ORIGINAL ARTICLE

Functional single nucleotide polymorphisms (SNPs) in the genes encoding the human deoxyribonuclease (DNase) family potentially relevant to autoimmunity

Junko Fujihara[1*], Misuzu Ueki[2*], Kaori Kimura-Kataoka[1], Reiko Iida[3], Haruo Takeshita[1], Toshihiro Yasuda [2]

[1] Department of Legal Medicine, Shimane University School of Medicine, Shimane, Japan
[2] Division of Medical Genetics and Biochemistry, [3]Division of Life Sciences, Faculty of Medical Sciences, University of Fukui, Fukui, Japan

[*]Contributed equally to this study.
E-mail: jfujihar@med.shimane-u.ac.jp

Keywords; Deoxyribonuclease (DNase) family; Functional SNPs; Genotype; Genetic distribution; Loss-of-function; Autoimmunity

ABSTRACT

Objective: **To continue our previous investigations, we have extensively investigated the function of the the 61, 41, and 35 non-synonymous single nucleotide polymorphisms (SNPs) in the human genes encoding *DNASE1*, *DNASE1L3*, and *DNASE2*, respectively, potentially relevant to autoimmune diseases.**

Methods: **The site-directed mutagenesis was employed to amino acid–substituted constructs corresponding to each SNP. The COS-7 cells were transfected with each vector and DNase activity was assayed by the single radial enzyme diffusion method. By using PolyPhen-2, changes in the DNase function of each non-synonymous SNP were predicted.** Genotyping of all the non-synonymous SNPs was performed in 3 ethnic groups including 14 populations using the polymerase chain reaction restriction fragment length polymorphism method.

Results: Expression analysis demonstrated these SNPs to be classified into 4 categories with regard to the effect on DNase activity: SNPs not affecting the activity level, ones reducing it, ones abolishing it, and ones elevating it. In particular, 9, 5, and 4 SNPs producing a loss-of-function variant of the enzymes in *DNASE1*, *DNASE1L3*, and *DNASE2*, respectively, were confirmed. SNPs producing DNase loss of function can be estimated by PolyPhen-2 to be “probably damaging” with a high accuracy of prediction. Almost all of these functional SNPs producing a loss of function or substantially low activity-harboring forms exhibited a mono-allelic distribution in all of the populations.

Conclusion: A minor allele of functional SNPs, despite the remarkably low genetic heterogeneity of the SNPs, might be a genetic risk factor for autoimmune diseases.

INTRODUCTION

Deoxyribonuclease I (DNase I; EC 3.1.21.1) and DNase I-like 3 (DNase 1L3) break down chromatin during apoptosis and/or necrosis (Counis et al., 2000; Napirei et al., 2006; Mizuta et al., 2006). The deoxyribonuclease II (DNase II; EC 3.1.21.1) present in lysosomes has been postulated to be involved in the degradation of endogenous DNA in apoptotic cells engulfed by macrophages (Odaka et al., 1999; Fischer et al., 2011) and exogenous DNA on the skin surface in mammalian cells (Mcllroy et al., 2000).

Triggered by nuclear antigens, the DNase-mediated clearance of cell debris resulting from cell death through apoptosis and/or necrosis might be primarily involved in the prevention of autoimmune conditions such as systemic lupus erythematosus (SLE) (Valle et al., 2008; Hedberg et al., 2011). Moreover, recent studies have shown that extracellular DNase could serve as an endogenous regulator of neutrophil extracellular traps and that DNase could improve the flow by DNase-dependent thrombolysis, preventing downstream injury by neutrophil extracellular traps (NETs) (Mangold et al., 2015).

Yasutomo et al. (2001) and Dittmar et al. (2007) have identified novel nonsense (p.Lys5Ter) and missense (p.Val111Met) mutations, respectively, that abolished and reduced DNase I activity in autoimmune-disease patients (Yasuda et al., 2010). Previous findings have indicated that the frequency of the homozygote for the *G* allele in single nucleotide polymorphism (SNP) p.Arg244Gln of *DNASE1* is much higher in patients with SLE who have the corresponding autoantibodies than in patients who do not have them (Shin et al., 2004). Recently, Al-Mayouf et al. (2011) found a null mutation resulting from a homozygous 1-bp deletion (c.643delT; p.Trp215GlyfsX2) and also identified homozygosity for a missense mutation (p.Arg206Cys) in an SLE patient, the latter producing a loss-of-function variant of DNase 1L3 Ueki et al., 2009). **Moreover, recent study has shown that the effect of the *HumDN1* VNTR polymorphism on *DNASE1* mRNA expression: increased *DNASE1* expression in SLE with 5/5, 3/4, and 3/5 genotypes and rheumatoid arthritis patients with a 3/4 genotype ($p = 0.02$) showed a significant increase in DNASE1 expression in Kuwaiti subjects (AlFadhli et al., 2014).**

It has also been reported that DNase I–deficient mice develop an SLE-like syndrome (Napirei et al., 2000) and that lupus-prone MRL and NZB/W F1 mice have impaired DNase 1L3 activity (Wilber et al., 2003).

On the other hand, Shin et al. reported that SNPs rs11085823, rs2293682, and rs4804209 and their haplotypes in *DNASE2* were associated with renal disorders in Korean patients with SLE (Shin et al., 2005). Moreover, it has been claimed that homozygous Caucasian carriers of the rs12609744 G allele, the rs11085823 G allele, and the rs7249143 A allele in *DNASE2* have an increased risk of rheumatoid arthritis (Rossol et al., 2008). These findings suggest that mutation and/or SNP in the genes encoding *DNASEI, DNASE1L3,* and *DNASE2,* resulting in variants that are inactive or have low activity, might be substantially responsible for the genetic background determining susceptibility to autoimmune diseases. Therefore, to clarify the genetic basis of the etiological role of endonucleases in autoimmune diseases, simultaneous evaluation of functional SNPs in *DNASEI, DNASE1L3,* and *DNASE2* is warranted.

Many SNPs in *DNASEI, DNASE1L3,* and *DNASEII* have been screened and are available on the NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP; January 2015) (Fig 1). We have previously investigated the effects of the non-synonymous SNPs in *DNASEI, DNASE1L3,* and *DNASE2* on catalytic activity, thereby evaluating the functionality of each SNP (Ueki et al., 2010, 2011, 2014a, 2014b; Kimura-Kataoka et al., 2013, 2014). Recently, many non-synonymous SNPs in *DNASEI, DNASE1L3,* and *DNASEII* were newly registered **(**11, 16, and 15 SNPs in DNases I, 1L3, and II, respectively**).** Comprehensive data on the biochemical-genetic aspects of these SNPs in *DNASE1, DNASE1L3,* and *DNASE2* potentially affecting their activity would be useful for clarifying their function in determining the genetically predisposed risks of autoimmune and other diseases.

In the present study, we developed genotyping techniques for all of these SNPs, investigated the distribution of each SNP in 14 populations, and determined the effect of amino acid substitution in the protein resulting from the SNPs on DNase activity. Furthermore, by using PolyPhen-2

(Polymorphism Phenotyping v2), the possible impact of the amino acid substitution resulting from each non-synonymous SNP on the function of the enzymes was predicted.

MATERIALS AND METHODS

DNA samples

Genomic DNA was extracted using a QIAamp DNA Mini Kit (Qiagen, Chatsworth, CA, USA) from blood or bloodstain samples randomly collected from healthy subjects (n = 1,752) derived from 14 populations from 3 ethnic groups. The Asian subjects included 110 Japanese (Shimane Prefecture, Japan), 352 Koreans (Busan, South Korea), 193 South Chinese (Shenyang and Guangzhou, China), 112 Mongolians (Ulaanbaatar, Mongolia), 153 Tibetans (Katmandu, Nepal), 35 Sri Lankan Tamils (Kandy, Sri Lanka), 48 Sri Lankan Sinhalese (Kandy, Sri Lanka), and 40 Tamangs (Kotyang, Nepal). The Caucasian subjects included 136 Turks (Adana area, Southern Turkey), 68 Germans (Munich, Germany) and 199 Mexicans (60 Mestizo, 88 Nahuas, and 51 Huicholes). The African subjects included 126 Ovambos (Bantusin, Namibia), 105 Ghanaians (Accra, Ghana), and 75 Xhosas (Cape Town, South Africa). The samples were the same as those used for previous studies (Ueki et al., 2009, 2010, 2011, 2014a, 2014b; Kimura-Kataoka et al., 2013, 2014). Written informed consent was obtained from each participant. The study was approved by the Human Ethical Committees of the Shimane University (approval number 1024 for the Human Genome and Genetics Analysis Study).

Construction of expression vectors encoding human DNases I, 1L3, and II and the amino acid–substituted forms corresponding to each non-synonymous SNP

Each SNP nomenclature was presented according to the recommendations for describing sequence variants (http://www.hgvs.org/mutnomen/examplesDNA.htlm); the sequences of DNases I, 1L3, and II (GenBank accession no. AB188151, NCBI Reference Sequence NM_004944.3, and GenBank

accession no. AB004574, respectively) have been used as the coding DNA Reference Sequence. Nucleotide and amino acid residues were numbered from the 5'-terminus of the translation initiation codon and the N-terminal amino acid residue of the precursor protein, respectively.

Each wild-type expression vector ligated into pcDNA3.1(+) (Invitrogen, San Diego, CA, USA) containing the entire coding sequence of each cDNA was constructed according to the methods described previously (Ueki et al., 2010, 2011, 2014a, 2014b; Kimura-Kataoka et al., 2013, 2014); since sequencing of the constructs showed the inserted DNA fragment to be derived from the predominant haplotype for the SNPs, the construct was used as a wild-type one. The site-directed mutagenesis using a KOD-Plus Mutagenesis Kit (Toyobo Co., Ltd., Osaka, Japan) with the wild-type construct as a template was employed to newly prepare 11, 16, and 15 amino acid–substituted constructs corresponding to each SNP of DNases I, 1L3, and II, respectively, in the same manner as that for other amino acid–substituted constructs in previous studies (Ueki et al., 2010, 2011, 2014a, 2014b; Kimura-Kataoka et al., 2013, 2014). In these constructs, the amino acid residue was replaced by the counterpart derived from a minor allele in each substitution site; e.g., the L38F construct of DNase II, in which the Leu residue at position 38 in the protein is replaced by Phe derived from the minor allele, corresponds to SNP p.Leu38Phe (rs373200195). Nucleotide sequences of all the constructs were confirmed by DNA sequence analysis. Two clones derived from each construct used for transfection were purified using the Plasmid Midi Kit (Qiagen).

