JCI insight

RESEARCH ARTICLE

A human *PSMB11* variant affects thymoproteasome processing and CD8+ T cell production

Izumi Ohigashi,[1] Yuki Ohte,[2] Kazuya Setoh,[3] Hiroshi Nakase,[1] Akiko Maekawa,[1] Hiroshi Kiyonari,[4] Yoko Hamazaki,[5] Miho Sekai,[5] Tetsuo Sudo,[6] Yasuharu Tabara,[3] Hiromi Sawai,[7] Yosuke Omae,[7] Rika Yuliwulandari,[8] Yasuhito Tanaka,[9] Masashi Mizokami,[10] Hiroshi Inoue,[11] Masanori Kasahara,[12] Nagahiro Minato,[5] Katsushi Tokunaga,[7] Keiji Tanaka,[13] Fumihiko Matsuda,[3] Shigeo Murata,[2] and Yousuke Takahama[1]

[1]Division of Experimental Immunology, Institute of Advanced Medical Sciences, University of Tokushima, Tokushima, Japan. [2]Laboratory of Protein Metabolism, Graduate School of Pharmaceutical Sciences, University of Tokyo, Tokyo, Japan. [3]Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [4]Animal Resource Development Unit and Genetic Engineering Team, RIKEN Center for Life Science Technologies, Kobe, Japan. [5]Department of Immunology and Cell Biology, Graduate School of Medicine, [6]Department of Nanobio Drug Discovery, Graduate School of Pharmaceutical Science, Kyoto University, Kyoto, Japan. [7]Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan. [8]Department of Pharmacology, Faculty of Medicine, YARSI University, Jakarta Pusat, Indonesia. [9]Department of Virology and Liver Unit, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan. [10]Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Japan. [11]Division of Genetic Information, Institute for Genome Research, University of Tokushima, Tokushima, Japan. [12]Department of Pathology, Graduate School of Medicine, Hokkaido University, Sapporo, Japan. [13]Laboratory of Protein Metabolism, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan.

The *Psmb11*-encoded β5t subunit of the thymoproteasome, which is specifically expressed in cortical thymic epithelial cells (cTECs), is essential for the optimal positive selection of functionally competent CD8+ T cells in mice. Here, we report that a human genomic *PSMB11* variation, which is detectable at an appreciable allele frequency in human populations, alters the β5t amino acid sequence that affects the processing of catalytically active β5t proteins. The introduction of this variation in the mouse genome revealed that the heterozygotes showed reduced β5t expression in cTECs and the homozygotes further exhibited reduction in the cellularity of CD8+ T cells. No severe health problems were noticed in many heterozygous and 5 homozygous human individuals. Long-term analysis of health status, particularly in the homozygotes, is expected to improve our understanding of the role of the thymoproteasome-dependent positive selection of CD8+ T cells in humans.

Introduction

CD8+ T cells are immune cells that play a central role in the defense against viral infection, intracellular pathogens, and malignant tumors. The thymus produces T cells, including CD8+ T cells. Thymic production of functionally competent CD8+ T cells is dependent on the TCR-mediated positive selection of immature thymocytes with a set of major histocompatibility complex (MHC) class I–associated self-peptides that are presented by cortical thymic epithelial cells (cTECs). The MHC class I–associated self-peptides displayed by cTECs are produced by the thymoproteasome, a form of proteasome specifically expressed in cTECs, because its unique catalytic subunit β5t, encoded by *Psmb11*, is exclusively and abundantly transcribed in cTECs (1–3). β5t deficiency in mice leads to the aberrant positive selection of CD8+ T cells, which are reduced in cellularity and are defective in TCR-mediated immune responses (4–6). β5t in humans is encoded by *PSMB11* in chromosome 14 and is also detectable specifically in cTECs (7, 8). However, the function of β5t-containing thymoproteasome in humans is unknown.

The present study describes a human genomic variation of *PSMB11* single nucleotide polymorphism, rs34457782, of which its minor allele variation, guanine (G) to adenine (A), is detectable at an appreciable

Conflict of interest: The authors have declared that no conflict of interest exists.

Submitted: March 1, 2017
Accepted: April 11, 2017
Published: May 18, 2017

Reference information:
JCI Insight. 2017;2(10):e93664.
https://doi.org/10.1172/jci.insight.93664.

Table 1. β5t amino acid sequences in different species in accordance with NCBI database

Species	Amino acid sequence
Homo sapiens	RLA*H***G***TTTLAFRFRH**G*
Pan troglodytes (Chimpanzee)	RLA*H***G***TTTLAFRFRH**G*
Pongo pygmaeus (Orangutan)	RLA*H***G***TTTLAFRFRH**G*
Callithrix jacchus (Marmoset)	RLA*H***G***TTTLAFRFRH**G*
Tarsius syrichta (Tarsier)	KLA*H***G***TTTLAFRFRH**G*
Loxodonta africana (Elephant)	KLA*H***G***TTTLAFRFCH**G*
Bos taurus (Cow)	RLA*H***G***TTTLAFRFRH**G*
Equus caballus (Horse)	RLA*H***G***TTTLAFRFRH**G*
Sus scrofa (Pig)	RLA*H***G***TTTLAFRFRD**G*
Oryctolagus cuniculus (Rabbit)	QLA*H***G***TTTLAFRFRH**G*
Canis familiaris (Dog)	RLA*H***G***TTTLAFRFRH**G*
Rattus norvegicus (Rat)	RLA*H***G***TTTLAFRFRH**G*
Mus musculus (Mouse)	RLA*H***G***TTTLAFRFRH**G*
Macropus eugenii (Wallaby)	ELA*H***G***TTTLAFRFRH**G*
Monodelphis domestica (Opossum)	DLA*H***G***TTTLAFRFQH**G*
Pteropus vampyrus (Bat)	RLA*H***G***TTTLAFRFRH**G*
Tursiops truncatus (Dolphin)	RLA*H***G***TTTLAFRFRH**G*
Ornithorhynchus anatinus (Platypus)	DLA*H***G***TTTLAFRFRH**G*
Anolis carolinessis (Anole lizard)	-LS*H***G***TTTLAFRCSH**G*
Xenopus tropicalis (African clawed toad)	---*H***G***TTTLAFIYSG**G*
Oryzias latipes (Medaka fish)	---*H***G***TTTLAFVFQG**G*
Takifugu rubripes (Fugu fish)	PLS*H***G***TTTLAFIFQG**G*
Danio rerio (Zebrafish)	TLS*H***G***TTTLGFAFQG**G*
Gasterosteus aculeatus (Stickleback fish)	PMS*H***G***TTTLAFLFQG**G*

Sequences corresponding to human amino acid positions 45–60 are shown. Glycine at position 49 is conserved in all species and shown in bold. Other conserved amino acids are shown in italics. Human and mouse are shown in bold.

allele frequency in humans. This minor allele variation alters the amino acid sequence of human β5t at the 49th amino acid, glycine (Gly), to serine (Ser), which affects the processing of nascent propeptides to catalytically active β5t proteins. The introduction of this G-to-A variation in the mouse genome by homologous recombination in embryonic stem cells (ES cells) revealed that the heterozygous mice showed reduced β5t expression in cTECs and the homozygous mice exhibited further reduced cellularity of $CD8^+$ T cells in vivo, suggesting that the production of $CD8^+$ T cells is affected in heterozygous and/or homozygous humans. A cohort study identified many heterozygotes and 5 individuals who are homozygous in this variation. So far, we know of no clear associations with severe health problems in those heterozygotes or homozygotes, at least indicating that this genomic variation does not cause apparently severe defects, including lethality, in humans. Long-term analysis of immune cells and health status in those heterozygotes and homozygotes, particularly in the homozygotes, is expected to improve our understanding of the role of thymoproteasome-dependent positive selection of $CD8^+$ T cells in humans.

Results

Single nucleotide variant rs34457782 in human populations. We previously showed that β5t-containing thymoproteasome is specifically expressed in cTECs in mice and humans (2, 7, 9) and is essential for the optimal positive selection of functionally competent $CD8^+$ T cells in mice (2, 4, 6). In order to explore the function of β5t in humans, we surveyed the NCBI public database for variations in the human *PSMB11* genome sequence and found a number of *PSMB11* single nucleotide variants that potentially affect the amino acid sequence of human β5t protein (Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.93664DS1). Among those variants, we noticed that the most common variant, rs34457782, carries a minor allele variation (G to A) that changes the 49th amino acid of β5t protein from Gly to Ser (G49S) at an appreciable allele frequency in human populations (Supplemental Table 2). The minor allele frequency appears to vary among human populations; it is 4.85% in Han Chinese in Beijing, 3.03% in Finnish in Finland and Kinh Vietnamese in Ho Chi Minh, and 0.00% in African Caribbeans in Barbados, Americans of African Ancestry in Southwestern US, and several other populations (Supplemental Table 2). In our own surveys of human genome samples previously collected at the University of Tokushima (10, 11), the minor allele frequency was 3.11% in 949 local healthy Japanese volunteers (Supplemental Table 3), whereas in the NCBI public database, the frequency was 2.40% in 104 Japanese people in Tokyo (Supplemental Table 2; Fisher's exact test indicated no significant differences in minor allele frequencies detected in Tokushima and Tokyo). These results indicate that the rs34457782 variant, which affects the amino acid sequence of β5t protein, is present at appreciable frequencies in various human populations.

