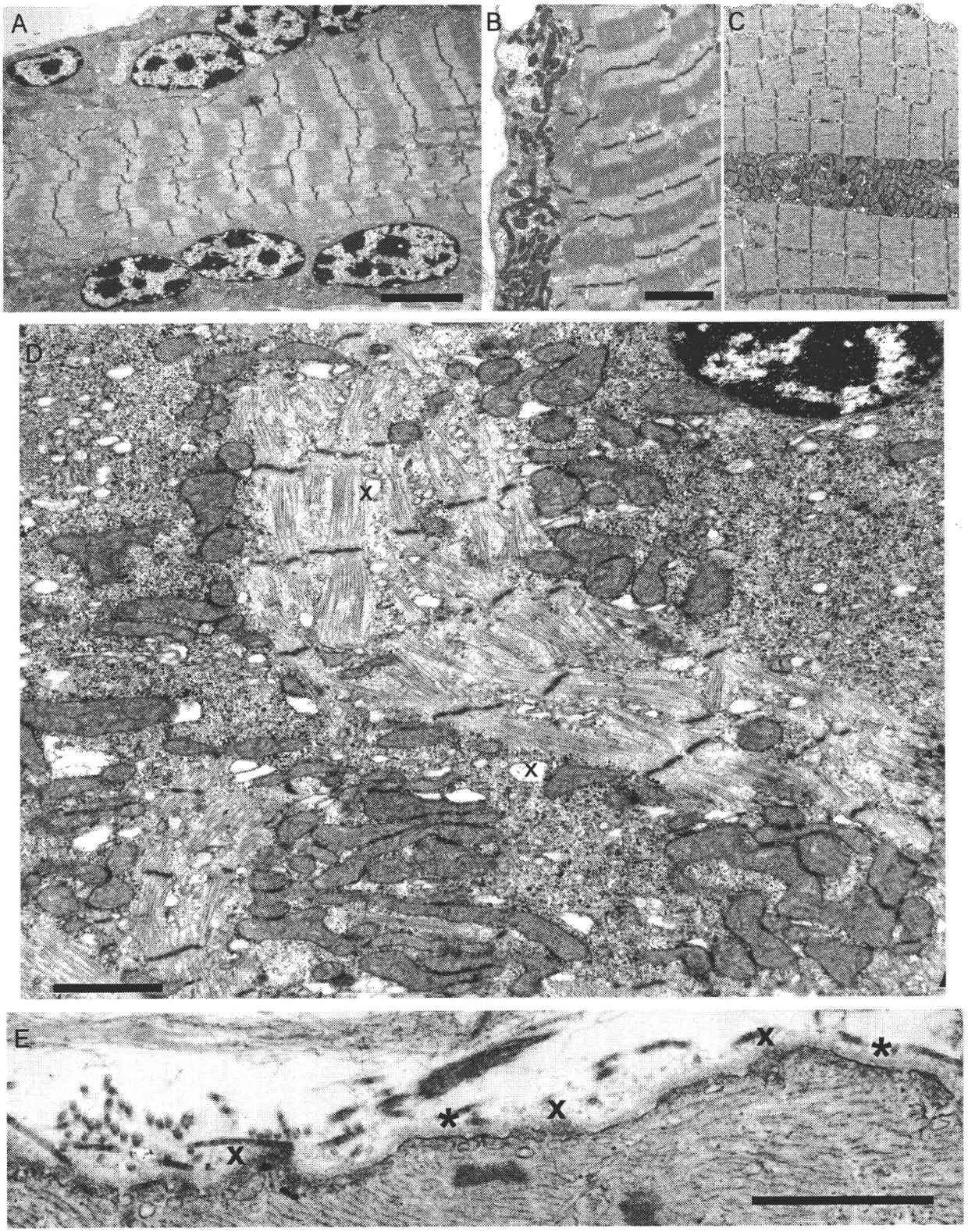
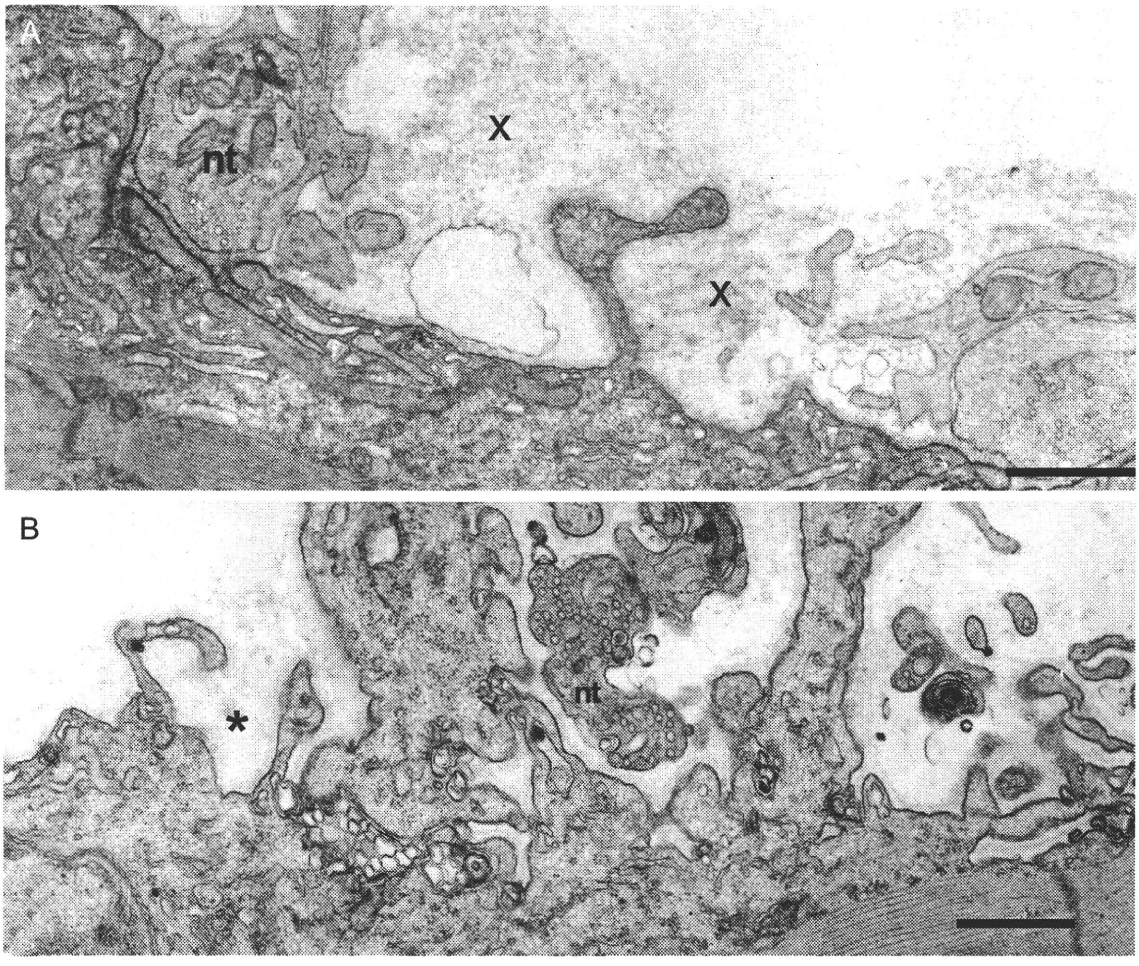
Figure 3 **Ultrastructural findings in abnormal muscle fibers of patient 2**



(A) Note subsarcolemmal rows of large nuclei harboring multiple prominent chromatin bodies. (B, C) Subsarcolemmal and intrafiber clusters of mitochondria surrounded by fiber regions devoid of mitochondria. (D) Aberrant and disrupted myofibrils surrounded by clusters of mitochondria intermingled with glycogen, ribosomes, and dilated vesicles (x). Note preapoptotic nucleus at upper right. (E) Focal sarcolemma defects due to gaps in the plasma membrane. Where the plasma membrane is absent, the overlying basal lamina is thickened (x). Small vesicles underlie the thickened basal lamina. Asterisks indicate segments of the preserved plasma membrane. Bars = 4 μm in (A), 3 μm (B, C), 1.4 μm in (D), 1 μm in (E).

Figure 4 **Abnormal endplate (EP) regions in patient 2**



(A) The imaged EP regions show partial occupancy of the postsynaptic region by the nerve terminal and remnants of degenerate folds (x). The nerve terminal (nt) occupies only part of the postsynaptic region. Degenerate remnants of the folds (x) appear over the simplified postsynaptic region from which folds were lost. Dark reaction product on postsynaptic membrane shows acetylcholine receptor localization with peroxidase labeled α-bungarotoxin. (B) Small nerve terminal occupies only part of a highly simplified postsynaptic region. Asterisk indicates remnants of basal lamina that surrounded preexisting folds. Bars = 1 μm.

