# cmgh

ORIGINAL RESEARCH

NFKB1 and *MANBA* Confer Disease Susceptibility to Primary Biliary Cholangitis via Independent Putative Primary Functional Variants

Yuki Hitomi,[1] Ken Nakatani,[1] Kaname Kojima,[2,3] Nao Nishida,[1,4] Yosuke Kawai,[1,2,3] Minae Kawashima,[1,5] Yoshihiro Aiba,[6] Masao Nagasaki,[2,3,7] Minoru Nakamura,[6,8,9] and Katsushi Tokunaga[1]

[1]Department of Human Genetics, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan; [2]Department of Integrative Genomics, Tohoku Medical Megabank Organization, [3]Graduate School of Medicine, [7]Graduate School of Information Sciences, Tohoku University, Sendai, Japan; [4]The Research Center for Hepatitis and Immunology, National Center for Global Health and Medicine, Ichikawa, Japan; [5]Japan Science and Technology Agency, Tokyo, Japan; [9]Headquarters of Primary Biliary Cholangitis (PBC) Research in National Hospital Organization Study Group for Liver Disease in Japan (NHOSLJ), [6]Clinical Research Center, National Hospital Organization, Nagasaki Medical Center, Omura, Japan; [8]Department of Hepatology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
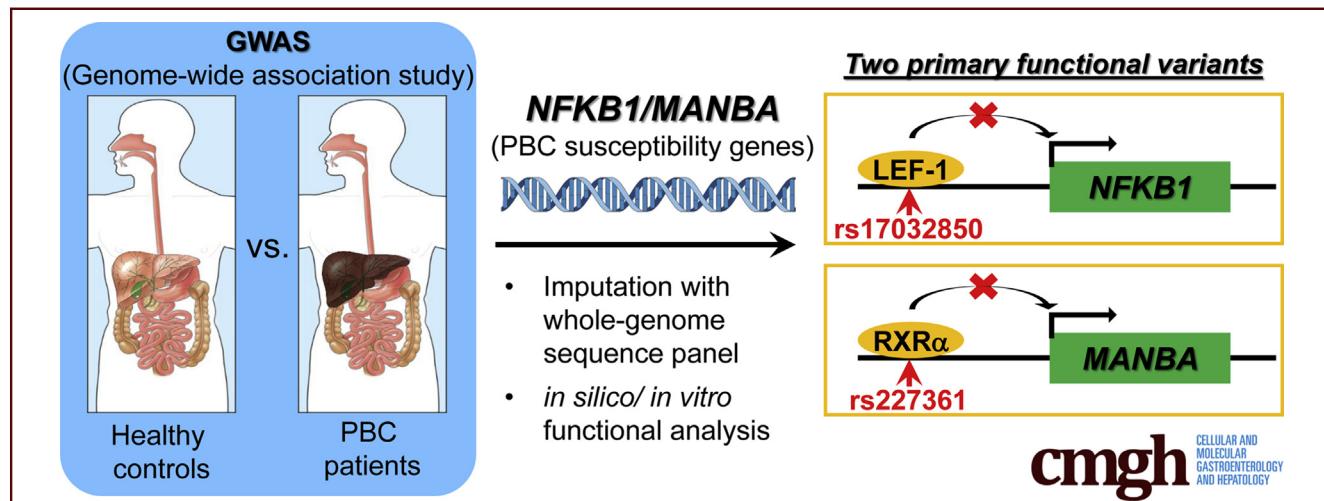
SUMMARY

We have identified novel putative primary functional variants for primary biliary cholangitis (PBC) in *NFKB1/MANBA*, and shown the molecular mechanism by which the variants confers susceptibility to PBC. This study indicates that *NFKB1*/lymphoid enhancer-binding factor 1 and *MANBA*/retinoid X receptor α have important roles in the development of PBC.

BACKGROUND & AIMS: Primary biliary cholangitis (PBC) is a chronic and cholestatic liver disease that eventually leads to cirrhosis and hepatic failure. We recently identified several susceptibility genes included *NFKB1* and *MANBA* for PBC in the Japanese population by genome-wide association study. However, the primary functional variants in the *NFKB1/MANBA* region and the molecular mechanism for conferring disease susceptibility to PBC have not yet been clarified.

METHODS: We performed high-density association mapping based on a single-nucleotide polymorphism (SNP) imputation analysis, using data from a whole-genome sequence reference panel of 1070 Japanese individuals and the previous genome-wide association study (1389 PBC patients, 1508 healthy controls). Among SNPs ($P < 5.0 \times 10^{-7}$) in the *NFKB1/MANBA* region, putative primary functional variants and the molecular mechanism for conferring disease susceptibility to PBC were identified by in silico/in vitro functional analysis.

RESULTS: Among the SNPs in the *NFKB1/MANBA* region, rs17032850 and rs227361, which changed the binding of transcription factors lymphoid enhancer-binding factor 1 (LEF-1) and retinoid X receptor α (RXRα), respectively, were identified as putative primary functional variants that regulate gene expression. In addition, expression-quantitative trait locus data and gene editing using a clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 system supported the potential role of rs17032850 and rs227361 in regulating *NFKB1* and *MANBA* expression, respectively.

CONCLUSIONS: We identified independent putative primary functional variants in *NFKB1/MANBA* and showed the distinct molecular mechanism by which each putative primary functional variant conferred susceptibility to PBC. Our approach was useful to dissect the pathogenesis not only of PBC, but also other digestive diseases in which *NFKB1/MANBA* has been reported as a susceptibility locus. *(Cell Mol Gastroenterol Hepatol 2019;7:515–532; https://doi.org/10.1016/j.jcmgh.2018.11.006)*

Keywords: PBC; GWAS; Disease Susceptibility Gene; e-QTL; CRISPR/Cas9.