G49S variation affects processing of human β5t protein. The rs34457782 variant alters the 49th amino acid Gly (G49) of β5t protein, which is a highly conserved amino acid among vertebrate species, along with neighboring amino acids (Table 1). Mature β5t protein possesses a triple threonine amino terminus that is responsible for the proteolytic activity of catalytic β subunits and is generated by the cleavage of the propeptide (12). G49 is located at the carboxyl terminus of the propeptide to be cleaved off for the generation of catalytically active β5t protein (Table 1). Previous studies on other β subunits suggested that the alteration of this Gly residue might affect the propeptide cleavage (13–15). In order to examine whether the G49S variation might actually affect the processing of β5t protein, human embryonic kidney 293T cells were

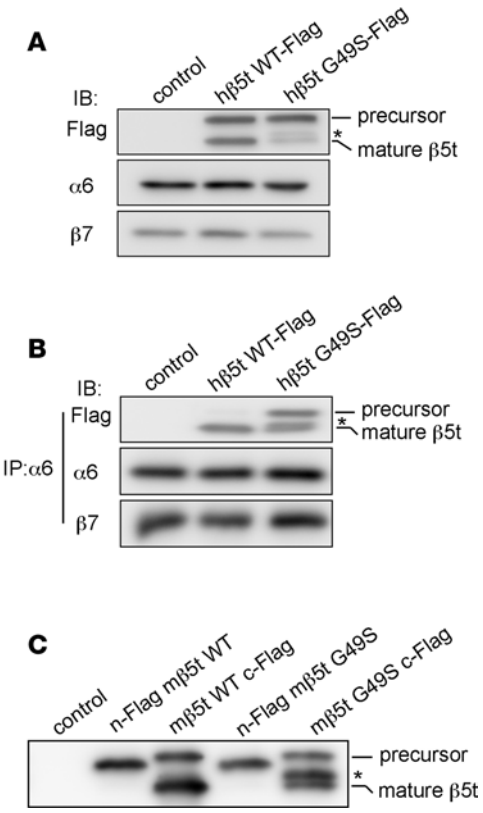

Figure 1. G49S variation affects processing of human and mouse β5t protein. (**A**) WT (hβ5t WT-Flag) and G49S (hβ5t G49S-Flag) human β5t tagged at C-terminus with the Flag epitope was transfected into HEK293T cells. Proteins extracted from transfected cells were immunoblotted with anti-Flag, anti-α6, and anti-β7 antibodies. (**B**) Proteins extracted from hβ5t WT-Flag and hβ5t G49S-Flag transfected cells were fractionated by glycerol gradient centrifugation, and the fractions containing the 26S proteasome were immunoprecipitated with anti-α6 antibody followed by immunoblotting with anti-Flag, anti-α6, and anti-β7 antibodies. (**C**) WT or G49S mouse β5t (mβ5t) tagged with the Flag epitope at either N-terminus or C-terminus was transduced in fibroblasts isolated from β5i-deficient embryos. Cells were treated with IFN-γ to induce the expression of β1i and β2i, and extracted proteins were immunoprecipitated and immunoblotted with anti-Flag antibody. See complete unedited blots in the supplemental material. Asterisks indicate additional signals with an intermediate molecular size.

transfected with a plasmid that was engineered to express FLAG-tagged human β5t protein of either WT or G49S variant sequence (Figure 1A). For the formation of functionally active proteasomes, the ordered incorporation of β subunits onto the assembled α-ring is followed by the cleavage of β subunit propeptides (12). The immunoblot analysis of FLAG-tagged β5t proteins in total cell lysates showed that both nascent precursor protein and processed mature β5t protein were detected when the cells were transfected with WT β5t–expressing plasmid (Figure 1A). However, parallel analysis of β5t-G49S–expressing cells revealed that the mature β5t protein was less abundant than the nascent precursor proteins (Figure 1A). In β5t-G49S–expressing cells, we also detected an additional FLAG signal with an intermediate molecular size (denoted by an asterisk, Figure 1A), possibly representing aberrantly processed abnormal β5t protein. Immunoblot analysis of proteins immunoprecipitated with anti-α6 antibody from the 26S proteasome fraction separated by glycerol gradient centrifugation revealed that the majority of β5t proteins incorporated in fully assembled proteasome complex in WT β5t–expressing cells were mature β5t protein rather than nascent precursor protein, whereas an appreciable fraction of β5t was the precursor protein in β5t-G49S–expressing cells (Figure 1B). An additional signal with an intermediate molecular size was also detectable (Figure 1B). These results indicate that human β5t protein with the G49S variation is inefficient in the processing of the propeptide to generate a mature β5t subunit that exposes the catalytically active triple threonine amino terminus. Unlike WT β5t protein, β5t-G49S protein in a fully assembled proteasome complex either remains unprocessed or undergoes abnormal processing.

Generation of β5t-G49S–knock-in mice. The inefficient processing of human β5t-G49S protein in cells prompted us to wonder whether humans carrying the G49S variants might have a problem in the processing of β5t protein in their cTECs and further have consequences in the generation of $CD8^+$ T cells. In order to begin addressing the effects of G49S variation in vivo, we next aimed to generate knock-in mice that carried the genome sequence corresponding to rs34457782, which encoded β5t-G49S protein in mice.

We decided that the introduction of the G49S variation in the mouse genome was a reasonable approach to explore the effect of this variation in humans in vivo because (i) the 49th amino acid, Gly, was highly conserved among vertebrate species, including humans and mice (Table 1), as well as among many mouse strains (Table 2), and (ii) the processing of mouse β5t-G49S protein expressed in mouse embryonic fibroblasts (MEFs) was inefficient and aberrant, similarly to the processing of human β5t-G49S protein (Figure 1C).

To generate knock-in mice that carried the genome sequence corresponding to rs34457782, a targeting vector was introduced into mouse ES cells for homologous recombination, and the ES

Table 2. β5t amino acid sequences at positions 45–60 in different mouse strains obtained from the Ensemble database

Mouse strain	Amino acid sequence
129S1/SvImJ	*RLAH***G***TTTLAFRFRHG*
BALB/cJ	*RLAH***G***TTTLAFRFRHG*
C57BL/6	*RLAH***G***TTTLAFRFRHG*
CBA/J	*RLAH***G***TTTLAFRFRHG*
DBA2/J	*RLAH***G***TTTLAFRFRHG*
C3H/HeJ	*RLAH***G***TTTLAFRFRHG*
CAST/EiJ	*RLAH***G***TTTLAFRFRHG*
FVB/NJ	*RLAH***G***TTTLAFRFRHG*
LP/J	*RLAH***G***TTTLAFRFRHG*
NOD/ShLtJ	*RLAH***G***TTTLAFRFRHG*
NZO/HlLtJ	*RLAH***G***TTTLAFRFRHG*
PWK/PhJ	*RLAH***G***TTTLAFRFRHG*
SPRET/EiJ	*RLAH***G***TTTLAFRFRHG*
WSB/EiJ	*RLAH***G***TTTLAFRFRHG*

Glycine at position 49 is conserved in all strains and shown in bold letters. Other conserved amino acids are highlighted in italicized letters.

