

Evaluation of All Nonsynonymous Single-Nucleotide Polymorphisms in the Gene Encoding Human Deoxyribonuclease I-Like 1, Possibly Implicated in the Blocking of Endocytosis-Mediated Foreign Gene Transfer

Misuzu Ueki,^{1,*} Kaori Kimura-Kataoka,^{2,*} Junko Fujihara,² Haruo Takeshita,²
Reiko Iida,³ and Toshihiro Yasuda¹

Many nonsynonymous single-nucleotide polymorphisms (SNPs) in the human deoxyribonuclease I-like 1 (DNase 1L1) gene, possibly implicated in the blocking of endocytosis-mediated foreign gene transfer, have been identified, but only limited population data are available and no studies have evaluated whether such SNPs are functional. Genotyping of all 21 nonsynonymous human DNase 1L1 SNPs was performed in 16 different populations representing three ethnic groups using the PCR-restriction fragment length polymorphism technique. All of the nonsynonymous SNPs, except for SNP p.Val122Ile in Caucasian populations, exhibited a monoallelic distribution in all of the populations. On the basis of alterations in the activity levels resulting from the corresponding amino acid substitutions, two activity-abolishing and four activity-reducing SNPs were confirmed to be functional. Although all of the nonsynonymous SNPs that affected the catalytic activity showed extremely low genetic heterogeneity, it seems plausible that a minor allele of six SNPs producing a loss-of-function or extremely low-activity variant could serve directly as a genetic risk factor for diseases. Especially, the amino acid residues in activity-abolishing SNPs were conserved in animal DNases 1L1. Furthermore, results of phylogenetic analysis suggest that DNase 1L1 might have appeared latest among the DNase I family during the course of molecular evolution.

Introduction

DEOXYRIBONUCLEASE I-LIKE 1 (DNase 1L1), also designated as DNase X/DNase Xib, was the first member of the DNase I (EC3.1.21.1) family, showing a high amino acid sequence similarity to DNase I, to be isolated (Parrish *et al.*, 1995; Coy *et al.*, 1996; Pergolizzi *et al.*, 1996; Los *et al.*, 2000). The gene encoding human DNase 1L1 is located in region q28 on the X chromosome. DNase 1L1 exhibits a unique tissue distribution distinct from the other members of the DNase I family, showing abundant expression in heart and skeletal muscle, where the latter DNases are scarcely expressed (Parrish *et al.*, 1995; Los *et al.*, 2000; Shiokawa *et al.*, 2001, 2005). An outstanding structural feature of mammalian DNase 1L1 is an extra-hydrophobic stretch in its C-terminal region, which is considered to serve as a signal sequence for the attachment of glycosylphosphatidylinositol (GPI), re-

sulting in membrane binding of the enzyme mediated by a GPI-anchor (Shiokawa *et al.*, 2007); the other members of the DNase I family are extra/intracellular secretory enzymes, whereas DNase 1L1 is expressed on the cell surface. On the basis of these features, it is plausible that DNase 1L1 might have distinct and unique physiological role(s) unlike those of other DNases. In this context, Shiokawa *et al.* (2007) have revealed that DNase 1L1 acts as a cell surface barrier to endocytosis-mediated transfer of foreign genes. It has also been postulated that DNase1L1 may be involved in some diseases; accumulation of the DNase 1L1 transcript has been observed in the vacuoles of muscle cells from patients with vacuolar myopathy, a form of acid maltase deficiency (Malferrari *et al.*, 1999); it has been suggested that lethal infantile Pompe disease is associated with deletion polymorphism in the DNase 1L1 gene (*DNASE1L1*) (Malferrari *et al.*, 2001); also, Singh *et al.* (2009) reported that a novel Alu-mediated Xq28

¹Division of Medical Genetics and Biochemistry, Faculty of Medical Sciences, University of Fukui, Fukui, Japan.

²Department of Legal Medicine, Shimane University School of Medicine, Shimane, Japan.

³Division of Life Sciences, Faculty of Medical Sciences, University of Fukui, Fukui, Japan.

*These two authors contributed equally to this work.

microdeletion partially deleted *DNASE1L1* in a patient with Barth syndrome. These findings suggest that genetic factors affecting *in vivo* DNase 1L1 activity such as single-nucleotide polymorphisms (SNPs) or deletions may have clinical implications in relation to the prevalence of various diseases.

Among the SNPs in *DNASE1L1*, we have focused on nonsynonymous SNPs potentially affecting the DNase activity through the corresponding amino acid substitution as a functional SNP; three nonsynonymous SNPs, p.Val122Ile (rs34952165), p.Gln170His (rs6643670), and p.Asp227Ala (rs5987256), have been examined in previous studies (Fujihara *et al.*, 2010; Ueki *et al.*, 2010a). Since then, a total of 18 nonsynonymous SNPs in the gene have been registered in the NCBI dbSNP database (www.ncbi.nlm.nih.gov/projects/SNP, May 2013) (Fig. 1). However, only limited population data are available for these SNPs. In this context, data pertaining to the contribution of *DNASE1L1* to Pompe disease have been inconsistent, possibly due to differences in the origins of the study populations (Malferrari *et al.*, 2001; Lichtenbelt *et al.*, 2006). To overcome this problem, there is a need to accumulate data on the genetic distributions of SNPs in many populations derived from different ethnic groups. In particular, it is still unclear whether individual SNPs in *DNASE1L1* may serve as a functional SNP (Wu *et al.*, 2012) affecting its *in vivo* activity of the enzyme. Therefore, comprehensive data on the biochemical-genetic aspects of these SNPs in *DNASE1L1* potentially affecting the *in vivo* enzyme activity would be useful for clarifying their functionality in determining genetic predisposition to disease.

In the present study, we investigated the genotype distribution of the 18 nonsynonymous SNPs in *DNASE1L1* in 16 different populations worldwide derived from three

ethnic groups, using a novel genotyping method for each SNP, and the effects of all these SNPs on the enzymatic activity, to evaluate the genetic heterogeneity and functionality of each SNP. Furthermore, multiple alignment analysis of the amino acid sequences of animal DNase 1L1, including phylogenetic analysis, was performed to examine the role of each SNP-related amino acid residue in the protein.

Materials and Methods

Samples from subjects

Genomic DNA was extracted using a QIAamp DNA Mini Kit (Qiagen, Chatsworth, CA) from blood or bloodstain samples randomly collected from healthy subjects ($n=1752$) derived from 16 different populations. The Asian populations included 110 Japanese (Shimane Prefecture), 352 Koreans (Busan, South Korea), 193 South Chinese (Shenyang and Guangzhou of China), 112 Mongolians (Ulaanbaatar, Mongol), 153 Tibetans (Katmandu of Nepal), 35 Sri Lanka Tamils (Kandy of Sri Lanka), 48 Sri Lanka Sinhalese (Kandy of Sri Lanka), and 40 Tamangs (Kotyang of Nepal); the Caucasian population included 136 Turks (Adana area, Southern Turkey), 68 Germans (Munich), and 199 Mexicans (60 Mestizos, 88 Nahuas, and 51 Huicholes); the African populations included 126 Ovambos (Bantusin, Namibia), 105 Ghanaians (Ghana), and 75 Xhosas (Cape Town, South Africa). Some of the samples were the same as those used for previous studies (Fujihara *et al.*, 2010; Ueki *et al.*, 2010a). Written informed consent was obtained from each participant. The study was approved by the Human Ethical Committees of the institute.