Primary biliary cholangitis (PBC) is a chronic and progressive cholestatic liver disease that is characterized histologically by chronic non-suppurative destructive cholangitis, ductopenia, interface hepatitis, fibrosis, and, eventually, biliary cirrhosis. Because of the existence of well-defined autoantigens (such as the E2 component of the pyruvate dehydrogenase complex [PDC-E2]), homogeneous clinical expression, and an overwhelming predominance in females, PBC is considered an organ-specific autoimmune disease.[1,2] The destruction of small bile ducts is mediated not only by adaptive immune responses, including CD4+ T cells, CD8+ T cells, and B cells, but also innate immune responses, including natural killer cells against biliary epithelial cells.[2–4]

The higher concordance rate for PBC in monozygotic twins compared with dizygotic twins, and the higher estimated relative sibling risk (λs), indicates the involvement of strong genetic factors in the development of PBC.[5,6] Previous genome-wide association studies (GWASs) and subsequent meta-analyses identified *HLA* and 31 non-*HLA* susceptibility regions (*IL12A, IL12RB2, STAT4, IRF5, MMEL1, SPIB, DENND1B, CD80, IL7R, CXCR5, TNFRSF1A, CLEC16A, NFKB1, RAD51L1, MAP3K7IP1, PLCL2, RPS6KA4, TNFAIP2, ELMO1, IRF8, TNFSF11, SH2B3, CRHR1, TYK2, IL1RL2, CCL20, DGKQ, C5orf30, LOC285626, TNFAIP3*, and chromosome 17q12-21) as susceptibility genes for PBC in patients of European descent.[7–14] In the Japanese population, we previously identified *HLA, TNFSF15, POU2AF1, PRKCB, NFKB1, MANBA, IL7R*, and the chromosome 17q12-21 locus as susceptibility regions for PBC by GWAS.[15,16]

The *NFKB1/MANBA* gene region, which is located on human chromosome 4q24, has been identified as a susceptibility region in several digestive disorders, such as ulcerative colitis (rs3774937, $P = 5 \times 10^{-14}$; rs3774959, $P = 4 \times 10^{-12}$)[17,18] and Crohn's disease (rs13126515, $P = 4 \times 10^{-8}$),[19] as well as PBC. Nuclear factor-κB (NF-κB) is a well-known transcription factor complex that includes p50 (NF-κB1 and its precursor p105), p52 (NF-κB2 and its precursor p100), Rel (c-Rel), RelA (p65), and RelB, and has important roles in apoptosis, lymphoid organogenesis, innate and adaptive immunity, and central and peripheral tolerance.[20] In *Nfkb1*-/- mice, proliferation of T cells and production of T-helper type 2 (Th2) cytokines reportedly are reduced. In addition, these *Nfkb1*-/- mice showed susceptibility to experimental autoimmune encephalomyelitis, typhlocolitis, and infection with Leishmania major, but resistance to asthma.[21–24] The protein product of *MANBA* on the other hand,

β-mannosidase, is the final exoglycosidase in the N-linked glycoprotein oligosaccharide catabolism pathway, and is localized to the lysosomal compartment.[25] β-mannosidosis is a rare lysosomal storage disease caused by the deficiency of β-mannosidase. The most severe cases of human β-mannosidosis show not only mental retardation, developmental delay, and dysmorphology, but also frequent infection.[26–31] Taken together, the evidence indicates that both *NFKB1* and *MANBA* may play important, but different, roles in the regulation of the immune system. However, largely because of the existence of a shared linkage disequilibrium (LD) block in this region (Figure 1*A*), primary functional variants in the *NFKB1/MANBA* loci and the molecular mechanisms by which these variants confer susceptibility to PBC have yet to be clarified.

In the present study, to identify putative primary functional variants in the *NFKB1/MANBA* region, we performed high-density association mapping based on single-nucleotide polymorphism (SNP) imputation analysis using the data from a whole-genome sequence reference panel of 1070 Japanese individuals[32] and our previous GWAS.[15,16] We next performed in silico and in vitro functional analyses to identify putative primary functional variants that alter gene expression. Finally, we attempted to identify the molecular mechanisms underlying disease susceptibility to PBC associated with these functional variants by expression-quantitative trait locus analyses and gene editing using a clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 system.

Results
SNP Imputation Analysis and High-Density Association Mapping

Only 16 SNPs located in *NFKB1/MANBA* had been genotyped in our previous GWAS using the Affymetrix (Santa Clara, CA) Axiom Genome-Wide ASI1 Array.[15,16] To analyze all SNPs located in this region, SNP imputation analysis and high-density association mapping for disease susceptibility to PBC were performed using a reference panel of 1070 Japanese individuals (1KJPN) and genotype data from our previous GWAS.[16,32]

Besides the SNPs which were significant to a level of $P < 5.0 \times 10^{-7}$ in our previous GWAS (rs230534, rs2866413, rs223489, and rs2903283),[16] an additional 131 SNPs were found to be significant to this level by SNP

Figure 1. LD structure around *NFKB1/MANBA* in the Japanese population. (*A*) SNPs in this locus for the analysis of LD were filtered by Minor allele frequencies (MAF) > 0.05 and has publications in the NCBI variation viewer, and 95 SNPs were selected. LDs of these SNPs were analyzed by LDlink using the 1000genome Japanese in Tokyo (JPT) data. (*B* and *C*) r^2 of each SNP with (*B*) rs230534, which was the *NFKB1* top-hit SNP with susceptibility to PBC and (*C*) rs2866413, which was the *MANBA* top-hit SNP with susceptibility to PBC, using the present genotyping data. (*B*) rs2866413 and (*C*) rs230534 are shown as *red dots*. Stronger colors indicate higher r^2 scores. Chr, chromosome.

imputation analysis, probably because of the existence of strong LD in this region (Figure 1 and Supplementary Table 1). Of these SNPs, 16 were located in the 5′ intergenic region of *NFKB1*, 67 were located in *NFKB1* introns, 1 was located in the 5′ intergenic region of *MANBA*, 2 were located in the coding region of *MANBA*, and 49 were located in *MANBA* introns (Supplementary Table 1 and Figure 2*A* and *B*).