Transient expression of the expression vectors and assay for DNase activity

COS-7 cells were maintained in Dulbecco's modified Eagle medium containing 1 mM L-glutamine, 50 U/ml penicillin, and 10% (v/v) fetal calf serum at 37ºC under 5% CO_2 in air. The cells were transiently transfected 4 times with 2 μg of each DNase I– or DNase II–related expression vector and 600 ng of pSV-β-galactosidase vector (Promega, Madison, WI, USA; 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. On the other hand, at 24 h after transfection of 2 μg of each DNase 1L3–related expression vector and pSV-β-galactosidase vector in the same manner as above, the cells were washed with phosphate-buffered saline and transferred to Dulbecco's modified Eagle medium containing 100-fold-diluted Insulin–Transferrin–Selenium-X (Invitrogen). The transfected cells were harvested after 72 h. Then the harvested cells were subjected to sonication using Bioruptor UCD-250 (COSMO BIO Co., Ltd., Tokyo, Japan) to prepare lysates for the subsequent assay. Each of the DNase activities in the transfected cells was assayed by the single radial enzyme diffusion (SRED) method using a LAS-3000 imaging analyzer (Fuji Film, Tokyo, Japan) as described previously (Ueki et al., 2010; 2011; 2014a; 2014b; Kimura-Kataoka et al., 2013; 2014). Transfection efficiencies were estimated by cotransfection with a pSV-β-galactosidase vector and subsequently assaying aliquots of cell lysates for β-D-galactosidase activity. After the DNase activity determined by the SRED method in each of the transfections was separately normalized by the activity of β-D-galactosidase determined in the same transfection, the mean activity of the amino acid–substituted form deviation derived from 4 transfections using 2 clones derived from each construct was expressed relative to that of the wild type; the relative activity is expressed as mean ± standard deviation.

Genotyping of each SNP in the DNase I, 1L3, and II genes by means of the PCR-RFLP method

In the present study, in order to determine the genotype of each SNP in the genes, a polymerase chain reaction (PCR) followed by a restriction fragment length polymorphism (RFLP) method was newly employed. For the corresponding substitution sites that neither suppressed nor created any known restriction enzyme recognition sites among these SNPs, we applied a mismatched PCR-amplification method (Yasuda et al., 1995) to genotyping. Incorporation of a deliberate mismatch close to the 3'-terminus of a PCR primer allows the creation of a recognition site for each enzyme. The design of the PCR primers used for the genotyping was based on the nucleotide sequences of the

human DNase I, 1L3, and II genes (GenBank accession no. D83195, EMBL accession no. AC137936, and NCBI Reference Sequence NC 000019.9, respectively). The sequences of the primer, annealing temperature, and restriction enzyme used for the analysis of each SNP are summarized in Supplementary Tables 1–3.

PCR amplification was performed in a 25-μl reaction mixture using approximately 5 ng of DNA. The reaction mixture contained a 1×buffer (15 mM Tris-HCl, pH 8.0, 50 mM KCl), 1.5 mM $MgCl_2$, 0.5 μM of each primer, 200 μM dNTPs, and 1.25 U of Taq polymerase (AmpliTaq Gold; Applied Biosystems, Foster City, CA, USA). PCR was performed with a protocol consisting of initial denaturation at 94ºC for 7 min, followed by 30 cycles with denaturation at 94ºC for 30 s, annealing at the temperature indicated in Supplementary Tables 1–3 for 30 s, and extension at 72ºC for 1 min, followed by a final extension at 72ºC for 7 min. Next, 2 μl of the PCR product obtained using a pair of primers for each SNP was digested with 5 U of each restriction enzyme (New England Biolabs, Ipswich, MA, USA) at 37ºC for 3 h in a final reaction mixture volume of 15 μl. The digested products (5 μl) were separated in an 8% polyacrylamide gel, and the patterns on the gels were visualized by silver staining, as described previously (Yasuda et al., 1995). The appearance of the expected product, as shown in Supplementary Tables 1–3, derived from the respective alleles in each SNP, allowed us to determine the genotypes easily; in SNP p.Leu38Phe, the amplified product derived from the *G* allele was completely digested with *MboI* to yield an 84-bp fragment, whereas that from the *A* allele did not yield such a fragment. The same procedure was employed for the other SNPs.

The validity of the genotyping results obtained by these methods was confirmed by a sequencing analysis of genomic DNAs derived from several representative subjects (n = 5). Genomic sequences, including the substitution site of each SNP, were determined by the dideoxy chain-terminating method with the BigDye® Terminator Cycle Sequencing Kit (Applied Biosystems) and the sequencing run on a Genetic Analyzer 310 (Applies Biosystems).

Other analytical methods

In order to predict a possible change in the DNase function through the amino acid substitution resulting from each non-synonymous SNP, PolyPhen-2 (Polymorphism Phenotyping 2) (Adzhubei et al., 2010) was used (http://genetics.bwh.harvard.edu/pph2/index.shtml). If the score is less than 0.5, 0.5, or more than 0.5, the predictions are benign, possibly damaging, or probably damaging, respectively.

Orthologs of human DNases I, 1L3, and II were surveyed using the Kyoto Encyclopedia of Genes and Genomes database (http://www.genome.jp/kegg/), in which the amino acid sequences of human counterparts were used as query sequences for Basic Local Alignment Search Tool (BLAST) searches of each genome database. Then multiple alignment analyses of the amino acid sequences of animal DNases were performed using DNASIS Pro v3.0 (Hitachi Solutions, Ltd., Tokyo, Japan).

The activity of each amino acid–substituted construct was compared with that of the wild type by means of the unpaired Student's *t*-test.

RESULTS AND DISCUSSION

Effect of the corresponding amino acid substitution derived from each of the non-synonymous SNPs on DNase activity

We have examined the effect of the amino acid substitution resulting from each of the non-synonymous SNPs in the genes encoding human DNases I, 1L3, and II on DNase activity. Expression vectors containing the entire coding sequence of respective cDNA derived from the predominant haplotype among these SNPs were prepared and used as the wild type; 11, 16, and 15 substituted vectors expressing the amino acid–substituted DNase protein encoded by the minor allele in each SNP of *DNASE1*, *DNASE1L3*, and *DNASE2*, respectively, were newly constructed and transiently expressed in COS-7 cells, and the resulting DNase activity in the transfected cells was determined by the SRED method. Consequently, as shown in Supplementary Tables 4–6, in addition

to the previous findings (Ueki et al., 2009, 2010, 2011, 2014a, 2014b; Kimura-Kataoka et al., 2013, 2014), compared with the activity level of the wild-type enzyme, we could reveal the effect of each amino acid substitution corresponding to total 61, 41, and 32 non-synonymous SNPs in *DNASE1*, *DNASE1L3*, and *DNASE2*, respectively, on each DNase activity. With regard to the effect on activity, these SNPs were classified into 4 categories: SNPs not affecting the activity level, ones reducing it, ones abolishing it, and ones elevating it (Table 1). It is noteworthy that, since each activity originating from the 11, 6, and 7 amino acid–substituted forms of *DNASE1*, *DNASE1L3*, and *DNASE2*, respectively, such as R244Term in *DNASE1*, R51C in *DNASE1L3*, and R314W in *DNASE2*, could not be detected under our assay conditions, the minor allele in the corresponding SNPs could be defined as a loss-of-function variant resulting in the disappearance of DNase activity. Thus, all of the non-synonymous SNPs as functional SNPs producing a loss of function or substantially low activity-harboring forms identified in *DNASE1*, *DNASE1L3*, and *DNASE2* were confirmed (Fig. 1).

Evaluation of the prediction of the functional effect of non-synonymous SNPs in DNASE1, DNASE1L3, and DNASE2 using PolyPhen-2

By using PolyPhen-2, possible changes in the DNase function through the amino acid substitution resulting from each non-synonymous SNP were predicted (Table 2). Based on the PolyPhen-2 scores, all of the abolishing SNPs could be predicted to be probably damaging except for the SNP rs562507670 (p.Thr307Asn) predicted to be possibly damaging in *DNASE2*; no abolishing SNPs were found in any of the SNPs that were predicted to be benign and/or possibly damaging. The amino acid substitutions resulting from the SNPs assumed to be benign and/or possibly damaging, especially in *DNASE1* and *DNASE1L3*, generally induced no alteration in the DNase activities. In contrast, many of the probably damaging SNPs were classified into SNPs reducing or abolishing the activity levels. These findings permitted us to demonstrate that at least the SNPs producing DNase loss-of-function can be estimated by PolyPhen-2 scores corresponding to “probably damaging” with

a high accuracy of prediction. PolyPhen-2 (Adzhbei et al., 2010) is a useful tool, available as software and via a Web server, for prediction of the possible impact of an amino acid substitution on the function of a human protein, being indispensable for the interpretation of genetic variants resulting from non-synonymous SNPs, especially such as the identification of functional SNPs. This study could provide evaluation on the accuracy of prediction of the functional impact on DNases attributed to an amino acid substitution corresponding to the non-synonymous SNPs using PolyPhen-2. Thus, the combination of the analysis of genetic variants of the genes encoding human DNases I, 1L2, and II corresponding to non-synonymous SNPs, together with an in-silico analysis to predict their possible functional impact, can facilitate their effects on DNase activity levels.

Genotype distribution of non-synonymous SNPs in the human genes encoding DNases I, 1L3, and II for 14 populations worldwide

A simple and novel genotyping procedure was developed using PCR-RFLP for all of the 11, 16, and 15 SNPs in *DNASE1*, *DNASE1L3*, and *DNASE2* (Supplementary Tables 1–3), respectively; those for the other SNPs have been developed previously (Ueki et al., 2009, 2010, 2011, 2014a, 2014b; Kimura-Kataoka et al., 2013, 2014). The distribution of the genotype and allele frequencies for non-synonymous SNPs in *DNASE1*, *DNASE1L3*, and *DNASE2* were determined in 14 populations, including 3 ethnic groups. Consequently, in addition to the other non-synonymous SNPs in these DNase genes examined previously in the same populations (Ueki et al., 2009, 2010, 2011, 2014a, 2014b; Kimura-Kataoka et al., 2013, 2014), the genetic distribution of all of the non-synonymous SNPs in these DNase genes that have already been registered in the database were clarified. Therefore, our population data provide valuable information regarding the genetic heterogeneity of *DNASE1*, *DNASE1L3*, and *DNASE2*.

All of the non-synonymous SNPs as functional SNPs producing a loss of function or substantially low activity-harboring forms identified in *DNASE1*, *DNASE1L3*, and *DNASE2* are summarized in

Table 3. Almost all of these SNPs were mono-allelic in all of the populations examined in this study; these findings allowed us to estimate that the minor allele frequency (MAF) for each SNP would be less than 0.0003. In *DNASE1*, the minor allele of only 2 SNPs (rs121912990 and rs148373909) have been found in the Japanese population (Yasuda et al., 1999); however, the corresponding alleles were not distributed in populations in this study. Furthermore, in *DNASE1L3*, we have previously revealed that only rs35677470 is polymorphic in Caucasians (Ueki et al., 2009). Therefore, with regard to the non-synonymous SNPs resulting in alterations of in vivo DNase activity, the human DNase genes showed remarkably low genetic diversity. These facts indicate that the human genes encoding DNases I, 1L3, and II have been well conserved at the activity level during the evolution of human populations, thereby avoiding any marked reduction of the enzyme activity in human populations.