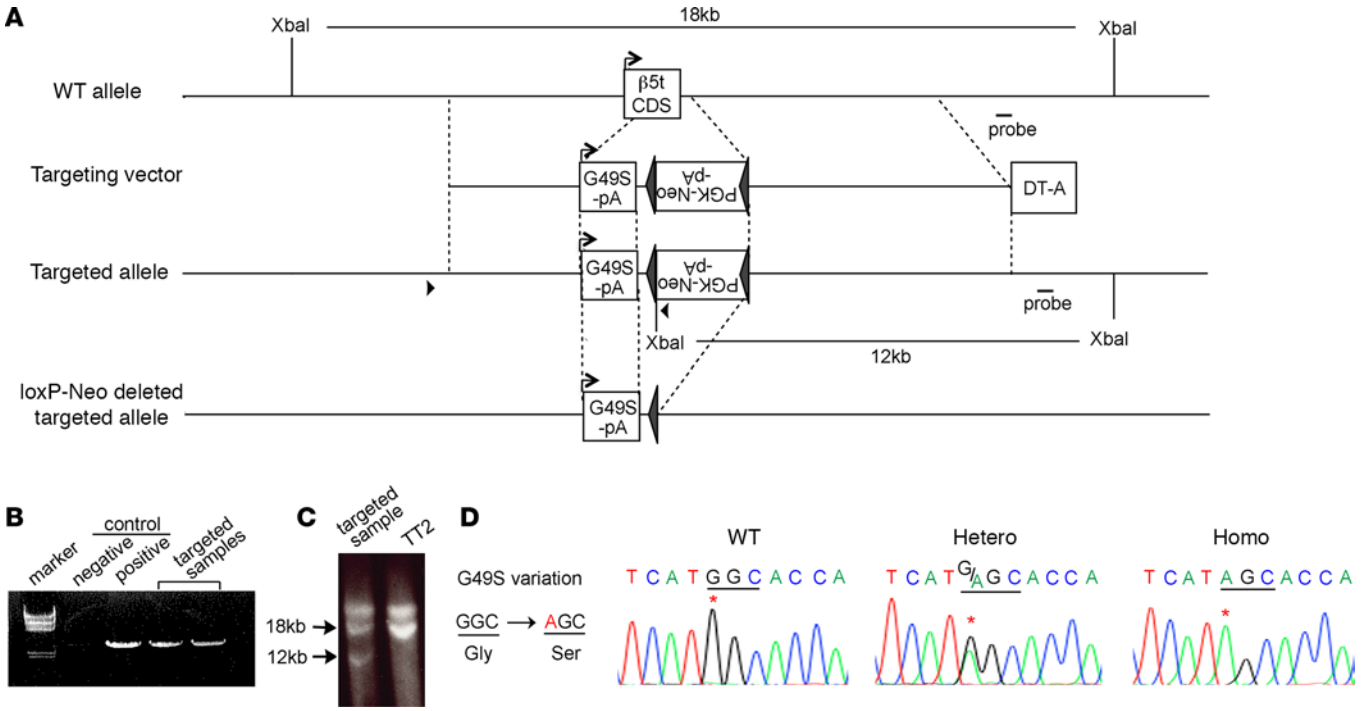

Figure 2. Generation of β5t-G49S–knock-in mice. (**A**) Schematic diagram of β5t genomic locus in mouse chromosome 14. The diagrams for targeting vector and targeted allele are also shown. Targeted mice were further crossed with CAG-Cre–transgenic mice to delete the PGK-Neo insertion. Arrowheads indicate the primers for genomic PCR. Probe for Southern blot analysis is also shown. DT-A, diphtheria toxin A; PGK-Neo, phosphoglycerate kinase I promoter–driven neomycin resistance gene. (**B**) Gel electrophoresis of PCR-amplified genomic DNA. Positions for primers are shown in **A**. Plasmid DNA containing the targeted fragment was used as positive control. (**C**) Southern blot analysis of XbaI-digested genomic DNA from targeted ES cells and control TT2 ES cells. Probe is shown in **A**. (**D**) The G-to-A 1 bp substitution at amino acid position 49 (Gly to Ser) of β5t in genomic DNA isolated from mouse tails in WT, heterozygous (hetero), and homozygous (homo) knock-in mice. See complete unedited blots in the supplemental material. Asterisks indicate nucleotides where 1-bp substitution is introduced.

cells were positively and negatively selected for the incorporation of the targeted allele (Figure 2A). The ES cells were screened by PCR analysis (Figure 2B), confirmed by Southern blot analysis (Figure 2C), and verified by sequence analysis (Figure 2D). Three founder mice were mated with CAG-Cre–transgenic mice for germline removal of the loxP-Neo-loxP cassette (Figure 2A). All of the 3 mouse strains generated in this study exhibited essentially identical phenotypes; here, we describe the following results from one strain of knock-in mice.

G49S variation in mice diminishes β5t expression in cTECs. β5t-G49S–knock-in mice, both heterozygotes and homozygotes, were born and grew to be fertile without apparent problems. Biochemical analysis of β5t-G49S–expressing cells indicated that the human and mouse β5t-G49S proteins were inefficient and aberrant in the processing to generate mature β5t protein (Figure 1), suggesting that in β5t-G49S–knock-in mice, the processing of mature β5t protein in cTECs might be inefficient and aberrant, which could in turn affect the production of $CD8^+$ T cells. Consequently, we examined the thymic architecture in β5t-G49S–knock-in mice, focusing on the expression of β5t proteins in cTECs.

H&E staining of the thymic sections showed that the corticomedullary architecture of the thymus was generated in heterozygous and homozygous β5t-G49S–knock-in mice, without noticeable abnormalities (Figure 3A). Flow cytometric analysis of liberase-digested thymic cells indicated that the numbers of thymic cells, including cTECs and medullary TECs (mTECs), were normal in heterozygous and homozygous β5t-G49S–knock-in mice (Figure 3B). Immunofluorescence analysis of the thymic sections revealed the expression of β5t proteins in the cortex and of Aire proteins in the medulla in heterozygous and homozygous β5t-G49S–knock-in mice (Figure 3C). However, the expression of β5t proteins in cTECs was significantly reduced in heterozygous ($P < 0.05$) and more severely in homozygous ($P < 0.01$) β5t-G49S–knock-in mice (Figure 3D), indicating that G49S variation affects the expression of β5t proteins for the optimal expression of β5t-containing thymoproteasomes in cTECs in vivo. In contrast, the amounts of MHC class I and class II molecules expressed in cTECs and mTECs were not altered in heterozygous and homozygous β5t-G49S–knock-in mice (Figure 3, E and F). These results indicate that

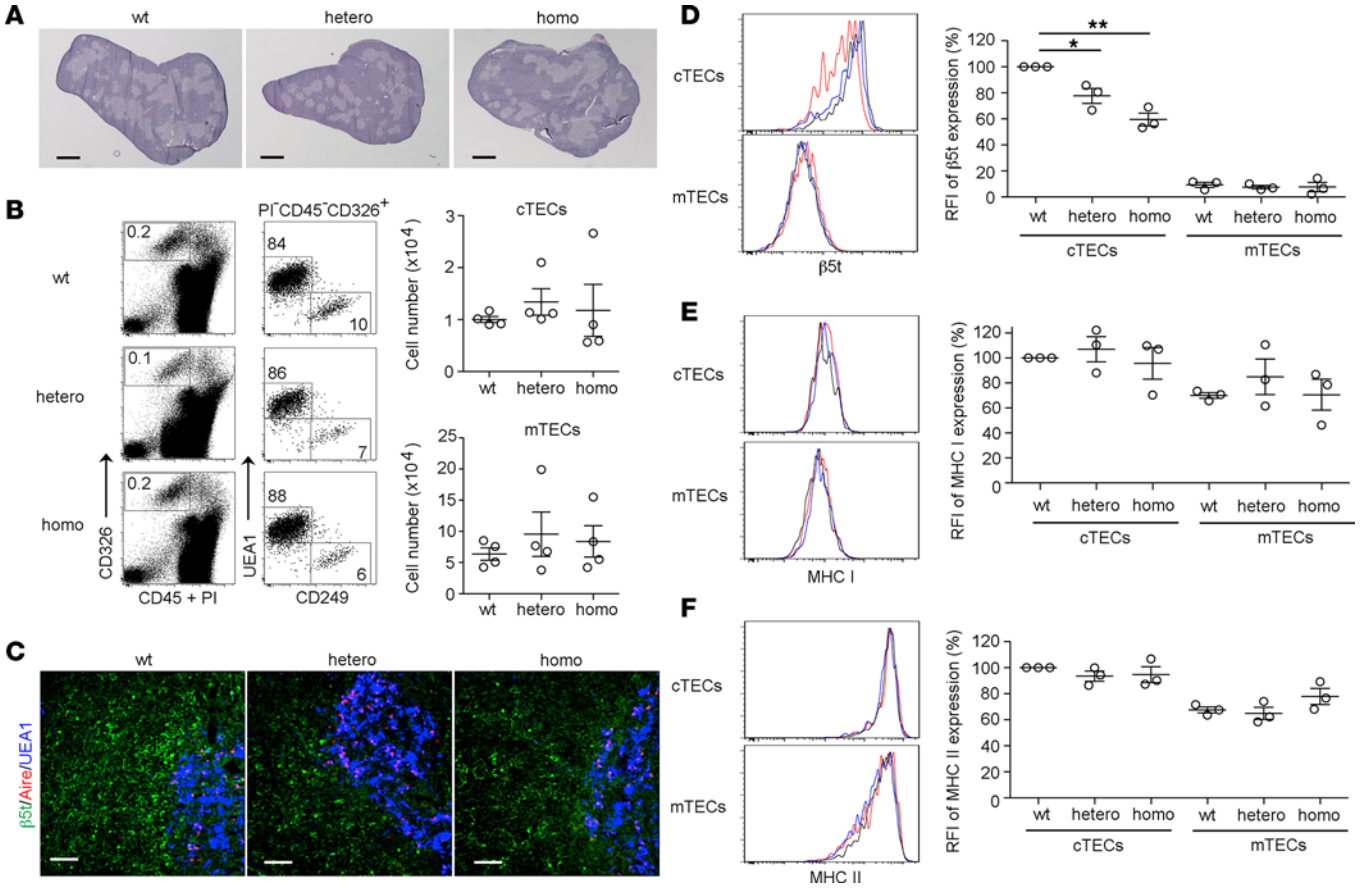

Figure 3. G49S variation in mice diminishes β5t expression in cTECs. (**A**) H&E staining of thymic sections from indicated mice at 5 weeks old. Representative data from 3 independent experiments are shown. Scale bars: 1 mm. (**B**) Flow cytometric analysis of enzyme-digested thymic cells isolated from 2-week-old mice of indicated genotypes. Dot plots show CD326 (EpCAM) and CD45 expression in total thymic cells (left), and UEA-1 reactivity and CD249 expression in propidium iodide⁻ (PI⁻) CD45⁻CD326⁺ viable thymic epithelial cells (middle). Graphs show cell number (means ± SEM, n = 4) of PI⁻CD45⁻CD326⁺UEA1⁻CD249⁺ cTECs and PI⁻CD45⁻CD326⁺UEA1⁺CD249⁻ mTECs. (**C**) Immunofluorescence analysis of β5t (green), Aire (red), and UEA-1 binding molecules (blue) in thymic sections from 5-week-old mice of indicated genotypes. Representative data from 3 independent experiments are shown. Scale bars: 75 μm. (**D**–**F**) Histograms show the expression of β5t (**D**), MHC class I (**E**), and MHC class II (**F**) in cTECs and mTECs of WT (black lines), heterozygous (blue lines), and homozygous (red lines) knock-in mice at 2 weeks old. Graphs show relative fluorescence intensity (RFI, n = 3) of β5t (**D**), MHC class I (**E**), and MHC class II (**F**) expression normalized to the mean fluorescence intensity measured in WT cells. *P < 0.05, **P < 0.01 by one-way ANOVA with Tukey's correction.