Although there was a strong LD block in the *NFKB1/MANBA* locus, LDs were not as strong between *NFKB1* SNPs and *MANBA* SNPs, which showed relatively stronger association with susceptibility to PBC (Figure 1*B* and *C*). Therefore, primary functional variants seemed to be located in both the *NFKB1* locus and the *MANBA* locus.

Identification of the Putative Primary Functional Variant in NFKB1

From the total of 83 *NFKB1* SNPs, we used the RegulomeDB database (http://www.regulomedb.org/index)[33] to identify those that were located in DNase hypersensitivity clusters and histone acetylation regions, and that were predicted to affect the binding of transcription factors. Twelve SNPs with RegulomeDB scores higher than 2c, which were supported by transcription factor binding data and DNase peaks, were selected as potential candidates

(Table 1) (rs747559 in the 5′ intergenic region; rs2272676 in intron 1; rs230532 in intron 2; rs230493, rs230504, and rs230519, and rs230520 in intron 5; rs227361 in intron 10; rs1598859, rs3774959, and rs17032850 in intron 11; and rs4648055 in intron 12). Among these 12 SNPs, 5 SNPs (Table 1) (rs2272676, rs230504, rs17032850, rs3774959, and rs4648055), which were located in DNase I hypersensitivity clusters and H3K27Ac markers in at least 1 cell type in the UCSC genome browser (http://genome.ucsc.edu/index.html), were selected as the final candidates.[34] In addition, the most significant SNP in the *NFKB1/MANBA* region identified from high-density association mapping (Table 1) (rs230534) also was selected as the final candidate in the *NFKB1* locus for the conferral of susceptibility to PBC (Figure 2*B* and *C*). The PICS algorithm (https://pubs.broadinstitute.org/pubs/finemapping/pics.php), which calculates the most likely causal SNPs given the observed association signal at a locus, indicated that all of the final candidates showed the PICS values of greater than 0.001 were indexed by rs230534 (ie, the top-hit SNP in the *NFKB1* locus) (Supplementary Table 2).[35]

To evaluate the effect of final-candidate SNPs on the binding affinity of transcription factors, electrophoretic mobility shift assays (EMSAs) were performed using biotin-labeled probes corresponding to the different alleles of each
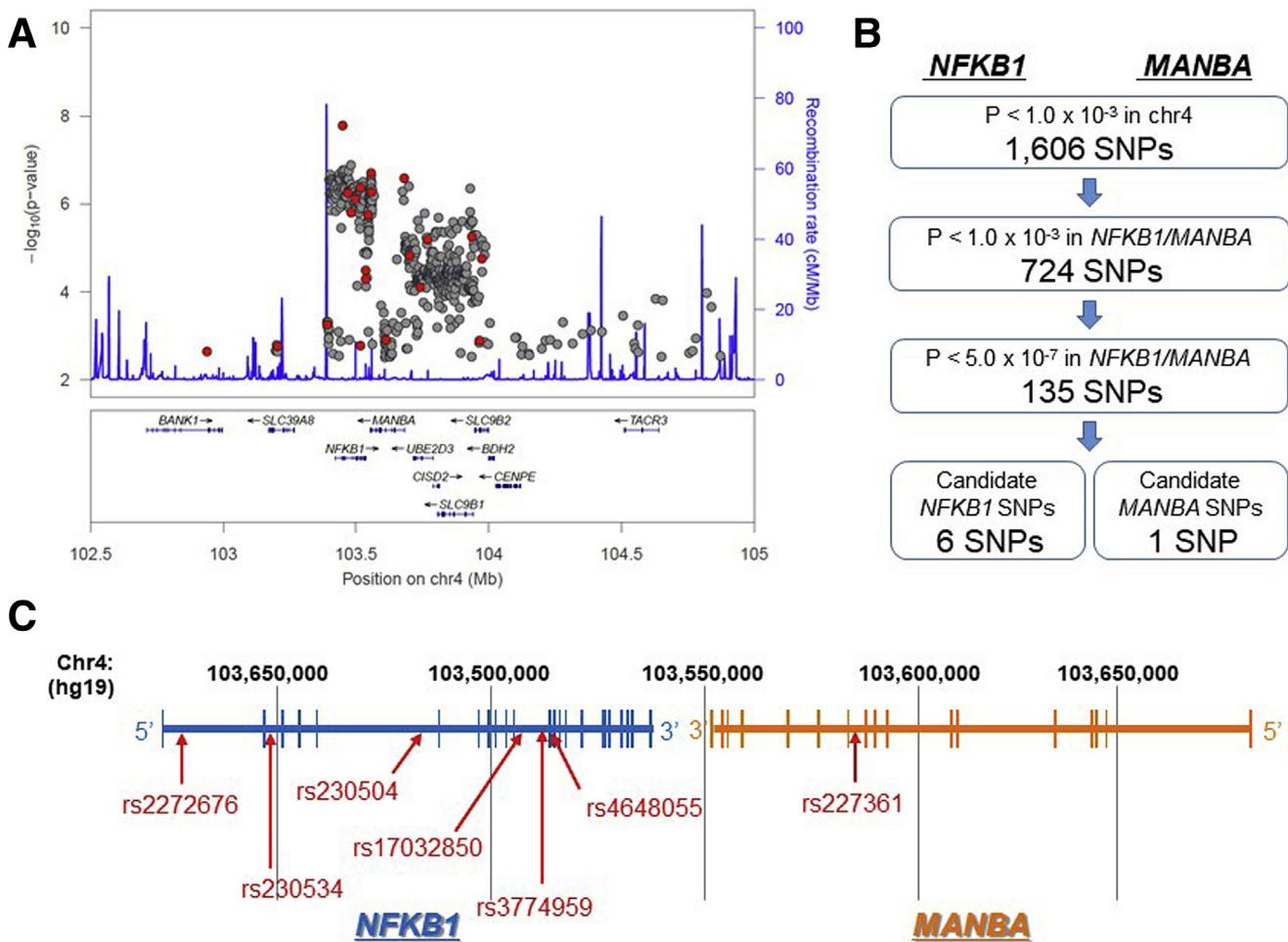
Figure 2. High-density association mapping for the selection of candidate putative primary functional variants in *NFKB1/MANBA* that conferred susceptibility to PBC. (*A*) Results of high-density association mapping in *NFKB1/MANBA* based on SNP imputation analysis. A whole-genome sequence reference panel of 1KJPN and genotype data from our previous GWAS (1389 PBC patients, 1508 healthy controls) were used. *Dots* show the *P* value of each SNP, and *red dots* show experimentally genotyped SNPs from GWAS. (*B*) Schematic of the selection process for candidate putative primary functional variants. There were 724 and 135 SNPs whose *P* values were less than 1.0×10^{-3} and 5.0×10^{-7}, respectively, that were located in *NFKB1/MANBA*. Of the 135 SNPs, rs230534, which showed the most significant association with susceptibility to PBC in the *NFKB1/MANBA* region, and the candidate putative primary functional variants located in a transcription regulatory element (*NFKB1*: rs2272676, rs230504, rs17032850, rs3774959, and rs7678055; *MANBA*: rs227361) were chosen. (*C*) The positions of the 7 candidate functional variants in *NFKB1/MANBA* are shown. Chr, chromosome.