It has been reported that the levels of serum DNase I activity are reduced in patients with SLE (Chitrabamrung et al., 1981; Dittmar et al., 2007; Sallai et al., 2005; Martinea-Valle et al., 2009). DNase 1L3 is present in serum, in addition to DNase I, and it was assumed that these DNases may be concerned with each other during DNA degradation, providing effective clearance after exposure or release from dying cells (Napirei et al., 2005, 2006, 2009). Furthermore, Hakkim et al. (2010) demonstrated that insufficient NET degradation by DNase I would allow NETs to persist and thus foster the presentation of chromatin-associated self-antigens, a process that may promote SLE. Since DNase II is directly implicated in engulfment-mediated DNA degradation, the failure of efficient DNA clearance due to the loss/reduction of enzyme function would result in autoimmune dysfunction (Kreiser et al., 2002; Kawane et al., 2003). These findings suggest that a remarkable decrease in the activity levels of these DNases might be related to disease pathogenesis. Therefore, functional SNPs producing genetic variants with remarkably reduced and/or abolished activity in the genes encoding these DNases should be considered as notable genetic factors involved in the abolishment or reduction of in vivo activity of the DNases, thereby leading to autoimmune dysfunction. However, there is little information about functional SNPs in the genes encoding these

DNases that might affect enzyme activity. In the present study, many of the functional SNPs affecting the activity of each DNase could be identified, irrespective of whether they showed polymorphic distribution worldwide; in particular, 9, 5, and 4 SNPs producing a loss-of-function variant of the enzymes in *DNASE1*, *DNASE1L3*, and *DNASE2*, respectively, possibly serving as a genetic risk factor for autoimmune diseases, were confirmed.

In order to clarify any clinical association of these functional SNPs in *DNASE1*, *DNASE1L3*, and *DNASE2* with the incidence of autoimmune diseases, the distribution of each SNP in various patient groups should be examined. Therefore, further studies will be required to examine the correlation between the levels of DNase activity in serum and the prevalence of autoimmune diseases for each SNP genotype.

DECLARATION OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported in part by JSPS KAKENHI Grant Number 26713025(Grants-in-Aid for Young Scientists A) to J.F and 25460864 to M.U.

REFERENCES

Adzhubei IA, Schmidt S, Peshkin L, et al. 2010. A method and server for predicting damaging missense mutations.Nat Methods, 7: 248-49.

AlFadhli S, Ghanem AA. 2014. Influence of *HumDN1* VNTR polymorphism on *DNASE1* expression in systemic lupus erythematosus and rheumatoid arthritis. Immunol Invest. 43:411-23.

Al-Mayouf SM, Sunker A, Adbwani R, et al. 2011. Loss-of-function variant in DNASE1L3 causes a familial form of systemic lupus erythematosus. Nat Genet 43, 1186–8.

Chitrabamrung S, Ribin RL, Tan EM. 1981.Serum deoxyribonuclease I and clinical activity in systemic lupus erythematosus. Rheumatol Int 1: 55–60.

Counis MF, Torriglia A, 2000. DNases and apoptosis. Biochem Cell Biol 78: 405–14.

Dittmar M, Poppe R, Bischofs C, et al., 2007. Impaired deoxyribonuclease I activity in monoglandular and polyglandular autoimmunity. Exp Clin Endocrinol Diabetes 115: 387–91.

Dittmar M, Bischofs Matheis N., et al., 2009..A novel mutation in the DNASE1 gene is related with protein instability and decreased activity in thyroid autoimmunity. J Antoimmun 32: 7–13.

Fischer H, Schez J, Szabo S, et al. 2011. DNase 2 is the main DNA-degrading enzyme of the stratum corneum. PLoS ONE 6: e17581.

Hakkim, A., Fürnrohr, B.G., Amann. K., Laube, B., Abed. U.A., Brinkmann, V., Herrmann, M., E. Voll, R., & Zychlinsky A. (2010) Impairment of neutrophil extracellular trap degradation is associated with lupus nephritis. *Proc Natl Acad Sci. USA* **107**, 9813–9818.

Hedberg A, Mortensen ES, Rekvig OP. 2011. Chromatin as a target antigen in human and murine lupus nephritis. Arthritis Res Therapy 13: 214–22.

Kawane K, Fukuyama H, Yoshida H, et al, 2003. Impaired thymic development in mouse embryos deficient in apoptotic DNA degradation.Nat Immunol 4: 138–44.

Kimura-Kataoka K, Ueki M, Takeshita H, et al., 2013.Seven nonsynonymous SNPs in the gene encoding human deoxyribonuclease II may serve as a functional SNP potentially implicated in

autoimmune dysfunction. Electrophoresis 34:3361-9.

Kimura-Kataoka K, Ueki M, Takeshita H., et al, 2014. Identification of the functional alleles of the nonsynonymous single-nucleotide polymorphisms potentially implicated in systemic lupus erythematosus in the human deoxyribonuclease I gene. DNA Cell Biol. 33: 492-502.

Kreiser R J, MacLea KS, Longnecker DS, et al. 2002. Deoxyribonuclease IIalpha is required during the phagocytic phase of apoptosis and its loss causes perinatal lethality.Cell Death Differ. 9: 956–62.

Mangold A , Alias S, Scherz T, et al., 2015.Coronary Neutrophil Extracellular Trap Burden and DNase Activity in ST-Elevation. Acute Coronary Syndrome are Predictors of ST-Segment Resolution and Infarct Size. Circ Res. 116: 1182-92.

Martinea-Valle F, Balada E, Ordi-Ros J, et al., 2009.DNase I activity in patients with systemic lupus erythematosus: relationship with epidemiological, clinical, immunological and therapeutical features. Lupus 18: 418–23.

Mcllroy D, Tanaka M, Sakahira H. et al., 2000 An auxiliary mode of apoptotic DNA fragmentation provided by phagocytes.Genes Dev 14: 549–58.

Mizuta R, Mizuta M, Araki S, et al., 2006. Action of apoptotic endonuclease DNase gamma on naked DNA and chromatin substrates. Biochem Biophys Res Commun 345: 560–7.

Napirei M, Karsunky M, Zevnik B, et al., 2000.Feature of systemic lupus erythematosus in DNase 1-deficient mice. Nat Genet 2:177–81.

Napirei M, Gültekin A, Kloeckl T, et al., 2006. Systemic lupus-erythematosus: deoxyribonuclease I in necrotic chromatin disposal. Int J Biochem Cell Biol 38: 297–306.

Napirei M, Ludwig S, Mezrhab J, et al., 2009.Murine serum nucleases – contrasting effects of plasmin and heparin on the activities of DNase 1 and DNase 1-like 3 (DNase 1 l3). FEBS J 276: 1059–73.

Napirei M, Wulf S, Eulitz D, et al., 2005. Comparative characterization of rat deoxyribonuclease 1 (Dnase1) and murine deoxyribonuclease 1-like 3 (DNase1 l3). Biochem J 389: 355–64.

Odaka C, Mizuochi T. 2003. Murine T cells expressing high activity of prolyl endopeptidase are susceptible to activation-induced cell death. J Immunol 163: 5346–52.

Rossol M, Pierer M, Arnold S, et al. 2008. Homozygosity for DNASE2 single nucleotide polymorphisms in the 5'-regulatory region is associated with rheumatoid arthritis. Ann Rheum Dis 68: 1498–503.

Sallai K, Nagy E, Derfalvy B, et al. 2005. Antinucleosome antibodies and decreased deoxyribonuclease I activity in sera of patients with systemic lupus erythematosus. Clin Diagn Lab Immunol 12: 56–9.

Shin HD, Park BL, Kim LH, et al. 2004. Common DNase I polymorphism associated autoantibody production among systemic lupus erythermatosus. Hum Mol Genet 13: 2343–50.

Shin HD, Park BL, Cheong H S, et al., 2005: A single nucleotide polymorphism in CAPN1 associated with marbling score in Korean cattle. J Hum Genet 50: 107–11.

Ueki M., Takeshita H, Fujihara J, 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: 20–4.

Ueki M, Takeshita H, Fujihara J, et al, 2010. 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: 92-8.

Ueki M, Fujihara J, Takeshita H, et al, 2011.Global genetic analysis of all single nucleotide polymorphisms in exons of the human deoxyribonuclease I-like 3 gene and their effect on its catalytic activity. Electrophoresis. 32:1465-72.

Ueki M, Kimura-Kataoka K, Takeshita H, et al, 2014a.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. 28: 1376-90.

Ueki M, Kimura-Kataoka K, Fujihara J, et al., 2014b. 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. DNA Cell Biol 33:79-87.

Valle FM, Balada E, Ordi-Ros J, et al. 2008. DNase I and systemic lupus erythematosus. Autoimmun Rev 7: 359–63.

Wilber A, O'Connor TP, Lu ML, et al., 2003.DNase1 l3 deficiency in lupus-prone MRL and NZB/W F1 mice. Clin Exp Immunol 134: 46–52.

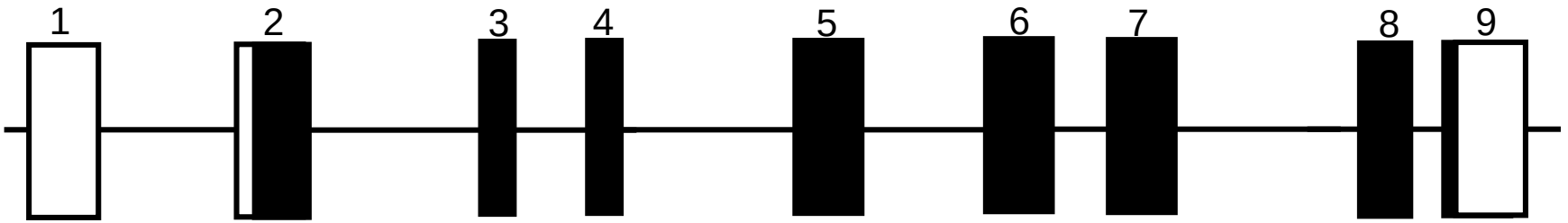
Yasuda T, Nadano D, Tenjo E, et al. 1995. Genotyping of human deoxyribonuclease I polymorphism by the polymerase chain reaction. Electrophoresis. 16:1889-93.

