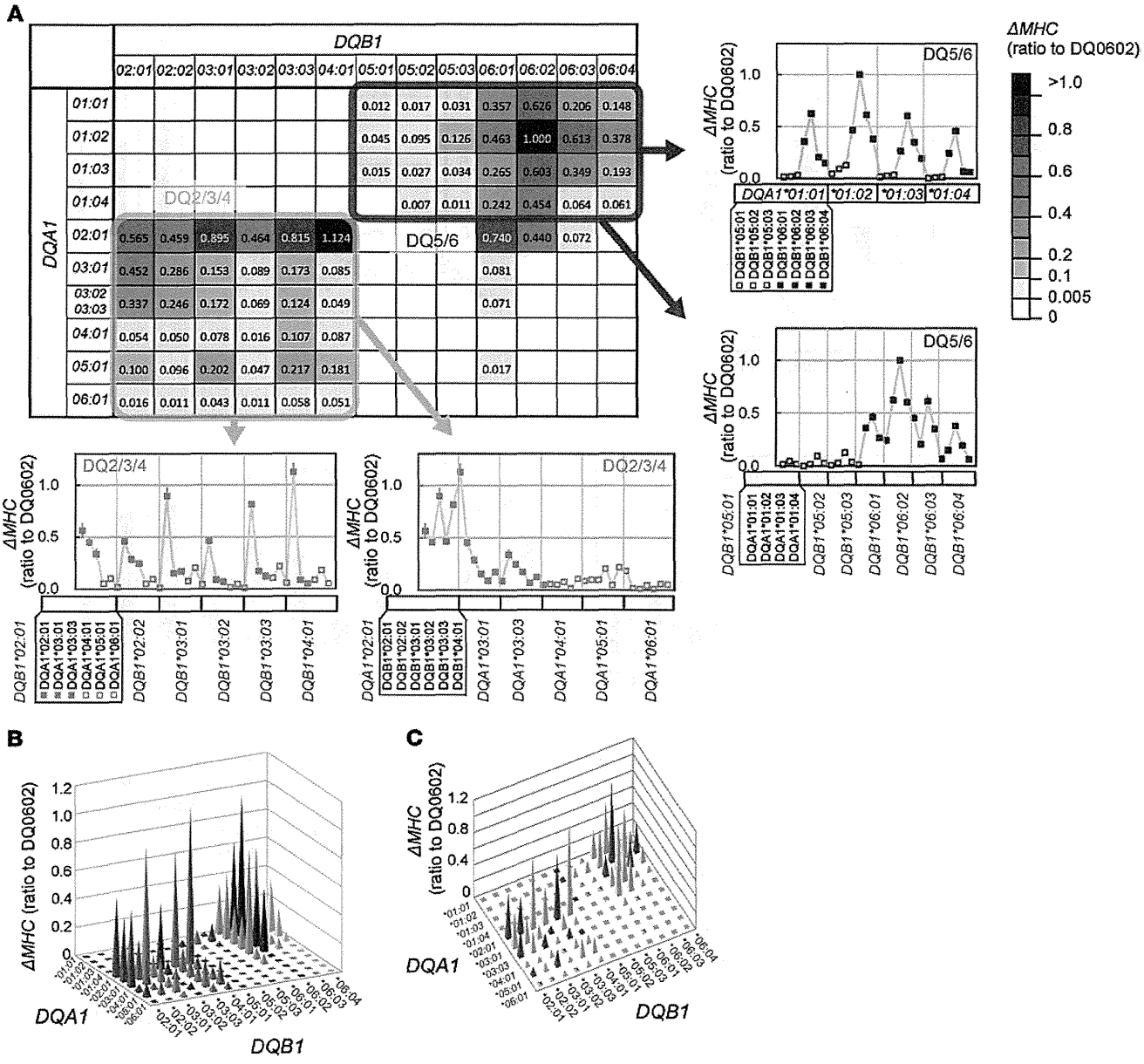

Figure 2. Δ*MHC* profile for the *HLA-DQ* allele products. (**A**) Δ*MHC* profile for HLA-DQ. Numbers and color scale indicate the Δ*MHC* (normalized to the Δ*MHC* of DQ0602). Δ*MHC* values below the threshold (<0.005 after normalization) are blank. Graphs display the Δ*MHC* value for each subgroup. Error bars represent the SEM. See also Supplemental Table 1. (**B** and **C**) Δ*MHC* profiles for the major *HLA-DQ* allele products (**B**) and the major *DQ* haplotype products (**C**). In **C**, the *DQ* haplotypes that are present at a haplotype frequency greater than 0.03 in at least 1 of the following populations are magenta colored: Congo Kinshasa Bantu (*n* = 90) (122, 135); Tunisian (*n* = 100) (122); South Korean (*n* = 324) (122); Indian Uttar Pradesh (*n* = 202) (122, 136); and European-American (*n* = 1,899) (133).

The Δ*MHC* profile was reproducible with pan-HLA II β mAbs BL-IA/6, TDR31.1 (Supplemental Figure 8, A–C), and IVA-12 (H. Miyadera, unpublished observations). The Δ*MHC* assay was not successful in the B cell lines due to the low efficiency of retroviral transduction (H. Miyadera, unpublished observation).

The effects of peptides on ΔMHC. To confirm that Δ*MHC* reflects the net stability of the MHC protein, Δ*MHC* was measured in the presence of high- and low-affinity peptides using the DQB1*06:02 peptide fusion constructs (designed according to ref. 66) (see Methods). Δ*MHC* increased 3.4-fold in the presence of insulin B_{1-15}, a strong binder of DQ0602 (30), and 1.5-fold in the presence of the class II–associated invariant chain peptide $(CLIP)_{81-107}$ (Figure 3A). Δ*MHC* decreased 0.6-fold in the presence of insulin B_{9-23}, a weak binder of DQ0602 (30), as well as in the presence of an artificial negative control peptide (GGSGGSGGSGGS) (GGS peptide) (Figure 3A), with which the interaction of the MHC protein with the peptide side chain may be limited. These data confirm that Δ*MHC* reflects the net stability of the pMHC.

The effect of endogenous peptides on the Δ*MHC* profile (Figure 2A) was estimated using DQB1-GGS peptide fusion constructs.

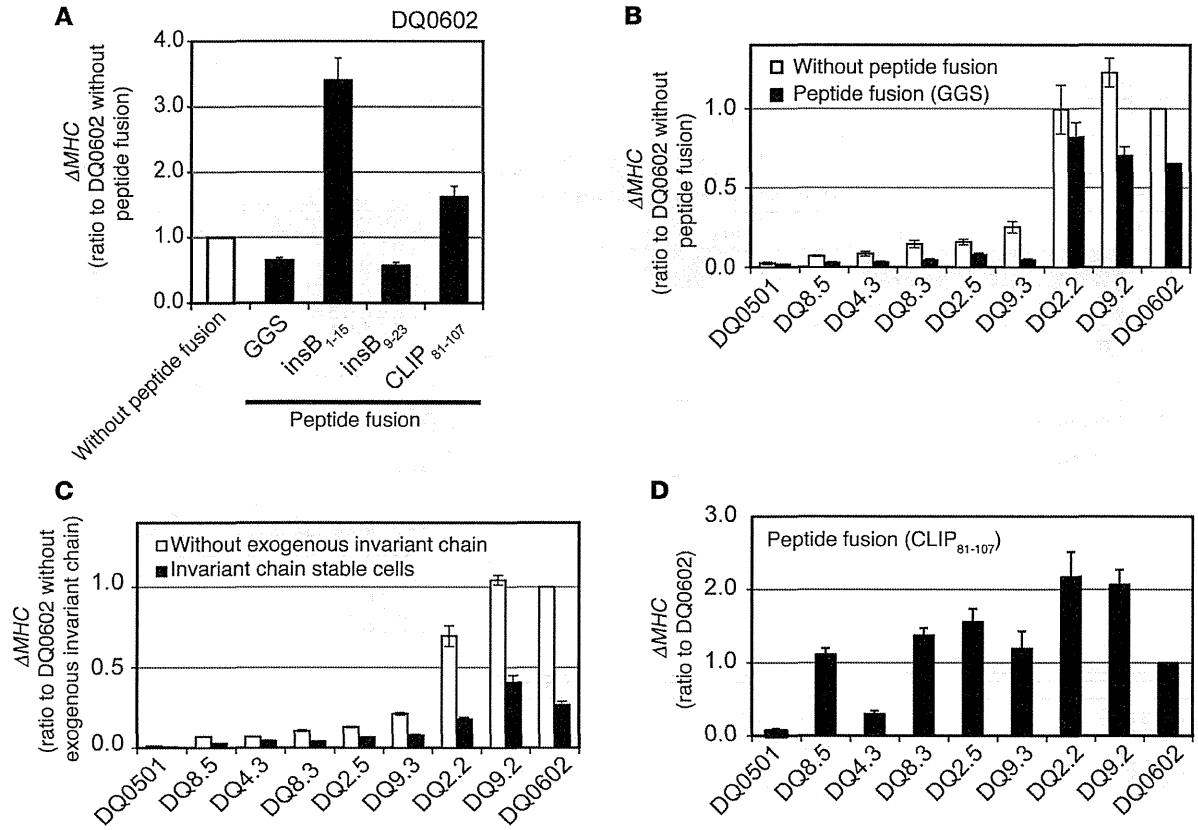

Figure 3. The effect of peptides on the ΔMHC. (**A**) ΔMHC profile for the DQ0602-peptide fusion constructs. *HLA-DQA1*01:02*–stable cells were transduced with pMXs-IG/*DQB1*06:02*, which carries the peptide sequence between the signal and the mature protein sequence (see Methods). (**B**) ΔMHC profile in the presence and absence of the artificial negative control peptide GGS. *HLA-DQA1*–stable cells were transduced with pMXs-IG/*DQB1* (without peptide fusion) (white bars) or with pMXs-IG/*DQB1*-GGS peptide (black bars). (**C**) ΔMHC profile for HLA-DQ in the presence (black bars) and absence (white bars) of the human invariant chain. ΔMHC was measured in NIH3T3 cells that stably expressed the human invariant chain (Supplemental Figure 10) and *HLA-DQA1*. (**D**) ΔMHC profile for the DQB1-$CLIP_{81-107}$ fusion constructs. Error bars represent the SEM ($n \geq 3$).

The presence of GGS peptide decreased ΔMHC for nearly all of the tested alleles; however, the overall allelic hierarchy in ΔMHC was maintained in the presence of the GGS peptide (Figure 3B, and Supplemental Figure 9, A and B). These data indicate that the HLA-DQ proteins in Figure 2A were loaded with endogenous peptides and were not empty heterodimers. The high ΔMHC values of DQ0602, DQ9.2, and DQ2.2 indicate that these HLA-DQ proteins are inherently stable. DQ2.2 showed relatively large interassay variation with and without GGS peptide (Figure 3B). DQ2.3 showed relatively high ΔMHC values, even in the presence of GGS peptide (Supplemental Figure 9B). It is currently unknown whether the GGS peptide occupies the peptide-binding groove of HLA-DQ.

MHC II is usually coexpressed with the invariant chain, which influences peptide loading and trafficking of the MHC II protein (67). To examine whether the lack of the human invariant chain affected cell-surface HLA-DQ expression, we measured ΔMHC in cells that abundantly expressed the human invariant chain (Supplemental Figure 10). In the presence of the human invariant chain, ΔMHC decreased by nearly half, possibly due to insufficiencies in cathepsins and HLA-DM. We found that the ΔMHC hierarchy was maintained in both the presence and absence of the human invariant chain (Figure 3C).

We analyzed the effect of CLIP on ΔMHC using the DQB1-$CLIP_{81-107}$ fusion constructs. The fusion of $CLIP_{81-107}$ greatly increased ΔMHC for certain alleles including DQ2.5 (Figure 3D), which is consistent with the high affinity of $CLIP_{94-104}$ for DQ2.5 (68). The variable effects of $CLIP_{81-107}$ on ΔMHC of the tested alleles indicate that the ΔMHC profile (Figure 2A) was not affected by CLIP-mediated stabilization.

Collectively, the ΔMHC profile (Figure 2A) does not appear to be biased by endogenous peptides, invariant chain, or CLIP. These observations support the possibility that the ΔMHC profile represents the allelic hierarchy in intrinsic HLA-DQ protein stability.

Polymorphic residues that regulate the ΔMHC of DQ5/6. We next searched for polymorphic sites that regulate ΔMHC. Supplemental Figures 11 and 12 show the pairwise comparisons of ΔMHC and the polymorphisms between the representative alleles (Supplemental Figure 11, A and B, and Supplemental Figure 12, A and B). For the highly conserved DQ5/6 subgroup, the responsible residues were identified through mutagenesis. For the highly polymorphic DQ2/3/4 subgroup, the major regulators of ΔMHC were identified through association analysis (see Supplemental Figure 13, A and B, for the amino acid sequence alignments of *HLA-DQ*).

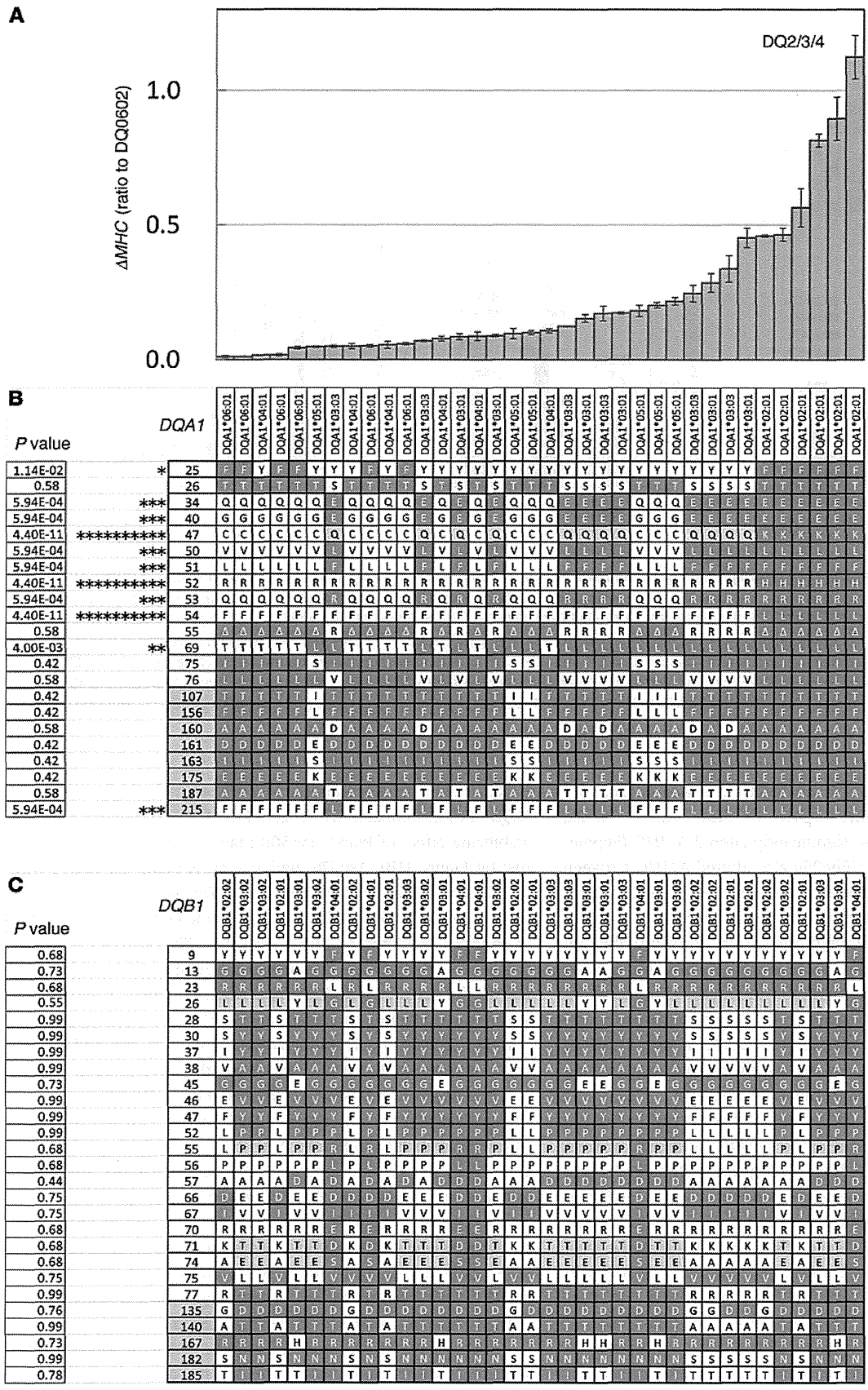

Figure 4. Association of amino acid variants in DQ2/3/4 with Δ*MHC*. (**A**) The HLA-DQ heterodimers in DQ2/3/4 are organized in the order of their Δ*MHC* values. Error bars represent the SEM. (**B** and **C**) Polymorphic variants in *HLA-DQA1* (**B**) and *-DQB1* (**C**) and their association with Δ*MHC*. Numbers indicate the amino acid residues. Residue numbers for the α2, β2, transmembrane, and cytosolic domains are shaded in brown. Variants identical to *DQA1*02:01-DQB1*04:01* (magenta) and other variants (white or gray). The association between each amino acid variant and Δ*MHC* was analyzed. The lowest *P* values at each site are indicated on the left with asterisks (2-tailed *t* test). The association table is presented in Supplemental Table 2.