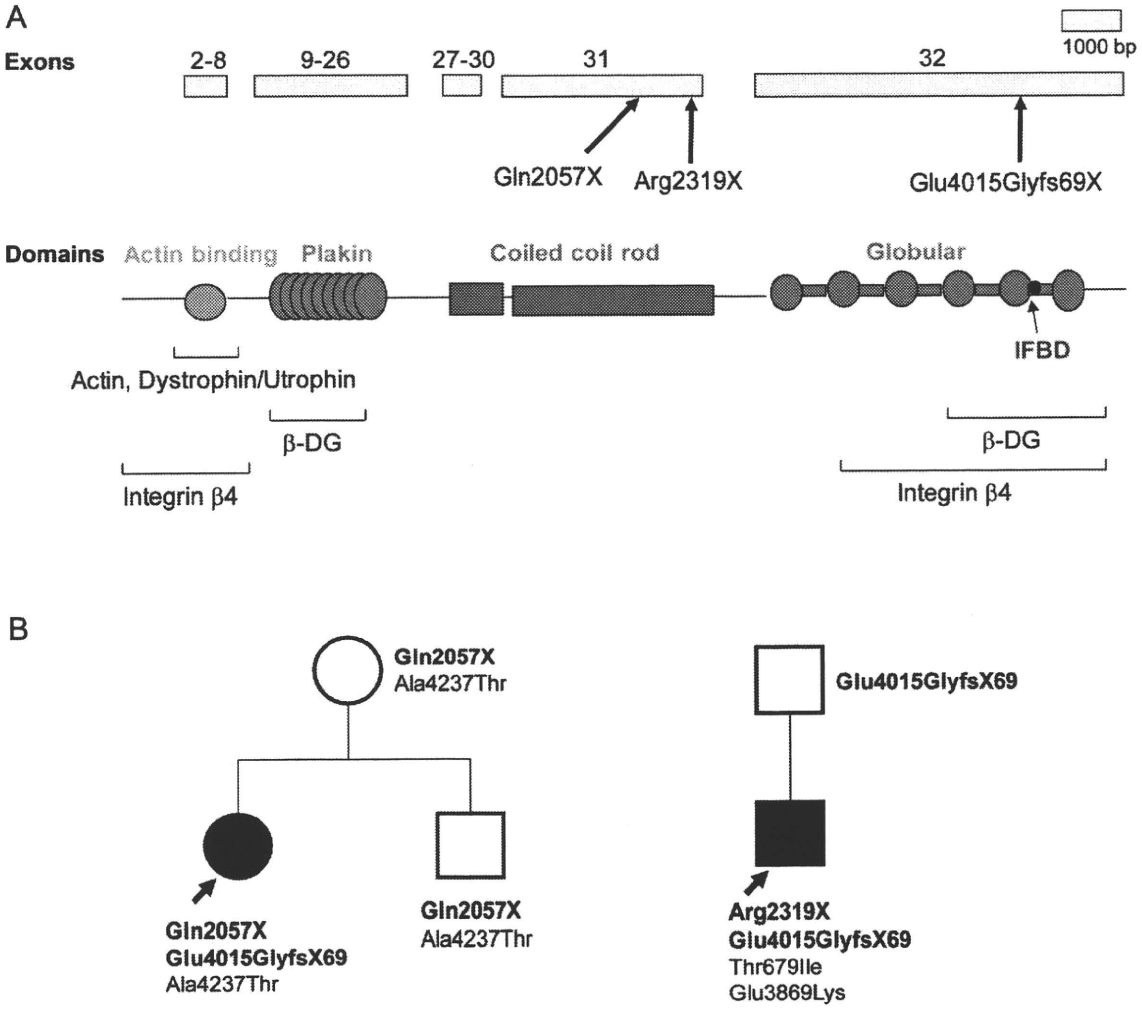
ated nerve fibers (not shown). In normal muscle, the anti-Rod Ab (figure 2E) was more reactive than the anti-C antibody (figure 2H) and stained type 1 fibers more intensely than type 2 fibers. In the patients, the anti-Rod Ab demonstrated loss of sarcoplasmic and trace sarcolemmal reactivity in type 1 fibers but, as noted by others,[19] type 2 fibers retained plectin positivity (figure 2, F and G, and figure e-2). In contrast, the C-20 Ab revealed no sarcoplasmic and only slight sarcolemmal reactivity in all fibers (figure 2I).

Muscle fiber ultrastructure in P2. ***Nuclear abnormalities.*** Consistent with light microscopy, numerous muscle fibers harbored subsarcolemmal rows or clusters of large ovoid nuclei containing 3 to 10 large, highly electron dense chromatin deposits (figure 3A). Nuclei in other fiber regions were of normal size and heterochromatic or euchromatic. Some nuclei harbored large clumps of heterochromatin with small islands of euchromatin indicating preapoptotic changes (figure 3D).

Mitochondrial abnormalities. In some fibers, the mitochondria were normally aligned with the Z disk and evenly distributed in the muscle fiber. In other fibers, the mitochondria aggregated into clusters under the sarcolemma (figure 3B), near nuclei, or in other fiber regions (figure 3C), leaving adjacent fiber regions depleted of mitochondria. Some mitochondrial aggregates were interspersed with glycogen granules, ribosomes, and dilated vesicles (figure 3D).

Myofibrillar abnormalities. Aberrant myofibrils appeared among dislocated organelles in both atrophic and nonatrophic fibers (figure 3D). Disintegration and streaming of Z disks were detected in some fiber regions with or without myofibrillar disarray. A few abnormal fiber regions harbored small nemaline rods. End-stage muscle fibers contained few mitochondria, remnants of

Figure 5 **Plectin domains and identified *PLEC* variants**



(A) Schematic representation of *PLEC* exons 2–32 indicating identified patient mutations and binding domains associated with C- and N-terminal regions of plectin.[39] (B) Family analysis of P1 and P2 shows transmission of pathogenic mutations (bold face) and polymorphisms.

Z disks with attached short filaments (Z brushes), and debris. Small membrane-bound vacuoles of different sizes that likely represent dilated components of the sarcoplasmic reticulum appeared in fiber regions with or without other abnormalities (figure 3D).

Sarcolemmal defects. The focal subsarcolemmal calcium deposits in scattered muscle fibers prompted us to search for sarcolemmal defects at the ultrastructural level and we detected these in some fibers of both patients. Where the plasma membrane was discontinuous, it was often covered by focally thickened basal lamina (figure 3E).

Endplate studies in P2. Synaptic contacts on the muscle fibers, visualized by the cholinesterase reaction, consisted of multiple small EP regions arrayed over an extended length of the muscle fiber (figure 2D). This finding has been observed in other patients with ongoing destruction and remodeling of the postsynaptic region.[31]

Colocalization of plectin and AChR by fluorescence microscopy showed strong expressions of plectin and AChR at normal EPs. At patient EPs AChR expression was not appreciably attenuated but plectin expression was barely perceptible (figure e-3).

On electron microscopy, some EP regions appeared normal but many had an abnormal conformation, displaying one or more of the following: partial occupancy of the postsynaptic region by nerve terminals, small nerve terminals, atrophic and remnants of degenerated junctional folds resulting in highly simplified postsynaptic regions, and nerve sprouts near degenerating EPs (figure 4 and figure e-4). Table e-1 shows the frequency of the observed conformational changes. Table e-2 shows morphometry revealing reduced size of presynaptic and postsynaptic EP components.

Genetic analysis. Genetic analysis was challenging owing to the very large size of the *PLEC* transcript

(~14 Kb vs ~11 Kb for dystrophin), the very large exons 31 (3381 bp) and 32 (6219 bp), multiple splice variants of exon 1, and polymorphisms that may be race-dependent. Eventually we detected 2 truncating mutations in each patient (figure 5). P1 harbors a stop codon mutation at nucleotide 6169 in exon 31 (c.6169C>T/p.Gln2057X), and a duplication at nucleotide 12043 in exon 32 that predicts 68 missense codons followed by a stop codon (c.12043dupG/p.Glu4015GlyfsX69). P2 harbors a previously reported nonsense mutation at nucleotide 6955 in exon 31 which generates a stop codon (c.6955C>T/p.Arg2319X),[16] and the same duplication mutation detected in P1 (figure 5B). Presence of the pathogenic mutation was confirmed at the cDNA level in both patients. Both stop codon mutations abrogate, and the c.12043dupG mutation disrupts, the IF binding site and one of the two β-dystroglycan and integrin β4 binding sites (figure 5A).

Both patients also harbored polymorphisms not listed in the SNP database (see figure 5B). In P1, p.Ala4237Thr was deemed a polymorphism because it appeared together with Gln2057X in the unaffected mother. In P2, p.Thr679Ile was present in 1 of 60 African Americans and p.Glu3869Lys in 1 of 100 African Americans, although both variants were absent in 200 Caucasians.