Yasuda, T., Takeshita, H., Iida ,R., Kogure, S., & Kishi, K. (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, Ueki M, Takeshita H 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 Cel Biol 42: 1216–25.

Yasutomo K, Horiuchi T, Kagami H, et al. 2001. Mutation of DNASE1 in people with systemic lupus erythematosus. Nat Genet; 28:313–4.

(A) ***DNASE1***



rs121912990
p.Lys5Ter

rs143865851*
p.Arg53Cys
rs142318540
p.Gln60Arg

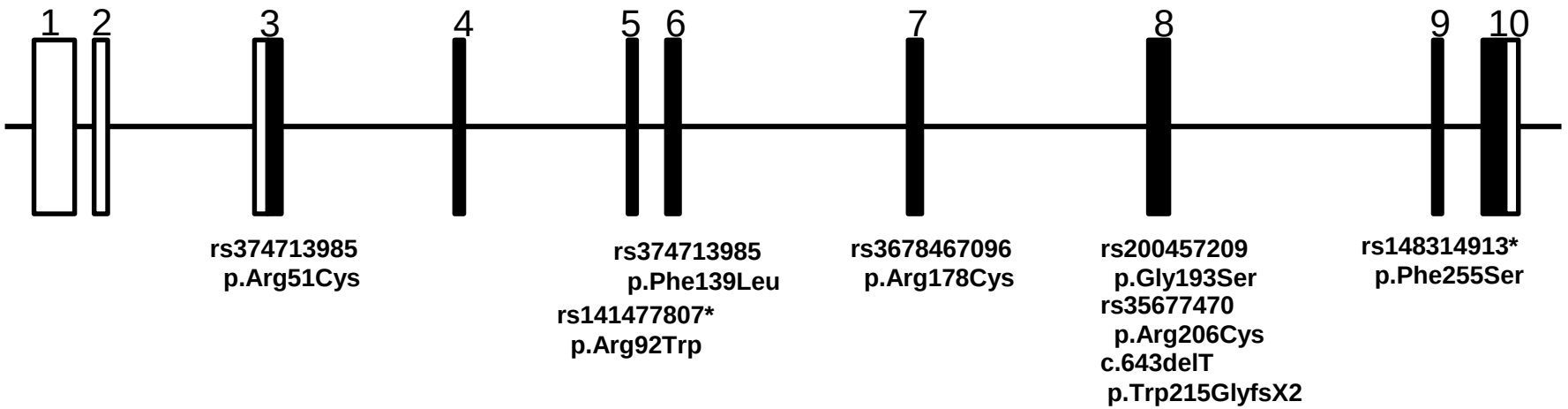
rs8176928
p.Arg107Gly

rs150621329
p.Arg133Gln
rs150621329
p.Arg133Leu

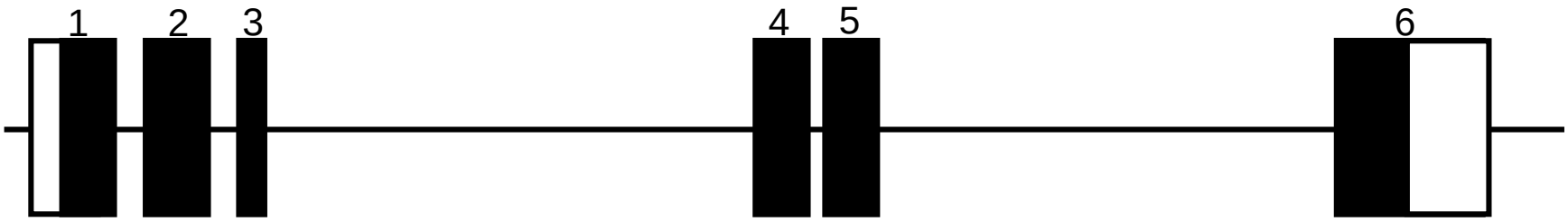
rs142644209
p.Asp190His
rs146236198
p.Asn192Ile
rs148373909*
p.Arg207Cys
rs8176940
p.Cys231Tyr
rs199826318
p.Tyr233Cys
rs139254891
p.Arg235Trp

rs201571412
p.Arg244Ter
rs8176939*
p.Ala246Pro

(B) ***DNASE1L3***


1
2
3
4
5
6
7
8
9
10
rs374713985
p.Arg51Cys
rs374713985
p.Phe139Leu
rs141477807*
p.Arg92Trp
rs3678467096
p.Arg178Cys
rs200457209
p.Gly193Ser
rs35677470
p.Arg206Cys
c.643delT
p.Trp215GlyfsX2
rs148314913*
p.Phe255Ser

(C) ***DNASE2***



rs562700068
p.Ala16Asp

rs2010309538
p.Tyr95Set

rs557707591*
p.Gly48Glu

rs34391539
p.Ala74ProfsV101

rs554251996*
p.Phe234Cys

rs35851337* **rs1061192**
p.Val300Met **p.Arg314Leu**

rs562507670 **rs1802335**
p.Thr307Asn **p.Gln338ter**

rs544319961 **rs35851337***
p.Arg314Trp **p.Val341Leu**

Table 1.

Classification of non-synonymous SNPs in the human genes encoding DNase family with regard to effect on the activity

DNase family	non-synonymous SNP (n)						Reference
	total	not affecting the activity	reducing the activity	abolishing the activity	elevating the activity	showing a genetic heterogeneity	
DNase I	61	27	14	9	11	9	this study
DNase 1L1	21	14	4	2	1	0	Ueki et al.
DNase 1L2	35	10	15	9	1	0	Ueki et al.
DNase 1L3	41	26	7	6	1	1	this study
DNase II	32	11	14	7	0	1	this study

Table 2

Evaluation on prediction of functional effect of non-synonymous SNPs in *DNASE1*, *DNASE1L3* and *DNASE2* using PolyPhen-2

	Effect on the activity[a]	Total SNPs	Prediction[b] using PolyPhen-2		
			Benign	Possibly damaging	Probably damaging
DNASE1	No effect[c]	38	19	10	9
	Abolishing	9	0	0	9
	Reducing	14	3	0	11
DNASE1L3	No effect[c]	27	14	5	8
	Abolishing	5	0	0	5
	Reducing	7	0	2	5
DNASE2	No effect	10	8	1	1
	Abolishing	4	0	1	3
	Reducing	14	7	1	6

[a]Effect of the amino acid substitution corresponding to each non-synonymous SNP was separately determined as shown in the text.

[b]Total numbers of SNPs categorized into three types predicted by PolyPhen-2 were presented.

[c]SNPs elevation the activity were included.

Evaluation on prediction of functional effect of non-synonymous SNPs in *DNASE1*, *DNASE1L3* and *DNASE2* using PolyPhen-2

	Prediction using PolyPhen-1[a]	Total SNPs	Effect on the activity[b]		
			No effect[c]	Abolishing	Reducing
DNASE1	Benign	22	19	0	3
	Possibly damaging	10	10	0	0
	Probably damaging	29	9	9	11
DNASE1L3	Benign	14	14	1	0
	Possibly damaging	7	5	0	2
	Probably damaging	13	8	5	5
DNASE2	Benign	15	8	0	7
	Possibly damaging	3	1	1	1
	Probably damaging	10	1	3	6

[a]Total numbers of SNPs categorized into three types predicted by PolyPhen-2 were presented.

[b]Effect of the amino acid substitution corresponding to each non-synonymous SNP was separately determined as shown in the text.

[c]SNPs elevation the activity were included.

Table 3

All the non-synonymous SNPs as functional SNPs producing a loss-of-function or substantially low activity-harboring forms identified in *DNASE1*, *DNASE1L3* and *DNASE2*; genetic heterogeneity, MAF (mutant allele frequency), heterozygosity, effect of the activity, Polyphen-2 score and conservation of the corresponding amino acid residue in animal DNases

SNP	Genetic heterogeneity[a)]	MAF[b)]	Heterozygosity[b)]	Activity[c)]	Prediction Polyphen-2 (score)[d)]	Conservation of the amino acid resides corresponding to SNP in animal DNases[e)]
			DNASE1			
rs121912990 p.Lys5Ter; c.13A>T	Japanese[f)]	– (<0.0003)	– (0.000)	n.d.		nonsense substitution in the signal sequence
rs143865851 p.Arg53Cys; c.157C>T	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.20±0.050 (P<0.001)	Probably damaging (1.000)	well conserved in the 32species; substituted by Gln (2), Glu (2), Thr, Ala, Gly or Met
rs142318540 p.Gln60Arg; c.179A>G	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	well conserved in 38species; substituted by Asp or Glu
rs8176928 p.Arg107Gly; c.319A>G	mono-allelic	0.0037 (<0.0003)	0.007 (0.000)	n.d.	Probably damaging (1.000)	well conserved in 38 species; substituted by Ile or Lys
rs150621329 p.Arg133Gln; c.398G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs150621329	mono-allelic	0.0005	0.001	n.d.	Probably	conserved in all the species

p.Arg133Leu; c.398G>T		(<0.0003)	(0.000)		damaging (1.000)	
rs142644209 p.Asp190His; c.568G>C	mono-allelic	– (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs146236198 p.Asn192Ile; c.575A>T	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs148373909 p.Arg207Cys; c.619C>T	Japanese [g)]	0.0014 (<0.0003)	0.003 (0.000)	0.12±0.022 (P<0.001)	Probably damaging (1.000)	well conserved in the higher species above chondrichthyes; substituted by Ser (2) or Glu
rs200620452 p.Thr229Met; c.686C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.20±0.037 (P<0.001)	Probably damaging (1.000)	not conserved; substituted by His, Ser, Asn (4), Ala or Gly
rs8176940 p.Cys231Tyr; c.692G>A	mono-allelic	– (<0.0003)	0.009 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs199826318 p.Tyr233Cys; c.692G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs139254891 p.Arg235Trp; c.703A>T	mono-allelic	– (<0.0003)	0.000 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs201571412 p.Arg244Ter; c.730C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.		nonsense substitution

rs8176939 p.Ala246Pro; c.736G>C	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.16±0.060 (P<0.001)	Probably damaging (0.979)	not conserved; substituted by Ser (7), Gly (8), Ile or Asp