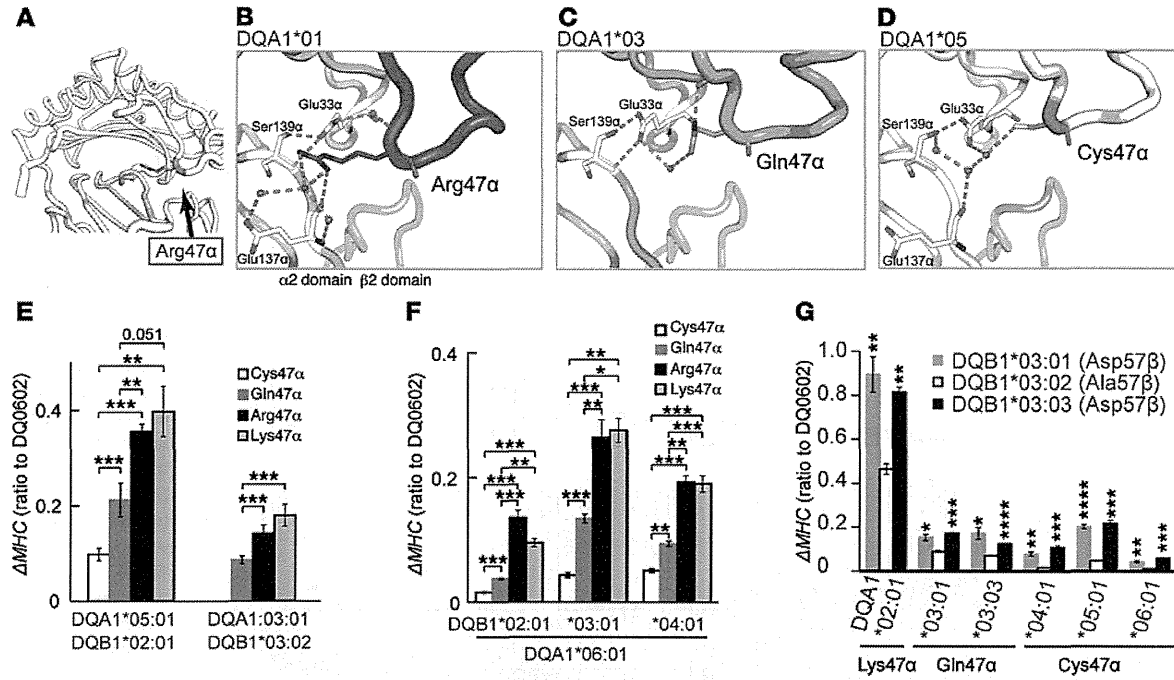

Figure 5. Stabilization/destabilization of HLA-DQ protein by 47α and 57β. (**A**) The location of Arg47α (magenta) in the protein structure of DQ0602 (PDB: 1uvq) (70). α1 domain (white), α2 domain (pink), and bound peptide (yellow). (**B–D**) Interdomain H-bonds formed between the α1 and α2/β2 domains in the presence of Arg47α in DQA1*01:02 (DQ0602) (PDB: 1uvq) (70) (**B**), Gln47α in DQA1*03 (DQ8.3) (PDB: 2nna) (69) (**C**), and Cys47α in DQA1*05 (DQ2.5) (PDB: 1s9v) (72) (**D**). α1 and α2 domains (magenta in **B**, orange in **C**, and pink in **D**); β2 domains (gray); water molecules (red spheres); and distances of less than 3.4 Å (green dots). (**E**) Effects of variants at 47α on the Δ*MHC* values of DQ2.5 and DQ8.3. (**F**) Effects of variants at 47α on the Δ*MHC* values of DQA1*06:01. (**G**) Differences in Δ*MHC* between DQB1*03:02 and *03:03, which differ at 57β, and between HLA-DQB1*03:01 and *03:03, which differ at 13β, 26β, and 45β in the β1 domain (Supplemental Figure 12B and 13B). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$ by 2-tailed *t* test. Error bars represent the SEM ($n \geq 3$).

In DQ5/6, the *HLA-DQA1*01* alleles differ at Asp/Gly2α, Tyr/Phe25α, Gln/Glu34α, and Lys/Arg41α. Of these, variants at 2α, 25α, and 34α, but not at 41α, significantly altered Δ*MHC* (Supplemental Figure 14, A–D). Tyr/Phe25α also altered Δ*MHC* between DQA1*04:01 and *06:01 (Supplemental Figure 14E). Asp2α, which is located outside of the peptide-binding groove, stabilized the HLA-DQ protein, possibly through the formation of interdomain H-bonds (Supplemental Figure 14, F and G). Tyr25α and Gln34α stabilized the HLA-DQ through the interdomain/intersubunit H-bonds and interactions with the peptide main chain (Supplemental Figure 14, H and I). The hierarchy in Δ*MHC* among *DQA1*01* alleles is perfectly explained by the stabilizing and destabilizing effects of 2α, 25α, and 34α (Supplemental Figure 14J).

For DQB1*06, we analyzed the effect of polymorphic residue on Δ*MHC* through mutagenesis at the sites that differed between DQB1*06:02 and *06:04 (9β, 30β, 57β, 70β, and 87β) (Supplemental Figure 14, K and L). The replacement of Tyr30β with His30β in DQB1*06:02 and of Tyr9β with Phe9β in DQB1*06:04 decreased Δ*MHC*. However, the substitution of Phe9β with Tyr9β in the presence of Tyr30β (in DQB1*06:02) and the substitution of His30β with Tyr30β in the presence of Tyr9β (in DQB1*06:04) did not affect Δ*MHC* (Supplemental Figure 14, K and L), indicating that Tyr9β and Tyr30β act complementarily to increase the Δ*MHC*. Tyr9β forms an H-bond with Asn72α, which interacts with the peptide main chain (Ser7p O and Glu9p N) in DQ8.3 (ref. 69 and Supplemental Figure 14M). Tyr30β interacts with the peptide main chain (Ser7p N) in DQ8.3 (ref. 69 and Supplemental Figure 14N). The higher Δ*MHC* of DQB1*06:02 compared with that of DQB1*06:03 suggests a greater stabilizing effect of Phe9β-Tyr30β than of Tyr9β-His30β (Supplemental Figure 11B). Asp57β mediates an H-bond/salt bridge with Arg79α and H-bonds with Ala10p N and Tyr37β (ref. 70 and Supplemental Figure 14O). Arg70β, which is predicted to interact with the p6 residue (71), increased Δ*MHC* relative to Gly70β in DQB1*06:04, but not in DQB1*06:02 (Supplemental Figure 14, K and L). Tyr/Phe87β did not affect Δ*MHC* (Supplemental Figure 14, K and L). Collectively, the hierarchy in Δ*MHC* among DQB1*06:02, *06:03, and *06:04 is shaped mainly through the variants at 9β, 30β, and 57β, the residues that alter the intrinsic stability of the MHC protein.

The large difference in Δ*MHC* between DQB1*05 and *06 is consistent with their difference in SDS stability (36) and was replicated with pan–HLA II β mAb IVA-12 and the DQw1-specific mAb Genox3.53 (H. Miyadera, unpublished observations). We determined that 14β (Leu in DQB1*05 and Met in DQB1*06), which projects its side chain toward the interface of the α1 and α2/β2 domains, is one of the responsible variants that diversifies Δ*MHC* between DQB1*05 and *06 (H. Miyadera, unpublished observations). Other variants that also contribute to the low Δ*MHC* value of DQB1*05 have not been fully elucidated in this study.

Polymorphic residues that regulate the ΔMHC of DQ2/3/4. To identify the residues that affect the Δ*MHC* value of DQ2/3/4, we conducted an association analysis between each amino acid variant and Δ*MHC* (Figure 4, A–C). In *HLA-DQA1*, variants at 47α

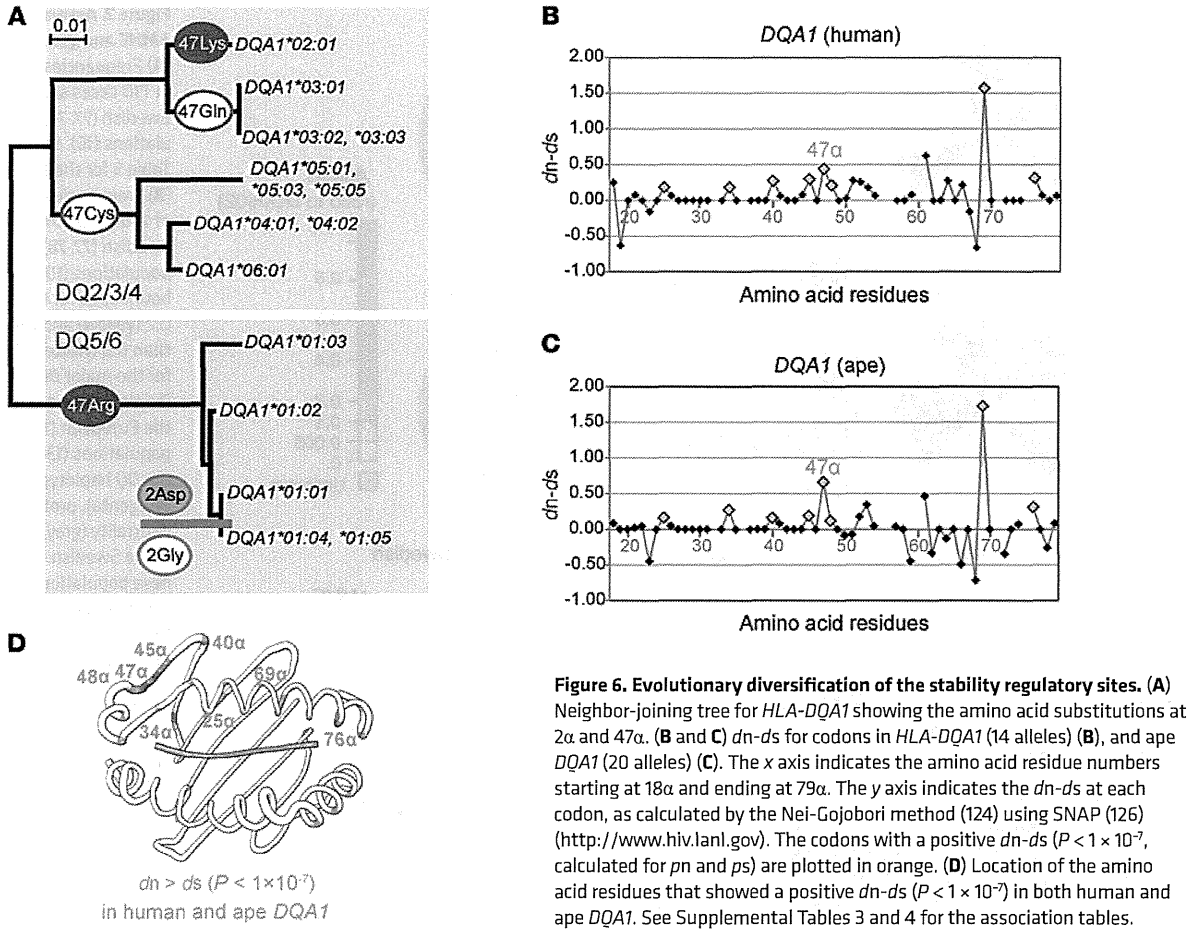

Figure 6. Evolutionary diversification of the stability regulatory sites. (A) Neighbor-joining tree for *HLA-DQA1* showing the amino acid substitutions at 2α and 47α. (**B** and **C**) *dn*-*ds* for codons in *HLA-DQA1* (14 alleles) (**B**), and ape *DQA1* (20 alleles) (**C**). The *x* axis indicates the amino acid residue numbers starting at 18α and ending at 79α. The *y* axis indicates the *dn*-*ds* at each codon, as calculated by the Nei-Gojobori method (124) using SNAP (126) (http://www.hiv.lanl.gov). The codons with a positive *dn*-*ds* ($P < 1 \times 10^{-7}$, calculated for *pn* and *ps*) are plotted in orange. (**D**) Location of the amino acid residues that showed a positive *dn*-*ds* ($P < 1 \times 10^{-7}$) in both human and ape *DQA1*. See Supplemental Tables 3 and 4 for the association tables.

(Lys vs. Cys and Gln), 52α (His vs. Arg), and 54α (Leu vs. Phe) were most strongly associated with Δ*MHC* ($P = 4.4 \times 10^{-11}$ by 2-tailed *t* test) (Figure 4B and Supplemental Table 2). In the structures of DQ2.5 and DQ8.3 (69, 72), 47α is located at the interface of the α1 and α2/β2 domains and appears to affect heterodimer stability. Arg52α, which does not participate in interdomain interaction, and Phe54α, which projects its side chain toward the outer surface or to the β1 domain, does not seem to explain the lower Δ*MHC* values of DQA1*03 and *05 compared with those of DQA1*02.

47α, which encodes Arg, Lys, Gln, and Cys, is the most variable site in *HLA-DQA1* (Supplemental Figure 13A). Unlike typical MHC polymorphisms, 47α is located outside the peptide-binding groove and TCR-recognition surface (Figure 5A and Supplemental Figure 14F). In DQA1*01:02, Arg47α forms extensive H-bonds with the α2 domain (70), which may be partially maintained by Lys47α (in DQA1*02), but not by Gln47α in DQ8.3 (69, 73) or by Cys47α in DQ2.5 (ref. 72 and Figure 5, B–D). The substitution of Cys47α with Lys47α in DQ2.5 increased the Δ*MHC* value by 4.1-fold (from 0.10 to 0.40) (Figure 5E), accounting for approximately 69% to 86% of the Δ*MHC* of DQ2.2 (Δ*MHC* = 0.46–0.57). The substitution of Gln47α with Lys47α in DQ8.3 increased the Δ*MHC* value by 2.1-fold (from 0.09 to 0.18) (Figure 5E), accounting for approximately 39% of the Δ*MHC* of DQ8.2 (DQA1*02:01-DQB1*03:02) (Δ*MHC* = 0.46). Mutagenesis at 47α in the least stable DQA1*06:01 confirmed greater stabilizing effects of Lys47α and Arg47α compared with those of Gln47α and Cys47α (Figure 5F). Therefore, DQ2.2 and DQ9.2, which carry Lys47α, are intrinsically more stable than the heterodimers formed by DQA1*03–*06.

In *HLA-DQB1*02*, **03*, and **04*, none of the variants were associated with Δ*MHC* (Figure 4C and Supplemental Table 2). Non-Asp/Asp57β is responsible for the differences in the Δ*MHC* within DQB1*03, but not the entire DQ2/3/4 that is attributable to *HLA-DQA1* (Figure 5G).

In both DQ5/6 and DQ2/3/4, the hierarchies in Δ*MHC* are generated mainly through stabilization and destabilization at the polymorphic sites that mediate the intersubunit or interdomain interactions or the interaction with the peptide main chain. These findings demonstrate that the major factor that determines the Δ*MHC* is the intrinsic stability of the MHC protein. Based on these and earlier findings, the Δ*MHC* values in Figure 2A were used as indicators of the intrinsic stability of HLA-DQ.