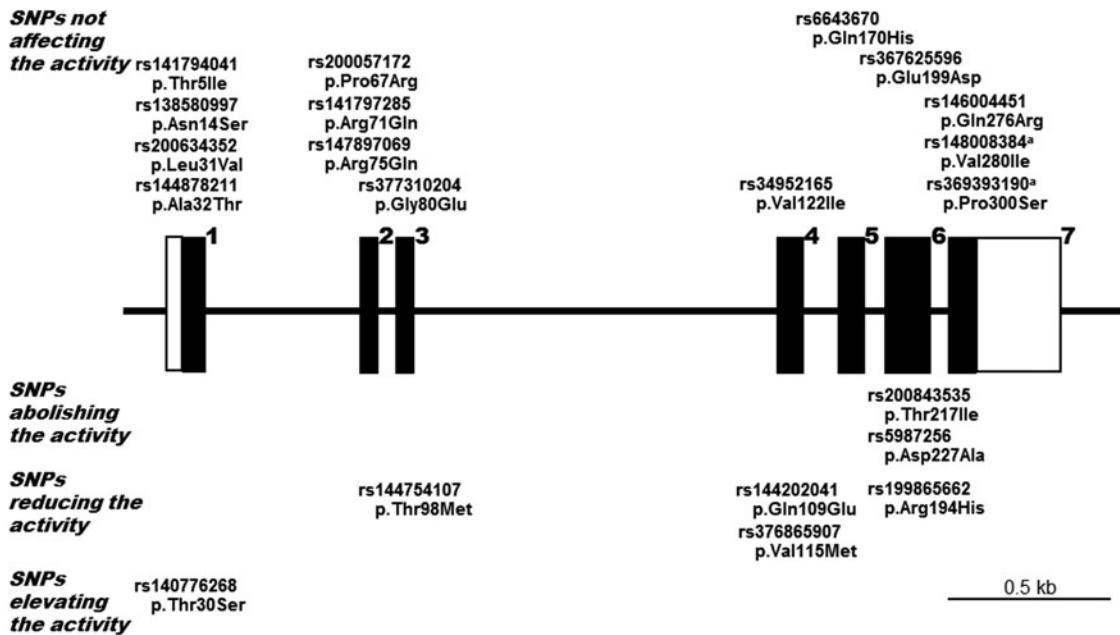


FIG. 1. Schematic representation of the genomic structure of the human *DNASE1L1* and the positions of the single-nucleotide polymorphisms (SNPs) examined. Each ID number in the NCBI database is shown. Exons are shown by solid boxes. According to alterations in the levels of the DNase 1L1 activity, 19 nonsynonymous SNPs were classified into 14 SNPs not affecting the activity, 2 abolishing the activity, 4 reducing the activity, and 1 elevating the activity; each of the SNPs in these four groups are shown at the sites of the exons where they are located. ^aThe effect of the amino acid substitution derived from SNPs p.Val280Ile and p.Pro300Ser, on the activity was not examined.

Construction of expression vectors encoding human DNase 1L1 and its variant forms corresponding to each SNP

Expression pcDNA3.1 (+) vectors (Invitrogen, San Diego, CA) inserted with the entire coding sequence of human DNase 1L1 cDNA was prepared from the total RNA of human skeleton muscle (Clontech, Palo Alto, CA) by reverse transcriptase-PCR amplification using an Expanded High Fidelity PCR System (Roche Diagnosis, Mannheim, Germany) with a set of two primers, 5'-GAATTCAGCCACACAGCCATGCACTACC-3' (sense) and 5'-CTCGAGCGCTCAGGCAGCAGGGCACAG-3' (antisense), corresponding to the sequences of cDNA from positions -13-10 and 900-932, respectively; inserted cDNAs was derived from the predominant haplotype of the SNPs in Japanese. Using this construct as a template, the DNase 1L1 Δ C construct, which was devoid of the nucleotide region corresponding to the C-terminal domain, including Ala277-Ala302, used for the activity assay was prepared using the KOD-Plus Mutagenesis kit (Toyobo Co., Ltd., Osaka, Japan) and used as a wild type. Nineteen amino acid-substituted constructs corresponding to each SNP were prepared by site-directed mutagenesis using a KOD-Plus Mutagenesis Kit with the wild-type construct as a template. In these constructs, the amino acid residue was replaced by the counterpart derived from a minor allele in each SNP; for example, the T5I construct of DNase 1L1, in which a Thr residue at position 5 in the protein is replaced by Ile derived from the minor allele, corresponds to SNP p.Thr5Ile. Furthermore, the five other amino acid-substituted constructs related to the SNPs were prepared in the same manner. Nucleotide sequences of all the constructs were confirmed by DNA sequence analysis. Purification of each construct used for transfection was performed using the Plasmid Midi kit (Qiagen).

Transient expression of the expression vectors and assay for the DNase 1L1 activity

COS-7 cells were maintained in Dulbecco's modified Eagle's medium containing 1 mM L-glutamine, 50 U/mL penicillin, 50 μ g/mL and 10% (v/v) fetal calf serum at 37°C under 5% CO₂ in air. The cells were transiently transfected four times with 2 μ g of each DNase 1L1-related expression vector and 600 ng of pSV- β -galactosidase vector (Promega, Madison, WI; for estimation of transfection efficiency) using Lipofectamine 2000 reagent (Invitrogen) according to the method described previously (Yasuda *et al.*, 2010). At 48 h after transfection, the cells were harvested. Then, the cells were subjected to sonication using Bioruptor UCD-250 (COSMO BIO Co., LTD, Tokyo, Japan) to prepare the cell lysates for the subsequent assay. DNase 1L1 activity in both the lysates derived from the transfected cells and their conditioned medium were assayed by the single radial enzyme diffusion (SRED) method using a LAS-3000 imaging analyzer (Fuji Film, Tokyo, Japan) according to our previous reports (Ueki *et al.*, 2010a). The activity of wild-type DNase 1L1 was defined as 1.0 and that of the amino acid-substituted form was expressed relative to the wild type. The relative activity is expressed as mean \pm standard deviation derived from four transfections.