G49S variation in mice decreases β5t expression in cTECs in heterozygotes and homozygotes, and the decrease is greater in the latter.

β5t-G49S variation in mice impairs CD8⁺ T cell production in thymus. The reduced expression of β5t proteins in cTECs suggested the possibility that the production of CD8⁺ T cells in the thymus might be affected in β5t-G49S–knock-in mice. To examine this possibility, we performed flow cytometric analysis of thymocytes from β5t-G49S–knock-in mice. The analysis of CD4, CD8, and TCRβ expression profiles indicated that the frequencies and numbers of CD4⁻CD8⁻, CD4⁺CD8⁺ (double positive, DP), and CD4⁺CD8⁻ TCRβ^{hi} thymocytes were unaltered in heterozygous and homozygous β5t-G49S–knock-in mice (Figure 4A). However, the cellularity of CD4⁻CD8⁺ TCRβ^{hi} thymocytes was significantly (P < 0.05) reduced in homozygous, but not heterozygous, β5t-G49S–knock-in mice (Figure 4A). Within DP thymocytes, the TCRβ^{hi}CD5^{intermediate} DP3 population, which is enriched with DP thymocytes that are destined to become CD4⁻CD8⁺ TCRβ^{hi} thymocytes (16), and which is likely enriched with the postselection CD8-lineage "stage 5" DP thymocytes (17), was specifically reduced in cellularity (Figure 4B). These results indicate that the production of CD4⁻CD8⁺ TCRβ^{hi} thymocytes is specifically reduced in homozygous, but not heterozygous, β5t-G49S–knock-in mice.

Consequently, the cellularity of CD4⁻CD8⁺ TCRβ^{hi} T cells, but not CD4⁺CD8⁻ TCRβ^{hi} T cells in the periphery, including the spleen, was reduced in homozygous, but not heterozygous, β5t-G49S–knock-in mice (Figure 4C). The reduction in cellularity of CD8⁺ T cells was associated with the reduced number

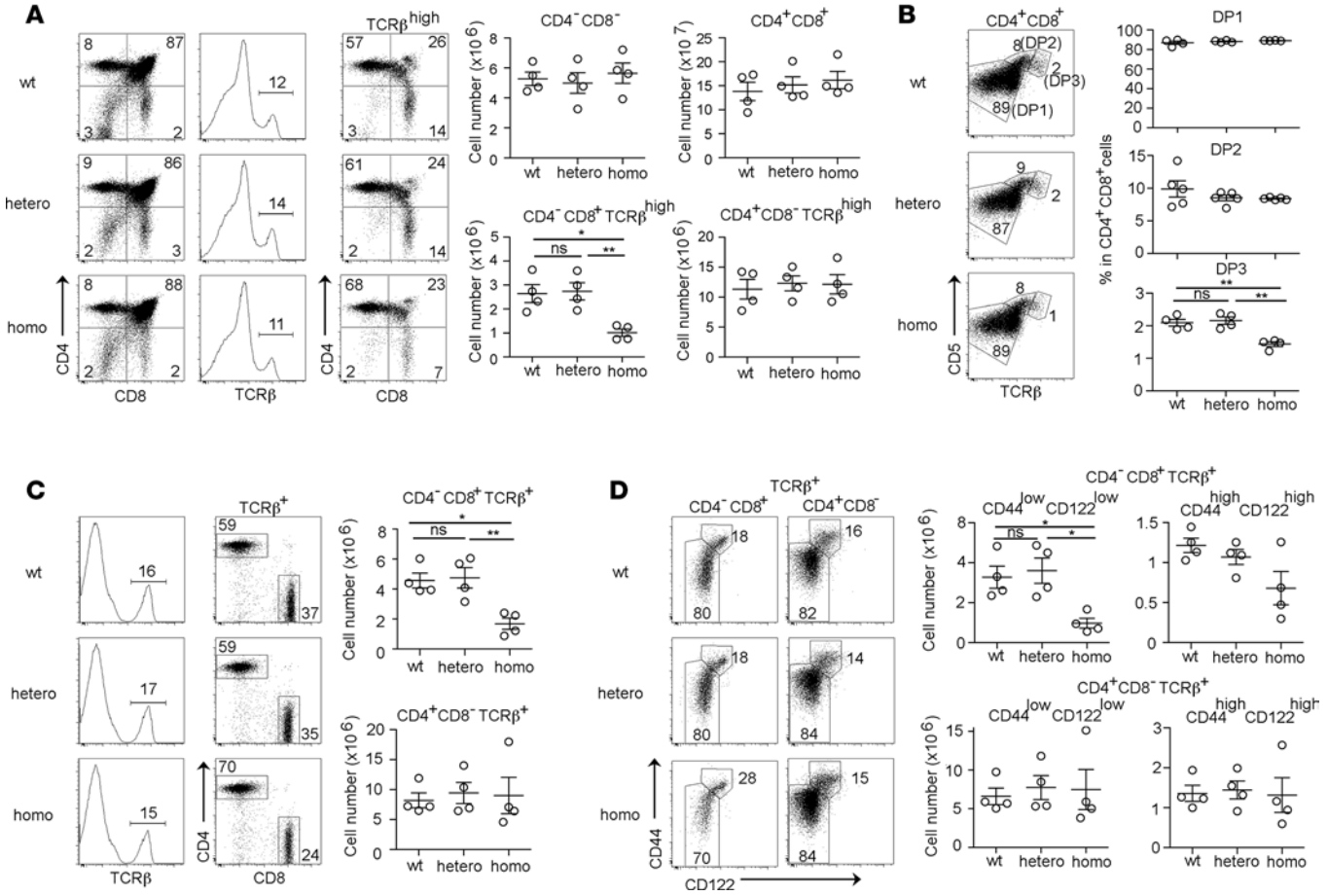

Figure 4. β5t-G49S variation in mice impairs CD8⁺ T cell production in thymus. (**A**) Flow cytometric analysis of thymocytes from 5-week-old mice of indicated genotypes. Shown are dot plots of CD4 and CD8 expression (left), TCRβ expression (middle) in PI⁻ viable cells, and dot plots of CD4 and CD8 expression in PI⁻TCRβ^{hi} cells (right). Graphs show the number (means ± SEM, *n* = 4) of indicated thymocyte populations. (**B**) Flow cytometric analysis of thymocytes from 5-week-old mice of indicated genotypes. Shown are dot plots of TCRβ and CD5 expression in CD4⁺CD8⁺ thymocytes. Boxes indicate TCRβ^{lo}CD5^{lo} (DP1), TCRβ^{intermediate}CD5^{hi} (DP2), and TCRβ^{hi}CD5^{intermediate} (DP3) CD4⁺CD8⁺ thymocytes. Graphs show the frequency (means ± SEM, *n* = 4) of indicated thymocyte populations. (**C**) Flow cytometric analysis of splenocytes from 5-week-old mice of indicated genotypes. Histograms show TCRβ expression in PI⁻ viable cells, and dot plots show CD4 and CD8 expression in PI⁻TCRβ^{hi} viable T cells. Graphs show the number (means ± SEM, *n* = 4) of CD4⁺CD8⁻ TCRβ^{hi} T cells and CD4⁻CD8⁺ TCRβ^{hi} T cells. (**D**) Flow cytometric analysis of splenocytes from 5-week-old mice of indicated genotypes. Shown are dot plots of CD44 and CD122 expression in CD4⁻CD8⁺ TCRβ^{hi} and CD4⁺CD8⁻ TCRβ^{hi} T cells. Graphs show the number (means ± SEM, *n* = 4) of indicated splenocyte populations. Numbers in dot plots and histograms indicate frequency of cells within indicated area. **P* < 0.05, ***P* < 0.01 by one-way ANOVA with Tukey's correction.

of CD44^{lo}CD122^{lo} naive CD8⁺ T cells (Figure 4D), similar to β5t-deficient mice (6), in agreement with the reduction in the thymic production of CD8⁺ T cells. Nevertheless, the numbers of CD8⁺ T cells in the thymus and the spleen of homozygous β5t-G49S–knock-in mice were reduced to approximately 40% of those in control mice (Figure 4, A and C), and the reduction was less severe than that in β5t-deficient mice (Figure 5). These results indicate that the production of CD8⁺ T cells is reduced in homozygous, but not heterozygous, β5t-G49S–knock-in mice, and the reduction in CD8⁺ T cell production is not as severe as that in β5t-deficient mice.

