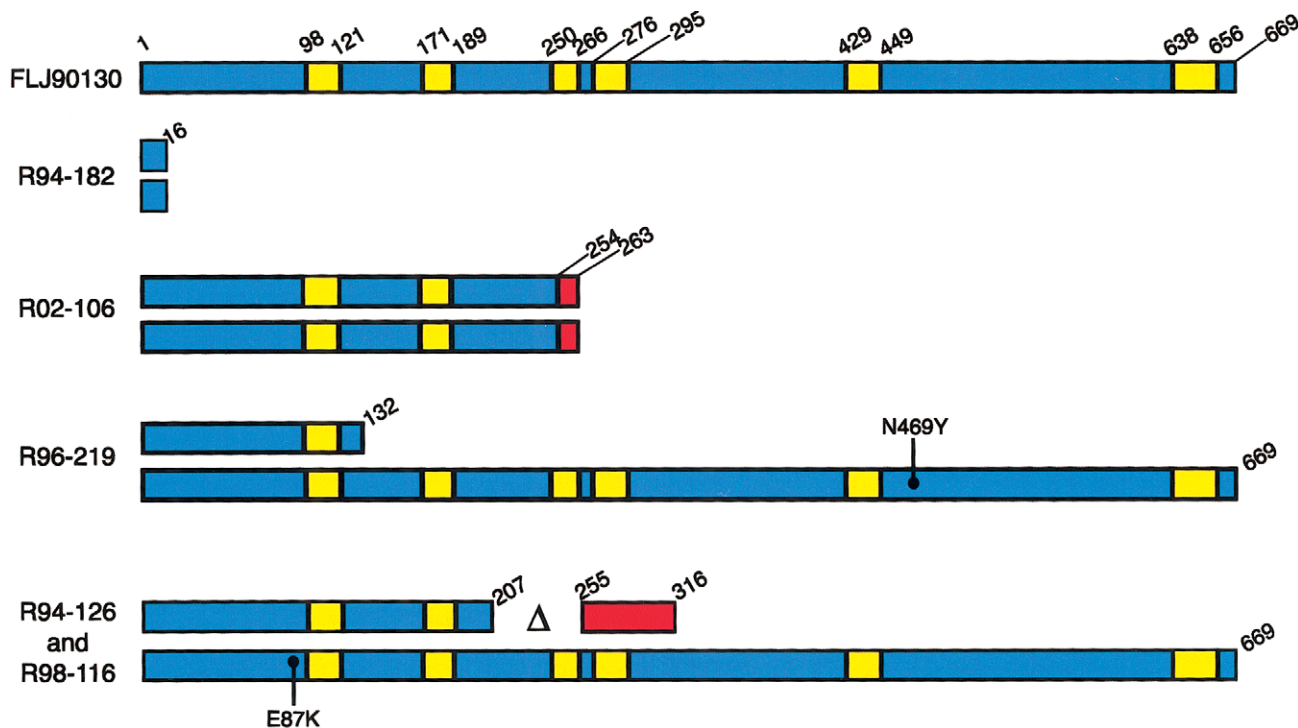


Figure 5 Structure of the FLJ90130 protein, with the six predicted transmembrane domains shown in yellow. The beginning and ending residues of each transmembrane segment are indicated. Shown below are the predicted consequences of each mutation on the protein sequence/structure in the three DMC and two SMC families studied. Where the mutation results in a frameshift, the portion of the protein beyond the frameshift is indicated in red. The locations of missense mutations are identified with a dot and the specific protein-sequence change.

ing that altered glycosaminoglycan function is a feature of DMC. Cartilage histology in both SMC and DMC have demonstrated thickened collagen fibrils, compatible with a possible deficiency in matrix glycoproteins (Horton et al. 1982; Engfeldt et al. 1983; Nakamura et al. 1997), which are known to regulate collagen fibril formation (Chandrasekhar et al. 1984; Vogel and Trotter 1987; Kuijer et al. 1988; Neame et al. 2000). Since normal proteoglycan biosynthesis is critical for elaboration of the extracellular matrix of cartilage and the brain, the hypothesis that *FLJ90130* encodes a protein that participates in proteoglycan metabolism is attractive. The high demand for proteoglycan synthesis in cartilage and brain could also explain why the phenotype is predominantly expressed in these tissues.

We conclude that SMC and DMC are allelic disorders that result from mutations in the *FLJ90130* gene. The data thus lay to rest the question of whether these two disorders result from mutations in the same gene. DMC appears to result from absence of the *FLJ90130* gene product, implying that at least one mutant allele in the SMC patients encodes a protein with partial function. The *FLJ90130* gene may encode a membrane protein that participates in proteoglycan metabolism, speculation that is supported by the recessive inheritance pat-

tern of the disease, the biochemical and histological findings in SMC and DMC cartilage, and the prominent role of proteoglycans in the two tissues most affected in these disorders, cartilage and the brain.