Genotyping of nonsynonymous SNPs in DNASE1L1 by the PCR-restriction fragment length polymorphism method

Genotyping assays for each of 18 nonsynonymous SNPs in *DNASE1L1* were separately performed using PCR followed by restriction fragment length polymorphism (RFLP) analysis (Yasuda *et al.*, 1995), in the same manner as the other SNPs (Ueki *et al.*, 2010a). Primers for the specific amplification of the DNA fragments encompassing a substitution site corresponding to each SNP were designed on the basis of the nucleotide sequences of the human *DNASE1L1* (NCBI Reference Sequence: NG 012884.1) (Supplementary Table S1; Supplementary Data are available online at www.liebertpub.com/dna). Nucleotide and amino acid residues of *DNASE1L1* were numbered from the 5'-terminus of the translation initiation codon and the N-terminal amino acid residue of the precursor protein, respectively. SNP nomenclature is based on the recommendations for the description of sequence variants (www.hgvs.org/mutnomen/examplesDNA.html, May 2013); the sequence of DNase 1L1 (GenBank: accession no. X90392) has been used as the coding DNA Reference Sequence.

PCR amplification was performed in a 25 μ L reaction mixture using \sim 5 ng of DNA. The reaction mixture contained a 1 \times buffer (15 mM Tris-HCl, pH 8.0, 50 mM KCl), 1.5 mM MgCl₂, 0.5 μ M of each primer, 200 μ M dNTPs, and 1.25 U of Taq polymerase (AmpliTaq Gold; Applied Biosystems, Foster City, CA). PCR was performed with a protocol consisting of initial denaturation at 94°C for 7 min, followed by 30 cycles with denaturation at 94°C for 30 s, annealing at 55°C-65°C (Supplementary Table 1) for 30 s, and extension at 72°C for 1 min, followed by a final extension at 72°C for 7 min. Two microliters of the PCR product obtained using each pair of primers was digested with 5 U each enzyme listed in Supplementary Table 1 (New England Biolabs, Ipswich, MA) at 37°C for 3 h in a final reaction mixture volume of 15 μ L to determine the genotype of each SNP. The digested products (5 μ L) were separated in an 8% polyacrylamide gel, and the patterns on the gels were visualized by silver or ethidium bromide staining, as described previously (Yasuda *et al.*, 2010).

Direct sequencing of the genomic DNA region including substitution site of each SNP in the DNase 1L1 gene for the representative subjects was performed by the dideoxy chain-terminating method with the BigDye[®] Terminator Cycle Sequencing kit (Applied Biosystems). The sequencing run was performed on a Genetic Analyzer 310 (Applied Biosystems).

Analytical methods

Each Ortholog of human DNases 1L1 and 1L2 in other available animal species were surveyed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (www.genome.jp/kegg, August 2013); the amino acid sequence of human DNase 1L1 or 1L2 (GenBank: accession no. BC063710) was used as query sequence for Basic Local Alignment Search Tool (BLAST) searches of each genome database. Multiple alignment analyses of the amino acid sequences of animal DNase 1L1 were performed using DNASIS Pro V3.0 (Hitachi Solutions, Ltd., Tokyo, Japan).

DNase 1L1 activity was compared between wild-type and variant enzymes by means of the unpaired, Student's *t*-test. Differences at $p < 0.05$ were considered to be statistically significant.

TABLE 1. SUMMARY ON EVALUATION OF ALL THE NONSYNONYMOUS SNPs IN *DNASE1L1* AS A FUNCTIONAL SNP; GENETIC HETEROGENEITY, EFFECT OF THE CORRESPONDING AMINO ACID SUBSTITUTION ON THE DNASE 1L1 ACTIVITY, AND AMINO ACID MULTIPLE ALIGNMENT OF ANIMAL *DNASE 1L1*

SNP	Polymorphism (this study)	Heterozygosity ^a	Effect ^b on the activity	Amino acid multiple alignment of 23 animal DNase 1L1 ^c
<i>SNPs not affecting the activity</i>				
rs141794041 p.Thr5Ile; c.14C>T	Monoallelic	0.001	0.82±0.05	Not conserved; substituted by Met, Val (3), Ala (4), Pro (2), Leu (4), or deletion (5)
rs138580997 p.Asn14Ser; c.41A>G	Monoallelic	0.002	0.98±0.23	Not conserved; substituted by Pro, Ser (5), Gly (6), Ala (2), Leu (3), or Val (2)
rs200634352 p.Leu31Val; c.91G>C	Monoallelic	N.D.	0.92±0.22	Well conserved in mammals; substituted by Val, Glu (2), Asp (3), Arg, Gln, or Lys
rs144878211 p.Ala32Thr; c.94G>A	Monoallelic	0.000	0.95±0.066	Not conserved; substituted by Thr (3), Ser (7), or Lys (3)
rs200057172 p.Pro67Arg; c.200C>G	Monoallelic	0.001	0.87±0.31	Not conserved; substituted by Leu, Ser (3), Gly, Asn (2), Lys (2), Tyr, or Val
rs141797285 p.Arg71Gln; c.212G>A	Monoallelic	0.000	0.77±0.22	Not conserved; substituted by Gln (3), Asp (3), Ser, His, Lys, or Glu
rs147897069 p.Arg75Gln; c.224G>A	Monoallelic	0.000	0.80±0.10	Not conserved; substituted by Ser (2), Gln, Lys, Ala, Gly, or deletion (2)
rs377310204 p.Gly80Glu; c.238G>A	Monoallelic	N.D.	1.22±0.39	Not conserved; substituted by Arg, Lys, Asp (3), Tyr, His (3), Asn, or Val
rs34952165 p.Val122Ile; c.366C>A	Caucasian	0.043	0.78±0.12	Not conserved; substituted by Ile (5), Leu (7), Thr (2), Ser, or Val (3)
rs6643670 p.Gln170His; c.510G>C	Monoallelic	N.D.	1.12±0.24	Not conserved; substituted by Lys (2), Gly, Asp (3), Asn, His, or Val
rs367625596 p.Glu199Asp; c.597G>C	Monoallelic	N.D.	0.96±0.28	Not conserved; substituted by Leu (3), Ala, Val (5), Arg (9), or Glu
rs146004451 p.Gln276Arg; c.827A>G	Monoallelic	0.004	1.10±0.13	Not conserved; substituted by Thr, Ala (6), Val, Leu (3), Cys, or deletion (7)
rs148008384 ^d p.Val280Ile; c.838G>A	Monoallelic	N.D.	n.e.	Not conserved; substituted by Ile (3), Pro, Cys (3), Glu, or deletion (6)
rs369393190 ^d p.Pro300Ser; c.898C>T	Monoallelic	N.D.	n.e.	Not conserved; substituted by Ile (7), Arg, Lys, Gly, or deletion (6)
<i>SNPs abolishing the activity</i>				
rs200843535 p.Thr217Ile; c.650C>T	Monoallelic	0.001	n.d.	Completely conserved
rs5987256 p.Asp227Ala; c.680A>C	Monoallelic	N.D.	n.d.	Completely conserved
<i>SNPs reducing the activity</i>				
rs144754107 p.Thr98Met; c.293C>T	Monoallelic	0.004	0.21±0.06 (<i>p</i> <0.001)	Not conserved: substituted by Lys (8) or Gln (10)
rs144202041 p.Gln109Glu; c.327G>A	Monoallelic	0.000	0.55±0.02 (<i>p</i> <0.001)	Not conserved; substituted by Glu (2), Arg, Ser (4), or Thr (3)
rs376865907 p.Val115Met; c.343G>A	Monoallelic	N.D.	0.37±0.10 (<i>p</i> <0.001)	Not conserved; substituted by Gln (5), Met, Leu (2), Tyr (3), Lys (2), His, or Glu
rs199865662 p.Arg194His; c.581G>A	Monoallelic	N.D.	0.52±0.061 (<i>p</i> <0.001)	Not conserved; substituted by His, Ser, Gln (3), Asp (3), or Lys (2)
<i>SNPs elevating the activity</i>				
rs140776268 p.Thr30Ser; c.88A>T	Monoallelic	0.000	2.85±0.68 (<i>p</i> <0.001)	Well conserved in mammals; substituted by Gly (9)