It has been suggested that expression of the rodless plectin transcript may mitigate the plectinopathy phenotype.[32] We therefore confirmed presence of the rodless domain by sequencing cDNA and used real-time PCR to compare the relative abundance of the rodless transcript in P1, P2, and 3 normal controls. The rodless transcript/full transcript ratio was 0.15 in P1, 0.32 in P2, and 0.22 ± 0.03 (mean ± SD) in 3 controls.

DISCUSSION Although each patient carries a nonsense mutation in *PLEC* exon 31 and an identical frameshift mutation in exon 32, the tempo of the disease was faster in P2 than in P1. In P2, EBS presented at age 6 weeks, MyS at age 3 years, and he lost ambulation by age 18 years. In P1, MyS presented at age 9 years, EBS at age 18 years, and she can still take a few steps at age 31 years. It has been suggested that expression of the rodless transcript can mitigate the phenotype in patients who carry mutations in the plectin rod domain.[32] However, real-time PCR indicates that expression of the rodless transcript was higher in the more severely affected P2. Thus in the patients studied by us the abundance of the rodless transcript was not a reliable indicator of the clinical phenotype.

Dislocation of the fiber organelles apparent at the ultrastructural level has multiple predictable consequences. Abnormal alignment and displacement of the myofibrils weakens or eliminates their contractile strength; separation of mitochondria from myofibrils renders energy delivery to contracting myofibrils inefficient. The eccentrically positioned large nuclei with multiple chromatin deposits may be dysfunctional or inefficient in their translational activities and in nuclear-cytoplasmic trafficking when not adjacent to organelles or fiber domains they subserve. Injury to the inadequately supported plasma membrane is evidenced by subsarcolemmal calcium deposits and sarcolemmal defects in a proportion of the fibers (see figure 2C, figure 3E, and figure e-1). Most of these were smaller than those in Duchenne dystrophy[27,33] but they still likely contribute to fiber injury and, if large, they may contribute to fiber necrosis.

Electron microscopy of the EPs indicates that plectin deficiency targets the junctional folds for destruction. When sarcomeres contract and relax, the extrajunctional sarcolemma bulges and relaxes but the junctional folds maintain a constant architecture.[34] This mandates enhanced rigidity of the junctional folds and renders them especially vulnerable to mechanical stress. Thus loss of cytoskeletal support of the junctional folds due to the plectin deficiency, as depicted in figure e-3, readily explains the progressive destruction of the folds. Destruction of junctional folds decreases the density or eliminates the voltage-gated Na^+ channels which are concentrated in troughs between the folds[35,36] and this increases the threshold for action potential generation.[37] Destruction of the folds also decreases the input resistance of the postsynaptic membrane and thereby the amplitude of synaptic potential.[38] Widening of the synaptic space reduces the concentration of acetylcholine before it reaches the junctional folds. The combination of these factors decreases quantal efficacy and compromises the safety margin of neuromuscular transmission. The relentless progression of the myasthenic symptoms in both patients implies that formation of new EP regions eventually fails to compensate for the ongoing destructive changes.

Why only a proportion of the muscle fibers is affected in a given muscle, and why some EPs are more severely affected than others, remain unanswered questions. It is uncertain that it can be attributed to preserved plectin expression in type 2 fibers, as shown by the 10F6 anti-Rod antibody. First, the anti-C terminal antibody showed nearly complete plectin deficiency in all muscle fibers in both patients. Second, the anti-rod domain antibody showed either no plectin reactivity in any fiber, or preserved immunoreactivity in type 1 instead of type 2 fibers.[19]

The fiber type specificities of the different anti-rod domain antibodies in plectinopathy are presently unexplained.

It is also unclear why some patients with EBS-MD have myasthenic symptoms and others do not. Possibly myasthenic weakness of the limb muscles is masked by the MD, or is overlooked in severely weak patients, but fatigable weakness of the oculobulbar muscles would be unlikely to go unrecognized. It also is not known whether some mutations are more prone to result in the MyS phenotype than others.

ACKNOWLEDGMENT

The authors thank Drs. Neill Graff-Radford and Nilufer Ertekin-Taner for anonymated DNA samples from African American control subjects.

DISCLOSURE

Dr. Selcen reports no disclosures. Dr. Juel has received research support from Alexion Pharmaceuticals, Inc. Dr. Hobson-Webb has served as a consultant for Novella Clinical Inc. and has received research support from Genzyme Corporation and AANEM Foundation. Dr. Smith reports no disclosures. Dr. Stickler has received research support from Agency for Toxic Substances and Disease Registry, Centers for Disease Control and Prevention (CDC) and has served as an expert witness in a legal proceeding. Ms. Bite reports no disclosures. Dr. Ohno has received Grants-in-Aid from the Japan Society for the Promotion of Science, the Ministry of Health, Labour and Welfare, and the Japan Science and Technology Agency. Dr. Engel serves as an Associate Editor of *Neurology*; receives publishing royalties for *Myology 3rd ed.* (McGraw-Hill, 2004); and has received research support from the NIH and the Muscular Dystrophy Association.

Received June 21, 2010. Accepted in final form September 21, 2010.

REFERENCES

1. Elliott CE, Becker B, Oehler S, Castanon MJ, Hauptmann R, Wiche G. Plectin transcript diversity: identification and tissue distribution of variants with distinct first coding exons and rodless isoforms. Genomics 1997;42:115–125.
2. Fuchs P, Zorer M, Rezniczek GA, et al. Unusual 5′ transcript complexity of plectin isoforms: novel tissue-specific exons modulate actin binding activity. Hum Mol Genet 1999;8:2461–2472.
3. Konieczny P, Fuchs P, Reipert S, et al. Myofiber integrity depends on desmin network targeting to Z-disks and costameres via distinct plectin isoforms. J Cell Biol 2008;181: 667–681.
4. Banwell BL, Russel J, Fukudome T, Shen X-M, Stilling G, Engel AG. Myopathy, myasthenic syndrome, and epidermolysis bullosa simplex due to plectin deficiency. J Neuropathol Exp Neurol 1999;58:832–846.
5. McLean W, Pulkkinen L, Smith F, et al. Loss of plectin causes epidermolysis bullosa with muscular dystrophy: cDNA cloning and genomic organization. Genes Dev 1996;10:1724–1735.
6. Smith FJ, Eady R, Leigh I, et al. Plectin deficiency results in muscular dystrophy and epidermolysis bullosa simplex. Nat Genet 1996;13:450–457.
7. Chavanas S, Pulkkinen L, Gache Y, et al. A homozygous nonsense mutation in the *PLEC1* gene in patients with epidermolysis bullosa simplex with muscular dystrophy. J Clin Invest 1996;98:2196–2200.
8. Pulkkinen L, Smith F, Shimizu H, et al. Homozygous deletion mutations in the plectin gene (PLEC1) in patients with epidermolysis bullosa simplex associated with late-onset muscular dystrophy. Hum Mol Genet 1996;5:1539–1546.
9. Gache Y, Chavanas S, Lacour J, et al. Defective expression of plectin/HD1 in epidermolysis bullosa simplex with muscular dystrophy. J Clin Invest 1996;97:2289–2298.
10. Mellerio J, Smith F, McMillan J, et al. Recessive epidermolysis bullosa simplex associated with plectin mutations: infantile respiratory complications in two unrelated cases. Br J Dermatol 1997;137:898–906.
11. Dang M, Pulkkinen L, Smith F, McLean W, Uitto J. Novel compound heterozygous mutations in the plectin gene in epidermolysis bullosa with muscular dystrophy and the use of protein truncation test for detection of premature termination codon mutations. Lab Invest 1998;78: 195–204.
12. Takizawa Y, Shimizu H, Rouan F, et al. Four novel plectin gene mutations in Japanese patients with epidermolysis bullosa and muscular dystrophy disclosed by heteroduplex scanning and protein truncation tests. J Invest Dermatol 1999;112:109–112.
13. Shimizu H, Takizawa Y, Pulkkinen L, et al. Epidermolysis bullosa simplex associated with muscular dystrophy: phenotype-genotype correlations and review of the literature. J Am Acad Dermatol 1999;41:950–956.
14. Rouan F, Pulkkinen L, Meneguzzi G, et al. Epidermolysis bullosa: novel and de novo premature termination codon and deletion mutations in the plectin gene predict late-onset muscular dystrophy. J Invest Dermatol 2000;114: 381–387.
15. Bauer JW, Rouan F, Kofler B, et al. A compound heterozygous one amino-acid insertion/nonsense mutation in the plectin gene causes epidermolysis bullosa simplex with plectin deficiency. Am J Pathol 2001;158:617–625.
16. Takahashi Y, Rouan F, Uitto J, et al. Plectin deficient epidermolysis bullosa simplex with 27-year-history of muscular dystrophy. J Dermatol Sci 2005;37:87–93.
17. Pfendner E, Rouan F, Uitto J. Progress in epidermolysis bullosa: the phenotypic spectrum of plectin mutations. Exp Dermatol 2005;14:241–249.
18. Bolling MC, Pas HH, De Visser M, et al. PLEC1 mutations underlie adult-onset dilated cardiomyopathy in epidermolysis bullosa simplex with muscular dystrophy. J Invest Dermatol 2010;130:1178–1181.
19. McMillan JR, Akiyama M, Rouan F, et al. Plectin defects in epidermolysis bullosa simplex with muscular dystrophy. Muscle Nerve 2007;35:24–35.
20. Rezniczeck GA, Walko G, Wiche G. Plectin defects lead to various forms of epidermolysis bullosa simplex. Dermatol Clin 2009;28:33–41.
21. Niemi K, Sommer H, Kero M, Kanerva L, Haltia M. Epidermolysis bullosa simplex associated with muscular dystrophy with recessive inheritance. Arch Dermatol 1988; 124:551–554.
22. Fine J-D, Stenn J, Johnson L, Wright T, Bock H, Horiguchi Y. Autosomal recessive epidermolysis bullosa simplex. Arch Dermatol 1989;125:931–938.
23. Doriguzzi C, Palmucci L, Mongini T, et al. Congenital muscular dystrophy associated with familial junctional epidermolysis bullosa letalis. Eur Neurol 1993;33:454–460.

