Hyperuricemia cosegregating with osteogenesis imperfecta is associated with a mutation in \textit{GPATCH8}

Hiroshi Kaneko · Hiroshi Kitoh · Tohru Matsuura · Akio Masuda · Mikako Ito · Monica Mottes · Frank Rauch · Naoki Ishiguro · Kinji Ohno

Abstract  Autosomal dominant osteogenesis imperfecta (OI) is caused by mutations in \textit{COL1A1} or \textit{COL1A2}. We identified a dominant missense mutation, c.3235G>A in \textit{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 \textit{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 \textit{PRPSAPI} and \textit{PRPSAP2} as candidate genes for hyperuricemia on chr 17 where \textit{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 \textit{ABCG2} and \textit{SLC22A12}. The same SNPs, however, were also detected in normouricemic individuals in three families. We then identified two missense SNVs in \textit{ZPBP2} and \textit{GPATCH8} on chromosome 17 that cosegregated with hyperuricemia in the Japanese family. \textit{ZPBP2} p.T69I was at the non-conserved region and was predicted to be benign by \textit{in silico} analysis, whereas \textit{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. \textit{GPATCH8} is only 5.8 Mbp distant from \textit{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.

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. (1979) distinguishes types I, II, III and IV (MIM# 166200, 166210, 259420, and 166220, respectively) on the basis of clinical and radiographic findings. 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, 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 z1 and z2 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 FKBP65 (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 PPILB (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 z1 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 (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 z1 chain and one copy of the z2 chain form a core triple helix comprising 338 uninterrupted Gly-X–Y triplet repeats, where X is often proline and Y is often hydroxyproline. Gly repeats at every third position are X–Y triplet repeats, where X is often proline and Y is often hydroxyproline. Gly repeats at every third position are X–Y triplet repeats, where X is often proline and Y is often hydroxyproline. Gly substitutions of the mutant amino acids (Bodian et al. 2008; Marini et al. 2007; 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 is 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).

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. (2009) sequenced exomes of 12 humans, including four unrelated individuals with autosomal dominant Free-
man-Sheldon syndrome (MIM# 193700) and eight Hap-
Map individuals. They successfully identified mutations in MYH3 (MIM# 160720) in all the affected individuals. Ng et al. (2010a) 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. (2010a) 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). Similarly, Lalonde et al. (2010) sequenced two unrelated fetuses with autosomal recessive Fowler syndrome (MIM# 225790) and de novo identified compound heterozygous mutations in FLVCR2 (MIM# 610865).

In a Japanese family with OI type I, hyperuricemia co
segregated 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). Under
excretion of urate is causally associated with mutations in UMOD (MIM# 191845) encoding uromodulin (Hart et al. 2007; Rauch and Glorieux 2004).
We genotyped all family members for three microsatellite markers flanking \textit{COL1A1} (D17S1293, −16 Mbp; D17S1319, −14 kbp; and D17S788, 2 Mbp). As no annotated microsatellite markers were available close to \textit{COLIA2}, we posted three new microsatellite markers to DDBJ (AB499843, −17 kbp; AB499844, 29 kbp; and AB499845, 123 kbp) and analyzed them in the family. We fluorescently labeled the \textit{S}‘ end of each forward primer with FAM (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 \textit{COL1A1}

After the microsatellite analysis suggested that \textit{COL1A1} was more likely to be a causative gene, we amplified all exons and flanking intronic regions of \textit{~}40 bp, as well as \textit{S}‘ and \textit{3}‘ UTRs of \textit{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 ran 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 \textit{COL1A1} gDNA using the Mutation Surveyor software version 2.61 (SoftGenetics). We numbered \textit{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 \textit{COL1A1} exon 33 is named \textit{COL1A1} exon 33–34 to match the exonic annotations of \textit{COLIA2}.

Allele-specific primer (ASP)-PCR to trace the \textit{COL1A1} mutation

We traced the mutation in family members using ASP-PCR. The wild-type ASP was 5′-TCCCCCGGGTGAC GTAGG-3′, and the mutant ASP was 5′-TCCCCCCTCC GTCTGTAAG-3′, where the mutated nucleotide is underlined, and an artificially introduced mismatch is shown in bold. The reverse primer was 5′-GCCACGGTGAC CCTTTATGC-3′.

Prediction of effects of mutations on pre-mRNA splicing

Five missense mutations in exon 45 of \textit{COL1A1} cause mild to lethal OI phenotypes (Fig. 2a) (Constantinou et al. 1989;
We first examined splicing 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), ESR-search (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 the effects of the mutations on splice site strength of the 18 sequence variants using two Web-based programs: the NetGene2 Server (http://www.cbs.dtu.dk/services/NetGene2/) (Bru-nak 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)12–18 primer (Invitrogen). 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 h, we harvested cells.