^aTaken from the dbSNP.

^bActivity of each amino acid substituted DNase 1L1 is expressed as relative one to that of wild-type enzyme; the activity (*n*=4) of each construct in the lysates derived from the transfected cells was compared with that of the wild-type enzyme (1.0±0.11).

^cMultiple alignment analysis of amino acid sequence was performed using 23 animal DNases 1L1 available in database. Number in parenthesis is the number of species in which the corresponding amino acid is substituted.

^dThe expression construct DNase 1L1 ΔC lacking the C-terminal hydrophobic domain was used for the activity assay; since the amino acid residues corresponding to SNPs, pVal280Ile and p.Pro300Ser, were located in the C-terminal hydrophobic domain, the corresponding amino acid-substituted constructs could not be prepared, and therefore, the effect of these substitutions on the activity was not examined.

N.D., not determined; n.d., not detected; SNP, single-nucleotide polymorphism; n.e., not examined.

Results

Effect of amino acid substitution resulting from each nonsynonymous SNP in DNASE1L1 on the expression of DNase activity

We have previously developed the SRED method for quantification of endonuclease activity, including DNase I, utilizing the radial diffusion of the enzyme in a thin agarose gel plate containing DNA as a substrate (Nadano *et al.*, 1993; Yasuda *et al.*, 1998). The SRED method allowed us to detect and quantify extremely low levels of the DNase activity. However, the SRED method can only be applied for assay of the DNase activity in a soluble state. Since DNase 1L1 is a GPI-anchored membrane DNase and its C-terminal hydrophobic domain serves as a signal sequence for GPI attachment (Shiokawa *et al.*, 2007), it is not possible to measure the activity of intact DNase 1L1 in cells with sufficient sensitivity to clarify the effects of amino acid substitutions corresponding to individual nonsynonymous SNPs in *DNASE1L1* on the activity. To overcome these problems, we constructed a DNase 1L1 Δ C form lacking the C-terminal hydrophobic domain, thus allowing the conversion of the GPI-anchored form of DNase 1L1 to a soluble form. This allowed us to quantify the DNase 1L1 with sufficient sensitivity using the conventional SRED method. An expression construct of DNase 1L1 Δ C containing the entire coding sequence of human DNase 1L1, but lacking the C-terminal hydrophobic domain, derived from the predominant haplotype of these SNPs was used as the wild type; 19 variant constructs of the DNase 1L1 protein derived from the minor allele in each SNP were prepared separately, transiently expressed in COS-7 cells, and the resulting DNase 1L1 activities in lysates derived from transfected cells (Table 1) and their conditioned medium (Fig. 2) were determined separately by the SRED method. Generally, alterations in the activity induced by amino acid substitution corresponding to each SNP in the lysates derived from the transfected cells were similar to those in the conditioned medium.

Among them, the levels of the DNase 1L1 activity derived from the T98M, Q109E, V115M, and R194H constructs were significantly lower than that of the wild-type enzyme. Substitution of Gln, Val, and Arg residues at positions 109, 115,

and 194, respectively, in the DNase 1L1 protein obviously reduced the activity to \sim 40%–50% of that of the wild-type enzyme, and furthermore, substitution of a Thr residue at position 98 reduced the activity greatly to \sim 20%. It was noteworthy that the two constructs, T217I and D227A, exhibited no DNase 1L1 activity under our assay conditions. Therefore, among the SNPs affecting the DNase 1L1 activity, the activity was found to be abolished by amino acid substitutions resulting from p.Thr217Ile and p.Asp227Ala. On the other hand, when compared to the activity level of the wild type, the activity level of T30S was significantly high; substitution of a Thr residue at position 30 in the protein elevated the DNase 1L1 activity to about twofold that of the wild-type enzyme. However, the levels of the DNase 1L1 activity derived from the other amino acid-substituted constructs were similar to that of the wild type, indicating that substitution of Ile5, Asn14, Leu31, Ala32, Pro67, Arg71, Arg75, Gly80, Val122, Gln170, Glu199, and Gln276 residues in the DNase 1L1 protein had little effect on the activity. These findings permitted us to demonstrate that, among the amino acid residues related to nonsynonymous SNPs in the DNase 1L1 protein, each of the amino acid residue at positions 98, 109, 115, and 194 might be involved in expression of the enzyme activity, and that furthermore, those at positions 217 and 227 might be indispensable.

From these analyses of the variant forms corresponding to each of the nonsynonymous SNPs, we were able to divide these SNPs into four classes according to the observed alterations in the enzyme activity levels resulting from the corresponding amino acid substitutions: those not affecting the activity, those abolishing it, those reducing it, and those elevating it; among the 19 nonsynonymous SNPs examined in *DNASE1L1*, 12 did not affect the activity, 2 abolished it, 4 reduced it, and 1 elevated it (Fig. 1).

Multiple alignment analysis of the amino acid sequences of animal DNase1L1 and possible involvement of the amino acid residues corresponding to individual SNPs in expression of the enzyme activity

Multiple alignment analysis of the amino acid sequences of animal DNase 1L1 is useful for assessing the roles of