SNP and nuclear extracts of either the HepG2 (human liver carcinoma) or Jurkat (human T lymphocyte) cell lines. Among the final candidates, a difference in mobility shift between the major allele (G allele) and the minor allele (C allele, PBC susceptible) of rs17032850 was detected (Figure 3*A* and *B*). In HepG2 cells, the thickness of the shifted band was different between major alleles and minor alleles (Figure 3*A*), whereas in Jurkat cells, the position of the shifted bands was different (Figure 3*B*). These band shifts were abrogated by incubation with a 200× concentration of a nonlabeled competitor probe (Figure 3*C*). However, no difference in mobility shift was detected for the other candidate SNPs (rs230534, rs2272676, rs230504, rs3774959, and rs7678055) (Figure 3*A* and *B*). These results indicated that the putative primary functional variant in *NFKB1* was rs17032850.

Molecular Features Associated With the PBC Susceptibility Allele of rs17032850

To determine differences in transcription efficiency between the major and PBC susceptibility alleles of rs17032850, luciferase reporter assays were performed in HepG2 and Jurkat cells. In HepG2, the luciferase activity of cells 24 hours after transfection with a reporter construct containing the C allele (ie, the PBC susceptibility allele) of rs17032850 was reduced significantly compared with those transfected with a plasmid containing the G allele (Figure 3*D* and *E*). However, in Jurkat cells, a difference in luciferase activity between the rs17032850 C allele and G allele was not observed (Figure 3*F*). This likely is owing to differences in transcription factor binding to these alleles between the different cell types.

Table 1. SNPs Associated With Susceptibility to PBC in the Japanese Population in NFKB1/MANBA by High-Density Association Mapping

SNP_ID[a]	GWAS or imputation[b]	Position (chromosome 4)[c]	P[d]	OR[e]	RegulomeDB[f]	UCSC[g]	Location
rs230534	GWAS	103449041	1.34E-08	1.37	4	△	*NFKB1* intron 2
rs35185023	Imputation	103448080	2.61E-08	1.37	6	×	*NFKB1* intron 2
rs2978641	Imputation	103575280	1.17E-07	1.33	6	×	*MANBA* intron 12
rs227277	Imputation	103594938	1.23E-07	1.34	6	×	*MANBA* intron 8
rs230531	Imputation	103450377	1.24E-07	1.34	6	×	*NFKB1* intron 2
rs11097786	Imputation	103406511	1.27E-07	1.34	6	×	5' of *NFKB1*
rs141936164	Imputation	103401723	1.37E-07	1.34	6	×	5' of *NFKB1*
rs228613	Imputation	103577465	1.42E-07	1.33	No data	×	*MANBA* intron 12
rs228612	Imputation	103575385	1.56E-07	1.33	No data	×	*MANBA* intron 12
rs62328536	Imputation	103399049	1.59E-07	1.34	4	×	5' of *NFKB1*
rs2272676	Imputation	103423326	1.87E-07	1.34	2b	○	*NFKB1* intron 1
rs2866413	GWAS	103557077	2.40E-07	1.32	No data	Nonsynonymous	*MANBA* exon 14 (T701M)
rs227368	Imputation	103611845	2.80E-07	1.33	5	Nonsynonymous	*MANBA* exon 6 (V253I)
rs230504	Imputation	103481561	3.58E-07	1.33	1d	△	*NFKB1* intron 5
rs3774959	Imputation	103511114	3.74E-07	1.32	1b	△	*NFKB1* intron 11
rs17032850	Imputation	103507703	3.74E-07	1.32	2b	△	*NFKB1* intron 11
rs4648055	Imputation	103515313	4.27E-07	1.32	1f	△	*NFKB1* intron 12
rs227361	Imputation	103586977	4.32E-07	1.32	1b	○	*MANBA* intron 10

[a]SNPs with underlines were the final candidate primary functional variants.
[b]Genotyped by our previous GWAS[16] or the imputed genotypes by the high-density association mapping in the present study.
[c]Position of the SNPs in hg19.
[d]P values calculated by the Pearson chi-square test for the allelic model.
[e]Odds ratio (OR) of minor allele from the 2-by-2 allele frequency table.
[f]Functinoal prediction scores of each SNP by the RegulomeDB database.
[g]Probability of the functional damages checked by the UCSC genome browser.