β5t-G49S variation in association studies in human populations. The results described so far indicated that the β5t-G49S variation caused the reduction in β5t expression and in CD8⁺ T cell production in mice, which prompted us to address whether β5t-G49S variation in humans might actually affect CD8⁺ T cell production and, more importantly, human health. As CD8⁺ T cells are important for immune responses toward infectious diseases, including viral infection, we next examined whether the rs34457782 variation might be associated with the susceptibility to hepatitis B, hepatitis C, and tuberculosis in human populations. We compared the minor allele frequency of the rs34457782 variation among Japanese healthy volunteers (417 samples), Japanese patients with hepatitis B (521 samples) and

Table 3. Susceptibility to hepatitis B and hepatitis C in humans carrying β5t-G49S variation

Genotype	Controls		Cases					
			Hepatitis B		Hepatitis C		Spontaneous clearance of hepatitis C	
	Number (*n* = 417)	Frequency (%)	Number (*n* = 521)	Frequency (%)	Number (*n* = 2,537)	Frequency (%)	Number (*n* = 785)	Frequency (%)
G/G	393	94.24	491	94.24	2,405	94.80	741	94.39
G/A	24	5.76	29	5.57	132	5.20	43	5.48
A/A	0	0	1	0.19	0	0	1	0.13
Major G allele	810	97.12	1,011	97.03	4,942	97.40	1,525	97.13
Minor A allele	24	2.88	31	2.98	132	2.60	45	2.87

Genotype, number, and allele frequency in 521 hepatitis B patients, 2,537 hepatitis C patients, 785 individuals who underwent spontaneous clearance of hepatitis C virus, and 417 healthy controls in Japan are shown. Statistical analyses performed by Fisher's exact test indicate no significant difference in minor allele frequency. The observed genotype distribution showed no significant deviation from the Hardy-Weinberg equilibrium (hepatitis B patients, $\chi^2 = 0.67$, degrees of freedom [df] = 1, $P = 0.41$; hepatitis C patients, $\chi^2 = 0.81$, df = 1, $P = 0.18$; individuals who underwent spontaneous clearance of hepatitis C, $\chi^2 = 0.21$, df = 1, $P = 0.65$; controls, $\chi^2 = 0.37$, df = 1, $P = 0.55$).

hepatitis C (2,537 samples), and Japanese individuals who underwent spontaneous clearance of hepatitis C virus (785 samples) (18, 19). We also compared the minor allele frequency of the rs34457782 variation between Indonesian healthy controls (514 samples) and Indonesian patients with pulmonary tuberculosis (573 samples) (20). We detected no association with this variation in any of those infectious diseases (Table 3 and Table 4), indicating that the minor allele of the rs34457782 variation is not associated with the susceptibility to infection with hepatitis B virus, hepatitis C virus, and Mycobacterium tuberculosis. These results suggest that humans who are heterozygous in the β5t-G49S variant are not defective in the immune responses to these infectious pathogens, in agreement with the results of the analysis of heterozygous β5t-G49S–knock-in mice showing that the heterozygotes are not impaired in the production of normal cellularity of $CD8^+$ T cells.

It should also be noted that the analysis described in Table 3 included 2 homozygotes in 4,260 Japanese people, indicating that the homozygous β5t-G49S variation is not lethal. Both of these 2 homozygotes experienced either hepatitis B or hepatitis C, even though the small number of homozygotes hampers the statistically valid analysis of the susceptibility of homozygotes to individual diseases.

Cohort study identifies 5 individuals homozygous in β5t-G49S variation. We finally examined the β5t-G49S variation in a cohort of people in Nagahama City, Japan (21, 22). Among 9,730 participants of the age range between 30 and 74 years old, 559 were heterozygous and 5 were homozygous in the β5t-G49S variation (Table 5). The minor allele frequency was 2.92%, similar to the frequencies in Japanese populations in Tokyo and Tokushima (Supplemental Table 2 and Supplemental Table 3). The presence of 5 homozygotes in the 9,730 volunteers (at $0.051\% = 2.3\%^2$ frequency) indicated that the β5t-G49S variation in humans is neither lethal nor severely excluded in the population. The frequency and cellularity of peripheral blood $CD4^+$ T cells and $CD8^+$ T cells, including naive $CD8^+$ T cells, were not significantly different between the major allele homozygotes (209 samples) and the minor allele heterozygotes (12 samples) (Figure 6A), indicating that $CD4^+$ and $CD8^+$ T cells are not reduced in cellularity in the heterozygotes. So far, the data on $CD8^+$ T cells are not available for the minor allele homozygotes. Nonetheless, as far as we are aware of, neither the heterozygotes nor the homozygotes were significantly associated with any severe diseases, including hepatitis B, hepatitis C, tuberculosis, and cancer (Figure 6B). Furthermore, we did not observe any increases in the frequency of empyema, periodontitis, or asthmas (Figure 6B). The questionnaire survey did not reveal any significant problems in cold, cold-associated coughs, or sputum without cold among the heterozygotes or the homozygotes (Figure 6C), although the baseline survey detected a significant increase in coughs after colds in the homozygotes (Figure 6C). Thus, no severe health problems were noted in the heterozygous and homozygous β5t-G49S variation detectable in the human cohort.

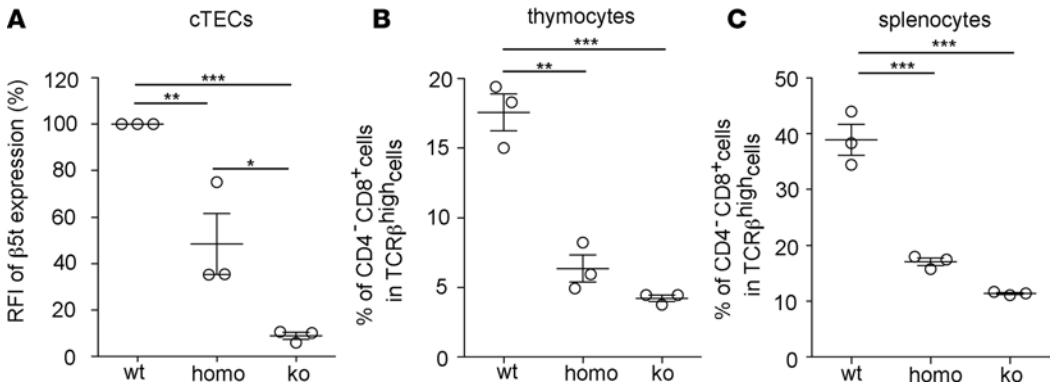

Figure 5. Less severe reduction in β5t expression and CD8⁺ T cells in β5t-G49S–knock-in mice than β5t-deficient mice. (**A**) Relative fluorescence intensity (RFI, means ± SEM, *n* = 3) of β5t expression in cTECs from 2-week-old β5t-G49S homozygous–knock-in mice (homo) and β5t-deficient mice (KO), normalized to the mean fluorescence intensity measured in WT cTECs. (**B** and **C**) Frequency (means ± SEM, *n* = 3) of $CD4^-CD8^+$ cells in $TCR\beta^{hi}$ cells of thymocytes (**B**) and splenocytes (**C**) in 6-week-old WT mice, β5t-G49S homozygous–knock-in mice (homo), and β5t-deficient mice (KO). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ by one-way ANOVA with Tukey's correction.

Discussion

The rs34457782 single nucleotide variant, which is present at an appreciable frequency in human populations, changes the 49th amino acid of β5t protein, Gly, to Ser (G49S). Our results indicate that human β5t protein with the G49S variation is inefficient and aberrant in the processing of the propeptide to generate the mature β5t subunit that exposes the catalytically active triple threonine amino terminus. Unlike WT β5t protein, an appreciable fraction of β5t-G49S protein incorporated in fully assembled proteasome complex either remains unprocessed or undergoes abnormal processing. The Gly at the 49th amino acid is highly conserved among vertebrate species, including humans and mice, and the processing of mouse β5t-G49S protein is also inefficient and aberrant, similarly to the processing of human β5t-G49S protein. By producing knock-in mice, we showed that the G49S variation in the mouse decreased β5t expression in cTECs in heterozygotes and homozygotes, with the decrease being greater in the latter. The reduction of β5t protein in cTECs might be due to the instability and/or the toxicity of β5t-G49S–containing proteasome complex in the cells. The production of $CD8^+$ T cells was significantly impaired in homozygous, but not heterozygous, β5t-G49S–knock-in mice. These results indicate that the G49S variation in the genome affects the expression of β5t-containing thymoproteasomes in cTECs and the thymic production of $CD8^+$ T cells in vivo, at least in mice.