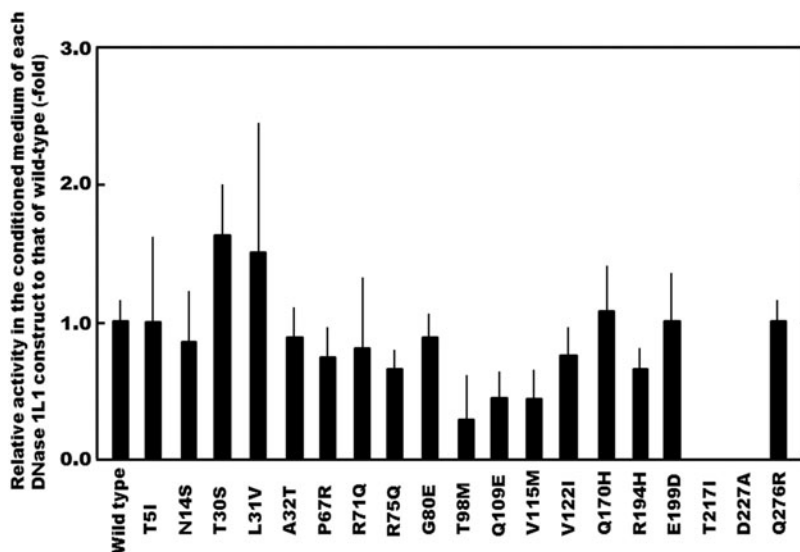


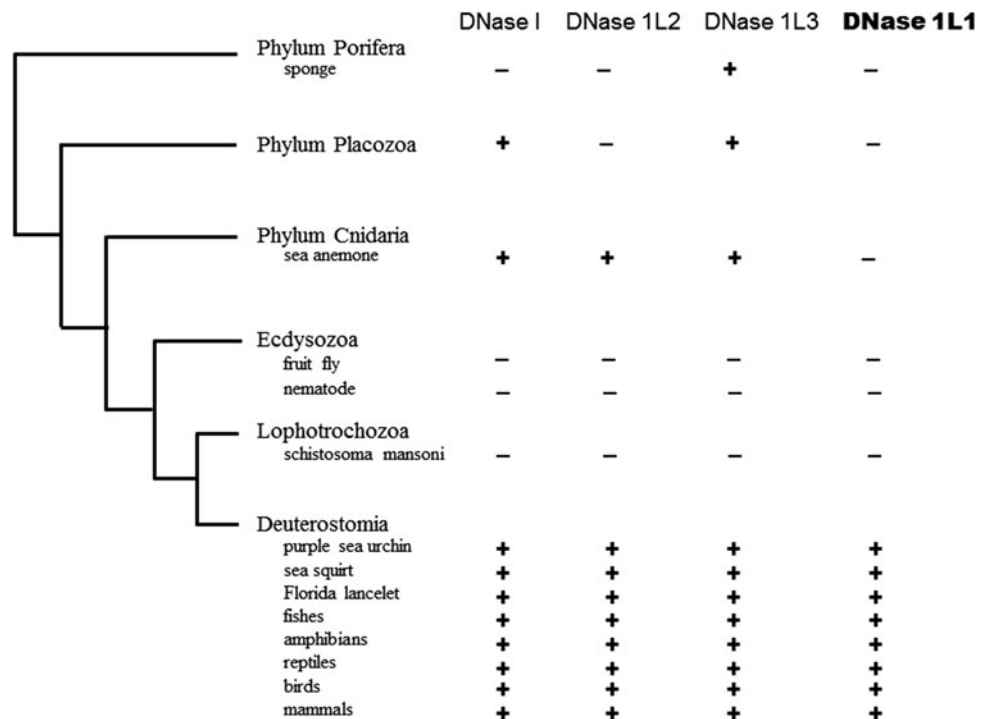
FIG. 2. Relative activity of variant constructs corresponding to nonsynonymous SNPs to that of the wild-type DNase 1L1. The DNase 1L1 activity in the conditioned medium derived from the cells transfected with each construct was assayed by the single radial enzyme diffusion (SRED) method (Ueki *et al.*, 2010a). The constructs, T217I and D227A, exhibited no DNase 1L1 activity detectable under our assay condition. The bars present SD ($n=4$).

individual SNP-related amino acid residues in the DNase 1L1 protein. Therefore, to survey orthologs of DNase 1L1, together with DNase 1L2, in Animalia organisms available in the KEGG database, a BLAST search of each genome database was performed using the amino acid sequence of human DNases 1L1 or DNase 1L2 as a query sequence. Orthologs of human DNase 1L1 with a low E-value were found in 23 organisms, demonstrating that DNase 1L1 is well distributed in vertebrates including not only mammals as reported previously (Shiokawa *et al.*, 2005) but also birds, reptiles, amphibians, and fish (Fig. 3). Furthermore, lancelets (Florida lancelet), ascidians (sea squirt), and echinoderms (purple sea urchin) were demonstrated to possess each ortholog of human DNases 1L1 and 1L2. In addition, an ortholog of human DNase 1L2 was also found in a cnidarian (sea anemone). Recently, we demonstrated that an ortholog of human DNase 1L3 was present in a poriferan sponge (*Amphimedon queenslandia*), representing the lowest class in Metazoa, whereas the placozoan *Trichoplax adhaerens*, probably situated in a molecular-phylogenetic position intermediate between Porifera and Animalia, possessed both DNases I and 1L3 (Ueki *et al.*, 2013b). On the other hand, an ortholog of human DNase 1L1 was found only in Deuterostomia but not in Porifera, placozoans and cnidarians; furthermore, similar to other DNases, DNase 1L1 was not distributed in any of the genome databases available for Protostomia, including insects and nematodes, among the Animalia, in contrast to another distinct type of DNase, DNase II, which has been identified even in non-metazoan organisms (MacLea *et al.*, 2003). This phylogenetic distribution of the DNase I family in Animalia allows us to assume that DNase 1L1 might have appeared later in the course of molecular evolution of the DNase I family than did DNases I, 1L2, and 1L3; DNase 1L1 might have appeared at the phylogenetic level of the Deuterostomia.

We then investigated whether the amino acid residues corresponding to the various SNPs were well conserved in animal DNase 1L1 by multiple alignment analysis of the amino acid sequences of DNase 1L1 from Animalia organisms available on the genome database (Supplementary Fig. S1). The C-terminal hydrophobic domain serving as a signal sequence for the GPI attachment characteristic of mammal DNase 1L1 (Shiokawa *et al.*, 2007) was deleted in the lower classes of Animalia, including Monotremata (platypus) and Marsupialia (opossum), suggesting that the localization of DNase 1L1 on the cell surface might have been acquired at the level of Marsupialia/Eutheria during the molecular evolution of animal DNase 1L1. There was no nonsynonymous SNP at the codon for the four amino acid residues crucially responsible for expression of the DNase activity, corresponding to His148, Thr222, Tyr226, and His265 in the DNase 1L1 protein (Shiokawa *et al.*, 2005). However, in SNP p.Asp227Ala, an Asp residue at position 227 served as a general base in its catalysis is replaced by Ala; in fact, the minor allele giving rise to occurrence of an Ala residue in the DNase 1L1 protein produced a loss-of-function form of the enzyme. It is worth noting that the amino acid residues related to loss of the enzyme activity through amino acid substitution derived from p.Thr217Ile and p.Asp227Ala were completely conserved in animal DNase 1L1, whereas those corresponding to SNPs not affecting the activity were not conserved (Table 1), being compatible with our finding that all of the amino acid residues in activity-abolishing SNPs were completely or well conserved in animal DNase family proteins (Ueki *et al.*, 2013a, 2013b).