By using the TRANSFAC database (Qiagen, Valencia, CA; http://www.gene-regulation.com/pub/databases.html), we predicted that the G allele of rs17032850, but not the C allele, constituted a lymphoid enhancer-binding factor 1 (LEF-1) binding motif (Figure 4*A*–*C*).[36] We first confirmed that LEF-1 is expressed in both HepG2 and Jurkat cells (Figure 4*D*). Subsequently, the binding of LEF-1 was investigated in a supershift assay using nuclear extracts of HepG2. Consistent with the prediction of transcription factor binding, the mobility shift associated with the G allele of rs17032850 was supershifted by pre-incubation with an anti–LEF-1 antibody before electrophoresis (Figure 4*E*).

To assess the influence of rs17032850 on endogenous gene expression, the expression levels of 11 genes located within 1 Mb of *NFKB1/MANBA* were compared (using the GTEx portal database; http://gtexportal.org/home/) for the different genotypes of rs17032850 in every organ in which the genes' expression levels were above the threshold for detection.[37] Among these 11 genes, *NFKB1* (muscle_skeletal: $P = .000062$) and *UBE2D3* (esophagus_mucosa: $P = .000013$) showed effects on expression that were associated significantly with the rs17032850 genotypes (Figure 5). These results indicated that *NFKB1* and *UBE2D3* were candidate effector genes of rs17032850, a PBC susceptibility allele that alters a putative LEF-1 binding site.

By expression-quantitative trait locus analysis, not only primary functional variants, but also other variations that show strong LD with the primary functional variants, are associated with gene expression levels. Therefore, the effect of rs17032850 on gene expression levels was assessed using rs17032850 genotype knock-in versions of the Jurkat cell lines constructed using a CRISPR/Cas9 system. Among the candidate effector genes, only *NFKB1* showed an expression level that differed significantly between rs17032850-C/C clones (n = 5) and rs17032850-G/G clones (n = 5) ($P = .016$; Mann–Whitney *U* test) (Figure 6*A* and *B*). In addition, the difference in *NFKB1* expression level between rs17032850-C/C clones and rs17032850-G/G clones was eliminated by knock-down of LEF-1 (Figure 6*C*). These results confirmed that rs17032850 is a primary functional variant of the *NFKB1* locus, such that the rs17032850 SNP alters LEF-1 binding to, and regulation of, the *NFKB1* gene.

Identification of the Putative Primary Functional Variant in MANBA

From the 52 total *MANBA* SNPs that showed P values less than 5.0×10^{-7} by SNP imputation analysis, candidate functional variants that were predicted to be located in transcription regulatory elements were selected. One SNP
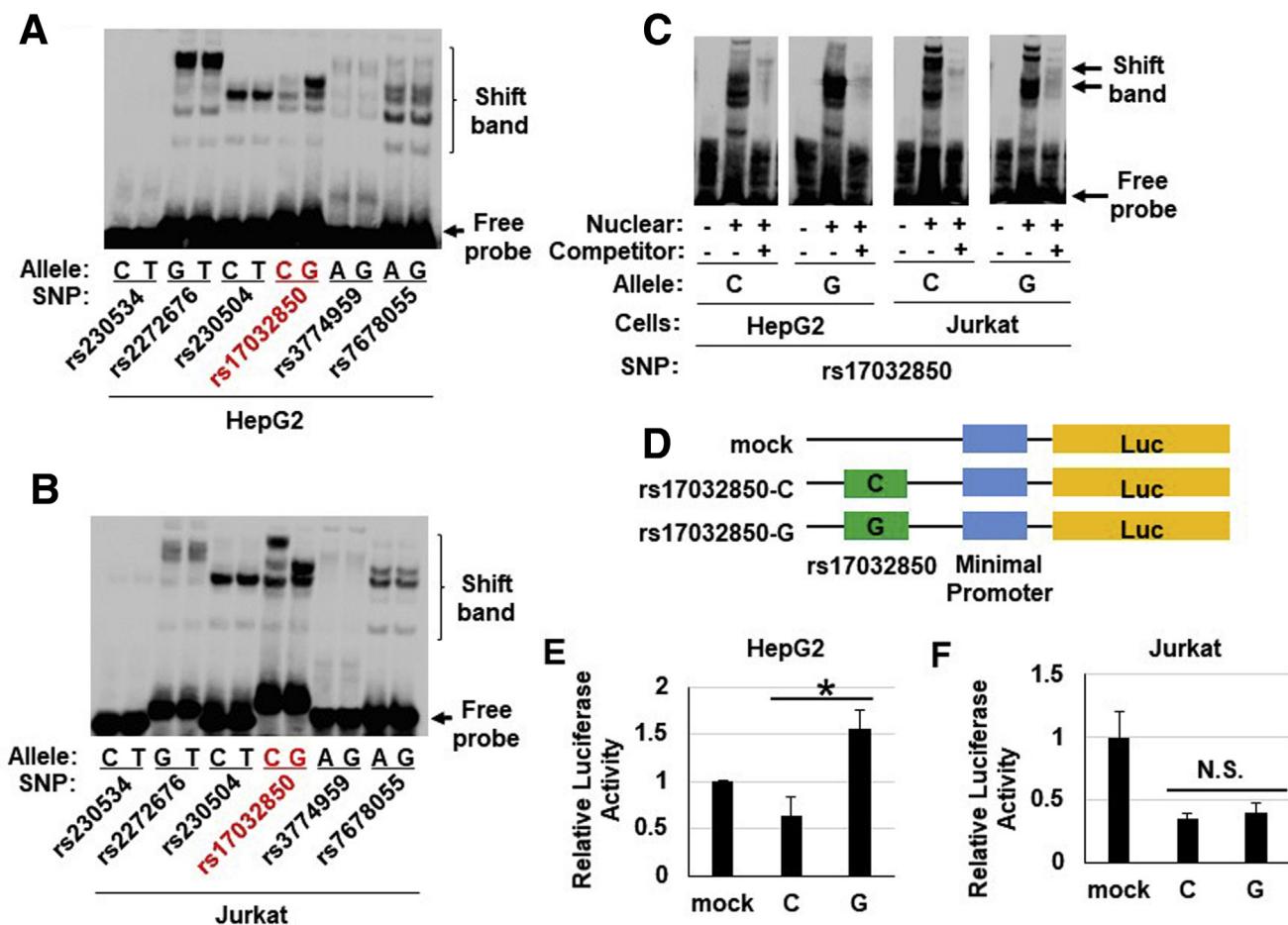
Figure 3. In vitro functional analysis of each candidate variant of *NFKB1*. (*A* and *B*) EMSA of each of the 6 candidate primary variants using biotin-labeled probes corresponding to the major and the minor alleles, and nuclear extracts of (*A*) HepG2 and (*B*) Jurkat. rs17032850 was the only variant to show a difference in mobility shift between the 2 alleles. (*C*) Competitor assay, using HepG2 and Jurkat nuclear extracts and a 200× concentration of unlabeled probe corresponding to either the C (ie, PBC susceptibility) or G alleles of rs17032850. (*D*) Outline of reporter plasmid constructs. PCR fragments of intron 11 of *NFKB1*, containing rs17032850, were subcloned into the pGL4.23 vector. (*E* and *F*) Transcription was measured by cellular luciferase activity, 24 hours after transfection of (*E*) HepG2 cells and (*F*) Jurkat cells. The luciferase activity of cells transfected with the PBC susceptibility allele (C allele) of rs17032850 was decreased compared with those transfected with the G allele. Three independent experiments with triplicate measurements were performed for each assay, and data represent means ± SD. *P < .05 (Student *t* test).