However, no apparent health problems, including the susceptibility to hepatitis B virus, hepatitis C virus, or Mycobacterium tuberculosis, were detected in humans heterozygous in the rs34457782 variant, which may not be surprising, given the detection of no apparent problems in the immune cells, including the unreduced cellularity of $CD8^+$ T cells, in the heterozygous β5t-G49S–knock-in mice. A publicly available database from

Table 4. Susceptibility to tuberculosis in humans carrying β5t-G49S variation

Genotype	Controls		Cases: Tuberculosis	
	Number (*n* = 514)	Frequency (%)	Number (*n* = 573)	Frequency (%)
G/G	505	98.25	567	98.95
G/A	9	1.75	6	1.05
A/A	0	0	0	0
Major G allele	1,019	99.12	1,140	99.48
Minor A allele	9	0.88	6	0.52

Genotype, number, and allele frequency in 573 tuberculosis patients and 514 healthy controls in Indonesia are shown. Statistical analyses performed by Fisher's exact test indicate no significant difference in minor allele frequency. The observed genotype distribution showed no significant deviation from the Hardy-Weinberg equilibrium (tuberculosis patients, $\chi^2 = 0.02$, degrees of freedom [df] = 1, $P = 0.90$; controls, $\chi^2 = 0.04$, df = 1, $P = 0.84$).

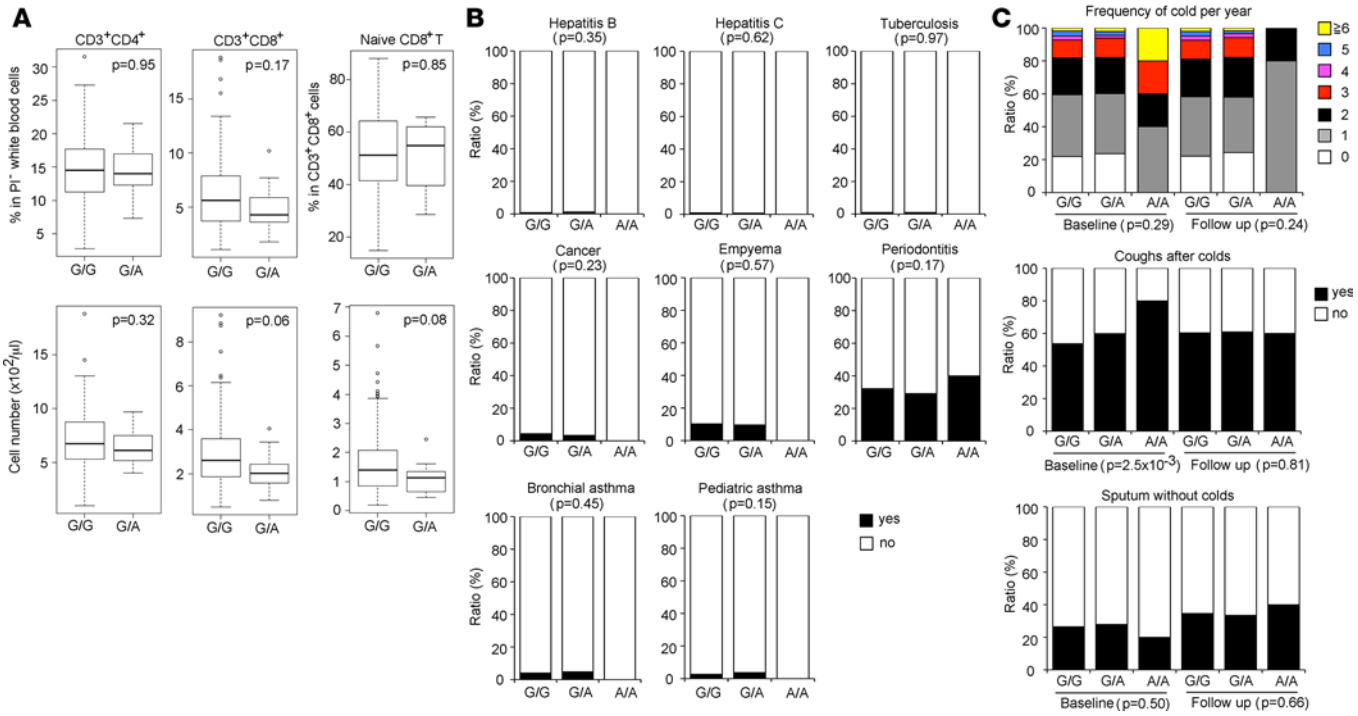

Figure 6. Nagahama study for β5t-G49S variation in humans. (**A**) Graph shows the frequency (top) and number (bottom) of indicated cell populations in peripheral white blood cells among 209 major allele homozygotes (G/G) and 12 minor allele heterozygotes (G/A). (**B** and **C**) Questionnaire survey of anamnesis was carried out in 9,166 major allele homozygotes (G/G), 559 minor allele heterozygotes (G/A), and 5 minor allele homozygotes (A/A) at the baseline assessment. The follow-up survey was carried out in 7,788 major allele homozygotes (G/G), 466 minor allele heterozygotes (G/A), and 5 minor allele homozygotes (A/A). Shown are the ratio of volunteers who experienced indicated diseases (filled bars in **B**) and the ratio of volunteers who caught colds for the indicated times per year (top), who experienced coughs after a cold (middle), and who noticed sputum without catching a cold (bottom) in 2 consecutive surveys (**C**). Logistic regression analyses and Jonckheere-Terpstra tests for the trend test indicated no statistically significant differences among the groups.

genome-wide association studies has revealed no apparent association of the β5t-G49S variation with any diseases so far examined (DisGeNET v4.0, http://www.disgenet.org/web/DisGeNET/menu/search). Indeed, we found that the number of $CD8^+$ T cells in heterozygous humans was not reduced. Nonetheless, the significant reduction in β5t expression in cTECs (approximately 80% of the amount in WT cTECs), which was detectable in heterozygous mice, might still affect TCR repertoire and T cell function in $CD8^+$ T cells produced in heterozygous humans.

On the other hand, it was rather surprising to readily find no severe health problems or no apparent association with various diseases in the 5 people who were identified as minor allele homozygotes in the cohort study. At the moment, the data on their $CD8^+$ T cells and other immune cells in those homozygotes are not available, although the data for homozygous β5t-G49S–knock-in mice indicate that the cellularity of their $CD8^+$ T cells is reduced to approximately 40% of those in the major-allele homozygotes. Thus, it is speculated that those homozygous people may also show reduced production of functionally optimal $CD8^+$ T cells and that they may experience reduced immune responses where $CD8^+$ T cells are essential.

Nevertheless, we would like to point out that the minor allele frequency of this variant is appreciable at 2.92% in a cohort in Japan (9,730 volunteers examined). The frequency of the 5 homozygotes found in the 9,730 volunteers (0.051%) suggests that the β5t-G49S variation has little or no negative impact on the human populations. It is important to point out that this frequency of the homozygotes is overwhelmingly larger than those of extremely rare variations (reported in less than 30 patients worldwide) known to affect MHC class I-dependent generation of $CD8^+$ T cells in humans (23–25). Those rare variations include the mutations deficient in the peptide transporter complex TAP and TAP-associated protein TAPASIN, in which patients often suffer from respiratory tract infection and skin granulomatous lesions but sometimes exhibit no clinical problems (23–25). Based on the results of this study, we would like to propose a careful and long-term analysis of health status in those rs34457782 minor allele populations, particularly in homozygotes, that carry the β5t-G49S variant, in hopes of better understanding the role of the thymoproteasome-dependent positive selection of $CD8^+$ T cells in humans.

Table 5. Nagahama study for β5t-G49S variation in humans

Genotype	Number	Frequency (%)
G/G	9,166	94.20
G/A	559	5.75
A/A	5	0.05
Major G allele	18,891	97.08
Minor A allele	569	2.92

Genomic DNA from 9,730 volunteers participating in the Nagahama study were genotyped. Genotype, number, and allele frequency are indicated. The observed genotype distribution showed no significant deviation from the Hardy-Weinberg equilibrium (χ^2 = 1.41, degrees of freedom [df] = 1, P = 0.24).

Methods

Public database analysis. Single nucleotide variants of human β5t-encoding sequence are available in the Single Nucleotide Polymorphism database (https://www.ncbi.nlm.nih.gov/snp/) hosted by the National Center for Biotechnology Information (NCBI, dbSNP Human Build 141). The G49S variation among various populations is available in the 1000 Genome database in the NCBI public website (https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/). Possible disease association of rs34457782 variant was analyzed using DisGeNET v4.0 (http://www.disgenet.org/web/DisGeNET/menu/search).

Human genome analysis. We initially performed a pilot sequencing of the β5t-encoding region in human genome samples obtained from 188 healthy volunteers collected at the University of Tokushima (Supplemental Table 3). After detecting 13 heterozygotes in the rs34457782 variant, we then examined this the in 949 healthy volunteers collected at the University of Tokushima (10, 11), by using TaqMan genotyping assay (Applied Biosystems) in accordance with the manufacturer's instructions.

Human genome analysis. SNP genotyping was conducted by using the TaqMan genotyping assay (Applied Biosystems) in accordance with the manufacturer's instructions (Table 3 and Table 4). We genotyped 521 Japanese patients with hepatitis B, 2,537 Japanese patients with hepatitis C, and 785 Japanese individuals with hepatitis C virus clearance, as well as 417 healthy Japanese patients. We also genotyped 573 Indonesian patients with tuberculosis and 514 healthy Indonesians.