---

DNASE1L3

rs374713985 p.Arg51Cys; c.151C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	well conserved in 24 species; substituted by Phe or Gly
rs141477807 p.Arg92Trp; c.274G>C	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.19±0.11 (P<0.001)	Probably damaging (1.000)	well conserved in all the species
rs374713985 p.Phe139Leu; c.151C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (0.994)	well conserved in 25 species; substituted by Val
rs367846709 p.Arg178Cys; c.532C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	not conserved; substituted by Gln (3), His (3), Lys (2), Val or Ala
rs200457209 p.Gly193Ser; c.577G>A	mono-allelic	– (<0.0003)	0.002 (0.000)	n.d.	Probably damaging (0.999)	not conserved; substituted by Ala (3) or Asp (4)
rs35677470 p.Arg206Cys; c.602C>T	Caucasian[g)]	0.0266 (0.0009)	0.052 (0.018)	n.d.	Probably damaging (1.000)	well conserved in 22 species; substituted by Val, Asp, Ser or Lys
not registrated p.Trp215GlyfsX2;c.643delT	Caucasian[f)]	–	–	n.d.		well conserved in all the species

rs148314913 p.Phe255Ser; c.764T>C	mono-allelic	– (<0.0003)	0.000 (0.000)	0.20±0.020 (P<0.001)	Probably damaging (0.999)	well conserved in 23 species; substituted by Tyr (2) or Val

---

DNASE2

rs562700068 p.Ala16Asp; c.47C>A	mono-allelic	0.0008 (<0.0003)	N.D. (0.000)	n,d.	Benign (0.421)	substitution in the signal sequence
rs557707591 p.Gly48Glu; c.143G>A	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	0.19±0.010 (P<0.001)	Probably damaging (1.000)	completely conserved
rs34391539 p.Ala74ProfsV101; c.220G>del	mono-allelic	– (<0.0003)	N.D. (0.000)	n,d.	Probably damaging (1.000)	frameshift mutation
rs201030953 p.Tyr95Ser; c.284A>C	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	completely conserved
rs554251996 p.Phe234Cys; c.701T>G	mono-allelic	0.0004 (<0.0003)	N.D. (0.000)	0.13±0.045 (P<0.001)	Probably damaging (0.994)	not conserved; substituted by Leu (2), Tyr, Ala or Trp (2)
rs35851337 p.Val300Met; c.898G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	0.13±0.015 (P<0.001)	Probably damaging (0.994)	well conserved; substituted by Ile (4)
rs562507670 p.Thr307Asn; c.920C>A	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Possibly damaging (0.923)	not conserved; substituted by Ala (23), Val (11) or Ile (4)

rs544319961 p.Arg314Trp; c.940C>T	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	completely conserved
rs1061192 p.Arg314Leu; c.941G>T	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	completely conserved
rs1802335 p.Gln338ter; c.1012C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.		nonsense mutation
rs35851337 p.Val341Leu; c.1021G>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.16±0.030 (P<0.001)	Probably damaging (0.995)	well conserved; substituted by Ala (4) or Ile (3)

---

a) Populations showing genetic heterogeneity for each of SNP in this study populations are shown.

b) Taken from the database. The numbers in parenthesis of each SNP were calculated based on the total subjects (n=1,752) examined in this study.

c) The values are expressed as relative activity of each amino acid-substituted construct in the cell lyzates to that of the wild-type ,representing the mean±SD (*n*=4). P-values in parenthesis were calculated on differences between the activities of the substituted and wild type enzyme by means of the unpaired, Student's *t*-test.

d) Effect of the amino acid substitution corresponding to each SNP on the activity was predicted using Polyphen-2.

e) Multiple alignment analysis on the amino acid sequence of animal DNases was performed; DNases I, 1L3 and II derived from 40, 26 and 55 animals, respectively, available in the genome database. The number in parenthesis is the number of species in which the corresponding amino acid is substituted.

f) The corresponding minor alleles of each SNP were found in the other populations [ ].

g) Found in our studies.

N.D., not determined in the database; －, no data available in the database; n.d., the activity derived from the corresponding amino acid substituted construct could not be detected under our assay conditions.

Supplementary Table 3

Primer sequence, annealing temperatures, and restriction enzymes for PCR-based genotyping for *DNASE2*[a]

Gene	Primer	Sense/ antisense	Sequences[b]	Annealing temp.(°C)	Restriction enzyme	Expected product[c] (bp)
rs562700068 p.Ala16Asp	562700068-1 562700068-2	Sense Antisense	5'-GCAGTCCCGACACAGATTCCTGGATCTCA-3' 5'-GGAGTCCCCGTAGCAGGTCAG-3'	65	*HaeIII*	A allele; 114 C allele; 91
rs373200195 p.Leu38Phe	373200195-1 373200195-2	Sense Antisense	5'-GACTGGTTCGTGGTCTACAAGCTGCCAGAT- 5'-TGTTGATGAGTGCCCTGCCGTCCCGCCAGC-	50	*MboI*	A allele; 113 G allele; 84
rs557707591 p.Gly48Glu	557707591-1 557707591-2	Sense Antisense	5'-CTGACCTGCTACGGGGACTCC-3' 5'-AGCTCTCGTCCAGATACTTGTACTGCGGC-3'	50	*HaeIII*	A allele; 124 G allele; 96
rs139356333 p.Arg84Trp	139356333-1 139356333-2	Sense Antisense	5'-TTCACCTGGCTGGTGTTGCTCC-3' 5'-GGCAGGGCACTCATCAACAG-3'	65	*MspI*	C allele; 83 G allele; 63
rs530686133 p.Ser145Thr	530686133-1 530686133-2	Sense Antisense	5'-GACTCTTCCATGCGTGGGCACACGAAGGGT- 5'-TGCCCGTAGGTACAGGCGCTATGAGGTCAA-	50	*Hpy166II*	C allele; 144 G allele; 116
rs559395836 p.Pro163Ser	559395836 -1 559395836 -2	Sense Antisense	5'-TACGGGCAGACCCTGCTCTGTGTGTCGATT-3' 5'-GTCGGGGAATTCCTGGGCAAAGATCCCTTC-3'	50	*HinfI*	C allele; 132 T allele; 105
rs577376992 p.Gln166Arg	577376992-1 577376992-2	Sense Antisense	5'-CTGCTCTGTGTGTCTTTTCCCTTCGCCC-3' 5'-GTCGGGGAATTCCTGGGCAAAGATCCCTTC-	50	*MspI*	A allele; 120 G allele; 93
rs374218185 p.Met170Thr	374218185-1 374218185-2	Sense Antisense	5'-CTGCTTGCCTGGAGGACAAGG-3' 5'-CCCTTCGCTCAGTTCTCGAACA-3'	50	*NlaIII*	G allele; 123 A allele; 103
rs367700946 p.Asp196Asn	367700946-1 367700946-2	Sense Antisense	5'-CCCTTGACCACATTCTCCGAGT-3' 5'-CATTTCCTCCTCCCCTTGTCC-3'	50	*HinfI*	T allele; 123 C allele; 104
rs565891293 p.Gly222Arg	565891293 -1 565891293 -2	Sense Antisense	5'-ACCTACCCCTGGGTCTATAACTACCAGCTG-3' 5'-TGAACTTGGCAAAGCTCTGGAAAACAGCCC-3'	65	*MspI*	A allele; 166 G allele; 134
rs375732425 p.Val224Ile	375732425-1 375732425-2	Sense Antisense	5'-GCCACACCAGTCACCCTGTCT-3' 5'-ACTCACATCCCAGGCCGGGCAT-3'	50	*NlaIII*	T allele; 123 C allele; 104
rs554251996 p.Phe234Cys	554251996-1 554251996-2	Sense Antisense	5'-TTGGAGAATGTGGTCAAGGGCCACCACGTT- 5'-TGCCAACCAGCCGGAGTACAGGTCATGTCCA-3'	50	*Hpy166II*	T allele; 144 G allele; 118
rs375552010 p.Ile279Thr	375552010-1 375552010-2	Sense Antisense	5'-CCGGCTGGTCCAGGGAAACAT-3' 5'-GCCCTTGGTACCAACCTGCAGGT-3'	50	*NlaIII*	A allele; 122 G allele; 100
rs562507670 p.Thr307Asn	562507670-1 562507670-2	Sense Antisense	5'-AACCAGATAGCTTTCCCTGGACCAGCCGGC-3' 5'-CTGGTTCCGATTCATGTCACCCACGACG-3'	50	*HpyCh4IV*	C allele; 120 A allele; 91
rs544319961 p.Arg314Trp	544319961-1 544319961-2	Sense Antisense	5'-AACCAGATAGCTTTCCCTGGACCAGCCGGC-3' 5'-ACCCCGTTGCTCCTCTCCCTGGTCCC-3'	50	*Hpy188III*	T allele: 138 C allele; 112
rs375137579 p.Pro339Leu	375137579-1 375137579-2	Sense Antisense	5'-GGGCTGGTAGTTCTTCACCACC-3' 5'-TGGACCTGCGTGGGTGACATGAA-3'	50	*MspI*	A allele; 123 G allele; 102

[a]Experimental information on the newly developed genotyping methods are shown; those on the other 12 SNPs have been described in ref. nos. 13 and 18

[b]The underlined residues indicate the mismatched nucleotide incorporated in each primer.

[c]Expected sizes of PCR-restriction fragment derived from each allele.

Supplementary Table 4

All the non-synonymous SNPs in *DNASE1*; genetic distribution, effect of the corresponding amino acid substitution on the activity, prediction of the effect of the amino acid substitution on the activity using PolyPhen-2, and conservation of the corresponding amino acid residue in animal DNases I

SNP	Genetic heterogeneity[a)]	MAF[b)]	Heterozygosity[b)]	Activity[c)]	Polyphen-2 prediction (score)[d)]	Conservation of the amino acid resides corresponding to SNP in 40 animal DNases I[e)]
			SNPs not affecting the activity			
rs8176927 p.Arg2Ser; c.6G>T	only African	0.0294 (0.0242)	0.057 (0.047)	1.17±0.17	Possibly damaging (0.483)	in the signal sequence
rs61741279 p.Gly3Asp; c.8G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.17±0.17	Benign (0.047)	in the signal sequence
rs148015097 p.Ala14Val; c.41C>T	mono-allelic	– (<0.0003)	0.001 (0.000)	1.10±0.35	Benign (0.066)	in the signal peptide
rs201161491[f)] p.Ile25Met; c.75T>G	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.12±0.21	Possibly damaging (0.685)	not conserved; substituted by Met, Val (2) or Leu (3)
rs141673463 p.Ala26Thr; c.76G>A	mono-allelic	– (<0.0003)	0.000 (0.000)	1.14±0.50	Probably damaging (0.990)	not conserved; substituted by Gly(11), Cys or Ser (4)
rs34907394 p.Glu35Asp; c.105G>C	mono-allelic	0.0037 (<0.0003)	0.007 (0.000)	1.23±0.28	Benign (0.001)	not conserved; substituted by Asp (14), Met (2), Thr (2), Ala, Asn, Gln, Arg or Val