Evolutional divergence at 47α. The extensive variation at the 47α stability regulatory site (Supplemental Figure 13A) suggests that 47α may have been the target of positive natural selection, as has been observed for the peptide-binding sites (74, 75). The variants at 47α may have appeared before or at the time of the divergence of the *HLA-DQA1* sublineages (Figure 6A). To determine whether positive selection has been operating at 47α, we compared the number of nonsynonymous substitutions per nonsynonymous sites (*dn*) with the number of synonymous substitutions per synonymous sites (*ds*)

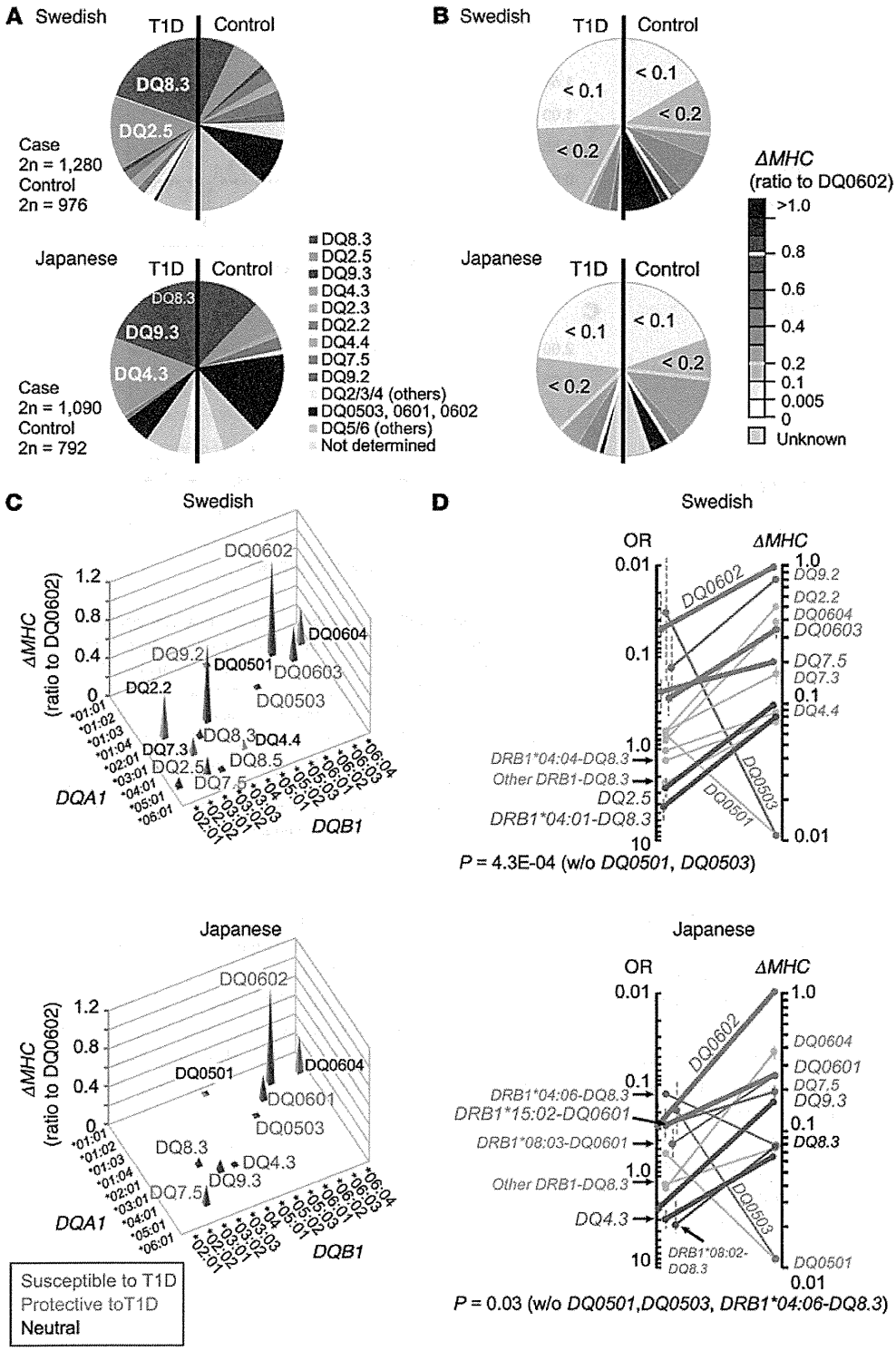

Figure 7. Association between the Δ*MHC* and genetic risk for T1D. (**A**) Frequencies of *DQ* haplotypes in T1D cases and controls in the Swedish (77, 78) and Japanese populations (10). See Supplemental Table 5 for the association table. (**B**) Frequencies of *DQ* haplotypes of various Δ*MHC* levels in the Swedish (77, 78) and Japanese populations (10). Lines indicate the boundaries for the Δ*MHC* less than 0.2 (yellow) and the Δ*MHC* greater than 0.8 (white). (**C**) Δ*MHC* profiles for the major *DQ* haplotypes (haplotype frequency >0.03) in the European (133) and Japanese populations (134). Colors indicate the *DQ* haplotypes that confer risk (magenta), protection (blue), or neutrality (gray) with regard to T1D in the Swedish (77, 78) and Japanese populations (10). (**D**) Relationship between the Δ*MHC* and the ORs for the major *DQ* haplotypes (haplotype frequency >0.03) in the Swedish and Japanese populations. The left axis indicates the OR, and the right axis indicates the Δ*MHC*. The dashed lines indicate a 95% CI for the ORs or standard errors for the Δ*MHC*. T1D risk (magenta), protective (blue), and neutral (gray) haplotypes. Haplotypes that are most strongly associated with susceptibility to or protection against T1D in each population are indicated by bold lines. *P* values indicate the association between the rankings in the Δ*MHC* and the ORs (Spearman's rank test).

for each codon in the *DQA1* of human, ape, and other mammals. The comparison of *dn* with *ds* detects a positive selection operating on each codon or gene. Under a null hypothesis of selective neutrality, the *dn* = *ds* is expected (ref. 76 and references therein).

In human and ape *DQA1*, the codons for 69α showed the highest *dn*-*ds* value (*dn*-*ds* = 1.573, $P = 5.17 \times 10^{-12}$ in human *DQA1*; *P* value calculated by the Wilcoxon signed-rank test for the proportion of synonymous differences per synonymous site [*ps*] and the proportion of nonsynonymous differences per nonsynonymous site [*pn*]; see Methods). The codons for 47α showed a *dn*-*ds* value that was one of the highest in *DQA1* (*dn*-*ds* = 0.443, $P = 1.63 \times 10^{-11}$ in human *DQA1*; *dn*-*ds* = 0.664, $P = 1.43 \times 10^{-23}$ in ape *DQA1*) (Figure 6, B and C, and Supplemental Tables 3 and 4), indicating that 47α has been subjected to positive diversifying selection. The codons for 47α in swine *DQA1* also showed a significantly positive *dn*-*ds* value (H. Miyadera, unpublished observations). The other residues that showed significantly positive *dn*-*ds* values ($P < 1.0 \times 10^{-7}$) in both human and ape *DQA1* were 25α, 34α, 40α, 45α, 48α, 69α, and 76α (Figure 6, B–D, and Supplemental Tables 3 and 4). Of these, 25α and 34α participate in peptide-binding and protein stability regulation, and 69α and 76α

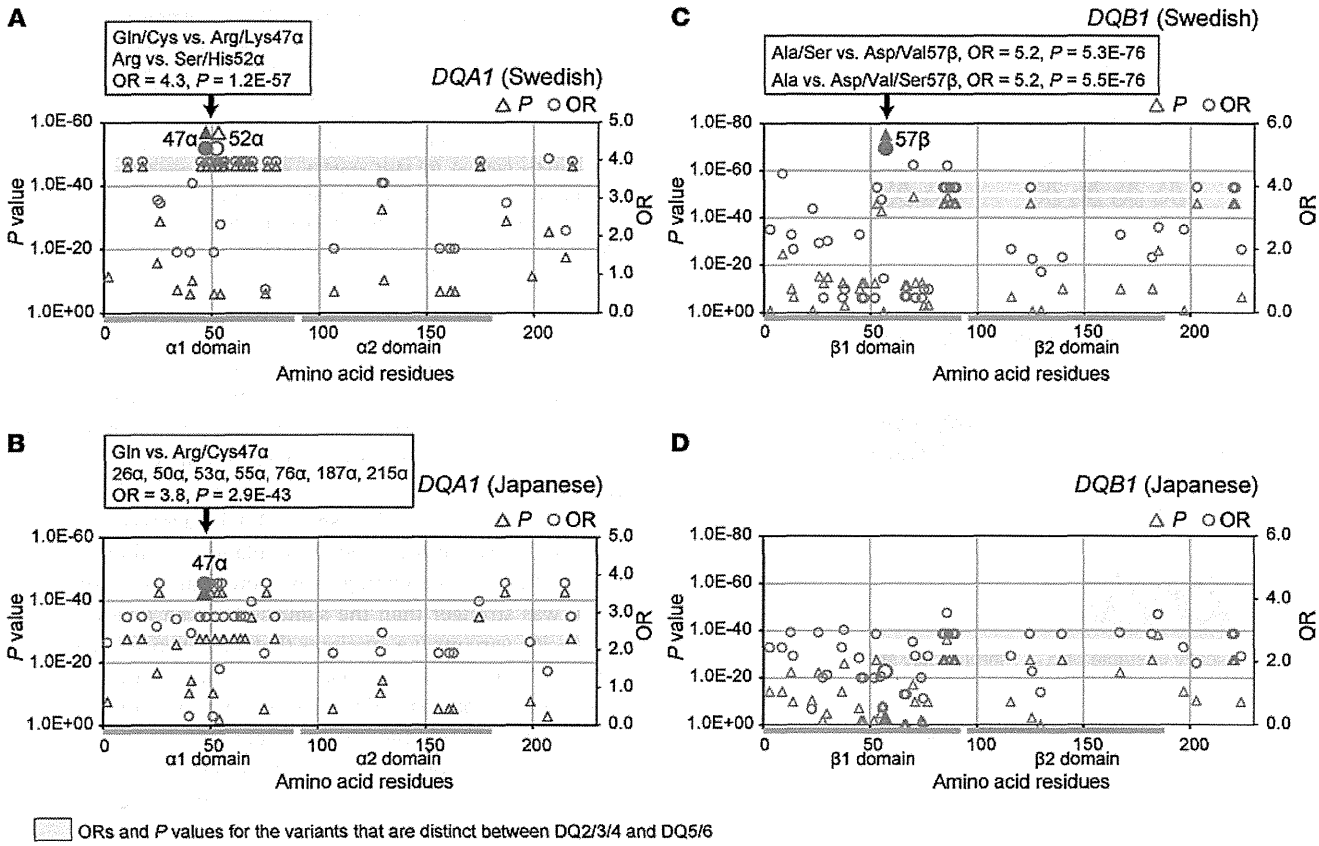

Figure 8. Association analyses between the amino acid variants in *HLA-DQ* and susceptibility to T1D. (**A**–**D**) Associations between the amino acid variants in *HLA-DQA1* (**A** and **B**) and *-DQB1* (**C** and **D**) and susceptibility to T1D in the Swedish (**A** and **C**) and Japanese populations (**B** and **D**). The ORs (circles) and the lowest *P* value (triangle) at each site (χ^2 test). ORs and *P* values for the variants that are distinct between DQ2/3/4 and DQ5/6 are shaded in orange. ORs for 2α and 199α in the Swedish population are not presented. See Supplemental Tables 6 and 7 for the association tables.

constitute the peptide-binding pockets. The significant excess of *dn* versus *ds* at 47α was not due to positive selection operating on antigen-binding sites, because the variants at 47α were not in linkage disequilibrium with the variants at the peptide-binding sites, such as 25α and 69α. The signature of positive selection at 47α independently of antigen-binding sites suggests that *HLA-DQ* has been evolving in favor of diversification in protein stability, in addition to increasing variations in the peptide-binding spectrum.

Autoimmune-susceptible DQ haplotypes encode unstable proteins. We next analyzed the relationship between Δ*MHC* and genetic risk for autoimmunity using the Δ*MHC* profile (Figure 2A) and case-control data for T1D in Swedish (77, 78) and Japanese populations (10). T1D is caused by the autoimmune destruction of insulin-producing β cells in the pancreas. Both *HLA-DR* and *HLA-DQ* confer a predisposition to T1D. The association of *DR-DQ* haplotypes with T1D has been extensively studied, and the risk hierarchy among the haplotypes has been established (59, 79, 80). In European and African-American populations, *DR3-DQ2.5* and *DR4-DQ8.3* are the risk haplotypes for T1D (59, 79–81). In the Japanese population, *DR9-DQ9.3* and *DR4-DQ4.3* are most strongly associated with T1D (8–10). *DR9-DQ9.3* confers a risk for T1D in the Filipino and Korean populations (8, 82, 83). *DQ0602* confers protection against T1D in Swedes and Japanese, among other populations (ref. 79, Figure 7A, and Supplemental Table 5).

The composition of the *DR-DQ* haplotypes differs greatly between the Swedish and Japanese populations (Figure 7A); however, the frequency of the *DQ* haplotypes is similar in the 2 populations when the haplotypes are subgrouped by Δ*MHC* (Figure 7B). The *DQ* haplotypes with low Δ*MHC* values (Δ*MHC* <0.2) are predominant in T1D cases, whereas those with high Δ*MHC* values (Δ*MHC* >0.8) are present at higher frequencies in controls than in T1D cases (Figure 7B). The steep hierarchy in Δ*MHC* among the *DQ* haplotypes appears to be correlated with the risk and protection for T1D (Figure 7C and Supplemental Figure 15, A and B).

Indeed, a clear inverse relationship exists between Δ*MHC* and susceptibility to T1D (estimated by the odds ratio [OR]) for the majority of *DQ* haplotypes, including those haplotypes that are most predisposing (*DQ2.5*, *DQ8.3*, *DQ9.3*, and *DQ4.3*) and protective (*DQ0602*, *DQ0603*, and *DQ9.2*). In the absence of *DQ0501* and *DQ0503* (in the Swedish and Japanese populations) and *DRB1*04:06-DQ8.3* (in the Japanese population), which encode unstable HLA-DQ proteins and are neutral or protective of T1D, the rankings in Δ*MHC* and in the ORs were inversely associated ($P = 4.3 \times 10^{-4}$ [Swedish]; $P = 0.03$ [Japanese], Spearman's rank correlation test) (Figure 7D). Although the CIs for the ORs are large, these data confirm an overall inverse relation between Δ*MHC* and genetic risk for T1D. The protective phenotype of the *DRB1*04:06-DQ8.3* haplotype may be ascribed to *DRB1*04:06*

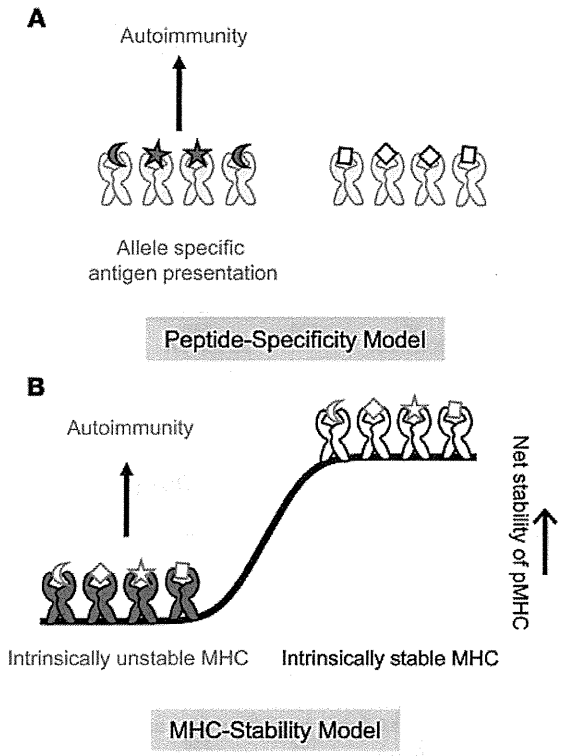

Figure 9. Hypothetical mechanisms of the *HLA*-autoimmunity association. (**A**) Peptide specificity model. This model postulates that the peptide-binding spectrum of the MHC determines the genetic association of *HLA* with autoimmunity. According to this model, the HLAs that are able to present the disease-relevant peptides (magenta) confer a risk for autoimmunity. Colors indicate the disease-relevant (magenta) and irrelevant peptides (white). (**B**) MHC stability model. This model postulates that intrinsically unstable MHC proteins (magenta), which form unstable MHC–self-epitope complexes through presentation of diverse self-peptides and are more likely to form unstable pMHC than intrinsically stable MHC, confer a risk for autoimmunity.

(59, 79, 80). The protective association of *DQ0503* may be attributable to *DR14* (84).

Destabilizing variants at 47α confer a predisposition to T1D. To address whether the association between Δ*MHC* and T1D is affected by the peptide-binding variants, we analyzed the association between each amino acid variant and T1D. In *HLA-DQA1*, the strongest association signals with T1D in the Swedish population were detected at 47α (Gln and Cys vs. Arg and Lys) and 52α (Arg vs. Ser and His) (OR = 4.33, $P = 1.2 \times 10^{-57}$, χ^2 test), which are the variants that differ between *HLA-DQA1*01, *02,* and **03-*06* (Figure 8A and Supplemental Table 6). The association of Arg52α with T1D has been noted previously (85, 86), but Arg52α may not affect the stability or peptide binding in the structures of DQ2.5 and DQ8.3 (69, 72). In the Japanese population, variants at 26α (Ser vs. Thr), 47α (Gln vs. Arg and Cys), 50α (Leu vs. Glu and Val), 53α (Arg vs. Lys and Gln), 55α (Arg vs. Gly and Del), 76α (Val vs. Leu and Met), 187α (Thr vs. Ala), and 215α (Leu vs. Phe), which distinguish *HLA-DQA1*03* from other alleles, were the most strongly associated with T1D (OR = 3.79, $P = 2.9 \times 10^{-43}$) (Figure 8B and Supplemental Table 6). The 2 populations share the risk variant Gln47α. Cys47α is not associated with T1D in the Japanese population due to the absence of *DQ2.5* (risk haplotype) and the presence of *DQ7.5* (protective haplotype).

Non-Asp57β has been known as a T1D risk variant in Europeans (87, 88) and in murine models (89). As expected from earlier studies, variants at 57β were strongly associated with T1D in the Swedish population (Ala and Ser vs. Asp and Val, OR = 5.22, $P = 5.3 \times 10^{-76}$; non-Asp vs. Asp, OR = 5.23, $P = 6.1 \times 10^{-68}$) but not in the Japanese population (Figure 8, C and D, and Supplemental Table 7). Non-Asp57β may confer a risk for T1D through an alteration in the binding preference for the p9 residue (90–94) and for the TCR (95). The protein-destabilizing effect of non-Asp57β may also contribute to T1D. However, the neutral phenotype of *DQ2.2* (non-Asp57β, Δ*MHC* = 0.46–0.57) (59, 80) and the susceptible phenotypes of *DQ9.3* and *DQ4.3* (Asp57β, Δ*MHC* = 0.05–0.12) suggest that the low Δ*MHC* value, rather than non-Asp57β, is indispensable to the shared mechanism of T1D risk across ethnicities.