Human epidemiological analysis. We analyzed a dataset of the Nagahama Prospective Cohort for Comprehensive Human Bioscience (the Nagahama Study) to examine possible associations between rs34457782 and clinical phenotypes in a human population (Figure 6 and Table 5). The Nagahama Study participants were recruited from 2008–2010 from community residents of Nagahama City, a suburban city of approximately 125,000 inhabitants in Shiga Prefecture, which is located in central Japan. Individuals who fulfilled the following criteria were included in this cohort, aged between 30 and 74 years old, living independently in the community, and having no apparent physical impairment or dysfunction. Among a total of 9,769 participants, 9,730 participants whose rs34457782 genotype was available were considered in this study.

Participants in the Nagahama cohort study were invited to a follow-up assessment 5 years after baseline evaluations, and 8,294 of the original 9,769 individuals participated in the follow-up survey. A dataset of the 8,259 participants whose rs34458872 genotype and clinical data were available was analyzed for the follow-up measurements.

DNA was extracted from peripheral blood specimen by a conventional phenol chloroform method, and rs34457782 genotype was analyzed by TaqMan assay using a predesigned primers and probe set (C_2111454_20, Applied Biosystems).

Clinical parameters — including the history of hepatitis B, hepatitis C, tuberculosis, any cancers, maxillary empyema, periodontitis diseases, bronchial asthma, and pediatric asthma — were obtained at the baseline measurement using a questionnaire. Frequency of a cold and related symptoms were asked by the following questions at both baseline and follow-up measurements: How many colds do you have per year? (times), Do you have a cough after a cold? (yes/no), and Do you bring up sputum when you do not have a cold? (yes/no).

Blood samples from the participants of the follow-up assessment in October 2014 were collected into tubes containing EDTA-2Na anticoagulant. Whole blood samples (50 μl) was directly stained with fluorescence-labeled antibodies and incubated at 4°C for 30 minutes. The antibodies used were Pacific Blue–labeled anti-CD3 antibody (clone UCHT1), APC-labeled anti-CD4 antibody (clone RPA-T4), APC-labeled anti-CD8a antibody (clone RPA-T8), Brilliant Violet 421-labeled anti-CD45RA antibody (clone HI100), APC-Cy7–labeled anti-CD45RO antibody (clone UCHL1), and Brilliant Violet 605-labeled anti-CCR7 antibody (clone G043H7) (antibodies from SONY). Mixtures of whole blood and antibodies were incubated with 1 ml of RBC lysis buffer (SONY) at room temperature for 15 minutes. The samples were preserved at 4°C, and propidium iodide (PI) was added 1 minute before the analysis. Data were acquired using a SONY SP6800 Spectral Analyzer, and results were analyzed with FlowJo software. Naive $CD8^+$ T cells were identified as $CCR7^+CD45RA^+CD45RO^-CD3^+CD8^+$ cells in the blood samples. The numbers of viable white blood cells, which were defined by PI^- cells, were measured by a flow cytometry based on electrical resistivity method (Sysmex XE-2100).

Western blot analysis. HEK293T cells (RIKEN Bioresource Center), MEFs (2), and Plat-E cells (26) were cultured in DMEM (Nacalai Tesque) supplemented with 10% FBS (Thermo Fischer Scientific), 100 units/ml penicillin, and 100 μg/ml streptomycin (Nacalai Tesque) at 37°C with 5% CO_2. Genes were transfected by the conventional calcium phosphate method. MEFs were isolated from β5i-deficient mice and immortalized by introducing SV40 large T antigen (27). To generate MEFs stably expressing mouse β5t, pMXs plasmids encoding Flag-tagged β5t WT or G49S were transfected into Plat-E cells to prepare retroviruses (26). β5i-deficient MEFs were infected with the retroviruses and selected with 4 μg/ml puromycin for β5t expression. To generate HEK293T cells stably expressing human β5t, HEK293T cells were transfected with a pIRES plasmid encoding 3× Flag-tagged β5t and selected with 4 μg/ml puromycin. To promote the formation of thymoproteasomes, MEFs were stimulated with mouse IFN-γ (Peprotech) at 50 units/ml for 96 hours. Antibodies against Flag-tag were purchased from Sigma-Aldrich. Other antibodies against proteasome subunits used in this study were described previously (2, 28).

Cells were lysed in ice-cold buffer containing 0.2% NP-40, 25 mM Tris-HCl (pH 7.5), 1 mM dithiothreitol, 2 mM ATP, and 5 mM $MgCl_2$, and clarified by centrifugation at 20,000 *g* for 15 minutes at 4°C. Glycerol gradient centrifugation was performed as described previously (29). The 26S fractions were determined by measuring the hydrolysis of succinyl-Leu-Leu-Val-Tyr-7-amido-4-methylcoumarin (Peptide Institute) (29). For immunoprecipitation, anti–Flag M2 agarose (Sigma-Aldrich) or anti-α6 antibodies bound to Protein G-Sepharose (GE Healthcare) were used. For Western blotting, samples were boiled with SDS sample buffer, separated by SDS-PAGE, and transferred to a polyvinylidene difluoride membrane. Membranes were incubated with Blocking One (Nacalai Tesque) with primary antibodies and then with secondary antibodies conjugated to horseradish peroxidase (Jackson ImmunoResearch, dilution 1:20,000) and visualized by Western Lightning Plus-ECL (Perkin Elmer). Images were obtained with LAS 4000 (GE Healthcare).

Generation of β5t-G49S–knock-in mice. C57BL/6 (B6) mice were obtained from SLC. β5t-deficient mice and CAG-Cre transgenic mice were described previously (2, 30, 31). The targeting vector was prepared by cloning β5t-containing genomic fragments amplified from B6 mouse genomic DNA into pCR-Blunt vector (Invitrogen). Point mutation at the 145th nucleotide (49th amino acid) of the β5t sequence was introduced with a PrimeSTAR mutagenesis basal kit (Takara). The fragment with the mutated β5t-coding sequence was subcloned into a plasmid containing a PGK-Neo cassette. The linearized targeting vector was introduced into TT2 ES cells (32) with 2 gRNAs targeting the β5t locus and the px335 plasmid, which encodes CRISPR and Cas9 nickase (33). The 2 gRNAs also target 5′ and 3′ arms of the targeting vector, and the target sequences are as follows; 5′ arm, ATGCGCGCAAGAGTCACCGA: 3′ arm, ATGGATGGAACCATGCGGGC. Targeted alleles were screened by genomic PCR analysis and Southern blot analysis. The primers used for genomic PCR analysis were 5′- GTGGATGTGAATGCTTACGC - 3′ and 5′- AGAGGTTCACTAGTACTGGC - 3′. Mutation was confirmed by genome sequencing of ES cells. Mice are available to the scientific community (accession no. CDB1255K, http://www2.clst.riken.jp/arg/mutant%20mice%20list.html). β5t-G49S–knock-in mice were crossed with CAG-Cre transgenic mice to remove PGK-Neo cassette. Genomic PCR for genotyping β5t-G49S–knock-in mice, in which PGK-Neo cassette was removed, was carried out by using primers as follows: 5′ - CATGCCACCCACCGTGATGC - 3′ on β5t-coding region and 5′ - GTCAGTCTTAGGACTGGAG - 3′ on β5t 3′ untranslated region. Targeted allele and WT allele were detected at 540 bp and 1.3 kb in size, respectively.

Southern blotting. Genomic DNA extracted from the spleen was digested with XbaI, electrophoresed in 1% agarose, and transferred to nylon membrane (Hybond-N^+, GE Healthcare). The probe was labeled with a PCR DIG Probe Synthesis Kit (Roche Diagnostics), and hybridization was detected using anti–DIG-AP Fab fragment (Roche Diagnostics), CDP-STAR (Roche Diagnostics), and Light Capture II (Atto).

Mouse genome sequencing. Genome DNA was PCR amplified and sequenced by using a Big Dye Terminator V3.1 cycle sequencing kit (Applied Biosystems) and analyzed by Genetic Analyzer 3500 (Applied Biosystems).

Thymic section analysis. Frozen thymuses embedded in OCT compound (Sakura Finetek) were sliced into 5-μm–thick sections. Thymic sections stained with H&E were examined under a light microscope. For immunofluorescence analysis, the thymuses were fixed in 4% (g/vol) paraformaldehyde and embedded in OCT compound. Frozen thymuses were sliced into 5-μm–thick sections and stained with anti-β5t antibody and UEA-1, followed by Alexa Fluor 488– and Alexa Fluor 546–conjugated secondary reagents, respectively, and with Alexa Fluor 660–conjugated anti-Aire antibody (eBioscience, clone 5H12). Images were analyzed with a TSC SP8 confocal laser-scanning microscope (Leica).

Flow cytometric analysis and isolation of thymic epithelial cells. For the analysis of thymocytes and splenocytes, cells were stained for the expression of CD4 (BioLegend, clone RM4-5), CD5 (BioLegend, clone 53-7.3), CD8 (eBioscience, clone 53-6.7), CD44 (eBioscience, clone IM7), CD122 (BioLegend, clone TM-β1), and TCRβ (BioLegend, clone H57-597). Multicolor flow cytometry and cell sorting were performed on FACSAriaII (BD Biosciences, clone 5H4). For the analysis of thymic epithelial cells, minced thymuses were digested with 1 unit/ml liberase (Roche Diagnostics) in the presence of 0.01% DNase I (Roche Diagnostics). Single-cell suspensions were stained for the expression of CD326 (EpCAM, BioLegend, clone G8.8), CD45 (eBioscience, clone 30-F11), CD249 (Ly51, eBioscience, clone 6C3), H-2K^b (BioLegend, clone AF6-88.5), and I-A^b (BioLegend, clone AF6-120.1), and for the reactivity with UEA-1 (Vector Laboratories).