rs369619441 [f)] p.Met38Thr; c.113T>C	mono-allelic	– (<0.0003)	N.D. (0.000)	1.43±0.41	Possibly damaging (0.903)	not conserved; substituted by Ile, Leu (3), Val (5), Ala (4) or Ser
rs370054136 [f)] p.Val44Ile; c.130G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	1.32±0.10	Benign (0.386)	not conserved; substituted by Ser (11), Ala (8), Thr (3), Leu (5), Arg, Met (3), Val (3) or Pro
rs140530129 p.Ala81Thr; c.241G>A	mono-allelic	0.0009 (<0.0003)	0.002 (0.000)	0.95±0.074	Benign (0.010)	not conserved; substituted by Gly (3), Lys, Ile, Asp (6), Thr (5), Ser (18) or His
rs190768401 p.Arg95Gln; c.284G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.14±0.39	Probably damaging (1.000)	well conserved in 35 species; substituted by His (3), Thr or Ala
rs199986334 [f)] p.Ala113Val; c.338C>T	only Caucasian	– (0.0026)	N.D. (0.006)	1.36±0.42	Benign (0.001)	not conserved; substituted by Val (26), Ile, Leu (3), Pro (6) or Ala
rs34923865 [f)] p.Tyr117Cys; c.590A>T	only Caucasian	0.0046 (0.0026)	0.009 (0.006)	1.29±0.37	Probably damaging (1.000)	well conserved in 35 species; substituted by Phe (4) or Glu
rs144059899 p.Asp129Asn; c.385G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	1.92±0.81	Probably damaging (1.000)	well conserved except reptiles (Gly or Thr)
rs76397583 p.Asn132Ser; c.395A>G	mono-allelic	– (<0.0003)	0.500 (0.000)	1.45±0.66	Benign (0.000)	not conserved; substituted by Ser (20), Ile (3), Pro, Met (2) or Glu (4)
rs374554105 [f)]	mono-allelic	–	N.D.	1.42±0.27		not conserved; substituted by Gln, Pro (21),

p.Arg143Trp; c.427C>T		(<0.0003)	(0.000)		Possibly damaging (0.876)	His, Leu (4), Lys (3), Asn, Ser (4) or Ala
rs140745748 p.Val172Ile; c.514G>A	mono-allelic	– (<0.0003)	0.000 (0.000)	1.21±0.71	Probably damaging (0.999)	well conserved in 39 species; substituted by Ala
rs74892550 p.Val185Ile; c.514G>A	mono-allelic	0.0023 (<0.0003)	0.005 (0.000)	1.49±0.36	Benign (0.000)	not conserved; substituted by Ile (23), Met (3) or Ala (4)
rs147093089 p.Met186Val; c.556A>G	mono-allelic	– (<0.0003)	0.000 (0.000)	0.78±0.41	Benign (0.000)	not conserved; substituted by Val (3), Leu (11), Phe (3) or Ile (4)
rs34186031 p.Pro219Thrr; c.655C>A	mono-allelic	– (<0.0003)	0.002 (0.000)	1.13±0.29	Benign (0.177)	not conserved; substituted by Ser (3), Gly (7) or Thr (3)
rs146249371 p.Ala232Gly; c.695C>G	mono-allelic	– (<0.0003)	0.000 (0.000)	1.29±0.12	Possibly damaging (0.774)	not conserved; substituted by Pro (8)
rs375616517 [f)] p.Arg235Lys; c.704G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	1.13±0.11	Probably damaging (0.997)	completely conserved
rs200149984 p.Val238Leu; c.712G>C	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.95±0.37	Possibly damaging (0.869)	not conserved; substituted by Ala (15) or Leu
rs148684969 p.Val247Iler; c.739G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.05±0.077	Benign (0.068)	not conserved; substituted by Ile (8), Thr, Leu, Tyr or Glu

rs200538894 p.Ser251Leu; c.752C>T	mono-allelic	– (<0.0003)	0.002 (0.000)	1.29±0.30	Probably damaging (1.000)	not conserved; substituted by Thr (6) or Met
rs142079857 p.Ala259Gly; c.777C>G	mono-allelic	– (<0.0003)	0.000 (0.000)	1.39±0.19	Benign (0.339)	not conserved; substituted by Val (3), Thr (3), Lys (5), Glu (6), Gln (2), Ser (2), Arg or Leu
rs201413861 p.Ala260Gly; c.779C>G	mono-allelic	– (<0.0003)	0.001 (0.000)	1.36±0.24	Possibly damaging (0.877)	not conserved; substituted by Glu (4), Thr (5), Asp (3), Asn, Val (2), Ser or Lys
rs8176924 p.Gly262Asp; c.785G>A	mono-allelic	– (<0.0003)	0.009 (0.000)	1.16±0.54	Possibly damaging (0.634)	not conserved; substituted by Arg (2), Asn (7), Lys (5), His (4) or Asp (2)

SNPs abolishing the activity

rs121912990 p.Lys5Ter; c.13A>T	only Japanese[g)]	– (<0.0003)	– (0.000)	n.d.		nonsense substitution in the signal sequence
rs142318540 p.Gln60Arg; c.179A>G	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	well conserved in 38species; substituted by Asp or Glu
rs8176928 p.Arg107Gly; c.319A>G	mono-allelic	0.0037 (<0.0003)	0.007 (0.000)	n.d.	Probably damaging (1.000)	well conserved in 38 species; substituted by Ile or Lys

rs150621329 p.Arg133Gln; c.398G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs150621329 p.Arg133Leu; c.398G>T	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs142644209 p.Asp190His; c.568G>C	mono-allelic	– (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs146236198 p.Asn192Ile; c.575A>T	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs8176940 p.Cys231Tyr; c.692G>A	mono-allelic	– (<0.0003)	0.009 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs199826318 p.Tyr233Cys; c.692G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs139254891 p.Arg235Trp; c.703A>T	mono-allelic	– (<0.0003)	0.000 (0.000)	n.d.	Probably damaging (1.000)	conserved in all the species
rs201571412 f) p.Arg244Ter; c.730C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.		nonsense substitution

---

SNPs reducing the activity

rs77254040 p.Gln31Glu; c.91C>G	only Japanese [g)]	0.0023 (<0.0003)	0.005 (0.000)	0.43±0.10 (P<0.001)	Benign (0.183)	not conserved; substituted by Arg (20), Lys (5), Glu (2), Ala, Asp or Phe
rs143865851 p.Arg53Cys; c.157C>T	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.20±0.050 (P<0.001)	Probably damaging (1.000)	well conserved in the 32species; substituted by Gln (2), Glu (2), Thr, Ala, Gly or Met
rs144227093 p.Tyr54Cys; c.161A>T	mono-allelic	– (<0.0003)	0.000 (0.000)	0.39±0.10 (P<0.001)	Probably damaging (1.000)	conserved in 39 species; substituted by His
rs45545238 p.Gln60His; c.180G>C	mono-allelic	0.0009 (<0.0003)	0.002 (0.000)	0.41±0.081 (P<0.001)	Probably damaging (1.000)	well conserved in 38species; substituted by Asp or Glu
rs143058517 p.Val111Met; c.331G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.24±0.019 (P<0.001)	Probably damaging (1.000)	well conserved in all the species above reptile; substituted by Ala, Leu (4), Met or Phe
rs138676148 p.Arg139Gly; c.415A>G	mono-allelic	– (<0.0003)	0.000 (0.000)	0.31±0.040 (P<0.001)	Benign (0.183)	not conserved; substituted by Lys (17), Trp, Met (7) or His
rs143407371 p.Ala157Val; c.470C>T	mono-allelic	– (<0.0003)	0.000 (0.000)	0.65±0.038 (P<0.05)	Probably damaging (0.963)	not conserved; substituted by Ser (5) or Thr (14)
rs146238243 p.Pro159Leu; c.476C>T	mono-allelic	– (<0.0003)	0.000 (0.000)	0.27±0.10 (P<0.001)	Probably damaging (0.999)	conserved in all the species
rs139424576 p.Asp167His; c.499G>C	mono-allelic	– (<0.0003)	0.000 (0.000)	0.52±0.073 (P<0.001)	Probably damaging (1.000)	not conserved; substituted by Asn (7), Ser (2) or Gln

rs143371936 p.Ala193Val; c.578C>T	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.63±0.078 (P<0.05)	Probably damaging (1.000)	well conserved in 39 species; substituted by Thr
rs148373909 p.Arg207Cys; c.619C>T	only Japanese [g)]	0.0014 (<0.0003)	0.003 (0.000)	0.12±0.022 (P<0.001)	Probably damaging (1.000)	well conserved in the higher species above chondrichthyes; substituted by Ser (2) or Glu
rs200620452 p.Thr229Met; c.686C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.21±0.037 (P<0.001)	Probably damaging (1.000)	not conserved; substituted by His, Ser, Asn (4), Ala or Gly
rs1053874 p.Gln244Arg; c.731A>G	polymorphic	– (0.436)	N.D. (0.492)	0.48±0.015 (P<0.001)	Benign (0.006)	not conserved; substituted by Arg (7), Leu (2), Met (5), Lys (4) or Thr
rs8176939 p.Ala246Pro; c.736G>C	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.16±0.060 (P<0.001)	Probably damaging (0.979)	not conserved; substituted by Ser (7), Gly (8), Ile or Asp

---

SNPs elevating the activity

rs147546841 p.Tyr46His; c.136T>C	mono-allelic	– (<0.0003)	0.000 (0.000)	1.85±0.63 (P<0.05)	Probably damaging (0.985)	not conserved; substituted by Val, Phe (3), Ile (10), Ser, Leu (3), Arg (2), Met or Thr (2)
rs140187838 [f)] p.Tyr46Ser; c.137A>C	mono-allelic	– (<0.0003)	0.001 (0.000)	2.65±0.43 (P<0.001)	Possibly damaging (0.954)	not conserved; substituted by Val, Phe (3), Ile (10), Ser, Leu (3), Arg (2), Met or Thr (2)

rs141801594 p.Asp120Asn; c.358G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.87±0.40 (P<0.05)	Benign (0.395)	well conserved in the 34 species; substituted by Asn, Lys, Ser (2) or Pro (2)
rs8176919 p.Gly127Arg; c.379G>A	only African	0.0184 (0.0094)	0.036 (0.018)	2.11±0.52 (P<0.01)	Probably damaging (1.000)	conserved except reptile (Ser)
rs139615062 p.Arg143Gln; c.428G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	1.76±0.43 (P<0.05)	Benign (0.011)	not conserved; substituted by Pro (21), His, Leu (4), Lys (3), Asn, Ser (4) or Ala
rs368307903 [f] p.Glu149Gln; c.445G>C	mono-allelic	– (<0.0003)	N.D. (0.000)	2.53±1.22 (P<0.05)	Benign (0.038)	not conserved; substituted by Ala, Asp (10), Lys, Asn (2) or Thr
rs1799891 p.Pro154Ala; c.460C>G	only Japanese [g]	0.0078 (<0.0003)	0.015 (0.000)	1.43±0.25 (P<0.05)	Benign (0.235)	well conserved in the 31 species; substituted by Ala (3), Ser (4), Val or Gly
rs201942334 p.Ile166Met; c.498C>G	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.45±0.10 (P<0.05)	Benign (0.269)	not conserved; substituted by Met, Val (5) or Leu (2)
rs150933932 p.Ala168Val; c.503C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	1.64±0.31 (P<0.05)	Possibly damaging (0.892)	well conserved in the 29 species; substituted by Ser (5), Gly, Ile or Asn
rs199540176 [f] p.Glu183Gln; c.547G>C	mono-allelic	– (<0.0003)	N.D. (0.000)	2.21±0.59 (P<0.001)	Benign (0.051)	not conserved; substituted by Asn (10), Thr (3) or Ser
rs138354028 p.Ser221Asn; c.662G>A	mono-allelic	– (<0.0003)	0.000 (0.000)	1.86±0.23 (P<0.01)	Benign (0.000)	not conserved; substituted by Thr (12), Asp (5), Lys, Asn (2), Ala, His or Cys

a) Populations showing genetic heterogeneity for each of SNP in this study populations are shown.