Except for 57β in the Swedish population, we found that no variant in the peptide-binding sites showed an association signal that was stronger than the signals observed for the sublineage-specific variants (i.e., variants that differ between DQ2/3/4 and DQ5/6) (Figure 8, A–D). These results indicate that (a) a certain combination of the peptide-binding variants confers a risk for T1D, or (b) the peptide-binding spectrum is not strongly associated with T1D. The observations that the peptide-binding spectra of DQ2/3/4 partially overlap and are partially unique (96–98) permit both mechanisms.

Collectively, the absence of a common association signal for T1D in the peptide-binding sites confirms the predisposing effect of HLA-DQ protein instability. Among the multiple factors that might contribute to T1D risk, instability of HLA-DQ may be one of the major components that determines the pathogenic potential of each *DR-DQ* haplotype.

Discussion

To uncover the mechanism of the *HLA*-autoimmunity association, the most logical approach would be to identify the functional variation(s) among *HLA* alleles that are correlated with disease risk. The present study revealed an inverse association between Δ*MHC* and T1D and confirmed the strong association signals for T1D at the protein-destabilizing variants. These results suggest a potential causal link between the intrinsic instability of HLA-DQ and T1D.

The *HLA*-autoimmunity association has been explained for decades by the allele-specific presentation of disease-relevant self-peptides (referred to as the peptide specificity model) (Figure 9A). Our study indicates that the intrinsic instability of the MHC protein may also be a major functional component of autoimmune susceptibility (i.e., the MHC stability model) (Figure 9B). The 2 mechanisms are not mutually exclusive, and neither mechanism alone sufficiently explains the *DR-DQ* haplotype association with T1D. The combination of the 2 parameters might generate the high level of allele specificity in the autoimmune disease risk.

The simplest mechanism by which an unstable MHC protein might confer autoimmune disease risk would be the contribution of MHC protein instability to the incomplete thymic negative selection. This possibility is consistent with the established concept of autoimmunity, in which the formation of an unstable pMHC

by low-affinity peptides permits thymic escape of self-reactive T cells (99, 100). Upon weak self-epitope presentation, intrinsically unstable MHC may form MHC–self-epitope complexes in the low-stability range and confer a risk for autoimmunity. The ability of intrinsically unstable MHC to form unstable pMHC with diverse self-epitopes, including the promiscuous binders, may explain the association of T1D risk *DR-DQ* haplotypes with a variety of autoimmune disorders (Table 1), as well as the involvement of promiscuous self-epitopes in allele-specific disease pathogenesis.

The intrinsically unstable MHC protein may be expressed either at a basal level or maximum level depending on the availability of high-affinity peptides and accessory molecules. The intrinsically unstable MHC protein is also expected to preferentially present high-affinity peptides on the cell surface. In contrast, intrinsically stable MHC may be expressed relatively constantly on the cell surface and present the peptides with wider affinity ranges. The potential variation in the expression patterns of each MHC type might affect the outcomes of thymic selection, peripheral activation, and subset development of T cells (101–105), processes that are controlled by cell-surface MHC density. The actual mechanism that links MHC instability to autoimmunity should involve diverse immunological processes and could be highly complex.

The *ΔMHC* profile provides a clue for dissecting the mechanism of the *DQ* haplotype association with T1D risk. The high risk for T1D of *DQ2.5/DQ8.3* heterozygotes (58, 59, 79, 80) may be ascribed to the instability of DQA1*05-DQB1*03:02 (DQ8.5) (*ΔMHC* = 0.05) (Figure 7C and Supplemental Figure 15, A and B) and its unique peptide-binding spectrum (106, 107).

DQ0602 and DQ9.2 may confer protective effects through abundant or sustained MHC expression. The high stability of DQ0602 and its ability to bind diabetogenic self-epitopes (30–32) are compatible with the proposed mechanisms of protection, such as the thymic deletion of self-reactive T cells (108) and the "affinity model" or "determinant capture" (106, 109, 110). The protective associations of *DRB1*15:01-DQB1*06:02* with T1D, autoimmune polyglandular syndrome (APS) types II and III, and selective IgA deficiency (Table 1) suggest a shared mechanism of protection among these disorders.

The neutral and protective phenotypes of *DQ0501*, *DQ0503*, *DQ7.3*, and *DQ7.5*, which encode unstable HLA-DQ proteins (Figure 7D and Supplemental Figure 15B), are not explained by the *ΔMHC* value. As established for the *DRB1*04-DQ8.3* haplotypes (10, 59, 79–81), the predisposing and protective phenotypes of *DR-DQ* haplotypes can be largely influenced by the *DR* allele. The instability of HLA-DQ could be the condition that permits the linked *DR* to exert pathogenic or protective effects. The increased risk for T1D of *DQ8.3/DQ0501* and *DQ8.3/DQ0604* heterozygotes (79, 80, 111, 112) (*DR4/DR13* in the Japanese population; refs. 8–10), which do not generate *trans* DQ heterodimers (Figure 2A), also remains difficult to explain by the existing hypothesis.

The dependency of MHC II on the invariant chain, CLIP, and HLA-DM are proposed as the risk factors for T1D (113–117). The affinity of HLA-DQ with $CLIP_{81-107}$ appears to be variable among the alleles (Figure 3D). The dependency of HLA-DQ on HLA-DM may also be variable due to polymorphisms at the 47α–56α segment (117). It will be intriguing to speculate how these interrelated parameters and their combinations might affect T1D risk.

In this study, *ΔMHC* was used as an indicator of the intrinsic stability of the MHC protein. Several lines of direct and indirect evidence, including the mutagenesis studies, have confirmed that the intrinsic stability of the MHC protein is the major factor in determining *ΔMHC*. In fibroblasts, the stabilization of MHC II promotes rapid transport of pMHC into the Golgi apparatus (118), and the MHC II proteins that weakly interact with peptides are protease sensitive and are degraded in the endosomes (50). In B cells, unassembled HLA-DR subunits can be degraded rapidly during intracellular transport (119, 120). Presumably, intrinsically less stable MHC proteins that are inefficient in both α and β subunit assembly and in the formation of the pMHC may generate a substantial amount of unassembled or unfolded subunits that are sensitive to proteolysis, resulting in reductions in both the number and lifetime of the pMHCs that reach the cell surface. However, because *ΔMHC* is measured by the cell-based assay, the possibility that *ΔMHC* is biased by unknown cellular components cannot be completely ruled out.

It should also be noted that the *ΔMHC* profile, which is measured in an engineered condition, does not represent the expression profile of HLA-DQ on professional APCs. As observed in B-LCLs, the *HLA-DQ* alleles that showed low *ΔMHC* values can be expressed at a high level (36), possibly through transcriptional upregulation and/or stabilization by high-affinity peptides.

Our study provides new insights into the molecular evolution of the MHC, which has been explained by diversification of the peptide-binding spectrum (ref. 121 and references therein). The *DQ* haplotypes that encode unstable and stable proteins are maintained at a high frequency in a variety of populations (allelefrequencies.net; ref. 122), indicating that the alleles encoding unstable HLA-DQ heterodimers are not selectively disadvantageous. Presumably, both the stable and unstable MHC may have functional advantages against important pathogens. The stability of HLA-DR may also be diverse (123); however, the scale of variation may be limited by the monomorphic nature of *HLA-DRA*.

Collectively, our study reveals an additional layer of functional hierarchy among the *HLA* alleles that has been generated through evolution and is associated with autoimmunity. These findings complement and extend the existing model of the *HLA*-autoimmunity association and suggest a mechanistic basis of autoimmune susceptibility.

Methods

Measurement of ΔMHC. Full-length cDNAs for *HLA-DQA1* and-*DQB1* were cloned from *HLA*-typed cell lines or peripheral blood samples from healthy individuals. The cDNAs were inserted into the retroviral vectors pMXs-puro or pMXs-IG (52) with an *Eco*RI site and a Kozak sequence in the 5′ terminus and a *Strep*-tag II (IBA GmbH) and a His-tag in the 3′ terminus of *HLA-DQA1* and *-DQB1*, respectively. The first nucleotide of the second amino acid of *HLA-DQA1* and *-DQB1* was changed to guanine to introduce a Kozak sequence. The *HLA-DQA1-Strep*-tag II and *HLA-DQB1*-His-tag were inserted into the pMXs-puro and pMXs-IG vectors, respectively, using *Eco*RI and *Not*I sites. To generate the retroviruses containing pMXs-puro/*DQB1* and pMXs-IG/*DQA1*, approximately 0.5–1 × 10^6 PLAT-E cells (53) were transfected with 1.5 μg plasmid and 4.5 μl Fu-GENE reagent (Roche Diagnostics), according to the manu-

facturer's instructions. *HLA-DQB1*–stable cells were established through the transduction of NIH3T3 cells with a retrovirus containing pMXs-puro/*DQB1* and selection with puromycin (6 μg/ml). To measure Δ*MHC*, the *HLA-DQB1*–stable cells were seeded at approximately 2×10^5 cells per well in a 12- or 24-well plate and cultured overnight. The cells were transduced with a retrovirus containing pMXs-IG/*DQA1* using 5–60 μl PLAT-E medium. Forty-eight hours after transduction, GFP and cell-surface MHC (in GFP-positive cells) levels were measured by flow cytometry (EPICS-XL; Beckman Coulter) using anti–HLA II β mAb (WR18) (65) (MorphoSys AG) or isotype control (mouse IgG2a [6H3]; Medical & Biological Laboratories Co. Ltd.) with phycoerythrin-conjugated anti-mouse IgG (Rockland Immunochemicals or Southern Biotechnology Associates Inc.). The MFI (GFP) was defined as the MFI for GFP-positive cells minus MFI for GFP-negative cells. The MFI (MHC) was defined as the MFI for anti–HLA II β [WR18] minus the MFI for the isotype control of the GFP-positive cells (Supplemental Figure 3A). The MFI (GFP) and MFI (MHC) were plotted to calculate the increase in the MHC relative to GFP (Δ*MHC*) (Figure 1D and Supplemental Figure 3B). To minimize interassay variation, the Δ*MHC* for each DQ heterodimer was normalized with the Δ*MHC* value of DQ0602, which was measured on the same day (Figure 1D and Supplemental Figure 4, A and B). The lower detection threshold of Δ*MHC* was set at 0.005 (after normalization to DQ0602). The assay was performed more than 3 times for each *HLA-DQA1* and *-DQB1* allelic pair. Site-directed mutagenesis was performed with QuikChange II (Agilent Technologies), according to the manufacturer's instructions. The following anti-HLA mAbs were purchased: BL-IA/6 (Santa Cruz Biotechnology Inc.) and TDR31.1 (Ancell Corporation). The mAb IVA12 was a gift of S. Kawai (Wakunaga Pharmaceutical Co.). The NIH3T3 cell line was obtained from the RIKEN Cell Bank.

The *HLA-DQB1*-peptide fusion construct was designed as described in Kozono et al. (66), with modification. Briefly, the sequences for insulin B_{1-15}, insulin B_{9-23}, $CLIP_{81-107}$, and an artificial GGS peptide (GGSGGSGGSGGS) were inserted between the signal sequence of *HLA-DQB1*06:02* and the mature protein region of *HLA-DQB1* via linkers (SGG and GGGGSIEGRGGGGSGSA at the N and C termini of the peptide, respectively). The fragments that encode the signal sequence and the peptide sequences were synthesized as double-stranded DNA and were ligated to the mature DQB1 sequence using the *Acc*III site. For the DQB1-peptide fusion constructs, Δ*MHC* was measured in *HLA-DQA1*–stable cells.

To establish NIH3T3 cells that stably express the human invariant chain, the full-length cDNA for the invariant chain (isoform b) was cloned from the B-LCL of the healthy individual and was inserted into pMXs-neo using *Eco*RI and *Not* I at the 5′ and 3′ termini of the invariant chain, respectively. NIH3T3 cells were transduced with the pMXs-neo/invariant chain and were selected with G418 (2 mg/ml) for 2 weeks. The invariant chain–stable cells were then transduced with pMXs-puro/*HLA-DQA1* and were selected with puromycin (6 μg/ml). Expression of the invariant chain was detected with anti–invariant chain mAb (LN2) (BioLegend). Intracellular staining was performed using the FIX & PERM Cell Fixation and Cell Permeabilization Kit (Invitrogen), according to the manufacturer's instructions.

Expression of HLA-DQ protein in insect cells. Full-length cDNAs for *HLA-DQA1* and *-DQB1* were inserted into pMT-V5/His (Invitrogen) with a Kozak sequence at the 5′ terminus and a *Strep*-tag II (IBA GmbH) and His-tag at the 3′ terminus of *HLA-DQA1* and *-DQB1*, respectively. The expression plasmids were cotransfected with pCoBlast (Invitrogen) into the *Drosophila melanogaster* Schneider cell line (S2) via calcium phosphate transfection (Invitrogen) or lipofection using Effectene (QIAGEN), according to the manufacturer's instructions. Stable polyclonal cells were obtained in the presence of blasticidin (25 μg/ml). For each *HLA-DQA1* and *-DQB1* allelic pair, stable cells were obtained from more than 3 independent transfections. The expression of HLA-DQ proteins was induced with 0.5 mM $CuSO_4$ for 48 hours and measured by flow cytometry (EPICS-XL; Beckman Coulter) with an anti–HLA II β mAb (WR18, FITC conjugate) (MorphoSys AG) or an isotype control (mouse IgG2a–FITC) (Beckman Coulter). 3′ Rapid amplification of cDNA ends (3′-RACE) was performed with the 3′-Full RACE Core Set (Takara Bio) using total RNA (1 μg) from the stable S2 cells as a template. Total RNA was primed for reverse transcription at 42°C with the Oligo dT-3 Sites Adaptor Primer (Takara Bio). 3′-RACE PCR was performed according to the manufacturer's instructions. Specifically, the internal primers for *HLA-DQA1* (p-50: 5′-ACTCTACCGCTGCTACCAATG-3′) and *HLA-DQB1* (p-48: 5′-ACGGTGTGCAGACACAACTAC-3′) were used as the 5′ primers, and the Oligo dT-3 Sites Adaptor Primer was used as the 3′ primer. The *Drosophila* ribosomal protein rp49 was amplified as a positive control with the primers p-266 (5′-ATGACCATCCGCCCAGCATAC-3′) and p-267 (5′-TGTGTATTCCGACCAGGTTAC-3′). For the purification of HLA-DQ proteins and Western blotting, the stably transfected S2 cells were harvested 48 hours after induction and stored at −80°C. The cells were thawed on ice and lysed in 30 mM sodium phosphate buffer (pH 8.0, 150 mM NaCl, 0.5% NP-40) in the presence of a protease inhibitor cocktail (Sigma-Aldrich) and Benzonase Nuclease (EMD Millipore, Merck KGaA) for 1 hour at 4°C. The lysate was centrifuged at 2,000 *g* for 10 minutes at 4°C, and the supernatant was subjected to purification. Purification by *Strep*-tag II and His-tag was performed with a *Strep*-Tactin Spin Column (IBA GmbH) and MagExtractor (Toyobo Co., Ltd.), respectively, according to the manufacturers' instructions. For Western blotting, the elution fractions were subjected to denaturation and boiling in the presence of SDS (2%) and reducing agents. *Strep*-Tactin-AP (IBA GmbH) and anti-His6 mAb-HRP (Roche Diagnostics) were used for the detection of DQα and DQβ, respectively.

Evolutionary analysis. The *dn*-*ds* for each codon was calculated in a pairwise manner by the Nei-Gojobori method (124) with Jukes-Cantor correction (125) using SNAP (126) (http://www.hiv.lanl.gov). The amino acids 18α–79α of the *DQA1* for humans (*HLA-DQA1*, 14 alleles) and apes (*Gogo-DQA1* [*Gorilla gorilla*], *Patr-DQA1* [*Pan troglodytes*], *Hyla-DQA1* [*Hylobates lar*], and *Popy-DQA1* [*Pongo pygmaeus*], 20 alleles), which represent all of the alleles that are registered in the IMGT/HLA and IMGT/IPD databases (EMBL-EBI) (127–130) and which were nonredundant in the 18α–79α region (i.e., at least 1 codon was different from the others), were used for the analyses. The codons with a deletion or insertion (55α and 56α in the human and ape *DQA1*) were excluded from the analysis. For the sequence pairs with codons encoding Met or Trp (48α, 66α, and 76α in human and ape *DQA1*), the *ps* was assumed to be 0 and was used to calculate the *ds*. For sequence pairs showing a *ps* greater than or equal to 0.75 (32α, 36α, 60α, 71α, and 75α in human and ape *DQA1*), the *ds* was not calculated (at these codons, the *ps* was greater than the *pn*, and

there were no codons with a $pn \geq 0.75$) (Supplemental Tables 3 and 4). To assess the possibility of positive selection (i.e., diversifying selection), the differences between the *p*s and *p*n as well as between the *d*s and *d*n were statistically tested for each codon using the Wilcoxon signed-rank test. A neighbor-joining phylogenetic tree was constructed for the nucleotide sequences of exon 2 of the *HLA-DQA1* alleles using the alleles with frequencies greater than 0.02 in worldwide populations (131). Molecular Evolutionary Genetics Analysis (MEGA) version 4 software (132) was used for the analysis.