with a RegulomeDB score higher than 2c was selected as a potential candidate (Table 1) (rs227361 in intron 10). We confirmed that this SNP is located in a DNase I hypersensitivity cluster and a H3K27Ac marker region for at least 1 cell type, using the UCSC genome browser, and this SNP consequently was selected as the final *MANBA* locus candidate for the conferral of PBC susceptibility (Figure 2*B* and *C*). The PICS algorithm indicated that rs227361 showed the highest PICS values indexed by rs2978641 (ie, the top-hit SNP in the *MANBA* locus) (Supplementary Table 3).[35]

By EMSA, a difference in mobility shift between the major allele (A allele) and the minor allele (G allele, PBC susceptible) of rs227361 was detected using nuclear extracts from HepG2 and Jurkat (Figure 7*A* and *B*). The bands shifted by the major and minor alleles were specifically diminished by incubation with a 200× concentration of a nonlabeled competitor probe (Figure 7*A* and *B*).

Among the *MANBA* SNPs which showed *P* values less than 5.0×10^{-7} by SNP imputation analysis, rs227368 (V253I) and rs2866413 (T701M), located in the 6th and 14th exons, respectively, also had the potential to be selected as candidate functional variants (Table 1). However, after SWISS model (https://swissmodel.expasy.org/) and Polyphen2 analysis (http://genetics.bwh.harvard.edu/pph2/), neither nonsynonymous substitution was predicted to have a deleterious effect on MANBA protein folding and function (Figure 8).[38,39] These results indicated that the only putative primary functional variant in *MANBA* was rs227361.