24. Engel AG. The muscle biopsy. In: Engel AG, Franzini-Armstrong C, eds. Myology, 3rd ed. New York: McGraw-Hill; 2004:681–690.
25. Engel AG, Lambert EH, Howard FM. Immune complexes (IgG and C3) at the motor end-plate in myasthenia gravis: ultrastructural and light microscopic localization and electrophysiologic correlations. Mayo Clin Proc 1977;52:267–280.
26. Sahashi K, Engel AG, Lambert EH, Howard FM Jr. Ultrastructural localization of the terminal and lytic ninth complement component (C9) at the motor end-plate in myasthenia gravis. J Neuropathol Exp Neurol 1980;39:160–172.
27. Bodensteiner JB, Engel AG. Intracellular calcium accumulation in Duchenne dystrophy and other myopathies: a study of 567,000 muscle fibers in 114 biopsies. Neurology 1978;28:439–446.
28. Gautron J. Cytochimie ultrastructurale des acétylcholinestérases. Microscopie 1974;21:259–264.
29. Engel AG. Quantitative morphological studies of muscle. In: Engel AG, Franzini-Armstrong C, eds. Myology, 2nd ed. New York: McGraw-Hill; 1994:1018–1045.
30. Engel AG, Lindstrom JM, Lambert EH, Lennon VA. Ultrastructural localization of the acetylcholine receptor in myasthenia gravis and in its experimental autoimmune model. Neurology 1977;27:307–315.
31. Engel AG, Lambert EH, Mulder DM, et al. A newly recognized congenital myasthenic syndrome attributed to a prolonged open time of the acetylcholine-induced ion channel. Ann Neurol 1982;11:553–569.
32. Nastsuga K, Nishie W, Akiyama M, et al. Plectin expression patterns determine two distinct subtypes of epidermolysis bullosa simplex. Hum Mutat 2010;31:308–316.
33. Mokri B, Engel AG. Duchenne dystrophy: electron microscopic findings pointing to a basic or early abnormality in the plasma membrane of the muscle fiber. Neurology 1975;25:1111–1120.
34. Ruff RL. Effects of length changes on Na^+ current amplitude and excitability near and far from the end-plate. Muscle Nerve 1996;19:1084–1092.
35. Flucher BE, Daniels MP. Distribution of Na^+ channels and ankyrin in neuromuscular junctions is complementary to that of acetylcholine receptors and the 43 kd protein. Neuron 1989;3:163–175.
36. Ruff RL, Whittlesey D. Na^+ current density and voltage dependence in human intercostal muscle fibers. J Physiol 1992;458:85–97.
37. Ruff RL, Lennon VA. How myasthenia gravis alters the safety factor of neuromuscular transmission. J Neuroimmunol 2008;15:13–20.
38. Martin AR. Amplification of neuromuscular transmission by postjunctional folds. Proc R Soc Lond B 1994;258:321–326.
39. Konieczny P, Wiche G. Muscular integrity: a matter of interlinking distinct structures via plectin. In: Lang NG, ed. The Sarcomere and Skeletal Muscle Disease. New York: Springer; 2009:165–175.

Hyperuricemia cosegregating with osteogenesis imperfecta is associated with a mutation in *GPATCH8*

Hiroshi Kaneko[1,2], Hiroshi Kitoh[2], Tohru Matsuura[1], Akio Masuda[1], Mikako Ito[1], Monica Mottes[3], Frank Rauch[4], Naoki Ishiguro[2], and Kinji Ohno[1,*]

[1]Division of Neurogenetics, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan

[2]Department of Orthopaedic Surgery, Nagoya University Graduate School of Medicine, Nagoya, Japan

[3]Department of Life and Reproduction Sciences, University of Verona, Verona, Italy

[4]Genetics Unit, Shriners Hospital for Children and McGill University, Montréal, Québec, Canada

*Correspondence to:

Kinji Ohno,

Division of Neurogenetics, Center for Neurological Diseases and Cancer,

Nagoya University Graduate School of Medicine,

65 Tsurumai, Showa-ku, Nagoya 466-8550, Japan.

e-mail: ohnok@med.nagoya-u.ac.jp

Abstract

Autosomal dominant osteogenesis imperfecta (OI) is caused by mutations in *COL1A1* or *COL1A2*. We identified a dominant missense mutation, c.3235G>A in *COL1A1* exon 45 predicting p.G1079S, in a Japanese family with mild OI. As mutations in exon 45 exhibit mild to lethal phenotypes, we tested if disruption of an exonic splicing *cis*-element determines the clinical phenotype, but detected no such mutations. In the Japanese family, juvenile-onset hyperuricemia cosegregated with OI, but not in the previously reported Italian and Canadian families with c.3235G>A. After confirming lack of a founder haplotype in three families, we analyzed *PRPSAP1* and *PRPSAP2* as candidate genes for hyperuricemia on chr 17 where *COL1A1* is located, but found no mutation. We next resequenced the whole exomes of two siblings in the Japanese family, and identified variable numbers of previously reported hyperuricemia-associated SNPs in *ABCG2* and *SLC22A12*. The same SNPs, however, were also detected in normouricemic individuals in three families. We then identified two missense SNVs in *ZPBP2* and *GPATCH8* on chromosome 17 that cosegregated with hyperuricemia in the Japanese family. *ZPBP2* p.T69I was at the non-conserved region and was predicted to be benign by *in silico* analysis, whereas *GPATCH8* p.A979P was at a highly conserved region and was predicted to be deleterious, which made p.A979P a conceivable candidate for juvenile-onset hyperuricemia. *GPATCH8* is only 5.8 Mbp distant from *COL1A1* and encodes a protein harboring an RNA-processing domain and a zinc finger domain, but the molecular functions have not been elucidated to date.

Keywords: osteogenesis imperfecta; hyperuricemia; type I collagen; splicing *cis*-element; exome resequencing; *GPATCH8*

Introduction

Osteogenesis imperfecta (OI) is a heritable connective tissue disorder characterized by bone fragility and low bone mass. Clinical severities are widely variable ranging from intrauterine fractures and perinatal lethality to very mild forms without fractures. Patients also exhibit associated features including blue sclera, dentinogenesis imperfecta, hyperlaxity of ligaments and skin, and hearing loss (Rauch and Glorieux 2004). The widely used classification initially described by Sillence *et al.* distinguishes types I, II, III, and IV (MIM# 166200, 166210, 259420, and 166220, respectively) on the basis of clinical and radiographic findings (Sillence et al. 1979). Recently, five additional types of V, VI, VII, VIII, and IX (MIM# 610967, 610968, 610682, 610915, and 259440, respectively) have been reported (Cabral et al. 2007; Glorieux et al. 2000; Glorieux et al. 2002; van Dijk et al. 2009; Ward et al. 2002). OI type I is the mildest form characterized by fractures with little or no limb deformity and normal or mildly short stature, whereas type II is a perinatal lethal form, mostly due to respiratory failure resulting from multiple rib fractures. Type III is characterized by progressive deformities and fractures that are often present at birth. Severities of types IV, V, VI and VII are between those of types I and III. Type VIII and IX carry features of both types II and III.