b) Taken from the database. The numbers in parenthesis of each SNP were calculated based on the total subjects (n=1,752) examined in this study.

c) The values are expressed as relative activity of each amino acid-substituted construct in the cell lyzates to that of the wild-type ,representing the mean±SD (*n*=4). P-values in parenthesis were calculated on differences between the activities of the substituted and wild type enzyme by means of the unpaired, Student's *t*-test.

d) Effect of the amino acid substitution corresponding to each SNP on the activity was predicted using PolyPhen-2.

e) Multiple alignment analysis on the amino acid sequence of 40 animal DNases I was performed; the number in parenthesis is the number of species in which the corresponding amino acid is substituted.

f) Determined in this study.

g) The corresponding minor alleles of each SNP were found in the other Japanese study populations [ ].

N.D., not determined in the database; －, no data available in the database; n.d., the activity derived from the corresponding amino acid substituted construct could not be detected under our assay conditions.

Supplementary Table 5

All the non-synonymous SNPs in *DNASE1L3*; genetic distribution, effect of the corresponding amino acid substitution on the activity, prediction of the effect of the amino acid substitution on the activity using PolyPhen-2, and conservation of the corresponding amino acid residue in animal DNases1L3

SNP	Genetic heterogeneity[a)]	MAF[b)]	Heterozygosity[b)]	Activity[c)]	PolyPhen-2 prediction (score)[d)]	Conservation of the amino acid residues corresponding to SNP in 26 animal DNases 1L3[e)]
			SNPs not affecting the activity			
rs370020945[f)] p.Leu12His; c.35T>A	mono-allelic	– (<0.0003)	N.D. (0.000)	0.83±0.38	Probably damaging (0.994)	in the signal sequence
rs142751723 p.Ser14Tyr; c.41C>A	mono-allelic	– (<0.0003)	0.000 (0.000)	0.73±0.30	Possibly damaging (0.845)	in the signal sequence
rs139865883 p.Arg22Lys; c.65G>A	mono-allelic	– (<0.0003)	0.001 (0.000)	0.90±0.22	Benign (0.000)	not conserved; substituted by Lys (7), Asn, Gln or Leu
rs188529894 p.Ile23Val; c.67A>G	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.0±0.29	Benign (0.022)	well conserved in 20 species; substituted by Leu (5) or Val
rs201136542[f)] p.Glu33Lys; c.98G>A	mono-allelic	– (<0.0003)	0.002 (0.000)	0.84±0.08	Possibly damaging (0.896)	not conserved; substituted by Ala (3), Gly, Asp, Arg (2), Gln, Lys (2) or Thr
rs148208826 p.Asp38His; c.112G>C	mono-allelic	0.0009 (<0.0003)	0.002 (0.000)	1.3±0.25	Possibly damaging (0.801)	not conserved; substituted by Asn (7), Arg (5), Lys (6) or Thr

rs145973332 p.Ala41Val; c.122C>T	mono-allelic	– (<0.0003)	0.000 (0.000)	0.90±0.20	Benign (0.000)	not conserved; substituted by Thr, Val (13) or Ile (2)
rs148406554 p.Cys52Gly; c.154T>G	mono-allelic	– (<0.0003)	0.000 (0.000)	1.1±0.25	Probably damaging (0.997)	well conserved in the higher species above fish; substituted by Tyr (4)
rs376894478 f) p.Val57Met; c.169G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	0.79±0.45	Probably damaging (0.960)	not conserved; substituted by Leu (17) or Ile (5)
rs149330798 p.Glu73Lys; c.217G>A	mono-allelic	– (<0.0003)	0.000 (0.000)	0.82±0.16	Benign (0.003)	not conserved; substituted by Gln (3), Lys, Asp, Ala, Thr or Ser
rs74350392 p.Gly82Arg; c.244G>C	mono-allelic	0.0101 (<0.0003)	0.020 (0.000)	0.68±0.21	Benign (0.012)	not conserved: substituted by Ser (3), Arg (2), Asn (2), Pro (2), Lys or Gln
rs143383107 p.Val88Met; c.262G>A	mono-allelic	– (<0.0003)	0.000 (0.000)	1.0±0.17	Probably damaging (0.984)	not conserved; substituted by Thr (2) , Phe, Ile, Ser, Leu (3) or Ala
rs12491947 p.Asn96Lys; c.288C>A	mono-allelic	0.0129 (<0.0003)	0.025 (0.000)	0.82±0.23	Benign (0.002)	not conserved; substituted by Lys (5), Thr, Ser (5) or Gln
rs372940270 f) p.Ser116Arg; c.346A>C	mono-allelic	– (<0.0003)	N.D. (0.000)	0.90±0.54	Benign (0.004)	not conserved; substituted by Lys (4), Tyr, Arg, Trp, Thr (6), Gln, Asp or Asn
rs369165467 f) p.Asp126Tyr; c.376G>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.71±0.27	Probably damaging (1.000)	not conserved; substituted by Gln, Asn, Gly or deletion (5)

rs201928908 [f)] p.Val146Ala; c.437T>C	mono-allelic	– (<0.0003)	N.D. (0.000)	1.10±0.14	Benign (0.145)	not conserved; substituted by Ala (3), Ile (3), Gly or Lys
rs142361820 p.Ile165Met; c.495C>G	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.90±0.20	Probably damaging (0.966)	well conserved in 22species; substituted by Met, Asn (2) or Leu
rs140654958 p.Val175Leu; c.523G>T	mono-allelic	– (<0.0003)	0.000 (0.000)	1.0±0.050	Benign (0.295)	not conserved; substituted by Ala (3),Ile (3) or Thr (2)
rs371048745 [f)] p.Ala181Val; c.542C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.72±0.20	Benign (0.017)	not conserved; substituted by Thr (6), Pro, Val (2), Ile, Ser (4), Asp, Met or Glu
rs147210152 p.Thr209Ile; c.626C>T	mono-allelic	– (<0.0003)	0.000 (0.000)	1.0±0.26	Benign (0.004)	well conserved in 21 species; substituted by Asn (2), Ser, Val or Met
rs369901098 [f)] p.Glu221Gln; c.651G>C	mono-allelic	– (<0.0003)	N.D. (0.000)	0.87±0.09	Possibly damaging (0.695)	not conserved; substituted by Gln, Asn (4), Thr (2), Val (3) or Ala (3)
rs76440799 p.Ile243Met; c.729C>G	mono-allelic	– (<0.0003)	0.000 (0.000)	0.82±0.10	Possibly damaging (0.771)	not conserved; substituted by Leu (10) or Met (2)
rs146444966 p.Tyr261Cys; c.533G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	1.5±0.51	Probably damaging (0.999)	not conserved; substituted by Phe (6) or Leu (2)
rs367807030 [f)]	mono-allelic	–	N.D.	0.76±0.13	Benign	not conserved; substituted by Ser (6), or Asp

p.Thr264Asn; c.861C>A		(<0.0003)	(0.000)		(0.048)	
rs113005222 p.Ala268Gly; c.803C>G	mono-allelic	– (<0.0003)	0.500 (0.000)	0.80±0.10	Probably damaging (1.000)	well conserved in 22species; substituted by Thr (3) or Asp
rs151161986 p.Arg285Lys; c.854G>A	mono-allelic	0.0009 (<0.0003)	0.002 (0.000)	1.1±0.076	Benign (0.000)	not conserved; substituted by Gly (4) , Ser (2), Val or Lys (3)
			SNPs abolishing the activity			
rs374713985 f) p.Arg51Cys; c.151C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	well conserved in 24 species; substituted by Phe or Gly
rs372683613 f) p.Phe139Leu; c.151C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (0.994)	well conserved in 25 species; substituted by Val
rs367846709 f) p.Arg178Cys; c.532C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	not conserved; substituted by Gln (3), His (3), Lys (2), Val or Ala
rs200457209 f) p.Gly193Ser; c.577G>A	mono-allelic	– (<0.0003)	0.002 (0.000)	n.d.	Probably damaging (0.999)	not conserved; substituted by Ala (3) or Asp (4)
rs35677470 p.Arg206Cys; c.602C>T	Caucasian	0.0266 (0.0009)	0.052 (0.018)	n.d.	Probably damaging (1.000)	well conserved in 22 species; substituted by Val, Asp, Ser or Lys
not registrated	Caucasian g)	–	–	n.d.		well conserved in all the species

p.Trp215GlyfsX2;c.643delT

SNPs reducing the activity

rs202183427 [f)] p.Arg3Trp; c.7C>T	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.63±0.28 (P<0.05)	Probably damaging (0.963)	in signal sequence
rs145888358 p.Val46Met; c.136G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.60±0.15 (P<0.05)	Possibly damaging (0.947)	not conserved; substituted by Arg (2), Thr, Ile (2), Ser, Leu or Lys
rs141477807 p.Arg92Trp; c.274G>C	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.19±0.11 (P<0.001)	Probably damaging (1.000)	well conserved in all the species
rs146805633 p.Val150Met; c.447G>A	mono-allelic	0.0005 (<0.0003)	0.001 (0.000)	0.63±0.17 (P<0.05)	Probably damaging (1.000)	well conserved in the mammals; substituted by Ala (4), Phe, Leu or Thr
rs147219402 p.Thr156Pro; c.466A>C	mono-allelic	– (<0.0003)	0.000 (0.000)	0.30±0.18 (P<0.001)	Probably damaging (1.000)	well conserved in 21 species ; substituted by Ala (5)
rs3772986 [f)] p.Arg178Leu; c.533G>T	mono-allelic	– (<0.0003)	0.009 (0.000)	0.49±0.01 (P<0.001)	Possibly damaging (0.929)	not conserved; substituted by Gln (3), His (3), Lys (2), Val or Ala
rs148314913 p.Phe255Ser; c.764T>C	mono-allelic	– (<0.0003)	0.000 (0.000)	0.20±0.020 (P<0.001)	Probably damaging (0.999)	well conserved in 23 species; substituted by Tyr (2) or Val

---

SNPs elevating the activity

rs376419946 [f] p.Ser302Arg; c.886C>G	mono-allelic	− (<0.0003)	N.D. (0.000)	1.90±0.27 (P<0.001)	Benign (0.255)	not conserved; substituted by Gly (2), Val (2), Ala (3), Thr, Arg, Asn, Tyr, Val, Pro or deletion (6)

a) Populations showing genetic heterogeneity for each of SNP in this study populations are shown.

b) Taken from the database. The numbers in parenthesis of each SNP were calculated based on the total subjects (n=1,752) examined in this study.

c) The values are expressed as relative activity of each amino acid-substituted construct in the cell lyzates to that of the wild-type ,representing the mean±SD (*n*=4). P-values in parenthesis were calculated on differences between the activities of the substituted and wild type enzyme by means of the unpaired, Student's *t*-test.

d) Effect of the amino acid substitution corresponding to each SNP on the activity was predicted using PolyPhen-2.

e) Multiple alignment analysis on the amino acid sequence of 26 animal DNases 1L3 was performed; the number in parenthesis is the number of species in which the corresponding amino acid is substituted.

f) Determined in this study.

g) The corresponding minor allele of each SNP was found in the other populations [ ].