To clarify the potential involvement of each amino acid residue corresponding to these SNPs in the expression of the enzyme activity, a series of amino acid-substituted constructs were examined (Fig. 4). Similar to substitution of the Thr217 residue by Ile (T217I) resulting from SNP. p.Thr217Ile, that

FIG. 3. Molecular-phylogenetic distribution of DNases 1L1 and 1L2 in animals. A Basic Local Alignment Search Tool (BLAST) search was performed in each genome database available in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (www.genome.jp/kegg/) using the amino acid sequence of human DNase 1L1 or 1L2 as a query sequence to survey each ortholog of the enzymes in Animalia. Results in distribution of DNases I and 1L3 were taken from our study (Ueki *et al.*, 2013b). Each ortholog of human DNase I family with low E-value could be found (+); could not (-).



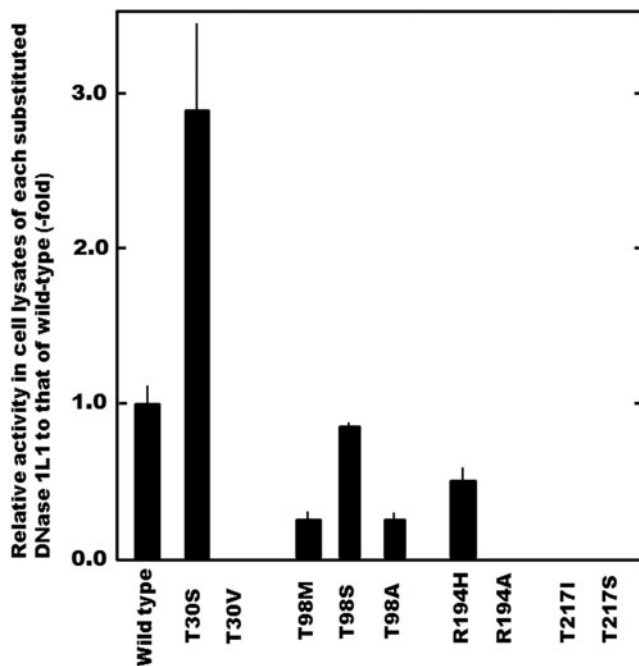


FIG. 4. Relative activity of amino acid-substituted variant constructs related to non-synonymous SNPs to that of the wild-type DNase 1L1. The DNase 1L1 activity in the lysates derived from the cells transfected with each construct was assayed by the SRED method (Ueki *et al.*, 2010a). The constructs, T30V and T217S, exhibited no DNase 1L1 activity detectable under our assay condition. The bars present SD ($n=4$).

by Ser (T217S) ablated the activity completely, even though it is a hydroxyl amino acid. Therefore, Thr itself located at position 217 in the protein, completely conserved in animal DNase 1L1, was found to be indispensable for the expression of the activity. Furthermore, replacement of Thr30 by a Val residue (T30V) abolished the DNase 1L1 activity, unlike the situation for the T30S construct, indicating that the hydroxyl group of the amino acid residue at position 30 in the protein, and not a hydrophobic residue, is required for expression of the enzyme activity. Since substitution of Thr30 by Ser elevated the DNase 1L1 activity, it seems plausible that a Ser residue is a more suitable hydroxyl amino acid than a Thr residue for expression of the DNase activity. With regard to SNPs reducing the activity, substitution of Thr98 by Ala (T98A) reduced the activity in the same manner as T98M, whereas that by Ser (T98S) did not greatly affect the activity. In other species, Lys or Gln residues are situated at position 98, in place of the Thr in human DNase 1L1. Also, substitution of Arg194 by Ala (R194A) reduced the activity more markedly than substitution by His (R194H). In this context, although these residues at position 194 are not conserved in animal DNases 1L1, substitution of several hydrophilic amino acids at the corresponding positions can all occur. Therefore, it seems plausible that at least a hydrophilic amino acid residue at positions 98 and 194 might be required for full expression of the enzyme activity. However, each amino acid residue corresponding to these SNPs reducing the activity was not necessarily well conserved in animal DNase 1L1, suggesting that the effect of amino acid substitutions resulting from these

SNPs on the enzyme activity may be attributable to the properties of individual amino acids.

A simple procedure for genotyping each SNP in DNASE1L1 using a PCR-RFLP method

A simple and novel genotyping procedure was developed using PCR-RFLP for all of the 18 SNPs in *DNASE1L1*, those for the other 3 SNPs having been developed previously (Ueki *et al.*, 2010a). As the substitution sites corresponding to the 12 SNPs in *DNASE1L1* neither suppressed nor created any known restriction enzyme recognition sites, we employed a mismatched PCR-amplification method for genotyping (Yasuda *et al.*, 1995); incorporation of a deliberate mismatch close to the 3'-terminus of a PCR primer allows the creation of a recognition site for each enzyme. After digestion of the amplified DNA fragment for each SNP, the appearance of the expected product, as shown in Supplementary Table 1, derived from the respective alleles in each SNP, allowed us to determine the genotypes easily. To confirm the validity of the genotyping results obtained by these methods, direct sequencing analysis of the genomic DNA region including the substitution site of each SNP in *DNASE1L1* derived from several representative subjects was performed. The genotyping results for each SNP using a PCR-RFLP method were completely congruent with those obtained by direct sequencing, indicating the validity of the novel genotyping methods developed in this study. Furthermore, the technical simplicity and lack of need for any specialized instruments makes the genotyping methods developed in this study advantageous over other methods.

Distribution of the 21 nonsynonymous SNPs in DNASE1L1 for 16 different populations worldwide

The distribution of the genotype and allele frequencies for 18 nonsynonymous SNPs in *DNASE1L1* were determined in 16 different populations, including three ethnic groups. Consequently, in addition to three nonsynonymous SNPs in *DNASE1L1* examined previously in the same populations (Fujihara *et al.*, 2010; Ueki *et al.*, 2010a), these results allowed us to clarify the genetic distribution of all of the nonsynonymous SNPs (total 21 SNPs) in *DNASE1L1* that have already been registered; the allele distributions of the SNPs in our study populations are shown in Supplementary Table S2.

All of the nonsynonymous SNPs in *DNASE1L1*, except for p.Val122Ile, were found to be distributed in a monoallelic manner; only the predominant allele for each SNP was found in all of the subjects in our study populations. In the NCBI dbSNP, no frequency data are available for p.Leu31Val, p.Gly80Glu, p.Gln170His, p.Arg194His, p.Glu199Asp, p.Asp227Ala, p.Val280Ile, and p.Pro300Ser. Since only the predominant allele was observed in each of these 8 SNPs in the 16 different populations, we were able to demonstrate that these SNPs exhibited extremely low genetic diversity with a heterozygosity of 0.000. Therefore, our population data provide valuable information about the genetic heterogeneity of *DNASE1L1*. A minor allele, A-allele, for SNP p.Val122Ile was distributed in German and Turk populations, albeit at a low frequency of 0.047, whereas only the predominant allele, G-allele, was observed in the other study populations, showing that p.Val122Ile is Caucasian-specific.