Association analysis between the ΔMHC and genetic risk for T1D. The case-control data for T1D in the Swedish population (77) were collected and genotyped by the T1D component of the 13th International Histocompatibility Workshop (78) and were deposited in the dbMHC database (http://www.ncbi.nlm.nih.gov/projects/gv/mhc/). For the Japanese population, published case-control data (10), which were collected and genotyped by the Committee on T1D of the Japanese Diabetes Society, were used for the analysis. All of the T1D subtypes in the Japanese population (acute, fulminant, and slowly progressive) (10) were combined. The *HLA-DQA1* alleles were manually estimated from the *DRB1-DQB1* haplotype data based on the *DRB1-DQA1-DQB1* haplotype frequency in the general European-American (133) and Japanese (134) populations. In the Japanese population, all of the *DRB1*08:02-DQB1*03:02* haplotypes were assumed to carry *HLA-DQA1*03*. The rare *DRB1-DQB1* haplotypes and *DRB1-DQB1* haplotypes that are linked to more than 2 *HLA-DQA1* alleles were excluded from the analysis. Closely related alleles, such as *HLA-DQA1*01:04* and **01:05*; *HLA-DQA1*05:01*, **05:03*, and **05:05*; and *HLA-DQB1*04:01* and **04:02*, were assumed to have identical *ΔMHC* values. Because not all of the *HLA-DQA1*03* and *-DQB1*02* alleles were genotyped at a 4-digit resolution in the case-control data, the *ΔMHC* values of DQA1*03 and of DQB1*02 were assumed to be the average of the *ΔMHC* values for DQA1*03:01, *03:02, and *03:03, and of DQB1*02:01 and *02:02, respectively, and were used for the association analyses. For the analysis of the association between the amino acid variants and genetic risk for T1D, the variants at 160α, which differ between *HLA-DQA1*03:01* and **03:03*, and at 135β, which differ between *HLA-DQB1*02:01* and **02:02*, were excluded from the analysis.

Statistics. Allelic differences in the *ΔMHC* values as well as associations between the amino acid variants and *ΔMHC* were analyzed by a 2-tailed *t* test. Differences between *p*s and *p*n and between *d*s and *d*n were analyzed by the Wilcoxon signed-rank test. Associations between the *DQ* haplotype frequency and susceptibility to T1D were analyzed by the χ^2 test with Bonferroni's correction. Spearman's rank test was used to evaluate associations between the ranking in the *ΔMHC* and ORs. A χ^2 test test was used to analyze associations between the amino acid variants and genetic risk for T1D. A *P* value of less than 0.05 was considered statistically significant.

Study approval. The use of cDNAs and cell lines from healthy individuals was approved by the ethics committee of the Graduate School of Medicine, The University of Tokyo, and written informed consent was obtained from all participants.

Acknowledgments

We acknowledge W.W. Kwok for his early work and valuable comments, without which this work could not have been achieved. We thank S. Harada and K. Kita for their valuable advice. We thank E.D. Mellins and E. Mignot for discussions. We are grateful to the T1D component of the 13th International Histocompatibility Workshop; the dbMHC Database (a public database funded by the NIH); and the Committee on T1D of the Japanese Diabetes Society for the case-control data. We thank T. Yabe for the *HLA*-genotyped cells and S. Kawai for the IVA12 mAb. We are grateful to the following groups for permission to use structural data in the figures: DQ2.5 (1s9v), C.-Y. Kim, H. Quarsten, C. Khosla, and L. M. Sollid; DQ8.3 (2nna), K.N. Henderson, J.A. Tye-Din, J. Rossjohn, and R. P. Anderson; and DQ0602 (1uvq), C. Siebold, B.E. Hansen, E.Y. Jones, and L. Fugger. We thank the RIKEN Cell Bank for the NIH3T3 cell line. This work was funded by grants from the Japan Society for the Promotion of Science (JSPS) KAKENHI (22133008, to K. Tokunaga; 22133006 and 18770106, to H. Miyadera; and 23133502, to J. Ohashi).

Address correspondence to: Hiroko Miyadera, Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Phone: 81.3.5841.3653; E-mail: miyadera-h@umin.net.

Jun Ohashi's present address is: Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo, Japan.

1. Jones EY, Fugger L, Strominger JL, Siebold C. MHC class II proteins and disease: a structural perspective. *Nat Rev Immunol.* 2006;6(4):271–282.
2. Price P, et al. The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. *Immunol Rev.* 1999;167:257–274.
3. de Bakker PI, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet.* 2006;38(10):1166–1172.
4. Fernando MM, et al. Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet.* 2008;4(4):e1000024.
5. Rioux JD, et al. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. *Proc Natl Acad Sci U S A.* 2009;106(44):18680–18685.
6. Weinstock C, et al. Autoimmune polyglandular syndrome type 2 shows the same HLA class II pattern as type 1 diabetes. *Tissue Antigens.* 2011;77(4):317–324.
7. National Human Genome Research Institute. Division of Genomic Medicine. A Catalog of Published Genome-Wide Association Studies. NIH Web site. http://www.genome.gov/gwastudies/. Updated November 9, 2014. Accessed November 10, 2014.
8. Kawabata Y, et al. Asian-specific HLA haplotypes reveal heterogeneity of the contribution of HLA-DR and -DQ haplotypes to susceptibility to type 1 diabetes. *Diabetes.* 2002;51(2):545–551.
9. Katahira M, Ishiguro T, Segawa S, Kuzuya-Nagao K, Hara I, Nishisaki T. Reevaluation of human leukocyte antigen DR-DQ haplotype and genotype in type 1 diabetes in the Japanese population. *Horm Res.* 2008;69(5):284–289.
10. Kawabata Y, et al. Differential association of HLA with three subtypes of type 1 diabetes: fulminant, slowly progressive and acute-onset. *Diabetologia.* 2009;52(12):2513–2521.
11. Hashimoto K, et al. Susceptibility alleles and haplotypes of human leukocyte antigen DRB1, DQA1, and DQB1 in autoimmune polyglandular syndrome type III in Japanese population. *Horm Res.* 2005;64(5):253–260.
12. Matsushita S, Takahashi K, Motoki M, Komoriya K, Ikagawa S, Nishimura Y. Allele specificity of structural requirement for peptides bound to HLA-DRB1*0405 and -DRB1*0406 complexes: implication for the HLA-associated susceptibility to methimazole-induced insulin autoimmune syndrome. *J Exp Med.* 1994;180(3):873–883.
13. Wucherpfennig KW, et al. Structural basis for

major histocompatibility complex (MHC)-linked susceptibility to autoimmunity: charged residues of a single MHC binding pocket confer selective presentation of self-peptides in pemphigus vulgaris. *Proc Natl Acad Sci U S A.* 1995;92(25):11935–11939.

14. Wall M, et al. High affinity for class II molecules as a necessary but not sufficient characteristic of encephalitogenic determinants. *Int Immunol.* 1992;4(7):773–777.
15. Valli A, et al. Binding of myelin basic protein peptides to human histocompatibility leukocyte antigen class II molecules and their recognition by T cells from multiple sclerosis patients. *J Clin Invest.* 1993;91(2):616–628.
16. Wucherpfennig KW, et al. Structural requirements for binding of an immunodominant myelin basic protein peptide to DR2 isotypes and for its recognition by human T cell clones. *J Exp Med.* 1994;179(1):279–290.
17. Yu B, Gauthier L, Hausmann DH, Wucherpfennig KW. Binding of conserved islet peptides by human and murine MHC class II molecules associated with susceptibility to type I diabetes. *Eur J Immunol.* 2000;30(9):2497–2506.
18. Hall FC, et al. Relationship between kinetic stability and immunogenicity of HLA-DR4/peptide complexes. *Eur J Immunol.* 2002;32(3):662–670.
19. Fairchild PJ, Wildgoose R, Atherton E, Webb S, Wraith DC. An autoantigenic T cell epitope forms unstable complexes with class II MHC: a novel route for escape from tolerance induction. *Int Immunol.* 1993;5(9):1151–1158.
20. Mason K, Denney DW, Denney DW Jr, McConnell HM. Myelin basic protein peptide complexes with the class II MHC molecules I-Au and I-Ak form and dissociate rapidly at neutral pH. *J Immunol.* 1995;154(10):5216–5227.
21. Fugger L, Liang J, Gautam A, Rothbard JB, McDevitt HO. Quantitative analysis of peptides from myelin basic protein binding to the MHC class II protein, I-Au, which confers susceptibility to experimental allergic encephalomyelitis. *Mol Med.* 1996;2(2):181–188.
22. Muraro PA, et al. Immunodominance of a low-affinity major histocompatibility complex-binding myelin basic protein epitope (residues 111-129) in HLA-DR4 (B1*0401) subjects is associated with a restricted T cell receptor repertoire. *J Clin Invest.* 1997;100(2):339–349.
23. Boyton RJ, et al. Glutamic acid decarboxylase T lymphocyte responses associated with susceptibility or resistance to type I diabetes: analysis in disease discordant human twins, non-obese diabetic mice and HLA-DQ transgenic mice. *Int Immunol.* 1998;10(12):1765–1776.
24. He XL, Radu C, Sidney J, Sette A, Ward ES, Garcia KC. Structural snapshot of aberrant antigen presentation linked to autoimmunity: the immunodominant epitope of MBP complexed with I-Au. *Immunity.* 2002;17(1):83–94.
25. Levisetti MG, Suri A, Petzold SJ, Unanue ER. The insulin-specific T cells of nonobese diabetic mice recognize a weak MHC-binding segment in more than one form. *J Immunol.* 2007;178(10):6051–6057.
26. Levisetti MG, Lewis DM, Suri A, Unanue ER. Weak proinsulin peptide-major histocompatibility complexes are targeted in autoimmune diabetes in mice. *Diabetes.* 2008;57(7):1852–1860.
27. Stadinski BD, et al. Chromogranin A is an autoantigen in type 1 diabetes. *Nat Immunol.* 2010;11(3):225–231.
28. Stadinski BD, Zhang L, Crawford F, Marrack P, Eisenbarth GS, Kappler JW. Diabetogenic T cells recognize insulin bound to IAg7 in an unexpected, weakly binding register. *Proc Natl Acad Sci U S A.* 2010;107(24):10978–10983.
29. Crawford F, et al. Specificity and detection of insulin-reactive CD4+ T cells in type 1 diabetes in the nonobese diabetic (NOD) mouse. *Proc Natl Acad Sci U S A.* 2011;108(40):16729–16734.
30. Ettinger RA, Kwok WW. A peptide binding motif for HLA-DQA1*0102/DQB1*0602, the class II MHC molecule associated with dominant protection in insulin-dependent diabetes mellitus. *J Immunol.* 1998;160(5):2365–2373.
31. Harfouch-Hammoud E, et al. Identification of peptides from autoantigens GAD65 and IA-2 that bind to HLA class II molecules predisposing to or protecting from type 1 diabetes. *Diabetes.* 1999;48(10):1937–1947.
32. Astill TP, Ellis RJ, Arif S, Tree TIM, Peakman M. Promiscuous binding of proinsulin peptides to type 1 diabetes-permissive and -protective HLA class II molecules. *Diabetologia.* 2003;46(4):496–503.
33. Carrasco-Marin E, Shimizu J, Kanagawa O, Unanue E. The class II MHC I-Ag7 molecules from non-obese diabetic mice are poor peptide binders. *J Immunol.* 1996;156(2):450–458.
34. Reizis B, et al. Molecular characterization of the diabetes-associated mouse MHC class II protein, I-Ag7. *Int Immunol.* 1997;9(1):43–51.
35. Nabavieh A, et al. Development of an I-Ag7-expressing antigen-presenting cell line: Intrinsic molecular defect in compact I-Ag7 dimer generation. *J Autoimmun.* 1998;11(1):63–71.
36. Ettinger RA, Liu AW, Nepom GT, Kwok WW. Exceptional stability of the HLA-DQA1*0102/DQB1*0602 αβ protein dimer, the class II MHC molecule associated with protection from insulin-dependent diabetes mellitus. *J Immunol.* 1998;161(11):6439–6445.
37. Sadegh-Nasseri S, Germain RN. How MHC class II molecules work: peptide-dependent completion of protein folding. *Immunol Today.* 1992;13(2):43–46.
38. Nelson CA, Petzold SJ, Unanue ER. Identification of two distinct properties of class II major histocompatibility complex-associated peptides. *Proc Natl Acad Sci U S A.* 1993;90(4):1227–1231.
39. Verreck FA, Termijtelen A, Koning F. HLA-DR β chain residue 86 controls DR αβ dimer stability. *Eur J Immunol.* 1993;23(6):1346–1350.
40. Nelson C, et al. Amino acid residues on the I-Ak α-chain required for the binding and stability of two antigenic peptides. *J Immunol.* 1996;156(1):176–182.
41. Nelson CA, Viner NJ, Young SP, Petzold SJ, Unanue ER. A negatively charged anchor residue promotes high affinity binding to the MHC class II molecule I-Ak. *J Immunol.* 1996;157(2):755–762.
42. Wu S, Gorski J, Eckels DD, Newton-Nash DK. T cell recognition of MHC class II-associated peptides is independent of peptide affinity for MHC and sodium dodecyl sulfate stability of the peptide/MHC complex. Effects of conservative amino acid substitutions at anchor position 1 of influenza matrix protein19-31. *J Immunol.* 1996;156(10):3815–3820.
43. Reizis B, Altmann DM, Cohen IR. Biochemical characterization of the human diabetes-associated HLA-DQ8 allelic product: Similarity to the major histocompatibility complex class II I-Ag7 protein of non-obese diabetic mice. *Eur J Immunol.* 1997;27(10):2478–2483.
44. Natarajan SK, Stern LJ, Sadegh-Nasseri S. Sodium dodecyl sulfate stability of HLA-DR1 complexes correlates with burial of hydrophobic residues in pocket 1. *J Immunol.* 1999;162(6):3463–3470.
45. Ettinger RA, Liu AW, Nepom GT, Kwok WW. β57-Asp plays an essential role in the unique SDS stability of HLA-DQA1*0102/DQB1*0602 αβ protein dimer, the class II MHC allele associated with protection from insulin-dependent diabetes mellitus. *J Immunol.* 2000;165(6):3232–3238.
46. Verreck FAW, et al. The generation of SDS-stable HLA DR dimers is independent of efficient peptide binding. *Int Immunol.* 1996;8(3):397–404.
47. Norcross MA, Bentley DM, Margulies DH, Germain RN. Membrane Ia expression and antigen-presenting accessory cell function of L cells transfected with class II major histocompatibility complex genes. *J Exp Med.* 1984;160(5):1316–1337.
48. Austin P, Trowsdale J, Rudd C, Bodmer W, Feldmann M, Lamb J. Functional expression of HLA-DP genes transfected into mouse fibroblasts. *Nature.* 1985;313(5997):61–64.
49. Sant AJ, Hendrix LR, Coligan JE, Maloy WL, Germain RN. Defective intracellular-transport as a common mechanism limiting expression of inappropriately paired class-Ii major histocompatibility complex alpha/beta chains. *J Exp Med.* 1991;174(4):799–808.
50. Ceman S, Wu S, Jardetzky TS, Sant AJ. Alteration of a single hydrogen bond between class II molecules and peptide results in rapid degradation of class II molecules after invariant chain removal. *J Exp Med.* 1998;188(11):2139–2149.
51. Arneson LS, Katz JF, Liu M, Sant AJ. Hydrogen bond integrity between MHC class II molecules and bound peptide determines the intracellular fate of MHC class II molecules. *J Immunol.* 2001;167(12):6939–6946.
52. Kitamura T, et al. Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. *Exp Hematol.* 2003;31(11):1007–1014.
53. Morita S, Kojima T, Kitamura T. Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.* 2000;7(12):1063–1066.
54. Giles RC, DeMars R, Chang CC, Capra JD. Allelic polymorphisms and transassociation of molecules encoded by the HLA-DQ subregion. *Proc Natl Acad Sci U S A.* 1985;82(6):1776–1780.
55. Nepom BS, Schwarz D, Palmer JP, Nepom GT. Transcomplementation of HLA genes in IDDM. HLA-DQ alpha- and beta-chains produce hybrid molecules in DR3/4 heterozygotes. *Diabetes.* 1987;36(1):114–117.
56. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association