Type I collagen is the most abundant bone protein. Most patients (>90%) with OI types I-IV have dominant or recessive mutation(s) in either of two genes, *COL1A1* (MIM# 120150) on chromosome (chr) 17q21.31-q22 and *COL1A2* (MIM# 120160) on chr 7q22.1 that encode the α1 and α2 chains of type I procollagen, respectively (Rauch and Glorieux 2004). A genetic cause of type V remains undetermined to date. Types VI to IX are caused by recessive mutations. Type VI is caused by mutations in *FKBP10* (MIM# 607063) encoding FK506-binding protein 65 (FKBP65) that is a chaperone in type I procollagen folding (Alanay et al. 2010). Type VII is caused by mutations in *CRTAP* (MIM# 605497) encoding cartilage-associated protein (CRTAP) (Morello et al. 2006). Type VIII is caused by mutations in *LEPRE1* (MIM# 610339) encoding prolyl 3-hydroxylase 1 (P3H1) (Cabral et al. 2007). Type IX is caused by mutations in *PPIB* (MIM# 123841) encoding cyclophilin B (CYPB) (van Dijk et al. 2009). CRTAP, P3H1, and CYPB form an intracellular collagen-modifying complex that 3-hydroxylates proline at position 986 in the α1 chain of type I collagen, which is essential for correct folding and stability of the collagen triple helix. Mutations in *CRTAP* and *LEPRE1* are also identified in severe OI phenotypes including type II (Baldridge et al. 2008; Morello et al. 2006). Recently, recessive mutations in *SERPINH1* (MIM# 600943) encoding a chaperone-like protein for collagens, heat shock protein 47, and in *SP7/Osterix* (MIM# 606633) encoding an osteoblast-specific transcription factor have been identified in patients with types III and IV, respectively (Christiansen et al. 2010; Lapunzina et al. 2010).

Two copies of the α1 chain and one copy of the α2 chain form a core triple helix comprised of 338 uninterrupted Gly-X-Y triplet repeats, where X is often proline and Y is often hydroxyproline. Gly repeats at every third position are essential for the stability of collagen because larger amino acids cannot be accommodated in the tightly packed core without disruption of the triple helix (Bodian et al. 2008). The most common mutations (>80%) affect one of the repeated Gly residues in the triple helix. More than 800 mutations in *COL1A1* and *COL1A2* are currently deposited in the human type I collagen mutation

database (http://www.le.ac.uk/genetics/collagen/) (Dalgleish 1997; Marini et al. 2007). Clinical phenotypes may be determined by the chain in which the Gly substitution occurs, the position of the mutation within the chain, and/or the nature of the mutant amino acids (Bodian et al. 2008; Rauch and Glorieux 2004), but we still cannot predict a clinical phenotype of a given mutation. On the other hand, mutations that create a premature stop codon within *COL1A1* mostly exhibit a milder OI type I. This is because a truncation mutation unlikely to have a dominant negative effect, but the abundance of type I collagen chain is half of the normal (Marini et al. 2007; Rauch and Glorieux 2004).

Pre-mRNA splicing is regulated by intronic and exonic splicing *cis*-elements. Both constitutively and alternatively spliced exons harbor exonic splicing enhancers (ESEs) and silencers (ESSs). Splicing *trans*-factors are expressed in a developmental stage-specific and tissue-specific manner, and their expressions tightly regulate alternative splicing of an exon carrying ESEs/ESSs. A mutation in the coding region is predicted to exert its pathogenicity by compromising a functional amino acid, but nonsense, missense, and even silent mutations potentially disrupt an ESE/ESS, thereby resulting in aberrant splicing (Cartegni et al. 2002; Cartegni et al. 2003). Indeed, more than 16% to 20% of exonic mutations are predicted to disrupt an ESE (Gorlov et al. 2003).

Exome resequencing is a powerful and efficient method to identify a novel gene associated with a rare monogenic disorder, especially when the number of unrelated patients or the number of family members of a patient are too small to apply linkage studies. Filtering against existing SNP database and the exomes of unaffected individuals can remove common variants to identify a causal gene. Ng *et al.* sequenced exomes of 12 humans, including four unrelated individuals with autosomal dominant Freeman-Sheldon syndrome (MIM# 193700) and eight HapMap individuals. They successfully identified mutations in *MYH3* in all the affected individuals (Ng et al. 2009). Ng *et al.* also sequenced exomes of four patients in three families with autosomal recessive Miller syndrome (MIM# 263750) and *de novo* identified compound heterozygous mutations in *DHODH* (MIM# 126064) (Ng et al. 2010b). Ng *et al.* additionally sequenced exomes of ten unrelated patients with autosomal dominant Kabuki syndrome (MIM# 147920) and *de novo* identified nonsense or frameshift mutations in *MLL2* (MIM# 602113) (Ng et al. 2010a). Similarly, Lalonde *et al.* sequenced two unrelated fetuses with autosomal recessive Fowler syndrome (MIM# 225790) and *de novo* identified compound heterozygous mutations in *FLVCR2* (MIM# 610865) (Lalonde et al. 2010).

In a Japanese family with OI type I, hyperuricemia cosegregated with OI. To our knowledge, association of hyperuricemia with OI has been reported in two families, in which two of three OI patients had gouty arthritis and hyperuricemia at young ages (Allen et al. 1955). Underexcretion of urate is causally associated with mutations in *UMOD* (MIM# 191845) encoding uromodulin (Hart et al. 2002) as well as with SNPs in three genes: *SLC2A9* (MIM# 606142) encoding glucose transporter 9 (Doring et al. 2008; Vitart et al. 2008), *ABCG2* (MIM# 603756) encoding ATP-binding cassette subfamily G member 2, an urate transporter (Dehghan et al. 2008; Kolz et al. 2009; Stark et al. 2009; Woodward et al. 2009), and *SLC22A12* (MIM# 607096) encoding URAT1, a renal urate-anion exchanger (Graessler et al. 2006; Tabara et al. 2010). On the other hand, overproduction of urate is caused by mutations in *PRPS1* (MIM#

311850) encoding PRPP synthetase I (Roessler et al. 1993), and *HPRT1* (MIM# 308000) encoding hypoxanthine guanine phosphoribosyltransferase 1 (Gibbs and Caskey 1987).

We here identified a dominant missense mutation, c.3235G>A in *COL1A1* exon 45 predicting p.G1079S, in a Japanese family with OI type I and hyperuricemia, and analyzed molecular bases of two clinical features. First, we tested if mild to lethal phenotypes of mutations in exon 45 of *COL1A1* are accounted for by preservation or disruption of an ESE/ESS element, and found that ESE/ESS elements are not involved in disease severities. Second, we traced a cause of the hyperuricemia by exome resequencing of two siblings and found that a missense mutation in *GPATCH8* encoding the G patch domain-containing protein 8 close to *COL1A1* cosegregated with hyperuricemia.

Patients and methods

Samples and ethical considerations

We obtained blood from each family member and isolated genomic DNA (gDNA) from 2 ml peripheral blood using the QIAamp DNA Blood Midi Kit (Qiagen) according to the manufacturer's instructions. We also obtained skin biopsy of an affected individual (II-3) of a Japanese family (F1) and cultured non-transformed fibroblasts for splicing and sequencing analysis. Informed consent was obtained from all family members. The studies have been approved by the Institutional Review Boards of the Nagoya University, the University of Verona, and the McGill University. Clinical features and mutational analyses have been previously reported in the Italian and Canadian families (Mottes et al. 1992; Roschger et al. 2008).

Microsatellite analysis of *COL1A1* and *COL1A2* to identify a mutation in the Japanese family

We genotyped all family members for three microsatellite markers flanking *COL1A1* (D17S1293, -16 Mbp; D17S1319, -14 kbp; and D17S788, 2 Mbp). As no annotated microsatellite markers were available close to *COL1A2*, we posted three new microsatellite markers to DDBJ (AB499843, -17 kb; AB499844, 29 kb; and AB499845, 123 kb) and analyzed them in the family. We fluorescently labeled the 5' end of each forward primer with 5FAM (Sigma-Aldrich), and amplified microsatellite markers with the HotStarTaq Plus Master Mix (Qiagen) using gDNA and primers indicated in Supplementary Table 1.

We mixed 1.5 μl of 20-times diluted PCR product with 0.5 μl of GeneScan-500 ROX Size Standard (Applied Biosystems) and 24.5 μl of formamide, and incubated the mixture at 95 °C for 3 min. The mixture was run by capillary electrophoresis on an ABI PRISM 310 Genetic Analyzer, and was analyzed with the GeneScan and GeneMapper software (Applied Biosystems).

Sequence analysis of *COL1A1*

After the microsatellite analysis suggested that *COL1A1* was more likely to be a causative gene, we amplified all exons and flanking intronic regions of ~40 bp, as well as 5' and 3' UTRs of *COL1A1* from gDNA of II-3 by PCR. We performed the dye terminator cycle sequencing reaction with the

GenomeLab DTCS Quick Start Kit (Beckman Coulter) and run on the CEQ 8000 Genetic Analysis System (Beckman Coulter) according to the manufacturer's instructions. We compared the chromatograms with the GenBank reference sequences of *COL1A1* gDNA using the Mutation Surveyor software version 2.61 (SoftGenetics). We numbered *COL1A1* mutations with the translation initiator methionine as amino acid +1, and the A of the ATG codon as nucleotide +1 according to the Human Genome Variation Society (http://www.hgvs.org/mutnomen/recs.html). We numbered exons according to the human type I collagen mutation database (http://www.le.ac.uk/genetics/collagen/), in which *COL1A1* exon 33 is named *COL1A1* exon 33-34 to match the exonic annotations of *COL1A2*.