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 h, we harvested cells.

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

To examine if the COL1A1 c.3235G>A mutation arose from a common founder in the Japanese, Italian (Mottes et al. 1992) and 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

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 × 10⁶ 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...
SNVs/indels. SNVs and indels were compared to dbSNP Build 132.

Analysis of variants in KRBA2, ZPBP2 and GPATCH8

To trace if variants in KRBA2, ZPBP2 and GPATCH8 co-segregated 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'-CGT GTCTTCAGCACAATA GG-3' and 5'-AGAAGCCGTA GCACCACCTC C-3', respectively. The reverse primers were 5'-GGCCCAATCCA TAA GTACAT-3' and 5'-CCCA TGATCTTCCT GGAG-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 ten 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 15–30 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 18 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.

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 Mbp) and an 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-year-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. His 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...
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. Among the 12 SNPs, rs2231142 in \textit{ABCG2} as well as rs3825016 and rs11231825 in \textit{SLC22A12} are previously reported and 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 \textit{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 \textit{KRBA2} at 17p13.1, c.206C>T in exon 3 of \textit{ZPBP2} (MIM# 608499) at 17q12 (Fig. 3a) and c.2935G>C in exon 8 of \textit{GPATCH8} at 17q21.31 (Fig. 3b). Capillary sequencing revealed that variants in \textit{ZPBP2} and \textit{GPATCH8} cosegregated with hyperuricemia in F1, but the \textit{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 \textit{ZPBP2} and \textit{GPATCH8}.

\textit{ZPBP2} c.206C>T and \textit{GPATCH8} c.2935G>C predict amino acid substitutions of p.T69I and p.A979P, respectively. Threonine 69 and the flanking amino acids of \textit{ZPBP2} are not conserved across mammalian species. Additionally, an SNP rs35591738 mutates the N-terminal proline at codon 68, and an SNP rs34272593 causes a frameshift at codon 70 (Fig. 3a). In contrast, alanine 979 and the flanking amino acids of \textit{GPATCH8} are in the serine-rich region and are highly conserved across mammalian species (Fig. 3b). PolyPhen-2 (Adzhubei et al. 2010) predicted that \textit{ZPBP2} p.T69I was benign with a score of 0.025 and \textit{GPATCH8} p.A979P was damaging with a score of 0.988, where 1.0 was the worst score. Similarly, SIFT (Kumar et al. 2009) predicted that \textit{ZPBP2} p.T69I was tolerated with a score of 0.38 and \textit{GPATCH8} p.A979P was damaging with a score of 0.00, where a score < 0.05 was predicted to be deleterious.
Table 2  Splice site strength of COL1A1 exon 45 predicted by the NetGene2 and the Splice Site Prediction by Neural Network

<table>
<thead>
<tr>
<th>Sequence variation</th>
<th>NetGene2</th>
<th>Splice Site Prediction by Neural Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Confidence</td>
<td>Score</td>
</tr>
<tr>
<td>Wild type</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>rs1800215</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>rs1800217</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>c.3226G&gt;A</td>
<td>–</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3226G&gt;A+rs1800215</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3226G&gt;A+rs1800217</td>
<td>–</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3226G&gt;T</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3226G&gt;T+rs1800215</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3226G&gt;T+rs1800217</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3235G&gt;A</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3235G&gt;A+rs1800215</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3235G&gt;A+rs1800217</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3244G&gt;T</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3244G&gt;T+rs1800215</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3244G&gt;T+rs1800217</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3253G&gt;A</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3253G&gt;A+rs1800215</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>c.3253G&gt;A+rs1800217</td>
<td>–</td>
<td>0.94</td>
</tr>
</tbody>
</table>

– symbol represents being identical to the wild-type

*In addition to the native splice acceptor site of 0.98, a cryptic splice acceptor site ‘AG’ is generated at c.3253_3254

Table 3  Twelve SNPs identified by exome resequencing in five out of ten genes associated with hyperuricemia