Molecular Features Associated With the PBC Susceptibility Allele of rs227361

To assess differences in transcription efficiency between the major and the PBC susceptibility alleles of rs227361,
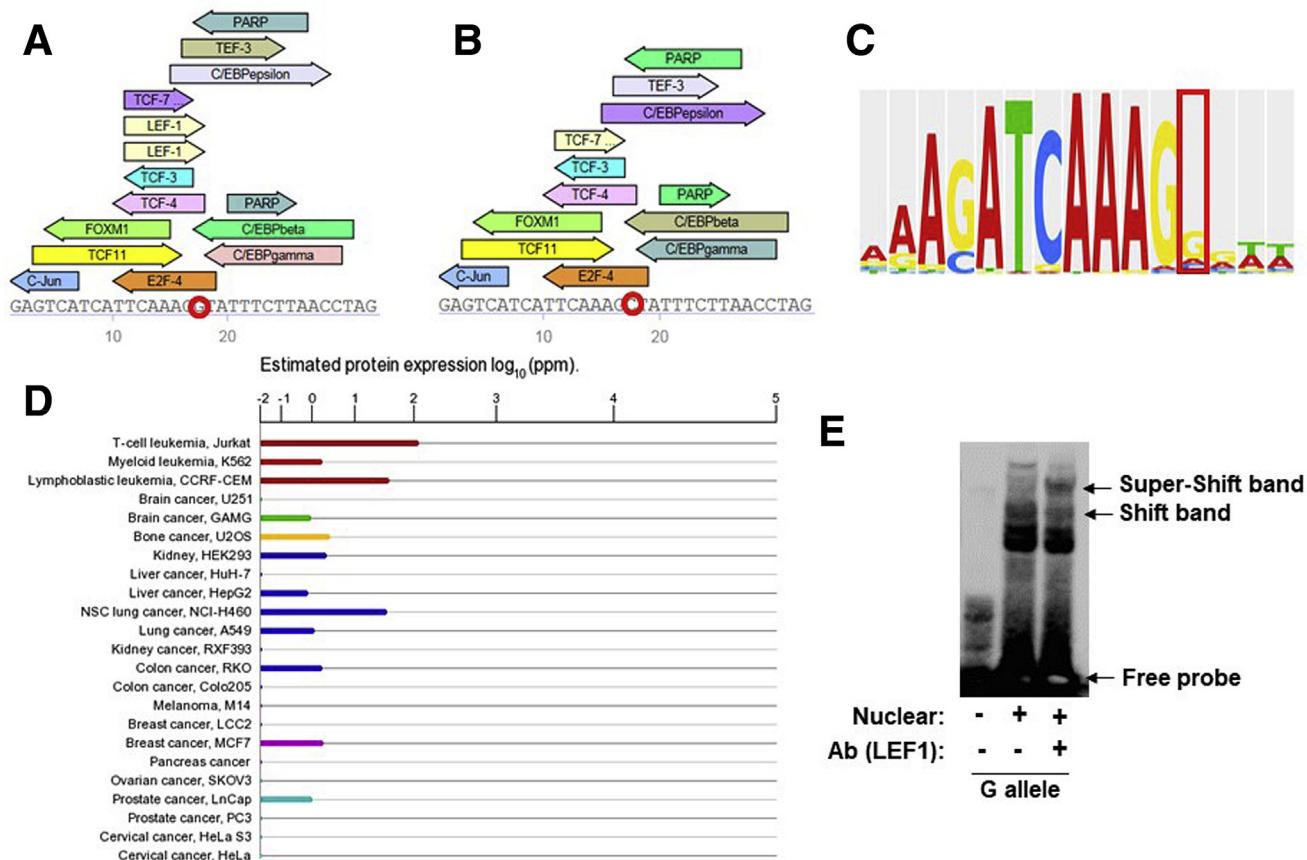
Figure 4. Identification of transcription factors targeting the G allele of rs17032850. (*A* and *B*) Prediction of transcription factor binding using the TRANSFAC database. (*A* and *B*) The G allele of rs17032850 produces a (*A*) LEF-1 binding site, however, this sequence is disrupted in the (*B*) C allele. (*C*) The binding motif of LEF-1, with the position of rs17032850 shown as a *rectangle*. (*D*) LEF-1 expression levels in cancer cell lines. Both Jurkat and HepG2 cells showed abundant expression of *LEF-1*. Data were obtained using GeneCards from the Weizmann Institute of Science (http://www.genecards.org/). (*E*) Supershift was observed after incubation of HepG2 cell nuclear extracts with an anti–LEF-1 antibody. Three independent experiments were performed in each assay. Ab, antibody.

luciferase reporter assays were performed in HepG2 and Jurkat cells. The luciferase activity of cells 24 hours after transfection of a reporter construct containing the G allele (ie, PBC susceptibility allele) of rs227361 was significantly lower compared with cells transfected with a plasmid containing the A allele in both cell lines (Figure 7*C*–*E*).

The G allele of rs227361, but not the A allele, was predicted using the TRANSFAC database to constitute a binding motif for the retinoid X receptor alpha (RXRα) (Figure 9*A*–*C*). RXRα forms a multiple-subunit corepressor complex that stabilizes repressive local chromatin structure, and blocks access of the transcriptional machinery to the promoter.[40] Therefore, RXRα is known as a nuclear receptor that binds to response elements in the promoter region of target genes, and inhibits transcription without an agonist ligand.[36,40] RXRα was found to be expressed in both the HepG2 and Jurkat cell lines (Figure 9*D*), and, hence, a supershift assay using the nuclear extract of HepG2 was performed to investigate the binding of RXRα with the G allele of rs227361. Concordant with the prediction, a supershift was found for the G allele of rs227361 after pre-incubation with an anti-RXRα antibody (Figure 9*E*).

The effect of rs227361 genotype on endogenous gene expression levels also was compared using the GTEx portal database as well as rs17032850.[37] Among these 11 genes, *NFKB1* (muscle skeletal: $P = 4.5 \times 10^{-6}$), *MANBA* (breast mammary tissue: $P = .000035$; small intestine terminal ileum: $P = .000079$; spleen: $P = 2.5 \times 10^{-6}$), *UBE2D3* (esophagus mucosa: $P = .000016$; lung: $P = .000043$; pancreas: $P = 8.9 \times 10^{-6}$; stomach: $P = 7.8 \times 10^{-6}$), *CISD2* (artery tibial: $P = 1.2 \times 10^{-8}$; nerve tibial: $P = .000045$; skin not sun exposed: $P = 5.3 \times 10^{-7}$; skin sun exposed: $P = 8.0 \times 10^{-11}$; thyroid: $P = 3.8 \times 10^{-12}$; uterus: $P = 6.9 \times 10^{-6}$), and *BDH2* (esophagus muscularis: $P = .000034$; muscle skeletal: $P = .000060$) showed changes in expression that were associated significantly with the genotypes of rs227361 (Figure 10). These results indicated that *NFKB1*, *MANBA*, *UBE2D3*, *CISD2*, and *BDH2* are candidate effector genes of rs227361, a PBC-susceptibility allele that has higher binding affinity for the inhibitory transcription factor RXRα.