All of the 21 nonsynonymous SNPs in the gene, except for SNP p.Val122Ile in Caucasian populations, exhibited a monoallelic distribution in all of the 16 different populations examined, including three ethnic groups, allowing us to conclude that with regard to the nonsynonymous SNPs *DNASE1L1* exhibits extremely low genetic heterogeneity. Overall, it could be assumed that for all of the nonsynonymous SNPs in *DNASE1L1*, only two haplotypes classified on SNP p.Val122Ile are distributed in Caucasian populations, whereas only one haplotype composed of the predominant allele in each SNP is distributed in other populations worldwide.

It is especially noteworthy that in six SNPs affecting the DNase activity, p.Thr98Met, p.Gln109Glu, p.Val115Met, p.Arg194His, p.Thr217Ile, and p.Asp227Ala, each of the alleles resulting in reduction/abolishment of the activity was not distributed worldwide, indicating that all of the nonsynonymous SNPs in *DNASE1L1* affecting the catalytic activity showed no genetic heterogeneity in any of the populations examined. These findings indicate that these SNPs in *DNASE1L1* may, in general, exert no influence on the levels of the DNase activity in human populations. Thus, it can be assumed that *DNASE1L1* is generally well conserved, and thereby full levels of the enzymatic activity in the context of these nonsynonymous SNPs are retained in human populations.

Discussion

In the present study, we have continued our ongoing investigation of the genetic and functional characteristics of nonsynonymous SNPs in *DNASE1L1* that may potentially serve as a cell surface barrier to endocytosis-mediated transfer of foreign genes. Although all of the 21 nonsynonymous SNPs in the gene, except for SNP p.Val122Ile in Caucasian populations, exhibited remarkably low genetic diversity in all of the 16 different populations examined, including three ethnic groups, it is worth noting that two activity-abolishing and four activity-reducing SNPs were confirmed to be functional (Fig. 1). To our knowledge, this is the first study to have comprehensively demonstrated the distribution of all of the nonsynonymous SNPs in *DNASE1L1* in different worldwide populations and also the effect of each SNP on the enzyme activity.

It has been demonstrated that, in the human DNase I family, subjects with a minor allele reducing the activity in nonsynonymous SNPs exhibit a lower level of *in vivo* DNase activity than those in subjects without the corresponding allele (Yasuda *et al.*, 1999; Yasutomo *et al.*, 2001; Dittmar *et al.*, 2009). Therefore, it seems plausible that subjects who are hetero/homozygous (female) and hemizygous (male) for each minor allele of p.Thr98Met, p.Gln109Glu, p.Val115Met, p.Arg194His, p.Thr217Ile, and p.Asp227Ala, producing DNase 1L1 variants with substantially low DNase activity or no activity as shown in Table 1, may exhibit lower levels of *in vivo* DNase 1L1 activity than subjects with other genotypes; in particular, a DNase 1L1-deficient phenotype is associated with the minor allele in the two latter SNPs producing a loss-of-function variant. The endogenous DNase I family has been noted as candidate group of endonucleases involved in the breakdown of chromatin during apoptosis and/or necrosis (Counis *et al.*, 2000; Mizuta *et al.*, 2006;

Napirei *et al.*, 2006; Fisher *et al.*, 2007). In this context, triggered by nuclear antigens, the clearance of cell debris resulting from cell death through apoptosis and/or necrosis might be primarily involved in the prevention of autoimmune conditions; the DNase I family has been especially highlighted for its possible involvement in the pathogenesis of autoimmune diseases (Tsukumo *et al.*, 2004; Valle *et al.*, 2008; Hedberg *et al.*, 2011). Therefore, we could postulate that a minor allele of activity-abolishing SNPs producing loss-of-function variants of the DNase I family would represent a direct genetic risk factor for autoimmune diseases (Ueki *et al.*, 2009; Yasuda *et al.*, 2010; Kimura-Kataoka *et al.*, 2013). In contrast, among the members of the DNase I family, no association of DNase 1L1 with diseases potentially attributable to a failure in the breakdown of chromatin during apoptosis and/or necrosis through reduction/abolishment of *in vivo* DNase activity has been reported. Although a 196-bp deletion in the 3'-untranslated region in exon 9 of the DNase 1L1 gene has been shown to be present at a high frequency in patients with Pomp disease (Malferriari *et al.*, 2001), and a partial deletion of the gene has been found in patients with Barth syndrome (Singh *et al.*, 2009), it remains unknown whether reduction/abolishment of *in vivo* DNase 1L1 activity may be involved in the etiology of these diseases. However, Shiokawa *et al.*, (2007) have suggested that suppression of foreign gene transfer by DNase 1L1 on the cell surface is probably important for maintaining healthy cell communities *in vivo*, thereby reducing the risk of malignant transformation. Therefore, it seems plausible that a minor allele of p.Thr217Ile and p.Asp227Ala, producing a loss-of-function DNase variant, might be a genetic risk factor for malignant transformation.

We have been conducting genetic and expression analyses of nonsynonymous SNPs in the DNase I family genes to evaluate whether these SNPs are functional and potentially implicated in various diseases. As a result, we have demonstrated that almost all of the nonsynonymous SNPs in the genes encoding other members of the human DNase I family, DNase I (Yasuda *et al.*, 2010; Ueki *et al.*, 2013b), I-like 2 (Ueki *et al.*, 2013a), and I-like 3 (Ueki *et al.*, 2009; 2013b), together with DNase II (Ueki *et al.*, 2010b; Kimura-Kataoka *et al.*, 2012, 2013), exhibit a monoallelic distribution in the same study populations, similar to those of DNase 1L1. Thus, with regard to nonsynonymous SNPs potentially resulting in alterations of *in vivo* DNase activity, the human DNase I family genes show remarkably low genetic diversity. These facts allow us to conclude that the human DNase family including DNase II has, on the whole, been well conserved at the protein level during the evolution of human populations.

Acknowledgments

DNA samples of Bloodstain samples of the Ovambo and Turkish populations were kindly provided by Dr. B. Brinkmann. Blood samples of Korean and Mongolian populations were kindly provided by Dr. K. Shiwaku. This study was supported in part by Grants-in-Aid from Japan Society for the Promotion of Science (25460864 to M.U., and 22249023 to T.Y.).

Disclosure Statement

No competing financial interests exist.