Allele-specific primer (ASP)-PCR to trace the *COL1A1* mutation

We traced the mutation in family members using ASP-PCR. The wild-type ASP was 5'-TCCCGCCGGTCCTGTAGG-3', and the mutant ASP was 5'-TCCCGCCGGTCCTGTAAG-3', where the mutated nucleotide is underlined, and an artificially introduced mismatch is shown in bold. The reverse primer was 5'-GCCACGGTGACCCTTTATGC-3'.

Prediction of effects of mutations on pre-mRNA splicing

Five missense mutations in exon 45 of *COL1A1* cause mild to lethal OI phenotypes (Fig. 2a) (Constantinou et al. 1989; Hartikka et al. 2004; Lund et al. 1997; Marini et al. 2007; Mottes et al. 1992; Roschger et al. 2008). We predicted effects on pre-mRNA splicing of eighteen sequence variants with or without each mutation in the presence or absence of each of two SNPs (rs1800215 and rs1800217) in exon 45 of *COL1A1* (Fig. 2b) using five web-based programs: ESEfinder 3.0 (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home) (Cartegni et al. 2003), ESRsearch (http://ast.bioinfo.tau.ac.il/) (Goren et al. 2006), FAS-ESS (http://genes.mit.edu/fas-ess/) (Wang et al. 2004), PESXs (http://cubweb.biology.columbia.edu/pesx/) (Zhang and Chasin 2004; Zhang et al. 2005), and RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/) (Fairbrother et al. 2002). We also predicted effects of the mutations on splice site strength of the eighteen sequence variants using two web-based programs: the NetGene2 Server (http://www.cbs.dtu.dk/services/NetGene2/) (Brunak et al. 1991; Hebsgaard et al. 1996) and the Splice Site Prediction by Neural Network (http://www.fruitfly.org/seq_tools/splice.html) (Reese et al. 1997).

Splicing analysis of fibroblasts of the OI patient (II-3)

We first examined splicing of *COL1A1* exon 45 in the patient's fibroblasts. We extracted total RNA from cultured fibroblasts of II-3 using Trizol reagent (Invitrogen), and synthesized cDNA using the oligo(dT)$_{12\text{-}18}$ primer (Invitrogen) and the ReverTra Ace reverse transcriptase (Toyobo). We examined skipping of *COL1A1* exon 45 using 5'-GGTTCCCCTGGACGAGAC-3' on exon 43 and 5'-TCCAGAGGGACCTTGTTCAC-3' on exon 47. We also sequenced RT-PCR products as described above to scrutinize splicing consequences. As skipping of exon 45 results in an in-frame deletion of 54 nucleotides, we did not downregulate the nonsense-mediated mRNA decay (NMD) before harvesting

cells.

COL1A1 minigene constructs

We amplified exons 44-46 and the intervening introns of *COL1A1* (Fig. 2b) by PCR. The PCR primers introduced a *Hin*dIII site and the Kozak consensus sequence of 5'-CCACCATG-3' to the 5' end, as well as a TAA stop codon and a *Bam*HI site at the 3' end of the PCR product, so that the minigene transcript is tolerant to NMD (Ohno et al. 2003). We inserted the PCR product into the pcDNA3.1(+) mammalian expression vector (Invitrogen), and confirmed lack of PCR artifacts by sequencing the entire insert. We next constructed 17 variant minigenes using the QuikChange site-directed mutagenesis kit (Stratagene) (Fig. 2b). We again confirmed presence of the introduced mutations and absence of artifacts by sequencing the entire inserts.

Splicing assays of *COL1A1* exon 45

We transfected 500 ng of a minigene construct into 50% confluent HEK293 cells in a 12-well plate using the FuGENE 6 transfection reagent (Roche Applied Science) according to the manufacturer's instructions. After 48 hours, we extracted total RNA from HEK293 cells, and synthesized cDNA as described above. To prevent amplification of the endogenous *COL1A1*, we used generic primers of 5'-TTAATACGACTCACTATAGGGAGACC-3' and 5'-TTAAGGAGGGCCAGGGGG-3' located on pcDNA3.1. Untransfected cells were used as a negative control.

Founder analysis of three families with *COL1A1* c.3235G>A

In order to examine if the *COL1A1* c.3235G>A mutation arose from a common founder in the Japanese, the Italian (Mottes et al. 1992), and the Canadian (Roschger et al. 2008) families, we genotyped three microsatellite markers flanking *COL1A1* described above, and sequenced an intragenic SNP (rs2075554) of *COL1A1* in the Italian and Canadian families (Fig. 1).

Sequence analysis of *PRPSAP1* and *PRPSAP2*

In order to seek for a responsible gene for hyperuricemia, we sequenced the entire coding regions of *PRPSAP1* and *PRPSAP2* on chr 17 using cDNA synthesized from cultured fibroblasts of II-3.

Resequencing of exome-enriched DNA

As we found no mutation in *PRPSAP1* and *PRPSAP2*, we enriched exonic regions of genomic DNA of II-2, II-3, and VIII-2 using the SureSelect human all exon kit v1 (Agilent Technologies) that covers 1.22% (34 Mbp) of the human genome. We sequenced 50 base pairs of each tag in a single direction using a quarter of a cell of the SOLiD 3 Plus system (Life Technologies) for each sample.

For II-2, II-3, and VIII-2, we obtained 79.1, 68.6, and 109.9 x 10^6 tags of 50-bp SOLiD reads and mapped 2.18 (69.6%), 1.87 (68.5%), and 3.37 (76.5%) Gbp to the human genome hg19/GRCh37, which yielded a mean coverage of 64.1, 55.1, and 99.0, respectively. Among the mapped tags, 70.1%,

72.0%, and 73.1% were located on the SureSelect exome probes. Among the 34-Mbp regions where the exome probes were designed, 3.4%, 3.6%, and 3.5% of nucleotides were not sequenced at all. Search for single nucleotide variants (SNVs) and indels with Bioscope 1.2.1 (Life Technologies) detected 52,436, 56,941, and 60,303 SNVs/indels. SNVs and indels were compared to dbSNP Build 132.

Analysis of variants in *KRBA2, ZPBP2, and GPATCH8*

In order to trace if variants in *KRBA2*, *ZPBP2*, and *GPATCH8* cosegregated with hyperuricemia, we analyzed all family members in the Japanese family (F1) by capillary sequencing.

We traced two variants, *ZPBP2* c.206C>T and *GPATCH8* c.2935G>C, in the Italian and Canadian families and 100 normal human genomes using ASP-PCR. The forward primers of *ZPBP2* and *GPATCH8* were 5'-CGTGTCTTCAGCACAAAATGG-3' and 5'-AGAAGCCGTAGCACCACT**C**C-3', respectively. The reverse primers were 5'-GGCCCAATCCATAAGTA**C**AT-3' and 5'-CCCATGATCTCTTCCTGGAG-3', respectively. The mutated nucleotide is underlined, and an artificially introduced mismatch is shown in bold.

To search for the identified variants in *ZPBP2* and *GPATCH8* in normal controls, we mapped 50 Tibetan exome reads (SRA accession number SRP002446) (Yi et al. 2010) to a 200-bp region spanning c.206C>T in exon 3 of *ZPBP2* and a 200-bp region spanning c.2935G>C in exon 8 of *GPATCH8* with the bowtie alignment tool version 0.12.7 (Langmead et al. 2009) using default parameters.

We analyzed amino acid conservations of *ZPBP2* and *GPATCH8* using the evolutionary annotation database, Evola, at the H-Inv DB (http://www.h-invitational.jp/evola/). We also predicted functional effects of amino acid substitutions using two web-based programs: PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) (Adzhubei et al. 2010) and SIFT (http://sift.jcvi.org/) (Kumar et al. 2009).

Results

Hyperuricemia cosegregated with OI type I in the Japanese family (F1)

In a Japanese family (F1), a father (age 56 years) and his three sons (ages 29, 26, and 23 years) had OI type I with blue sclera, dentinogenesis imperfecta, and joint laxity (Fig. 1). Two sons (II-1 and II-3) had histories of more than 10 fractures before age 13 years, but the father (I-1) and another son (II-2) experienced no bone fracture. One son (II-1) had hearing loss from age 10 years likely due to fractures or deformities of small bones in the middle ear, and had hip joint deformities due to repeated femoral fractures. Interestingly, all the affected members had hyperuricemia of ~8 mg/dl that was diagnosed at ages 15 to 30 years. One son (II-3) had a gout attack, and the other two (I-1 and II-1) had urinary stones. Hyperuricemia is currently well controlled by medication in all the members.