of celiac disease to a particular HLA-DQ α/β heterodimer. *J Exp Med*. 1989;169(1):345–350.
57. Ronningen KS, Markussen G, Iwe T, Thorsby E. An increased risk of insulin-dependent diabetes mellitus (IDDM) among HLA-DR4,DQw8/DRw8,DQw4 heterozygotes. *Hum Immunol*. 1989;24(3):165–173.
58. Khalil I, Deschamps I, Lepage V, al-Daccak R, Degos L, Hors J. Dose effect of cis- and trans-encoded HLA-DQ αβ heterodimers in IDDM susceptibility. *Diabetes*. 1992;41(3):378–384.
59. Koeleman BPC, et al. Genotype effects and epistasis in type 1 diabetes and HLA-DQ trans dimer associations with disease. *Genes Immun*. 2004;5(5):381–388.
60. Kwok W, Schwarz D, Nepom B, Hock R, Thurtle P, Nepom G. HLA-DQ molecules form α-β heterodimers of mixed allotype. *J Immunol*. 1988;141(9):3123–3127.
61. Kwok W, Thurtle P, Nepom G. A genetically controlled pairing anomaly between HLA-DQα and HLA- DQβ chains. *J Immunol*. 1989;143(11):3598–3601.
62. Kwok W, Kovats S, Thurtle P, Nepom G. HLA-DQ allelic polymorphisms constrain patterns of class II heterodimer formation. *J Immunol*. 1993;150(6):2263–2272.
63. Raymond CK, et al. Ancient haplotypes of the HLA class II region. *Genome Res*. 2005;15(9):1250–1257.
64. Moore K, Cooper SA, Jones DB. Use of the monoclonal antibody WR17, identifying the CD37 gp40-45 Kd antigen complex, in the diagnosis of B-lymphoid malignancy. *J Pathol*. 1987;152(1):13–21.
65. Monos DS, Czanky E, Ono SJ, Radka SF, Kappes D, Strominger JL. L cells expressing DQ molecules of the DR3 and DR4 haplotypes: reactivity patterns with mAbs. *Immunogenetics*. 1995;42(3):172–180.
66. Kozono H, White J, Clements J, Marrack P, Kappler J. Production of soluble MHC class II proteins with covalently bound single peptides. *Nature*. 1994;369(6476):151–154.
67. Busch R, Cloutier I, Sekaly RP, Hammerling GJ. Invariant chain protects class II histocompatibility antigens from binding intact polypeptides in the endoplasmic reticulum. *EMBO J*. 1996;15(2):418–428.
68. Fallang L-E, et al. Complexes of Two Cohorts of CLIP Peptides and HLA-DQ2 of the Autoimmune DR3-DQ2 Haplotype Are Poor Substrates for HLA-DM. *J Immunol*. 2008;181(8):5451–5461.
69. Henderson KN, et al. A structural and immunological basis for the role of human leukocyte antigen DQ8 in celiac disease. *Immunity*. 2007;27(1):23–34.
70. Siebold C, et al. Crystal structure of HLA-DQ0602 that protects against type 1 diabetes and confers strong susceptibility to narcolepsy. *Proc Natl Acad Sci U S A*. 2004;101(7):1999–2004.
71. Reichstetter S, et al. Mutational analysis of critical residues determining antigen presentation and activation of HLA-DQ0602 restricted T-cell clones. *Hum Immunol*. 2002;63(3):185–193.
72. Kim CY, Quarsten H, Bergseng E, Khosla C, Sollid LM. Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. *Proc Natl Acad Sci U S A*. 2004;101(12):4175–4179.
73. Lee KH, Wucherpfennig KW, Wiley DC. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. *Nat Immunol*. 2001;2(6):501–507.
74. Hughes AL, Nei M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*. 1988;335(6186):167–170.
75. Hughes AL, Nei M. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. *Proc Natl Acad Sci U S A*. 1989;86(3):958–962.
76. Yang Z, Bielawski JP. Statistical methods for detecting molecular adaptation. *Trends Ecol Evol*. 2000;15(12):496–503.
77. Pugliese A, et al. The 13th International Histocompatibility Working Group for Type 1 Diabetes (T1D) Joint Report. In: Hansen JA, editor. *Immunobiology of the Human MHC: Proceedings of the 13th International Histocompatibility Workshop and Congress*. Vol 1. Seattle, Washington, USA: IHWG Press; 2007:788–796.
78. Steenkiste A, et al. 14th International HLA and Immunogenetics Workshop: report on the HLA component of type 1 diabetes. *Tissue Antigens*. 2007;69(suppl 1):214–225.
79. Thomson G, et al. Relative predispositional effects of HLA class II DRB1-DQB1 haplotypes and genotypes on type 1 diabetes: a meta-analysis. *Tissue Antigens*. 2007;70(2):110–127.
80. Erlich H, et al. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. *Diabetes*. 2008;57(4):1084–1092.
81. Noble JA, Johnson J, Lane JA, Valdes AM. HLA class II genotyping of African American type 1 diabetic patients reveals associations unique to African haplotypes. *Diabetes*. 2013;62(9):3292–3299.
82. Park Y, et al. Common susceptibility and transmission pattern of human leukocyte antigen DRB1-DQB1 haplotypes to Korean and Caucasian patients with type 1 diabetes. *J Clin Endocrinol Metab*. 2000;85(12):4538–4542.
83. Bugawan TL, et al. The association of specific HLA class I and II alleles with type 1 diabetes among Filipinos. *Tissue Antigens*. 2002;59(6):452–469.
84. Redondo MJ, et al. DR- and DQ-associated protection from type 1A diabetes: comparison of DRB1*1401 and DQA1*0102-DQB1*0602. *J Clin Endocrinol Metab*. 2000;85(10):3793–3797.
85. Khalil I, et al. A combination of HLA-DQ beta Asp57-negative and HLA DQ alpha Arg52 confers susceptibility to insulin-dependent diabetes mellitus. *J Clin Invest*. 1990;85(4):1315–1319.
86. Sanjeevi CB, Landin-Olsson M, Kockum I, Dahlquist G, Lernmark Å. The combination of several polymorphic amino acid residues in the DQα and DQβ chains forms a domain structure pattern and is associated with insulin-dependent diabetes mellitus. *Ann N Y Acad Sci*. 2002;958:362–375.
87. Todd JA, Bell JI, McDevitt HO. HLA-DQβ gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature*. 1987;329(6140):599–604.
88. Morel PA, Dorman JS, Todd JA, McDevitt HO, Trucco M. Aspartic acid at position 57 of the HLA-DQ beta chain protects against type I diabetes: a family study. *Proc Natl Acad Sci U S A*. 1988;85(21):8111–8115.
89. Acha-Orbea H, McDevitt HO. The first external domain of the nonobese diabetic mouse class II I-A β chain is unique. *Proc Natl Acad Sci U S A*. 1987;84(8):2435–2439.
90. Kwok WW, Domeier ME, Johnson ML, Nepom GT, Koelle DM. HLA-DQB1 codon 57 is critical for peptide binding and recognition. *J Exp Med*. 1996;183(3):1253–1258.
91. Kwok WW, Domeier ML, Raymond FC, Byers P, Nepom GT. Allele-specific motifs characterize HLA-DQ interactions with a diabetes-associated peptide derived from glutamic acid decarboxylase. *J Immunol*. 1996;156(6):2171–2177.
92. Oiso M, Nishi T, Ishikawa T, Nishimura Y, Matsushita S. Differential binding of peptides substituted at putative C-terminal anchor residues to HLA-DQ8 and DQ9 differing only at β57. *Hum Immunol*. 1997;52(1):47–53.
93. Suri A, Vidavsky I, van der Drift K, Kanagawa O, Gross ML, Unanue ER. In APCs, the autologous peptides selected by the diabetogenic I-Ag7 molecule are unique and determined by the amino acid changes in the P9 pocket. *J Immunol*. 2002;168(3):1235–1243.
94. Suri A, Walters JJ, Gross ML, Unanue ER. Natural peptides selected by diabetogenic DQ8 and murine I-Ag7 molecules show common sequence specificity. *J Clin Invest*. 2005;115(8):2268–2276.
95. Yoshida K, Corper AL, Herro R, Jabri B, Wilson IA, Teyton L. The diabetogenic mouse MHC class II molecule I-Ag7 is endowed with a switch that modulates TCR affinity. *J Clin Invest*. 2010;120(5):1578–1590.
96. Sidney J, del Guercio M-F, Southwood S, Sette A. The HLA molecules DQA1*0501/B1*0201 and DQA1*0301/B1*0302 share an extensive overlap in peptide binding specificity. *J Immunol*. 2002;169(9):5098–5108.
97. Johansen BH, Jensen T, Thorpe CJ, Vartdal F, Thorsby E, Sollid LM. Both α and β chain polymorphisms determine the specificity of the disease-associated HLA-DQ2 molecules, with β chain residues being most influential. *Immunogenetics*. 1996;45(2):142–150.
98. van de Wal Y, Kooy YM, Drijfhout JW, Amons R, Papadopoulos GK, Koning F. Unique peptide binding characteristics of the disease-associated DQ(α1*0501, β1*0201) vs the non-disease-associated DQ(α1*0201, β1*0202) molecule. *Immunogenetics*. 1997;46(6):484–492.
99. Liu GY, Fairchild PJ, Smith RM, Prowle JR, Kioussis D, Wraith DC. Low avidity recognition of self-antigen by T cells permits escape from central tolerance. *Immunity*. 1995;3(4):407–415.
100. James EA, Kwok WW. Low-affinity major histocompatibility complex-binding peptides in type 1 diabetes. *Diabetes*. 2008;57(7):1788–1789.
101. Kim DT, Rothbard JB, Bloom DD, Fathman CG. Quantitative analysis of T cell activation: role of TCR/ligand density and TCR affinity. *J Immunol*. 1996;156(8):2737–2742.
102. Baumgartner CK, Ferrante A, Nagaoka M, Gorski J, Malherbe LP. Peptide-MHC class II complex stability governs CD4 T cell clonal selection. *J Immunol*. 2010;184(2):573–581.

103. O'Garra A, Gabrysova L, Spits H. Quantitative events determine the differentiation and function of helper T cells. *Nat Immunol.* 2011;12(4):288–294.
104. Corse E, Gottschalk RA, Allison JP. Strength of TCR-peptide/MHC interactions and in vivo T cell responses. *J Immunol.* 2011;186(9):5039–5045.
105. Gottschalk RA, et al. Distinct influences of peptide-MHC quality and quantity on in vivo T-cell responses. *Proc Natl Acad Sci U S A.* 2012;109(3):881–886.
106. Eerligh P, et al. Functional consequences of HLA-DQ8 homozygosity versus heterozygosity for islet autoimmunity in type 1 diabetes. *Genes Immun.* 2011;12(6):415–427.
107. van Lummel M, et al. The type 1 diabetes associated HLA-DQ8-trans dimer accomodates a unique peptide repertoire. *J Biol Chem.* 2012;287(12):9514.
108. Schmidt D, Verdaguer J, Averill N, Santamaria P. A mechanism for the major histocompatibility complex-linked resistance to autoimmunity. *J Exp Med.* 1997;186(7):1059–1075.
109. Nepom GT. A unified hypothesis for the complex genetics of HLA associations with IDDM. *Diabetes.* 1990;39(10):1153–1157.
110. Deng H, et al. Determinant capture as a possible mechanism of protection afforded by major histocompatibility complex class II molecules in autoimmune disease. *J Exp Med.* 1993;178(5):1675–1680.
111. Resic-Lindehammer S, et al. Temporal trends of HLA genotype frequencies of type 1 diabetes patients in Sweden from 1986 to 2005 suggest altered risk. *Acta Diabetologica.* 2008;45(4):231–235.
112. Carlsson A, et al. Low risk HLA-DQ and increased body mass index in newly diagnosed type 1 diabetes children in the Better Diabetes Diagnosis study in Sweden. *Int J Obes.* 2011;36(5):718.
113. Patil NS, et al. Rheumatoid arthritis (RA)-associated HLA-DR alleles form less stable complexes with class II-associated invariant chain peptide than non-RA-associated HLA-DR alleles. *J Immunol.* 2001;167(12):7157–7168.
114. Hou T, et al. An insertion mutant in DQA1*0501 restores susceptibility to HLA-DM: implications for disease associations. *J Immunol.* 2011;187(5):2442–2452.
115. Busch R, De Riva A, Hadjinicolaou AV, Jiang W, Hou T, Mellins ED. On the perils of poor editing: regulation of peptide loading by HLA-DQ and H2-A molecules associated with celiac disease and type 1 diabetes. *Expert Rev Mol Med.* 2012;14:e15.
116. Busch R, et al. Achieving stability through editing and chaperoning: regulation of MHC class II peptide binding and expression. *Immunol Rev.* 2005;207:242–260.
117. Zhou Z, Jensen PE. Structural characteristics of HLA-DQ that may impact DM editing and susceptibility to Type-1 diabetes. *Front Immunol.* 2013;4:262.
118. Zhong G, Castellino F, Romagnoli P, Germain RN. Evidence that binding site occupancy is necessary and sufficient for effective major histocompatibility complex (MHC) class II transport through the secretory pathway redefines the primary function of class II-associated invariant chain peptides (CLIP). *J Exp Med.* 1996;184(5):2061–2066.
119. Cotner T. Unassembled HLA-DR beta monomers are degraded rapidly by a nonlysosomal mechanism. *J Immunol.* 1992;148(7):2163–2168.
120. Cotner T, Pious D. HLA-DR beta chains enter into an aggregated complex containing GRP-78/BiP prior to their degradation by the pre-Golgi degradative pathway. *J Biol Chem.* 1995;270(5):2379–2386.
121. Apanius V, Penn D, Slev PR, Ruff LR, Potts WK. The nature of selection on the major histocompatibility complex. *Crit Rev Immunol.* 1997;17(2):179–224.
122. Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res.* 2010;39(Database issue):D913–D919.
123. Devaux B, Wilson KJ, Aguilar B, Jorgensen B, Rothbard JB. Differential stability of HLA-DR alleles independent of endogenous peptides. *J Immunol.* 1995;155(4):1921–1929.
124. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol.* 1986;3(5):418–426.
125. Jukes TH, Cantor CR. Evolution of protein molecules. In: Munro HN, ed. *Mammalian Protein Metabolism.* New York, New York, USA: Academic Press; 1969:21–132.
126. Korber B. HIV signature and sequence variation analysis. In: Rodrigo AG, Learn GH, eds. *Computational Analysis of HIV Molecular Sequences.* Dordrecht, Netherlands: Kluwer Academic Publishers; 2000:55–72.
127. Ho CS, et al. Nomenclature for factors of the SLA system, update 2008. *Tissue Antigens.* 2009;73(4):307–315.
128. Robinson J, Mistry K, McWilliam H, Lopez R, Parham P, Marsh SG. The IMGT/HLA database. *Nucleic Acids Res.* 2010;39(Database issue):D1171–D1176.
129. de Groot NG, et al. Nomenclature report on the major histocompatibility complex genes and alleles of Great Ape, Old and New World monkey species. *Immunogenetics.* 2012;64(8):615–631.
130. Robinson J, Halliwell JA, McWilliam H, Lopez R, Marsh SG. IPD — the Immuno Polymorphism Database. *Nucleic Acids Res.* 2013;41(Database issue):D1234–D1240.
131. Gonzalez-Galarza FF, Christmas S, Middleton D, Jones AR. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res.* 2011;39(Database issue):D913–D919.
132. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol.* 2007;24(8):1569–1599.
133. Klitz W, et al. New HLA haplotype frequency reference standards: high-resolution and large sample typing of HLA DR-DQ haplotypes in a sample of European Americans. *Tissue Antigens.* 2003;62(4):296–307.
134. Hashimoto M, et al. Gene frequencies and haplotypic associations within the HLA region in 916 unrelated Japanese individuals. *Tissue Antigens.* 1994;44(3):166–173.
135. Renquin J, et al. HLA class II polymorphism in Aka Pygmies and Bantu Congolese and a reassessment of HLA-DRB1 African diversity. *Tissue Antigens.* 2001;58(4):211–222.
136. Agrawal S, Khan F, Bharadwaj U. Human genetic variation studies and HLA class II loci. *Int J Immunogenet.* 2007;34(4):247–252.
137. Spurkland A, Sollid LM, Polanco I, Vartdal F, Thorsby E. HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. *Hum Immunol.* 1992;35(3):188–192.
138. Ferreira RC, et al. High-density SNP mapping of the HLA region identifies multiple independent susceptibility loci associated with selective IgA deficiency. *PLoS Genet.* 2012;8(1):e1002476.
139. Sawcer S, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.* 2011;476(7359):214–219.
140. Tsuchiya N. Genetics of ANCA-associated vasculitis in Japan: a role for HLA-DRB1*09:01 haplotype. *Clin Exp Nephrol.* 2012;17(5):628–630.
141. Mignot E, et al. Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. *Am J Hum Genet.* 2001;68(3):686–699.

NEOPLASIA
www.neoplasia.com

Volume 16 Number 8 August 2014 pp. 656–664 656

Late Cornified Envelope Group I, a Novel Target of p53, Regulates PRMT5 Activity[1]

Zhenzhong Deng*,†, Koichi Matsuda†, Chizu Tanikawa†, Jiaying Lin*,†, Yoichi Furukawa‡, Ryuji Hamamoto* and Yusuke Nakamura*

*Section of Hematology/Oncology, Department of Medicine, The University of Chicago, Chicago, IL 60637 USA; †Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; ‡Division of Clinical Genome Research, Advanced Clinical Research Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Abstract

p53 is one of the most important tumor suppressor genes involved in human carcinogenesis. Although downstream targets of p53 and their biologic functions in cancer cells have been extensively investigated, it is still far from the full understanding. Here, we demonstrate that *Late Cornified Envelope Group I* (*LCE1*) genes, which are located in the *LCE* gene clusters encoding multiple well-conserved stratum-corneum proteins, are novel downstream targets of p53. Exogenous p53 overexpression using an adenoviral vector system significantly enhanced the expression of *LCE1* cluster genes. We also observed induction of LCE1 expressions by DNA damage, which was caused by treatment with adriamycin or UV irradiation in a wild-type p53-dependent manner. Concordantly, the induction of *LCE1* by DNA damage was significantly attenuated by the knockdown of p53. Among predicted p53-binding sites within the *LCE1* gene cluster, we confirmed one site to be a p53-enhancer sequence by reporter assays. Furthermore, we identified LCE1 to interact with protein arginine methyltransferase 5 (PRMT5). Knockdown of LCE1 by specific small interfering RNAs significantly increased the symmetric dimethylation of histone H3 arginine 8, a substrate of PRMT5, and overexpression of LCE1F remarkably decreased its methylation level. Our data suggest that LCE1 is a novel p53 downstream target that can be directly transactivated by p53 and is likely to have tumor suppressor functions through modulation of the PRMT5 activity.