Statistics. Statistical comparison in the analysis of mice was performed with one-way ANOVA. Fisher's exact test in a 2-by-2 contingency table was used to examine the association between rs34457782 and each phenotype (Table 3 and Table 4). In the analysis of the Nagahama cohort dataset, a logistic regression model was adopted for the association analysis between rs34457782 genotype and binomial values, while a Jonckheere-Terpstra test was used for the trend test between the genotype and frequency of catching a cold in a year (Figure 6). Only *P* values less than 0.05 were considered significant.

Study approval. All mouse experiments were performed with consent from the Animal Experimentation Committee of the University of Tokushima (T28-58), the Institutional Animal Care and Use Committee of RIKEN Kobe Branch (AH13-03), and the Institutional Animal Care Committee of the Graduate School of Pharmaceutical Sciences, the University of Tokyo (approval no. M25-19).

Regarding the human genome analysis described in Supplemental Table 3, all study procedures were approved by the Ethics Committee for Human Genome and Gene Research of the University of Tokushima. Written informed consent was obtained from all participants prior to enrollment in the study.

Regarding the human association studies described in Table 3 and Table 4, all study protocols conformed to relevant ethical guidelines as reflected in the a priori approval by the ethics committee of the Faculty of Medicine, The University of Tokyo, and by the ethics committees of all participating universities and hospitals. Written informed consent was obtained from all patients participating in this study, and all samples were anonymized.

Regarding the human cohort studies described in Figure 6 and Table 5, all study procedures were approved by the ethics committee of the Kyoto University Graduate School of Medicine and by the Nagahama Municipal Review Board. Written informed consent was obtained from all participants.

Author contributions

IO, KS, YH, MK, K. Tokunaga, K. Tanaka, FM, SM, and Y. Takahama designed the study; IO, Y. Ohte, HN, AM, and HK performed the experiments; KS, Y. Tabara, YH, MS, TT, NM, HS, Y. Omae, RY, Y. Tanaka, MM, HI, K. Tokunaga, and FM conducted the human analysis; IO, KS, YH, K. Tokunaga, K. Tanaka, FM, SS, and Y. Takahama wrote the manuscript.

Acknowledgments

This study was supported by grants from MEXT-JSPS (24111004, 23249025, and 16H02630 to Y. Takahama, 25860361 and 15K19130 to IO, 25221102 to SM, and 26000014 to K. Tanaka). We also thank Mitsuo Itakura for the advice on human genome analysis and Emi Ikeda and Chiyomi Inoue for technical assistance.

Address correspondence to: Yousuke Takahama at Division of Experimental Immunology, Institute of Advanced Medical Sciences, University of Tokushima, 3-18-15 Kuramoto, Tokushima 770-8503, Japan. Phone: 81.88.633.9452; E-mail: takahama@genome.tokushima-u.ac.jp.

1. Klein L, Kyewski B, Allen PM, Hogquist KA. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol*. 2014;14(6):377–391.
2. Murata S, et al. Regulation of CD8+ T cell development by thymus-specific proteasomes. *Science*. 2007;316(5829):1349–1353.
3. Takahama Y, Takada K, Murata S, Tanaka K. β5t-containing thymoproteasome: specific expression in thymic cortical epithelial cells and role in positive selection of CD8+ T cells. *Curr Opin Immunol*. 2012;24(1):92–98.
4. Nitta T, et al. Thymoproteasome shapes immunocompetent repertoire of CD8+ T cells. *Immunity*. 2010;32(1):29–40.
5. Xing Y, Jameson SC, Hogquist KA. Thymoproteasome subunit-β5T generates peptide-MHC complexes specialized for positive

selection. *Proc Natl Acad Sci USA*. 2013;110(17):6979–6984.
6. Takada K, et al. TCR affinity for thymoproteasome-dependent positively selecting peptides conditions antigen responsiveness in CD8(+) T cells. *Nat Immunol*. 2015;16(10):1069–1076.
7. Tomaru U, et al. Exclusive expression of proteasome subunit {beta}5t in the human thymic cortex. *Blood*. 2009;113(21):5186–5191.
8. Ströbel P, et al. Corticomedullary differentiation and maturational arrest in thymomas. *Histopathology*. 2014;64(4):557–566.
9. Ripen AM, Nitta T, Murata S, Tanaka K, Takahama Y. Ontogeny of thymic cortical epithelial cells expressing the thymoproteasome subunit β5t. *Eur J Immunol*. 2011;41(5):1278–1287.
10. Hamada D, et al. Association between single-nucleotide polymorphisms in the SEC8L1 gene, which encodes a subunit of the exocyst complex, and rheumatoid arthritis in a Japanese population. *Arthritis Rheum*. 2005;52(5):1371–1380.
11. Takata Y, et al. Genetic association between the PRKCH gene encoding protein kinase Ceta isozyme and rheumatoid arthritis in the Japanese population. *Arthritis Rheum*. 2007;56(1):30–42.
12. Hirano Y, et al. Dissecting beta-ring assembly pathway of the mammalian 20S proteasome. *EMBO J*. 2008;27(16):2204–2213.
13. Seemuller E, Lupas A, Baumeister W. Autocatalytic processing of the 20S proteasome. *Nature*. 1996;382(6590):468–471.
14. Chen P, Hochstrasser M. Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. *Cell*. 1996;86(6):961–972.
15. Schmidtke G, et al. Analysis of mammalian 20S proteasome biogenesis: the maturation of beta-subunits is an ordered two-step mechanism involving autocatalysis. *EMBO J*. 1996;15(24):6887–6898.
16. Saini M, Sinclair C, Marshall D, Tolaini M, Sakaguchi S, Seddon B. Regulation of Zap70 expression during thymocyte development enables temporal separation of CD4 and CD8 repertoire selection at different signaling thresholds. *Sci Signal*. 2010;3(114):ra23.
17. Kimura MY, et al. Timing and duration of MHC I positive selection signals are adjusted in the thymus to prevent lineage errors. *Nat Immunol*. 2016;17(12):1415–1423.
18. Tanaka Y, et al. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet*. 2009;41(10):1105–1109.
19. Nishida N, et al. Genome-wide association study confirming association of HLA-DP with protection against chronic hepatitis B and viral clearance in Japanese and Korean. *PLoS One*. 2012;7(6):e39175.
20. Yuliwulandari R, et al. Association of HLA-A, -B, and -DRB1 with pulmonary tuberculosis in western Javanese Indonesia. *Hum Immunol*. 2010;71(7):697–701.
21. Terao C, et al. A genome-wide association study of serum levels of prostate-specific antigen in the Japanese population. *J Med Genet*. 2014;51(8):530–536.
22. Tabara Y, et al. The causal effects of alcohol on lipoprotein subfraction and triglyceride levels using a Mendelian randomization analysis: The Nagahama study. *Atherosclerosis*. 2017;257:22–28.
23. Hanna S, Etzioni A. MHC class I and II deficiencies. *J Allergy Clin Immunol*. 2014;134(2):269–275.
24. Hanalioglu D, Ayvaz DC, Ozgur TT, van der Burg M, Sanal O, Tezcan I. A novel mutation in TAP1 gene leading to MHC class I deficiency: Report of two cases and review of the literature [published online ahead of print February 2, 2017]. *Clin Immunol*. https://doi.org/10.1016/j.clim.2017.01.011.
25. Yabe T, et al. A subject with a novel type I bare lymphocyte syndrome has tapasin deficiency due to deletion of 4 exons by Alu-mediated recombination. *Blood*. 2002;100(4):1496–1498.
26. Kitamura T. New experimental approaches in retrovirus-mediated expression screening. *Int J Hematol*. 1998;67(4):351–359.
27. Zhu JY, Abate M, Rice PW, Cole CN. The ability of simian virus 40 large T antigen to immortalize primary mouse embryo fibroblasts cosegregates with its ability to bind to p53. *J Virol*. 1991;65(12):6872–6880.
28. Uechi H, Hamazaki J, Murata S. Characterization of the testis-specific proteasome subunit α4s in mammals. *J Biol Chem*. 2014;289(18):12365–12374.
29. Akahane T, Sahara K, Yashiroda H, Tanaka K, Murata S. Involvement of Bag6 and the TRC pathway in proteasome assembly. *Nat Commun*. 2013;4:2234.
30. Araki K, Araki M, Miyazaki J, Vassalli P. Site-specific recombination of a transgene in fertilized eggs by transient expression of Cre recombinase. *Proc Natl Acad Sci USA*. 1995;92(1):160–164.
31. Sakai K, Miyazaki Ji. A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *Biochem Biophys Res Commun*. 1997;237(2):318–324.
32. Yagi T, et al. A novel ES cell line, TT2, with high germline-differentiating potency. *Anal Biochem*. 1993;214(1):70–76.
33. Ran FA, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154(6):1380–1389.