Heteroallelic c.3235G>A mutation in *COL1A1* in the Japanese family (F1)

Genotypes of three microsatellite markers flanking *COL1A1* cosegregated with the OI

phenotype in F1 (Fig. 1), whereas genotypes of three markers flanking *COL1A2* did not (data not shown). We thus sequenced the entire exons and the flanking noncoding regions, and identified a heteroallelic c.3235G>A mutation in exon 45 and a heteroallelic G/A SNP (rs2075554) in intron 11 of *COL1A1*. The c.3235G>A mutation predicts p.G1079S. We genotyped c.3235G>A in family members by ASP-PCR, and found that all affected members were heterozygous for c.3235G>A (Fig. 1).

Phenotypic variability of osteogenesis imperfecta is not accounted for by disruption of splicing *cis*-elements

In addition to c.3235G>A, four more mutations and two SNPs have been reported in *COL1A1* exon 45 with variable phenotypes ranging from mild type I to perinatal lethal type II (Fig. 2a) (Constantinou et al. 1989; Hartikka et al. 2004; Lund et al. 1997; Marini et al. 2007; Mottes et al. 1992; Roschger et al. 2008). We thus hypothesized that disruption or *de novo* generation of a splicing *cis*-element determines the clinical phenotype. The five mutations and two SNPs in *COL1A1* exon 45 were predicted to affect 16 putative splicing *cis*-elements by ESEfinder, ESRsearch, and PESXs (Table 1). FAS-ESS and RESCUE-ESE predicted no splicing *cis*-elements. All the five mutations with or without two SNPs in *COL1A1* exon 45 variably but slightly weaken acceptor and/or donor splice site strengths according to the NetGene2 (Table 2). The Splice Site Prediction by Neural Network also predicted that c.3235G>A generates a weak cryptic splice acceptor site in *COL1A1* exon 45 (Table 2). We first examined cultured fibroblasts of II-3 by RT-PCR and found that the *COL1A1* c.3235G>A mutation did not induce aberrant splicing of *COL1A1* (data not shown). NMD was unlikely to have masked aberrant splicing, because we observed heterozygous peaks at c.3235G>A in sequencing the RT-PCR product. We next constructed eighteen *COL1A1* minigenes with or without each of the five mutations in the presence or absence of each of the two SNPs (Fig. 2b). RT-PCR analysis of transfected HEK293 cells showed that all minigene constructs gave rise to a single fragment of 336 bp, indicating that splicing was not affected in any mutations or SNPs (Fig. 2c).

Japanese (F1), Italian (F2), and Canadian (F3) families with *COL1A1* c.3235G>A share no founder haplotype

We previously reported *COL1A1* c.3235G>A in the Italian and Canadian families (Marini et al. 2007; Roschger et al. 2008). Although we have not measured serum urate concentrations in these families, gout or urinary stone has not been documented in either family, which suggests that hyperuricemia is not simply due to c.3235G>A.

In order to pursue if a gene responsible for hyperuricemia is on the same chr as *COL1A1*, we looked for a founder haplotype for c.3235G>A in three families by genotyping three microsatellite markers flanking *COL1A1* (D17S1293, -16 Mbp; D17S1319, -14 kbp; and D17S788, 2 Mb) and a SNP (rs2075554) in intron 11 of *COL1A1*. The analysis revealed that each family carried a unique haplotype and shared no founder haplotype (Fig. 1). Thus, the mutation is likely to have occurred independently in three ethnic groups. Alternatively, c.3235G>A is an ancient founder mutation, and subsequent multiple

recombinations and divergence of microsatellite repeats have obscured a founder effect. In either case, lack of a found haplotype supports the notion that a gene responsible for hyperuricemia is potentially but not necessarily linked to *COL1A1*.

Hyperuricemia is not caused by mutations in *PRPSAP1* or *PRPSAP2*

We thus first looked into candidate genes for hyperuricemia on chr 17 where *COL1A1* (17q21.31-q22) is located. Two candidate genes for hyperuricemia are on chr 17: *PRPSAP1* (MIM# 601249) at 17q24-q25 encoding PAP39 (Ishizuka et al. 1996) and *PRPSAP2* (MIM# 603762) at 17p12-p11.2 encoding PAP41 (Katashima et al. 1998). PAP39 and PAP41 are subunits of phosphoribosylpyrophosphate (PRPP) synthetase that leads to urate production. No mutation has been reported in either gene in any diseases. We sequenced cDNAs of *PRPSAP1* and *PRPSAP2* in II-3, but found no mutation in either gene. As we detected no heterozygous SNPs in *PRPSAP1* and *PRPSAP2*, a mutant allele carrying a premature stop codon might have been missed due to mRNA degradation by NMD.

Resequencing of exomes reveals hyperuricemia-associated SNPs

We next traced a causative gene for hyperuricemia in two siblings (II-2 and II-3) by exome resequencing with the SureSelect human all exon kit v1 (Agilent) and with the SOLiD 3 Plus sequencer (Life Technologies).

We similarly analyzed an unrelated Japanese male (22 years old) with OI type I and hyperuricemia (VIII-2 in F4 in Fig. 1). His hyperuricemia was by chance detected at age 14 years when he had fractures. Hyperuricemia has been well controlled by medication since then. Exome resequencing of VIII-2 disclosed a novel heteroallelic c.577G>T mutation in *COL1A2* exon 12 predicting p.G193C. We confirmed the mutation by capillary sequencing. Family samples were not available for our analysis.

In F1, a probability that three siblings inherited an identical allele from the father is $(1/2)^3$ = 12.5%, which indicates that the causative gene is anywhere in the 12.5% regions of the entire genome. We first sought for a mutation responsible for hyperuricemia in 24 genes (see Table 3 and Suppl. Table 2) that are involved in urate metabolisms and excretion, but found none in either patient. We then scrutinized SNPs in ten genes that are known to be associated with hyperuricemia and identified 12 SNPs (Table 3). In this analysis, we excluded SNPs with a minor allelic frequency of 0.01 or less. Interestingly, among the 12 SNPs, rs2231142 in *ABCG2* as well as rs3825016 and rs11231825 in *SLC22A12* are previously reported markers of hyperuricemia, which will be addressed in the discussion. We traced the three SNPs in F1, F2, and F3 by capillary sequencing, and found that variable dosages of these SNPs were observed in hyperuricemic as well as in normouricemic individuals (Fig. 1).

A missense mutation in *GPATCH8* is likely to lead to hyperuricemia in the Japanese family (F1)

After eliminating SNPs in the dbSNP132 database, only three non-synonymous variants remained shared between II-2 and II-3 on chr 17; c.602A>G in exon 2 of *KRBA2* at 17p13.1, c.206C>T in

exon 3 of *ZPBP2* (MIM# 608499) at 17q12 (Fig. 3a), and c.2935G>C in exon 8 of *GPATCH8* at 17q21.31 (Fig. 3b). Capillary sequencing revealed that variants in *ZPBP2* and *GPATCH8* cosegregated with hyperuricemia in F1, but the *KRBA2* variant did not. These two variants were not detected in F2, F3, F4, 100 normal human individuals, or exomes of 50 Tibetans (Yi et al. 2010). In addition, exome-capture resequencing of VIII-2 detected no mutations in *ZPBP2* and *GPATCH8*.

ZPBP2 c.206C>T and *GPATCH8* c.2935G>C predict amino acid substitutions p.T69I and p.A979P, respectively. Threonine 69 and the flanking amino acids of *ZPBP2* are not conserved across mammalian species. Additionally, a SNP rs35591738 mutates the N-terminal proline at codon 68, and a SNP rs34272593 causes a frameshift at codon 70 (Fig. 3a). In contrast, alanine 979 and the flanking amino acids of *GPATCH8* are in the serine-rich region and are highly conserved across mammalian species (Fig. 3b). PolyPhen-2 (Adzhubei et al. 2010) predicted that *ZPBP2* p.T69I is benign with a score of 0.025 and *GPATCH8* p.A979P is damaging with a score of 0.988, where 1.0 is the worst score. Similarly, SIFT (Kumar et al. 2009) predicted that *ZPBP2* p.T69I is tolerated with a score of 0.38 and *GPATCH8* p.A979P is damaging with a score of 0.00, where a score < 0.05 is predicted to be deleterious.

Discussion

Phenotypic variabilities of OI mutations

We identified a heteroallelic c.3235G>A mutation in *COL1A1* exon 45 in a Japanese family with mild OI type I. The c.3235G>A mutation has been previously reported in six families with OI types I (Hartikka et al. 2004; Mottes et al. 1992; Roschger et al. 2008) and IV (Marini et al. 2007). *COL1A1* exon 45 encodes six of the 338 Gly-X-Y triplet repeats. Four additional mutations have been reported in exon 45 (Constantinou et al. 1989; Lund et al. 1997; Marini et al. 2007), and all substitute Ser or Cys for Gly (Fig. 2a) with mild to lethal phenotypes. Among the five mutations, two mutations introducing Cys result in a lethal type II, whereas three mutations introducing Ser give rise to milder types I, III, and IV. This notion, however, cannot be applied to the other exons according to the human type I collagen mutation database (http://www.le.ac.uk/genetics/collagen/).