Neoplasia (2014) 16, 656–664

Introduction

p53 is the most frequently mutated tumor suppressor gene involved in human cancers [1,2]. Its tetramer protein product can activate the transcription of a number of target downstream genes and mediate a variety of biologic functions through the transcriptional regulation of those targets [3]. To elucidate the critical roles of p53 in human carcinogenesis, we and others have attempted to identify p53 target genes through multiple approaches. We have mainly applied the expression profile analysis after the exogenous introduction of wild-type p53 into cancer cells using the adenovirus vector system and identified more than 50 p53 downstream candidate genes [4]. Among them, we have performed the functional analysis of more than a dozen of target genes including *p53R2*, *p53AIP1*, and *p53RDL1* [5–8]. Here, we report the characterization of the Late Cornified Envelope Group I (LCE1) as a novel downstream target of p53.

The *LCE* clusters contain multiple well-conserved genes encoding stratum-corneum proteins [9,10] and are located on chromosome 1q21 in a region called as the epidermal differentiation complex [11,12]. This region is enriched for genes, which are expressed during

Address all correspondence to: Yusuke Nakamura, MD, PhD, Section of Hematology/Oncology, Department of Medicine, The University of Chicago, 5841 S. Maryland Ave., MC2115 Chicago, IL 60637. E-mail: ynakamura@bsd.uchicago.edu

[1] This article refers to supplementary materials, which are designated by Tables W1 to W6 and Figures W1 to W6 and are available online at www.neoplasia.com.

Received 10 June 2014; Revised 26 July 2014; Accepted 29 July 2014

1476-5586/14
http://dx.doi.org/10.1016/j.neo.2014.07.008

epidermal differentiation, including *loricrin*, *involucrin*, *filagrin*, the small proline-rich protein genes, and the *LCE* genes [9,13]. In mice, members in the LCE1 group are expressed in the relatively late stage of epithelial development and incorporated into the cornified envelope through cross-linking by transglutaminases [10]. In addition, real-time quantitative polymerase chain reaction (qPCR) analysis demonstrated that human *LCE1* and *LCE2* genes were primarily expressed in skin, whereas LCE4 and LCE5 gene expressions were undetectable in any human tissues examined [9]. In general, physiological functions of LCE proteins, especially their involvement in human cancer are still largely unknown.

Protein arginine methyltransferases (PRMTs) constitute of a large family of enzymes having the arginine methyltransferases activity responsible for catalyzing the formation of monomethyl arginine, asymmetric dimethyl arginine, and symmetric dimethyl arginine (SDMA) [14]. PRMT5 is one of the most well-characterized family members with SDMA activity and catalyzes formation of SDMA in proteins with a glycine and arginine-rich motif [15]. PRMT5 was reported to regulate various cellular functions including apoptosis, Golgi structure, pluripotency, cell growth, and snRNP biosynthesis [16–18]. One important key marker of the PRMT5 activity is the symmetrical dimethylation of histone 3 arginine 8 (H3R8me2s) level. Through hypermethylation of histone H3R8 around the promoter regions, PRMT5 could cause the transcriptional silencing of cell cycle regulator genes [19,20]. Since overexpression of PRMT5 has been reported in various types of human cancer, including melanoma, leukemia, lymphoma, glioma, as well as ovarian, breast, prostate, and lung cancers [16,21–23], this enzyme is considered as a good molecular target for development of novel cancer therapy [16].

In the present study, we demonstrate that LCE is a novel direct target of p53, can interact with PRMT5, and might modulate histone H3 methylation by PRMT5. This mechanism may be important for the interplay of two important cancer-related genes, *p53* and *PRMT5*, and our findings could indicate a possible role of LCE1 in human carcinogenesis.

Materials and Methods

Microarray Analysis

Replication-deficient recombinant adenovirus designed to express wild-type p53 (Ad-p53) or LacZ (Ad-LacZ) was generated and purified, as previously described [5,24]. Microarray analysis was then carried out as previously described [1,4,24]. In brief, poly(A)+ RNAs were isolated from U373MG cells at different time points after infection with Ad-p53 or Ad-LacZ. Each RNA sample was labeled and hybridized to a microarray consisting of 36,864 genes or Expressed Sequence Tags (ESTs) (http://www.ncbi.nlm.nih.gov/geo/index.cgi, Accession No. GSE14953).

Cell Culture and Transfection

Human cell lines, U373MG (glioblastoma), H1299 (lung carcinoma), and Human Embryonic Kidney 293T (HEK293T), were purchased from American Type Culture Collection (ATCC, Manassas, VA). HCT116 $p53^{-/-}$ and HCT116 $p53^{+/+}$ cell lines were obtained from Dr Bert Vogelstein (Johns Hopkins University, Baltimore, MD). U373MG glioblastoma cells and H1299 lung cancer cells were infected with Ad-p53 or Ad-LacZ at various multiplicity of infection (MOI) conditions and incubated at 37°C until the time of harvest. HEK293T cells were transfected with HA-Mock (empty vector of pCAGGSnHC) or HA-LCE1F (pCAGGSnHC-LCE1F) using FuGENE6. For protein-protein interaction experiments, HEK293T cells were transfected with HA-Mock or HA-LCE1F and FLAG-Mock or FLAG-PRMT5 using FuGENE6. To examine the co-localization, HCT116 $p53^{+/+}$ cells were transfected by HA-Mock or HA-LCE1F using FuGENE6. For the gene reporter assay, U373MG (mutated p53) and H1299 (p53 null) cells were transfected with a reporter plasmid and either a Mock empty vector of pcDNA3.1+ or a wild-type p53 expression pcDNA3.1+ vector in combination with a pRL-CMV vector using FuGENE6 [25,26]. Small interfering RNAs (siRNAs) that were commercially synthesized by Sigma-Aldrich (St Louis, MO) were transfected with Lipofectamine RNAiMAX reagent (Life Technologies, Carlsbad, CA). Two siRNA oligonucleotides of *LCE1A-F* were designed to target the region commonly conserved among *LCE1A-F*. Sequences of oligonucleotides are shown in Table W1. Western blot or real-time qPCR was applied to validate the efficiency of overexpression or knockdown experiments.

DNA-Damaging Treatments

Cells were seeded 24 hours before treatment. When cells reached 60% to 70% confluency, cells were incubated with 1 μg/ml adriamycin for 2 hours followed by further incubation of the drug-free medium. Then, the cells were harvested at different time points as indicated in the figure legends. For UV irradiation (UVR) experiments, the cells were washed twice with phosphate-buffered saline (PBS) and exposed to UV rays at different doses using XL-1500 Spectrolinker (Spectronics, Westbury, NY; peak emission, 254 nm). Cells were harvested 36 hours after UVR treatment.

Real-Time qPCR

Total RNA was isolated from cultured cells using RNeasy mini-spin column kits (Qiagen, Hilden, Germany) according to the manufacturer's procedure. cDNAs were synthesized with the SuperScript Pre-amplification System (Life Technologies). Real-time qPCR was conducted using the SYBR Green I Master on the LightCycler 480 (Roche Applied Science, Mannheim, Germany) according to the manufacturer's instructions. The primer sequences used in this manuscript are shown in Tables W1 and W2. Primers for *LCE1A*, *LCE1C*, *LCE1E*, and *LCE1F* were described previously [9]. Primers for *LCE1B*, *LCE1D LCE3A*, *LCE3B*, *LCE3C*, *LCE4A*, *XP33*, and *C1orf45* genes were designed by us. Except *LCE1D*, the specificity of all primers was confirmed by DNA sequencing of amplicons. For graphic representation of transcript data, all expression of target genes was shown relative to the housekeeping gene β_2-microglobulin expression in the same sample.

Prediction of Putative p53-Binding Sites

DNA sequences of an entire genomic region of *LCE1* including 10 kb of the 5′ upstream sequence were downloaded from the University of California Santa Cruz (UCSC) website (http://genome.ucsc.edu/), and the putative p53-binding sites (p53BSs) were screened according to the following criteria; at least 80% matched with the 20 nucleotides of consensus sequence RRRCWWGYYY__RRRCWWGYYY (R, purine; W, A, or T; Y, pyrimidine); we started the screening of 11 consensus sequences without any spacers between the two halves of p53BSs (Figure W1) and confirmed one site that was likely to be a direct p53-binding sequence.

Gene Reporter Assay

DNA fragments including each of the potential p53BSs of the *LCE1* gene cluster were amplified by KOD-Plus-DNA polymerase

(Toyobo, Osaka, Japan) and subcloned into the pGL3-promoter (pGL3-pro) vector. The primers for amplification are indicated in Table W1. To make a series of mutant vectors, a point mutation "T" was inserted into the site of the fourth and the fourteenth nucleotide "C" and into the seventh and the seventeenth nucleotide "G" of the consensus p53-BS using the KOD-Plus-Mutagenesis Kit (Toyobo). Since a functional p53BS is known in the *Fas* promoter region, a wild-type Fas promoter construct, pGL3-Fas, was used as a positive control. U373MG (mutated p53) and H1299 (p53 null) cells were plated in 12-well culture plates (5×10^4 cells per well) 24 hours before co-transfection of 125 ng of a reporter plasmid and either 125 ng of a Mock vector or a wild-type p53 expression vector in combination with 25 ng of a pRL-CMV vector. Cells were rinsed with PBS 36 hours after transfection and lysed in 250 μl of lysis buffer. Twenty microliters and 5 μl of lysates from U373MG and H1299, respectively, were sequentially measured using the PGD-S Dual Luciferase assay system following the manufacturer's procedure (Toyo Ink, Tokyo, Japan). The firefly luciferase value was normalized by the Renilla luciferase activity.

Expression Plasmid Construction

An entire coding sequence of PRMT5 was amplified using cDNA generated from mRNA of HEK293T cells and cloned into pCAGGSn3FC (PRMT5) vector. The entire coding sequences of *LCE1A-F* were amplified by the use of cDNA generated from mRNA that was extracted from HCT116 $p53^{+/+}$ cells after 70 J/m^2 of UVR treatment. Due to unknown reasons, we have been unsuccessful in cloning the *LCE1D* cDNA. An HA-epitope tag was placed at the C-terminus of the pCAGGSnHC vector and three-tandem FLAG epitope tags were placed at the C-terminus of the pCAGGSn3FC vector. The DNA sequences of expression constructs for LCE1 (pCAGGSnHC-LCE1) and PRMT5 (pCAGGSn3FC-PRMT5) were confirmed by DNA sequencing using ABI PRISM 3730XL Genetic Analyzer (Life Technologies).

Antibodies

The following primary antibodies were deployed: rabbit anti-HA (Y-11; Santa Cruz Biotechnology, Santa Cruz, CA; dilution used in WB: 1:1000), rabbit anti-PRMT5 (07-415; Millipore, Billerica, MA; dilution used in WB: 1:1000, ICC: 1:400), rabbit anti-p53 (sc6243, Santa Cruz; dilution used in WB: 1:1000), mouse anti-p53 [Ab-1, Calbiochem, San Diego, CA, dilution used in chromatin immunoprecipitation (ChIP): 1:100], rabbit anti-H3R8me2s (ab130740, Abcam, Cambridge, United Kingdom; dilution used in Western blot (WB): 1:1000), mouse anti–α-tubulin (clone DM1A, Millipore; dilution used in WB: 1:1000), rat anti-HA (3F10, Roche; dilution used in Immunocytochemistry (ICC): 1:800), and anti-FLAG (F7425, Sigma-Aldrich; dilution used in ICC: 1:1000).

Immunocytochemistry

Forty-eight hours after transfection with HA-Mock (empty vector of pCAGGSnHC) or pCAGGSnHC-HA-LCE1F into HEK293T or HCT116 $p53^{+/+}$ cells in four-well chambers, the cells were fixed by 1.7% formaldehyde or 4% paraformaldehyde in PBS and permeabilized with 0.2% Triton X-100 in PBS. After covering with blocking solution (3% BSA in 0.2% Triton X-100) for 1 hour at room temperature, cells were incubated with a rat anti-HA antibody and a rabbit PRMT5 antibody overnight under humidified atmosphere at 4°C. Further, the cells were stained with fluorescence-conjugated secondary antibodies and then counterstained with 4′,6-diamidino-2-phenylindole (DAPI) in VECTASHIELD Mounting Media (HT1200; Vector Laboratories, Burlingame, CA).

Immunoprecipitation and Mass-Spectrometric Analysis

Forty-eight hours after transfection with HA-Mock or HA-LCE1F, HEK293T cells were lysed in lysis buffer [50 mM Tris-HCl (pH 8.0), 0.4% NP-40, and 150 mM NaCl] containing Protease Inhibitor Cocktail Set III (Calbiochem). Whole-cell lysates were pre-cleared by incubation with normal mouse IgG (sc2025, Santa Cruz Biotechnology) and recombinant protein G-sepharose 4B (Life Technologies) at 4°C for 1 hour, followed by incubation with anti-HA agarose (A2095, Sigma-Aldrich) overnight. The proteins were separated in Mini-PROTEAN 5% to 20% gradient sodium dodecyl sulfate–polyacrylamide gel electrophoresis precast gels (Bio-Rad Laboratories, Hercules, CA) and stained with the silver-staining kit (Life Technologies). Protein bands that were specifically observed in the cell extracts transfected with HA-LCE1F compared with Mock were excised and analyzed by liquid chromatography–tandem mass spectrometry as previously described [27]. For the co-immunoprecipitation experiments, HEK293T cells were co-transfected with HA-LCE1F and FLAG-PRMT5 or FLAG-Mock as a control. An HA pull-down experiment is exactly the same as mentioned above. As for FLAG pull-down experiment, whole-cell lysates were pre-cleared by incubation with normal mouse IgG (sc2025, Santa Cruz Biotechnology) and recombinant protein G-sepharose 4B (Life Technologies) at 4°C for 1 hour, followed by incubation with mouse anti-FLAG M2 agarose (F3165, Sigma-Aldrich) overnight. Rat anti-HA (3 F10, Roche) and rabbit anti-Flag (F7425, Sigma-Aldrich) antibodies were used for Western blot. For the histone methylation analysis, histones were extracted by the histone purification mini kit (40026, Active Motif, Carlsbad, CA) following the manufacturer's procedure strictly.

ChIP Assay

ChIP assays were performed using ChIP Assay kit (17-295; Millipore) according to the manufacturer's protocol [28]. Briefly, HCT116 $p53^{+/+}$ cells were treated with 1 μg/ml adriamycin for 2 hours, and cells were cultured in the drug free-medium for 24 hours. Then, cells were harvested and the fragment of p53 and chromatin complexes was immnoprecipitated with an anti-p53 antibody (Ab-1, Calbiochem). After DNA fragments bound to p53 were eluted out, an aliquot was subjected to real-time qPCRs. Protein A agarose/Salmon Sperm DNA (16-157; Millipore) was used as a negative control. Primers were designed to amplify the region containing the p53BS2-binding site, and their sequence information is shown in Table W1.

Results

Identification of a Novel p53 Downstream Target

We had previously performed expression profile analysis to compare U373MG glioblastoma cells infected with wild-type p53 (Ad-p53) with those infected with LacZ (Ad-LacZ) to screen possible downstream genes that are regulated by p53 and identified more than 50 genes that were likely to be induced by wild-type p53. Among them, we confirmed that the transcriptional levels of LCE1B and LCE1C were elevated more than seven-fold higher in the cells infected with wild-type p53 than those with LacZ (data not shown). *LCE1B* and *LCE1C* belong to the *LCE* gene cluster containing multiple well-conserved genes that encode stratum-corneum proteins.