<table>
<thead>
<tr>
<th>Ch</th>
<th>Gene</th>
<th>Position</th>
<th>Nuc.</th>
<th>Amino acid</th>
<th>AF</th>
<th>dbSNP</th>
<th>II-2</th>
<th>II-3</th>
<th>VIII-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AGL</td>
<td>100,336,361</td>
<td>C&gt;T</td>
<td>Syn.</td>
<td>0.7</td>
<td>rs2230306</td>
<td>T/T</td>
<td>–/T</td>
<td>–/T</td>
</tr>
<tr>
<td>4</td>
<td>ABCG2</td>
<td>89,034,551</td>
<td>G&gt;A</td>
<td>Syn.</td>
<td>0.02</td>
<td>rs35622453</td>
<td>–/–</td>
<td>–/–</td>
<td>–/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>89,052,323</td>
<td>C&gt;A</td>
<td>Q141K</td>
<td>0.31</td>
<td>rs2231142</td>
<td>–/A</td>
<td>A/A</td>
<td>–/–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>89,061,114</td>
<td>G&gt;A</td>
<td>V12M</td>
<td>0.19</td>
<td>rs2231137</td>
<td>–/A</td>
<td>–/–</td>
<td>–/A</td>
</tr>
<tr>
<td>4</td>
<td>SLC2A9</td>
<td>9,909,923</td>
<td>C&gt;T</td>
<td>P350L</td>
<td>0.33</td>
<td>rs2280205</td>
<td>–/T</td>
<td>–/T</td>
<td>–/–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9,922,130</td>
<td>G&gt;A</td>
<td>R294H</td>
<td>0.72</td>
<td>rs3735591</td>
<td>–/A</td>
<td>–/–</td>
<td>–/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9,998,440</td>
<td>G&gt;A</td>
<td>Syn.</td>
<td>0.54</td>
<td>rs10939650</td>
<td>–/A</td>
<td>A/A</td>
<td>–/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,022,981</td>
<td>G&gt;A</td>
<td>G25R</td>
<td>0.43</td>
<td>rs2276961</td>
<td>–/A</td>
<td>–/–</td>
<td>–/–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,027,542</td>
<td>G&gt;A</td>
<td>A17T</td>
<td>0.06</td>
<td>rs6820230</td>
<td>–/–</td>
<td>–/–</td>
<td>–/A</td>
</tr>
<tr>
<td>11</td>
<td>SLC22A1</td>
<td>64,359,286</td>
<td>C&gt;T</td>
<td>Syn.</td>
<td>0.21</td>
<td>rs3825016</td>
<td>–/T</td>
<td>T/T</td>
<td>–/T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>64,360,274</td>
<td>C&gt;T</td>
<td>Syn.</td>
<td>0.81</td>
<td>rs11231825</td>
<td>–/T</td>
<td>–/–</td>
<td>–/T</td>
</tr>
<tr>
<td>12</td>
<td>PFKM</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>UMOD</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>G6PC</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>HPRT1</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>PRPS1</td>
<td>n.d.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>MAOA</td>
<td>43,591,036</td>
<td>G&gt;T</td>
<td>Syn.</td>
<td>0.3</td>
<td>rs6323</td>
<td>–/–</td>
<td>T/T</td>
<td>T/T</td>
</tr>
</tbody>
</table>

*The gene is associated with purine metabolism

*The gene is associated with renal excretion of urate

– symbol represents in the patients’ genotypes mean being identical to the reference nucleotides

*Syn synonymous nucleotide change, AF allelic frequency of the changed nucleotide, n.d. no SNPs are detected
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 16–20% 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) experienced many fractures, whereas the father (I-1) and another son (II-2) had 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 was classified as OI type IV, whereas his daughter as 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 type 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 10 candidate genes that are associated with hyperuricemia, purine metabolism and renal excretion, and found 12 SNPs in 10 genes (Table 3). Among them, three SNPs in ABCG2 (rs2231142) and SLC22A12 (rs3825016 and rs11231825) are previously reported markers for hyperuricemia and/or gout.
Dehghan et al. (2008) 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). Woodward et al. (2009) showed a significant association between rs2231142 and hyperuricemia (OR = 1.68) in a population-based study of 14,783 individuals. Kolz et al. (2009) demonstrated that rs2231142 elevated the serum urate concentration more strongly in men than in women by meta-analysis of 28,141 individuals. Stark et al. (2009) analyzed 683 patients with gout and indicated a significant association between rs2231142 and gout (OR = 1.37). Although rs2231142 is an attractive causative SNP, our normouricemic subjects were also heterozygous for rs2231142.

Graeser et al. (2006) 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. Tabara et al. (2010) analyzed 1,526 normal Japanese individuals retrospectively and longitudinally, and clarified that rs11231825 was associated with reduced urate excretion and with future development of hyperuricemia. 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 considering the 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 OA.

Acknowledgments We would like to thank the families for their participation in this study. We are grateful to Dr. Kunio Ibara 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


Constantinou JD, 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. The human type I collagen gene has a single base mutation that substitutes glycine 904 of the alpha 1(I) chain of type I procollagen. J Clin Invest 83:574–584