N.D., not determined in the database; −, no data available in the database; n.d., the activity derived from the corresponding amino acid substituted construct could not be detected under our assay conditions.

Supplementary Table 6

All the non-synonymous SNPs in *DNASE2*; genetic distribution, effect of the corresponding amino acid substitution on the activity, prediction of the effect of the amino acid substitution on the activity using PolyPhen-2, and conservation of the corresponding amino acid residue in animal DNases II

SNP	Genetic heterogeneity[a)]	MAF[b)]	Heterozygosity[b)]	Activity[c)]	PolyPhen-2 prediction (score)[d)]	Conservation of the amino acid resides corresponding to SNP in 55 animal DNases II[e)]
			SNPs not affecting the activity			
rs149049935 p.Leu35Val; c.103C>G	mono-allelic	– (<0.0003)	0.001 (0.000)	0.97±0.13	Probably damaging (0.995)	well conserved; substituted by Ile (2) or Pro
rs373200195[f)] p.Leu38Phe; c.112C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.73±0.15	Possibly damaging (0.953)	not conserved; substituted by Ala, His (14), Asn, Gln (6), Glu (7), Ser (5), Val, Met, Arg (2), Ile or Lys (3)
rs36075196 p.Arg39Ile; c.116G>T	mono-allelic	– (<0.0003)	0.050 (0.000)	1.40±0.25	Benign (0.139)	not conserved; substituted by Ser (18), Thr (7), Asn (3), Pro (2), Asp (2), Lys (2), His (6), Gly, Glu or Tyr
rs112348773 p.Ala45Gly; c.134C>G	mono-allelic	0.0020 (<0.0003)	N.D. (0.000)	1.34±0.22	Benign (0.026)	not conserved; substituted by Val (3), Pro (6), Ser (2), Thr (7), Phe, Leu (4), Glu (2), Asp, His, Tyr or Ile (3)
rs539187149 p.Glu56Gln; c.211C>T	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	0.70±0.10	Benign (0.054)	not conserved; substituted by Gln (3), Ala (4), Lys (8), Gly (4), Ser (11), Val (2) or Ala

rs150023166 p.Pro71Ser; c.211C>T	mono-allelic	0.0014 (<0.0003)	N.D. (0.000)	1.38±0.20	Benign (0.003)	not conserved; substituted by Ser (17), Lys, Thr (17), Pro (3), Leu (3), Asn, Glu, Ala or Gln
rs530686133 [f)] p.Ser145Thr; c.434G>C	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	1.06±0.13	Benign (0.037)	not conserved; substituted by Arg, Thr (3), Asn, Tyr (4), Asp (2), Gly, Ala, Gln, Phe or Val
rs16978744 p.His204Arg; c.611A>G	mono-allelic	0.0062 (<0.0003)	N.D. (0.000)	1.20±0.13	Benign (0.047)	not conserved; substituted by Thr (2), Tyr, Gln (6), Leu, Gly (2), Asn (3), Arg, Ala (2), Lys (5), Val or Ser
rs202170563 p.Val206Ile; c.616G>A	only Korean	0.0012 (0.0022)	N.D. (0.0044)	0.84±0.13	Benign (0.006)	not conserved; substituted by Leu, Ile (3), Ala (2), Thr, Arg, His (2) or deletion (11)
rs375732425 [f)] p.Val224Ile; c.670G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	1.00±0.22	Benign (0.004)	not conserved; substituted by Phe, Thr (21), Ser, Met (2), Ala (2), Glu (6), Gln (2), Asn (7), Lys or Asp
SNPs abolishing the activity						
rs562700068 [f)] p.Ala16Asp; c.47C>A	mono-allelic	0.0008 (<0.0003)	N.D. (0.000)	n,d.	Benign (0.421)	signal sequence
rs34391539 p.Ala74ProfsV101; c.220G>del	mono-allelic	– (<0.0003)	N.D. (0.000)	n,d.		frameshift

rs201030953 p.Tyr95Ser; c.284A>C	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	completely conserved
rs562507670 f) p.Thr307Asn; c.920C>A	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Possibly damaging (0.923)	not conserved; substituted by Ala (23), Val (11) or Ile (4)
rs544319961 f) p.Arg314Trp; c.940C>T	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	completely conserved
rs1061192 p.Arg314Leu; c.941G>T	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	n.d.	Probably damaging (1.000)	completely conserved
rs1802335 p.Gln338ter; c.1012C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	n.d.		nonsense

---

SNPs reducing the activity

rs557707591 f) p.Gly48Glu; c.143G>A	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	0.21±0.05 (P<0.001)	Probably damaging (1.000)	completely conserved
rs139356333 p.Arg84Trp; c.250C>T	mono-allelic	– (<0.0003)	0.000 (0.000)	0.61±0.19 (P<0.01)	Probably damaging (0.997)	not conserved; substituted by Gln (7), Lys (4), Glu, Ser (7), Asn (4), Thr, Met, Gly or Asp (2)
rs559395836 f)	mono-allelic	0.0004	N.D.	0.47±0.060	Benign	not conserved; substituted by Ser (4), Gln,

p.Pro163Ser; c.487C>T		(<0.0003)	(0.000)	(P<0.001)	(0.016)	Lys, Asn (3), Gly or Asp (2)
rs577376992 [f)] p.Gln166Arg; c.497A>G	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	0.24±0.027 (P<0.001)	Benign (0.169)	not conserved; substituted by Glu (5), Arg (3), His, Ser, Asp, Asn (3) or Phe
rs374218185 [f)] p.Met170Thr; c.509T>C	mono-allelic	– (<0.0003)	N.D. (0.000)	0.21±0.054 (P<0.001)	Benign (0.003)	not conserved; substituted by Ile (46), Val (5) or Leu (2)
rs1472370333 p.Gln192Pro; c.575A>C	mono-allelic	– (<0.0003)	0.002 (0.000)	0.32±0.079 (P<0.005)	Benign (0.107)	not conserved; substituted by Arg (2), Lys (5), Ala (2), Val, Thr (2), Ser (9), Pro or Asn
rs367700946 [f)] p.Asp196Asn; c.586G>A	mono-allelic	– (<0.0003)	N.D. (0.000)	0.59±0.12 (P<0.005)	Benign (0.004)	not conserved; substituted by Thr (2), Asn (4), Tyr (9), His (3), Val (4), Glu (6), Ala (5), Ser (3) or Gly
rs565891293 [f)] p.Gly222Arg; c.665G>A	mono-allelic	0.0006 (<0.0003)	N.D. (0.000)	0.25±0.033 (P<0.001)	Probably damaging (0.999)	not conserved; substituted by Lys (2), Asp, Asn or Ser
rs554251996 [f)] p.Phe234Cys; c.701T>G	mono-allelic	0.0004 (<0.0003)	N.D. (0.000)	0.13±0.020 (P<0.001)	Probably damaging (0.994)	not conserved; substituted by Leu (2), Tyr, Ala or Trp (2)
rs375552010 [f)] p.Ile279Thr; c.836T>C	mono-allelic	– (<0.0003)	N.D. (0.000)	0.64±0.17 (P<0.01)	Benign (0.314)	not conserved; substituted by Thr (14), Val (7), Leu (4) or Val (5)
rs35851337	mono-allelic	–	N.D.	0.13±0.015	Probably	well conserved; substituted by Ile (4)

p.Val300Met; c.898G>A		(<0.0003)	(0.000)	(P<0.001)	damaging (1.000)	
rs138534991 p.Ala332Asp; c.995C>A	mono-allelic	0.0002 (<0.0003)	N.D. (0.000)	0.58±0.035 (P<0.01)	Benign (0.022)	not conserved; substituted by Ser (3), Thr (5), Asn (2), Val (6), Gln, Arg, Leu, Asp, Pro, Phe or deletion (6)
rs375137579 f) p.Pro339Leu; c.1016C>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.51±0.18 (P<0.01)	Possibly damaging (0.646)	well conserved; substituted by Ser (10), Thr (3), Ala, Gln (2), Gly, Asn, Asp, Glu, Lys or Arg
rs342056336 p.Val341Leu; c.1021G>T	mono-allelic	– (<0.0003)	N.D. (0.000)	0.16±0.030 (P<0.001)	Probably damaging (0.995)	well conserved; substituted by Ala (4) or Ile (3)

a) Populations showing genetic heterogeneity for each of SNP in this study populations are shown.

b) Taken from the database. The numbers in parenthesis of each SNP were calculated based on the total subjects (n=1,752) examined in this study.

c) The values are expressed as relative activity of each amino acid-substituted construct in the cell lyzates to that of the wild-type ,representing the mean±SD (*n*=4). P-values in parenthesis were calculated on differences between the activities of the substituted and wild type enzyme by means of the unpaired, Student's *t*-test.

d) Multiple alignment analysis on the amino acid sequence of 40 animal DNases I was performed; the number in parenthesis is the number of species in which the corresponding amino acid is substituted.

e) Effect of the amino acid substitution corresponding to each SNP on the activity was predicted using PolyPhen-2.

f) Determined in this study.

N.D., not determined in the database; －, no data available in the database; n.d., the activity derived from the corresponding amino acid substituted construct could not be detected under our assay conditions.