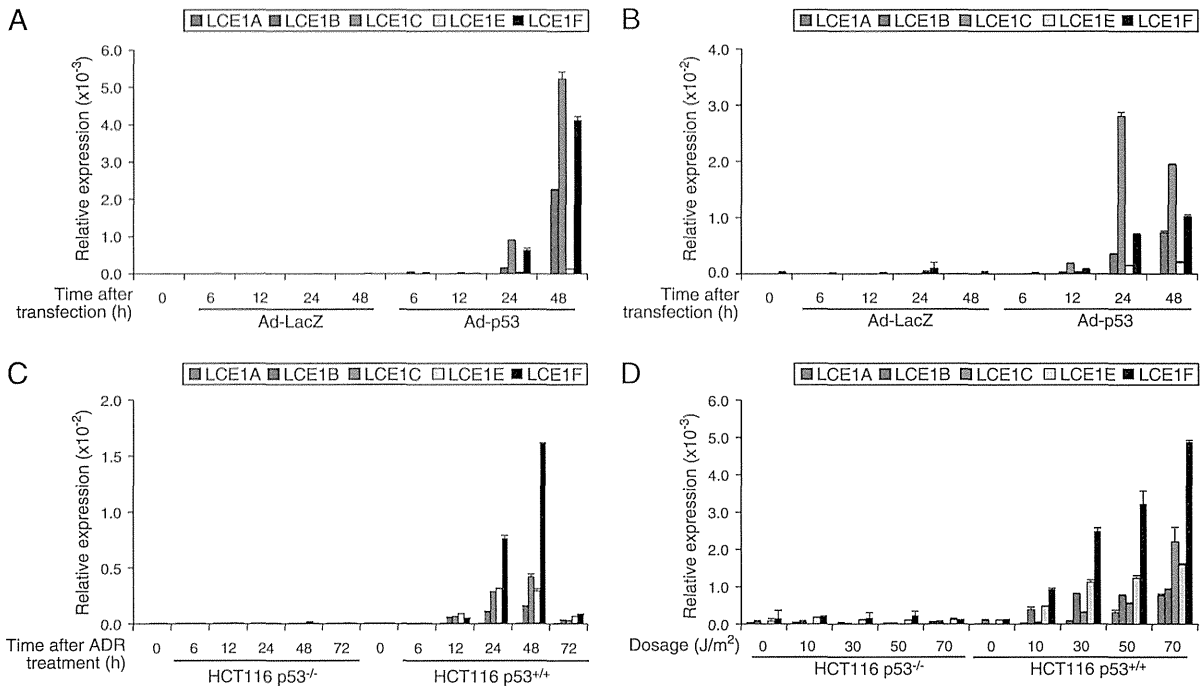

Figure 1. *LCE1* is upregulated by *p53*. Real-time qPCR analysis of *LCE1* mRNA expression in U373MG (*p53 mutant*) (A) and H1299 (*p53 null*) (B) cells at indicated time points after infection with Ad-p53 or Ad-LacZ at 8 MOIs. (C) Endogenous p53-dependent induction of *LCE1* mRNA transcription in $p53^{-/-}$ and $p53^{+/+}$ HCT116 cells after treatment with adriamycin (1 μg/ml) at various time points. (D) Cells were treated with UVR at different doses and harvested 36 hours after UVR. β_2-Microglobulin was used for the normalization of expression levels.

Within the *LCE* cluster, multiple genes form "groups" at chromosome 1q21 and are known to respond "group-wise" to various environmental stimuli like calcium and UV light [9], suggesting that other *LCE* members that were not included in our microarray might also be regulated by p53. We first compared the sequence similarity of the *LCE1B* and *LCE1C* transcripts with other members in the *LCE* cluster (Table W3) and selected genes showing high similarity (>80%) for further validation by real-time qPCR because the microarray results might reflect the cross-hybridization of other LCE members that were possibly induced by p53. Interestingly, expression levels of *LCE1* group genes were significantly increased in both U373MG (Figure 1*A*) and H1299 (Figure 1*B*) cells after 8 MOIs of Ad-p53 infection (Table W4), although the induction levels are very different. However, other genes in the LCE cluster were not induced by wild-type p53 introduction. In addition, endogenous p53, which was activated by adriamycin and UVR, could also induce expressions of the *LCE1* group genes in HCT116 $p53^{+/+}$ cancer cells but not in HCT116 $p53^{-/-}$ cancer cells (Figure 1, *C* and *D*), indicating that the LCE1 group genes can respond group-wise to the genotoxic stress condition in a p53-dependent manner [9].

LCE1 Is a Direct Target of p53

We then attempted to clarify whether *LCE1* is directly or indirectly regulated by p53. First, 11 putative p53BSs of 20 nucleotides with at least 80% match to the consensus p53-binding sequence (see Materials and Methods section) were selected from the *LCE1* cluster region (Figure W1*A*). We then constructed serial reporter vectors containing each of the 11 predicted p53BSs. Luciferase activity of p53BS2, p53BS6, and p53BS7 was likely to be enhanced by co-transfection of each luciferase vector and a p53 wild-type expression vector into U373MG (Figure W1*B*) and H1299 cells (Figure W1*C*) in comparison with those with the mock vector. Hence, these three p53BSs were further investigated by replacement of critical nucleotides in the core consensus sequence with other nucleotides (Figure 2*A*). As a result, we observed that substitutions in either of the p53BS2, p53BS6, and p53BS7 sequences significantly diminished the enhancement of the luciferase activity in both U373MG and H1299 cells (Figure 2, *B* and *C*).These results imply that p53 directly binds to these p53BSs in the *LCE1* cluster and regulates their transcriptions.

We subsequently transfected siRNA targeting p53 or that commonly targeting LCE1A-F into HCT116 $p53^{+/+}$ cells and then the cells were treated with adriamycin. As shown in Figure 2*D*, induction of *LCE1* genes by adriamycin treatment was significantly attenuated in the cells transfected with si-p53, also supporting that *LCE1* family genes are p53-direct targets. Additionally, ChIP analysis showed the direct binding of p53 to the *LCE1* cluster region in HCT116 $p53^{+/+}$ cells after adriamycin treatment (Figure W2). We then examined the expression of *LCE1F* in 82 cancer cell lines by real-time qPCR (Table W5) and confirmed that *LCE1F* was significantly downregulated in p53 mutant cells compared with p53 wild-type cells in colon and lung cancer cell lines (Figure W3). This is consistent with the induction of *LCE1F* by p53 in HCT116 colon and H1299 lung cancer cells as shown in Figure 1.

LCE1F Interacts with PRMT5

To further analyze possible biologic functions of LCE1F in cancer cells, we constructed an expression vector designed to express full-

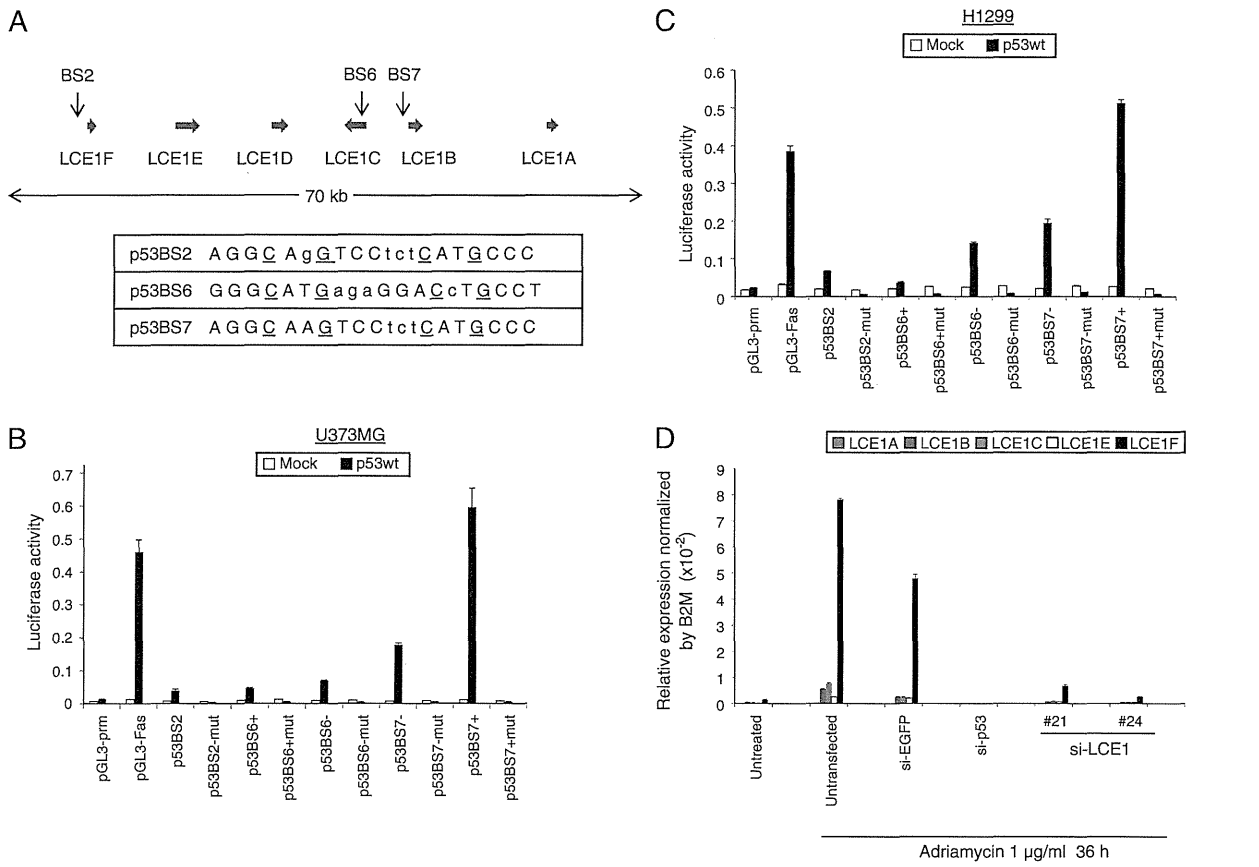

Figure 2. Identification of p53BSs in the *LCE1* cluster genes. (A) In the *LCE1* cluster genes, we identified 11 potential p53BSs that have at least 80% match to the consensus p53-binding sequence without any spacer nucleotides between the two halves of p53BSs. Among them, three possible p53 enhancer sequences are shown in the figure. Bold horizontal arrows indicate the locations and relative sizes of each *LCE1* gene. Vertical arrows show the potential p53BS locations. Identical nucleotides to the p53-binding sequence are written in capital letters. The underlined cytosine and guanine were substituted with thymine to introduce a mutant-type reporter vector at each p53BS. (B, C) Reporter assays of wild-type p53 and mutant-type p53 in U373MG (B) and H1299 (C) cells. Cells were co-transfected with a Mock (empty pcDNA3.1 +) or a wild-type p53 vector and each reporter plasmid, and luciferase assays were conducted 36 hours after transfection. Results are shown as the firefly luciferase activity normalized by the Renilla luciferase activity with 1 SD. (D) Real-time qPCR of *LCE1* family genes. HCT116 $p53^{+/+}$ cells were transfected with siRNAs targeting EGFP, p53, or LCE1 for 24 hours and then treated with 1 μg/ml adriamycin. Total RNA was extracted from the cells 36 hours after adriamycin treatment.

length LCE1F protein (pCAGGS-nHC-LCE1F) and transfected it into HEK293T cells. After confirmation of the LCE1F protein expression by Western blot analysis (Figure W4*A*), we performed immunocytochemical analysis to examine the subcellular localization of LCE1F. As shown in Figure W4*B*, LCE1F was strongly stained in both nucleus and cytoplasm.

In addition, we attempted to identify the interacting protein(s) of LCE1F. An HA-tagged LCE1F expression vector or an empty mock vector was transfected into HEK293T cells, and cell lysates were immunoprecipitated with anti–HA-conjugated agarose beads. Protein samples were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis, and the gel was silver stained. We found a 70-kDa protein, which was strongly stained only in HA-LCE1F–overexpressing cells, conducted liquid chromatography–tandem mass spectrometry analysis of this protein, and identified it to be an arginine methyltransferase, PRMT5. This result was confirmed using an anti–PRMT5-specific antibody (Figures 3*A* and W5). Reversibly, cell lysates of HA-LCE1F–overexpressed cells were immunoprecipitated with an anti-PRMT5 antibody, and the immunoprecipitates were blotted with an anti-HA antibody, which also validated the interaction between LCE1F and PRMT5 (Figure 3*B*). Moreover, we co-transfected either FLAG-PRMT5 or FLAG-Mock with HA-LCE1F or HA-Mock, and cell lysates were immunoprecipitated with an anti-HA antibody (Figure 3*C*) or an anti-FLAG antibody (Figure 3*D*). Subsequent Western blot analysis clearly indicated the interaction between PRMT5 and LCE1F. These results imply that LCE1F interacts with PRMT5 in the cells. Interestingly, we confirmed that all LCE1 family proteins examined could interact with PRMT5 (Figure W6), suggesting that the interaction with PRMT5 is likely to be a common characteristic among LCE1 family members.

LCE1F Suppresses PRMT5 Activity

Since PRMT5 was reported to regulate the transcription of various genes through the methylation of arginine 8 on histone H3 (H3R8), which is considered to play an important role in human carcinogenesis [16,20,21], we further investigated the biologic

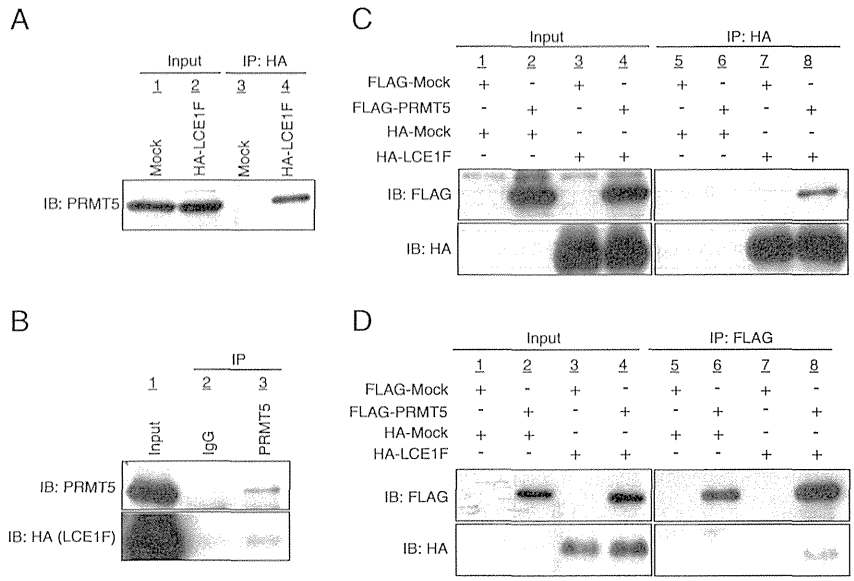

Figure 3. Interaction between LCE1F and PRMT5. (A) HEK293T cells were transfected with either a Mock (empty) or an HA-LCE1F expression vector and immunoprecipitated with anti-HA agarose. Samples were immunoblotted with an anti-PRMT5 antibody (Millipore). (B) HA-LCF1F–transfected HEK293T cells were immunoprecipitated with an anti-PRMT5 antibody or normal rabbit IgG (negative control). Samples were immunoblotted with an anti-PRMT5 or anti-HA antibody. (C, D) HEK293T cells were co-transfected with either FLAG-Mock or FLAG-PRMT5 and either HA-Mock or HA-LCE1F. Cells were immunoprecipitated with anti-HA agarose (C) or anti-FLAG antibody (D), and samples were immunoblotted with anti-FLAG or anti-HA antibodies.

significance of the interaction of PRMT5 and LCE1 members. We first performed immunocytochemical analysis and detected co-localization of LCE1F and PRMT5 in both the nucleus and the cytoplasm of the HCT116 $p53^{+/+}$ cells as shown in Figure 4. We then examined the methylation status of H3R8 after standardizing the quantity of PRMT5 using the anti-PRMT5 antibody. Using an antibody for symmetric dimethylation of H3R8 (H3R8me2s), we conducted Western blot analysis after transfection with siRNAs

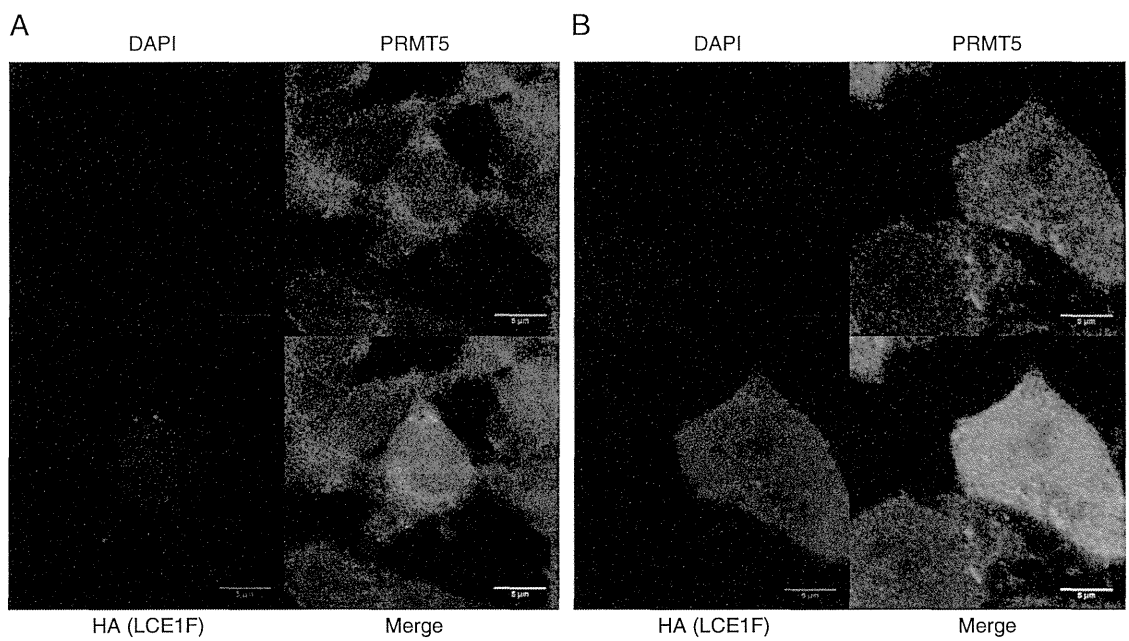

Figure 4. Co-localization of LCE1F and PRMT5. HCT116 $p53^{+/+}$ cells were transfected with HA-LCE1F. Cells were fixed by 1.7% formaldehyde and stained with an anti-HA antibody (red) and an anti-PRMT5 antibody (green).