Acknowledgments

We thank the families, for their participation in this study, and Robert Pogue, for assistance with bioinformatic analysis of gene structure and protein function. This work was supported in part by grant HD22657 from the National Institutes of Health.

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

- ClustalW, <http://www.ebi.ac.uk/clustalw/index.html> (for protein sequence comparisons)
- GenBank, <http://www.ncbi.nlm.nih.gov/GenBank/> (for *FLJ90130* cDNA [accession number AK074611]; *FLJ90130* genomic [accession numbers NT_010874.1, NT_033905.2, and AC100778]; proteins homologous to *FLJ90130* in *M. musculus* [accession number BC038276], *D. melanogaster* [accession number NM_136587], *A. gambiae* [accession

number AAAB01008980], *C. elegans* [accession number NM_064091], and *A. thaliana* [accession number NM_100301])

Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/> (for Dyggve-Melchior-Clausen dysplasia [MIM 223800] and Smith-McCort dysplasia [MIM 607326])

References

- Burns C, Powell BR, MD, Hsia YE, Reinker K. Dyggve-Melchior-Clausen syndrome: report of 7 patients and review of the literature. *J Pediatr Orthop* (in press)
- Chandrasekhar S, Kleinman HK, Hassell JR, Martin GR, Termine JD, Trelstad RL (1984) Regulation of type I collagen fibril assembly by link protein and proteoglycans. *Coll Relat Res* 4:323–337
- Dyggve HV, Melchior JC, Clausen J (1962) Morquio-Ullrich's disease: an inborn error of metabolism? *Arch Dis Child* 37: 525–534
- Ehtesham N, Cantor RM, King LM, Reinker K, Powell BR, Shanske A, Unger S, Rimoin DL, Cohn DH (2002) Evidence that Smith-McCort dysplasia and Dyggve-Melchior-Clausen dysplasia are allelic disorders that result from mutations in a gene on chromosome 18q12. *Am J Hum Genet* 71: 947–951
- Engfeldt B, Bui T-H, Eklof O, Hjerpe A, Reinholt FP, Ritzen EM, Wilkstrom B (1983) Dyggve-Melchior-Clausen Dysplasia: morphological and biochemical findings in cartilage growth zones. *Acta Paediatr Scand* 72:269–274
- Hofmann K, Stoffel W (1993) TMbase—a database of membrane spanning protein segments. *Biol Chem Hoppe-Seyler* 374:166–170
- Horton WA, Scott CI (1982) Dyggve-Melchior-Clausen syndrome: a histochemical study of the growth plate. *J Bone Joint Surg* 64:408–415
- International Nomenclature Group on Constitutional Disorders of Bone (2002) International nosology and classification of constitutional disorders of bone. *Am J Med Genet* 113: 65–77
- Kaufman RL, Rimoin DL, McAlister WH (1971) The Dyggve-Melchior-Clausen syndrome. *Birth Defects Orig Artic Ser* 7: 144–149
- Kuijjer R, van de Stadt RJ, de Koning MH, van Kampen GP, van der Korst JK (1988) Influence of cartilage proteoglycans on type II collagen fibrillogenesis. *Connect Tissue Res* 17: 83–97
- Maquat LE (2000) Nonsense-mediated RNA decay in mammalian cells: a splicing-dependent means to down-regulate the levels of mRNAs that prematurely terminate translation. In: Sonenberg N, Hershey JWB, Mathews MB (eds) *Translational control of gene expression*. Cold Spring Harbor Press, Cold Spring Harbor, NY, pp 849–868
- Nakamura K, Kurokawa T, Nagano A, Nakamura S, Taniguchi K, Hamazaki M (1997) Dyggve-Melchior-Clausen syndrome without mental retardation (Smith-McCort dysplasia): morphological findings in the growth plate of the iliac crest. *Am J Med Genet* 72:11–17
- Neame PJ, Kay CJ, McQuillan DJ, Beales MP, Hassell JR (2000) Independent modulation of collagen fibrillogenesis by decorin and lumican. *Cell Mol Life Sci* 57:859–863
- Smith R, McCort JJ (1958) Osteochondrodystrophy (Morquio-Brailsford type): occurrence in three siblings. *Calif Med* 88:55–59
- Spranger JW, Bierbaum B, Herrmann J (1976) Heterogeneity of Dyggve-Melchior-Clausen dwarfism. *Hum Genet* 33: 279–287
- Thauvin-Robinet C, El Ghouzzi V, Chemaitilly W, Dagoneau N, Boute O, Viot G, Megarbane A, Sefiani A, Munnich A, Le Merrer M, Cormier-Daire V (2002) Homozygosity mapping of a Dyggve-Melchior-Clausen syndrome gene to chromosome 18q21.1. *J Med Genet* 39:714–717
- Vogel KG, Trotter JA (1987) The effect of proteoglycans on the morphology of collagen fibrils formed in vitro. *Coll Relat Res* 7:105–114