References

- Counis, M.F., *et al.* (2000). DNases and apoptosis. *Biochem Cell Biol* **78**, 405–414.
- Coy, J.F., *et al.* (1996). Isolation, differential splicing and protein expression of a DNase on the human X chromosome. *Cell Death Differ* **3**, 199–206.
- Dittma, M., *et al.* (2009). A novel mutation in the DNASE1 gene is related with protein instability and decreased activity in thyroid autoimmunity. *J Autoimmun* **32**, 7–13.
- Fisher, H., *et al.* (2007). DNase 1L2 degrades nuclear DNA during corneocyte formation. *J Invest Dermatol* **127**, 24–30.
- Fujihara, J., *et al.* (2010). Global analysis of single nucleotide polymorphisms in the exons of human deoxyribonuclease I-like 1 and 2 genes. *Electrophoresis* **31**, 3552–3557.
- Hedberg, A., *et al.* (2011). Chromatin as a target antigen in human and murine lupus nephritis. *Arthritis Res Therapy* **13**, 214–222.
- Kimura-Kataoka, K., *et al.* (2012). Genetic and expression analysis of SNPs in the human deoxyribonuclease II: SNPs in the promoter region reduce its *in vivo* activity through decreased promoter activity. *Electrophoresis* **33**, 2852–2858.
- Kimura-Kataoka, K., *et al.* (2013). Seven nonsynonymous single nucleotide polymorphisms (SNPs) in the gene encoding human deoxyribonuclease II may serve as a functional SNP potentially implicated in autoimmune dysfunction. *Electrophoresis* (In press).
- Lichtenbelt, K.D., *et al.* (2006). Frequency of the deletion polymorphism of DNASE1L1 in 137 patients with acid maltase deficiency (Pompe disease). *Exp Mol Pathol* **80**, 308–309.
- Los, M., *et al.* (2000). Functional characterization of DNase X, a novel endonuclease expressed in muscle cells. *Biochemistry* **39**, 7365–7373.
- MacLea, K.S., *et al.* (2003). A family history of deoxyribonuclease II: surprise from *Trichinella spiralis* and *Burkholderia pseudomallei*. *Gene* **305**, 1–12.
- Malferrari, G., *et al.* (1999). Molecular characterization of a novel endonuclease (Xib) and possible involvement in lysosomal glycogen storage disorders. *Exp Mol Pathol* **66**, 123–130.
- Malferrari, G., *et al.* (2001). Deletion polymorphism of DNASE1L1, an X-linked DNase i-like gene, in acid maltase deficiency disorders. *Exp Mol Pathol* **70**, 173–174.
- Mizuta, R., *et al.* (2006). Action of apoptotic endonuclease DNase gamma on naked DNA and chromatin substrates. *Biochem Biophys Res Commun* **345**, 560–567.
- Nadano, D., *et al.* (1993). Measurement of deoxyribonuclease I activity in human tissues and body fluids by a single radial enzyme-diffusion method. *Clin Chem* **39**, 448–452.
- Napirei, M., *et al.* (2006). Systemic lupus-erythematosus: deoxyribonuclease I in necrotic chromatin disposal. *Int J Biochem Cell Biol* **38**, 297–306.
- Parrish, J.E., *et al.* (1995). A muscle-specific DNase 1-like gene in human Xq28 region. *Hum Mol Genet* **4**, 1557–1564.
- Pergolizzi, R., *et al.* (1996). Cloning of a gene encoding a DNase I-like endonuclease in the human Xq28 region. *Gene* **168**, 267–270.
- Shiokawa, D., *et al.* (2001). Characterization of human DNase I family endonucleases and activation of DNase γ during apoptosis. *Biochemistry* **40**, 143–152.
- Shiokawa, D., *et al.* (2005). Physical and biochemical properties of mammalian DNase X proteins: non-AUG translation initiation of porcine and bovine mRNAs for DNase X. *Biochem J* **392**, 511–517.
- Shiokawa, D., *et al.* (2007). DNase X is a glycosylphosphatidylinositol-anchored membrane enzyme that provides a barrier to endocytosis-mediated transfer of a foreign gene. *J Biol Chem* **282**, 17132–17140.
- Singh, H.R., *et al.* (2009). A novel Alu-mediated Xq28 microdeletion ablates TAZ and partially deletes DNL1L1 in a patient with Barth syndrome. *Am J Med Genet part A* **149A**, 1082–1085.
- Tsukumo, S., *et al.* (2004). DNase I in pathogenesis of systemic lupus erythematosus. *Clin Immunol* **113**, 14–18.
- Ueki, M., *et al.* (2009). Caucasian-specific allele in non-synonymous single nucleotide polymorphisms of the gene encoding deoxyribonuclease I-like 3, potentially relevant to autoimmunity, produces an inactive enzyme. *Clin Chim Acta* **407**, 20–24.
- Ueki, M., *et al.* (2010a). Genetic and expression analysis of all non-synonymous single nucleotide polymorphisms in the human deoxyribonuclease I-like 1 and 2 genes. *Electrophoresis* **31**, 2063–2069.
- Ueki, M., *et al.* (2010b). Genetic and expression analysis of all 7 non-synonymous single nucleotide polymorphisms in the human deoxyribonuclease II gene, with potential relevance to autoimmunity. *Clin Chim Acta* **411**, 92–98.
- Ueki, M., *et al.* (2013a). Five non-synonymous SNPs in the gene encoding human deoxyribonuclease I-like 2 implicated in terminal differentiation of keratinocytes reduce or abolish its activity. *Electrophoresis* **34**, 456–462.
- Ueki, M., *et al.* (2013b). Evaluation of all non-synonymous single nucleotide polymorphisms (SNPs) in the genes encoding human deoxyribonuclease I and I-like 3 as a functional SNP potentially implicated in autoimmunity. *FEBS J* (In press).
- Valle, F.M., *et al.* (2008). DNase I and systemic lupus erythematosus. *Autoimmun Rev* **7**, 359–363.
- Wu, J.-R., *et al.* (2012). Molecular basis for population variation: From SNPs to SAPs. *FEBS Lett* **586**, 2841–2845.
- Yasuda, T., *et al.* (1995). Genotyping of human deoxyribonuclease I polymorphism by the polymerase chain reaction. *Electrophoresis* **16**, 1889–1893.
- Yasuda, T., *et al.* (1998). Activity measurement for deoxyribonucleases I and II with pictogram sensitivity based on DNA/SYBR Green I fluorescence. *Anal Biochem* **255**, 274–276.
- Yasuda, T., *et al.* (1999). A new allele, DNASE1*6, of human deoxyribonuclease I polymorphism encodes an Arg to Cys substitution responsible for its instability. *Biochem Biophys Res Commun* **260**, 280–283.
- Yasuda, T., *et al.* (2010). A biochemical and genetic study on all non-synonymous single nucleotide polymorphisms of the gene encoding human deoxyribonuclease I potentially relevant to autoimmunity. *Int J Biochem Cell Biol* **42**, 1216–1225.
- Yasutomo, K., *et al.* (2001). Mutation of DNASE1 in people with systemic lupus erythematosus. *Nat Genet* **28**, 313–314.

Address correspondence to:

Haruo Takeshita, MD, PhD

Department of Legal Medicine

Shimane University School of Medicine

Izumo

Shimane 693-8501

Japan

E-mail: htakeshi@med.shimane-u.ac.jp

Received for publication November 1, 2013; accepted November 11, 2013.