Being prompted by a report that more than 16% to 20% of exonic mutations disrupt an ESE (Gorlov et al. 2003), we asked if a mutation disrupting an ESE in exon 45 causes skipping of an inframe exon 45 and exhibits a severe dominant negative phenotype. Three web-based programs predict that all the five mutations and the two SNPs affect 16 putative exonic splicing *cis*-elements (Table 1). We thus constructed and analyzed 18 minigenes carrying all possible combinations of the five mutations and two SNPs (Fig. 2b), but found that none affected pre-mRNA splicing (Fig. 2c). Our analysis suggests that the currently available algorithms of splicing *trans*-factors cannot efficiently predict splicing *cis*-elements. This is likely because the recognition motifs of splicing *trans*-factors are mostly determined by *in vitro* SELEX experiments. A recently developed technique, the high throughput sequencing coupled to crosslinking immunoprecipitation method (HITS-CLIP), enables us to extensively determine RNA segments recognized by a specific splicing *trans*-factor *in vivo* (Licatalosi et al. 2008; Yeo et al. 2009). Accumulation of knowledge with the HITS-CLIP technology will enable us to construct dependable

algorithms to efficiently predict splicing *cis*-elements.

In addition to the phenotypic variability among similar mutations in the same exon, the same mutation often exhibits variable phenotypes although the variability is usually less (Lund et al. 1996). This is also true for our families. In the Japanese family (F1), two sons (II-1 and II-3) experience many fractures, whereas the father (I-1) and another son (II-2) have no history of fractures. In the Italian family (F2), the son suffered from many fractures, but his affected mother did not (Mottes et al. 1992). In the Canadian family (F3), the father is classified as OI type IV, whereas his daughter is OI type I (Roschger et al. 2008). Variable clinical phenotypes of the c.3235G>A mutation is likely due to differences in environmental factors or to SNPs in disease-modifying genes, but the molecular bases have not been elucidated in any types of OI.

Molecular basis of hyperuricemia

In F1, hyperuricemia cosegregated with OI type 1. Although no known genes causing hyperuricemia are on chr 17 where *COL1A1* is located, two candidate genes of *PRPSAP1* and *PRPSAP2* are on chr 17. Capillary sequencing of these genes, however, detected no mutation. We thus employed exome-capture resequencing analysis of two siblings in F1 to search for a responsible gene for hyperuricemia. We first looked into SNPs in the 24 candidate genes that are associated with hyperuricemia, purine metabolism, and renal excretion, and found 18 SNPs in seven genes (Table 3 and Suppl. Table 2). Among them, three SNPs in *ABCG2* (rs2231142) and *SLC22A12* (rs3825016 and rs11231825) are previously reported markers for hyperuricemia and/or gout.

Dehghan *et al.* report that rs2231142 in *ABCG2* is associated with gout by a genome-wide association study (OR = 1.74 and 1.71 in white and black participants, respectively) (Dehghan et al. 2008). Woodward *et al.* show a significant association between rs2231142 and hyperuricemia (OR = 1.68) by a population-based study of 14,783 individuals (Woodward et al. 2009). Kolz *et al.* demonstrate that rs2231142 elevates the serum urate concentration more strongly in men than in women by meta-analysis of 28,141 individuals (Kolz et al. 2009). Stark *et al.* analyzed 683 patients with gout and indicated a significant association between rs2231142 and gout (OR = 1.37) (Stark et al. 2009). Although rs2231142 is an attractive causative SNP, our normouricemic subjects are also heterozygous for rs2231142.

Graessler et al. analyzed 389 German individuals with primary hyperuricemia and found that rs3825016 and rs11231825 in *SLC22A12* were significantly associated with reduced fractional excretion of urate in the kidney (Graessler et al. 2006). Tabara *et al.* analyzed 1,526 normal Japanese individuals retrospectively and longitudinally, and clarified that rs11231825 is associated with reduced urate excretion and with future development of hyperuricemia (Tabara et al. 2010). Again, although the two SNPs are attractive causes of hyperuricemia, we observe variable dosages of these SNPs even in our normouricemic subjects.

We next looked into neighboring genes of *COL1A1* without asking functions of the gene products, and identified that two missense variants in *ZPBP2* and *GPATCH8* cosegregated with the *COL1A1* mutation in F1. Neither variant was detected in 300 normal alleles or in dbSNP132. *ZPBP2*

p.T69I, however, is unlikely to be pathogenic for three reasons: lack of conservation in mammals; two missense/frameshifting SNPs at or close to the variant site (Fig. 3a); and the benign predicted outcome by PolyPhen-2 and SIFT. On the other hand, p.A979P in *GPATCH8* substitutes an amino acid in the highly conserved serine-rich region (Fig. 3b), and the substitution is predicted to damage the structure and function of the protein by *in silico* analysis. *GPATCH8* encodes the G patch domain-containing protein 8 that harbors both an RNA processing domain and a zinc finger domain. *GPATCH8* is expressed in a wide variety of human tissues including skeletal muscles, brain, heart, pancreas, liver, and kidney (McKinney et al. 2004). Functions of the *GPATCH8* gene product, however, have not been studied to date. The p.A979P variant in *GPATCH8* is highly likely to be associated with hyperuricemia in F1, but it may also cause another yet unidentified phenotype that cosegregates with OI.

Acknowledgments

We would like to thank the families for their participation in this study. We are grateful to Dr. Kunio Ihara at the Center for Gene Research of Nagoya University for the SOLiD sequencing analysis and Keiko Itano for technical assistance. This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and the Ministry of Health, Labor, and Welfare of Japan.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 7:248-249

Alanay Y, Avaygan H, Camacho N, Utine GE, Boduroglu K, Aktas D, Alikasifoglu M, Tuncbilek E, Orhan D, Bakar FT, Zabel B, Superti-Furga A, Bruckner-Tuderman L, Curry CJ, Pyott S, Byers PH, Eyre DR, Baldridge D, Lee B, Merrill AE, Davis EC, Cohn DH, Akarsu N, Krakow D (2010) Mutations in the Gene Encoding the RER Protein FKBP65 Cause Autosomal-Recessive Osteogenesis Imperfecta. Am J Hum Genet 87:572-573

Allen GE, Rogers FB, Lansbury J (1955) Osteogenesis imperfecta tarda with hyperuricemia and gout; report of three cases. Am J Med Sci 230:30-32

Baldridge D, Schwarze U, Morello R, Lennington J, Bertin TK, Pace JM, Pepin MG, Weis M, Eyre DR, Walsh J, Lambert D, Green A, Robinson H, Michelson M, Houge G, Lindman C, Martin J, Ward J, Lemyre E, Mitchell JJ, Krakow D, Rimoin DL, Cohn DH, Byers PH, Lee B (2008) CRTAP and LEPRE1 mutations in recessive osteogenesis imperfecta. Hum Mutat 29:1435-1442

Bodian DL, Madhan B, Brodsky B, Klein TE (2008) Predicting the clinical lethality of osteogenesis imperfecta from collagen glycine mutations. Biochemistry 47:5424-5432

Brunak S, Engelbrecht J, Knudsen S (1991) Prediction of human mRNA donor and acceptor sites from the DNA sequence. J Mol Biol 220:49-65

Cabral WA, Chang W, Barnes AM, Weis M, Scott MA, Leikin S, Makareeva E, Kuznetsova NV, Rosenbaum KN, Tifft CJ, Bulas DI, Kozma C, Smith PA, Eyre DR, Marini JC (2007) Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat Genet 39:359-365

Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3:285-298

Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31:3568-3571

Christiansen HE, Schwarze U, Pyott SM, AlSwaid A, Al Balwi M, Alrasheed S, Pepin MG, Weis MA, Eyre DR, Byers PH (2010) Homozygosity for a missense mutation in SERPINH1, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. Am J Hum Genet 86:389-398

Constantinou CD, Nielsen KB, Prockop DJ (1989) A lethal variant of osteogenesis imperfecta has a single base mutation that substitutes cysteine for glycine 904 of the alpha 1(I) chain of type I procollagen. The asymptomatic mother has an unidentified mutation producing an overmodified and unstable type I procollagen. J Clin Invest 83:574-584

Dalgleish R (1997) The human type I collagen mutation database. Nucleic Acids Res 25:181-187

Dehghan A, Kottgen A, Yang Q, Hwang SJ, Kao WL, Rivadeneira F, Boerwinkle E, Levy D, Hofman A, Astor BC, Benjamin EJ, van Duijn CM, Witteman JC, Coresh J, Fox CS (2008) Association of