Next, the effect of rs227361 on gene expression levels was assessed using the rs227361 genotype knock-in versions of Jurkat cell lines constructed using a CRISPR/Cas9

A
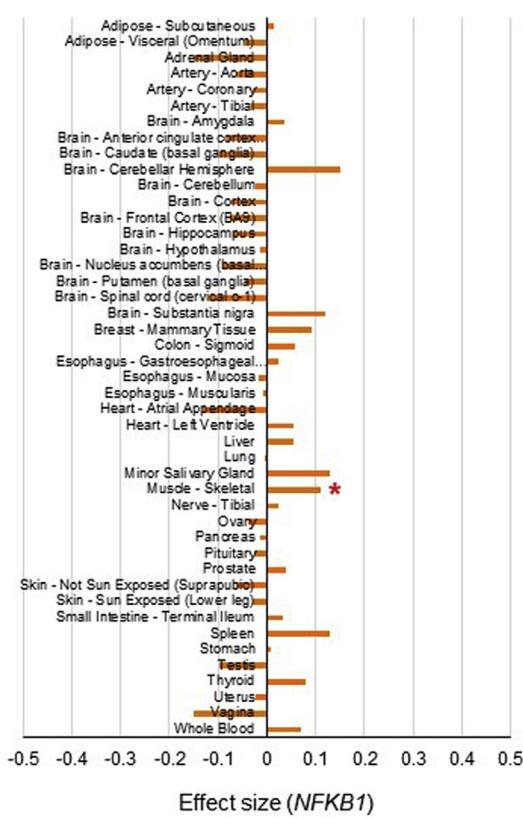
B
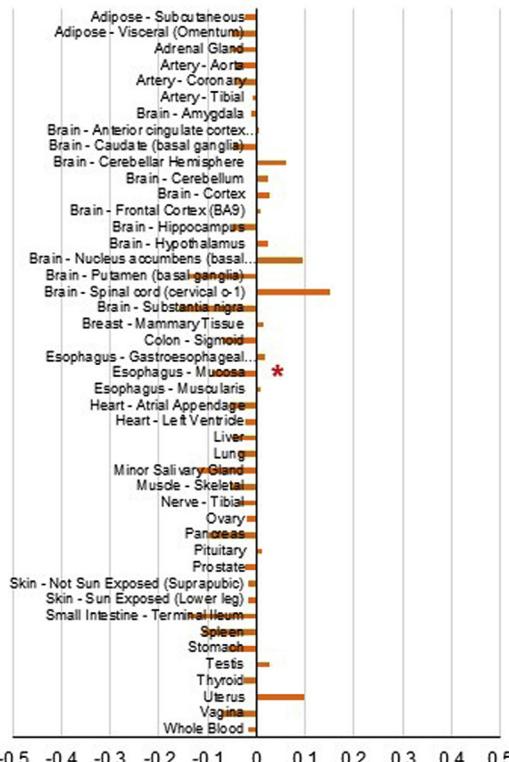
*P<1.0E-04

Figure 5. rs17032850 genotypes are associated with differences in endogenous *NFKB1* and *UBE2D3* expression levels. Comparison between rs17032850 genotype and the endogenous expressions of (*A*) *NFKB1* and (*B*) *UBE2D3* in all tissues registered in the GTEX database. Statistical significance levels after Bonferroni multiple comparison correction: $P = .000108$ (0.05/[11 genes × 45 organs]).

system. Among the candidate effector genes, only *MANBA* showed an expression level that differed significantly between rs227361-A/A clones (n = 12) and rs227361-G/G clones (n = 10) ($P = .00037$; Mann–Whitney *U* test)

(Figure 11*A–E*). In addition, the difference in *MANBA* expression level between rs227361-A/A clones and rs227361-G/G clones was eliminated by the knock-down of RXRα (Figure 11*F*). These results confirmed that rs227361

A

B

C
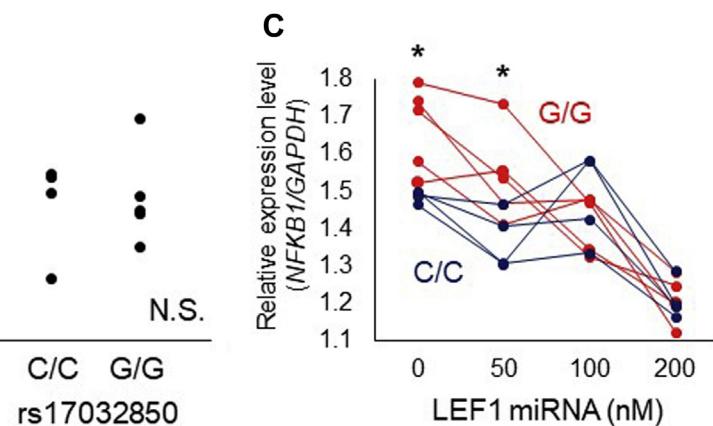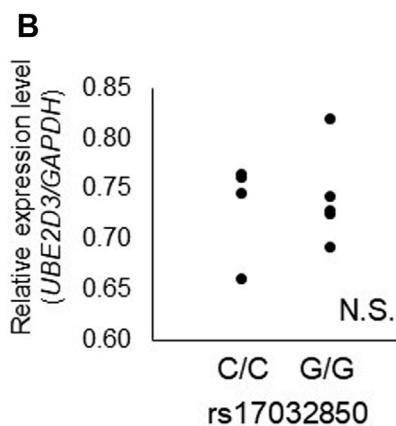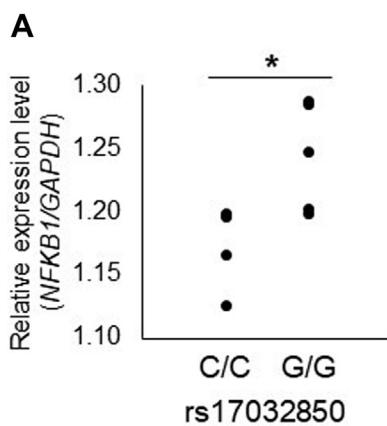
Figure 6. The effects of rs17032850 on gene expression in the knock-in clones of the Jurkat cell line generated using a CRISPR/Cas9 system. (*A and B*) Expression levels of (*A*) *NFKB1* and (*B*) *UBE2D3* when comparing between rs17032850-C/C knock-in clones and rs17032850-G/G knock-in clones. (*C*) *NFKB1* expression levels in each knock-in clone after the knock-down of LEF-1. *$P < .05$ (Mann–Whitney *U* test). miRNA, microRNA.
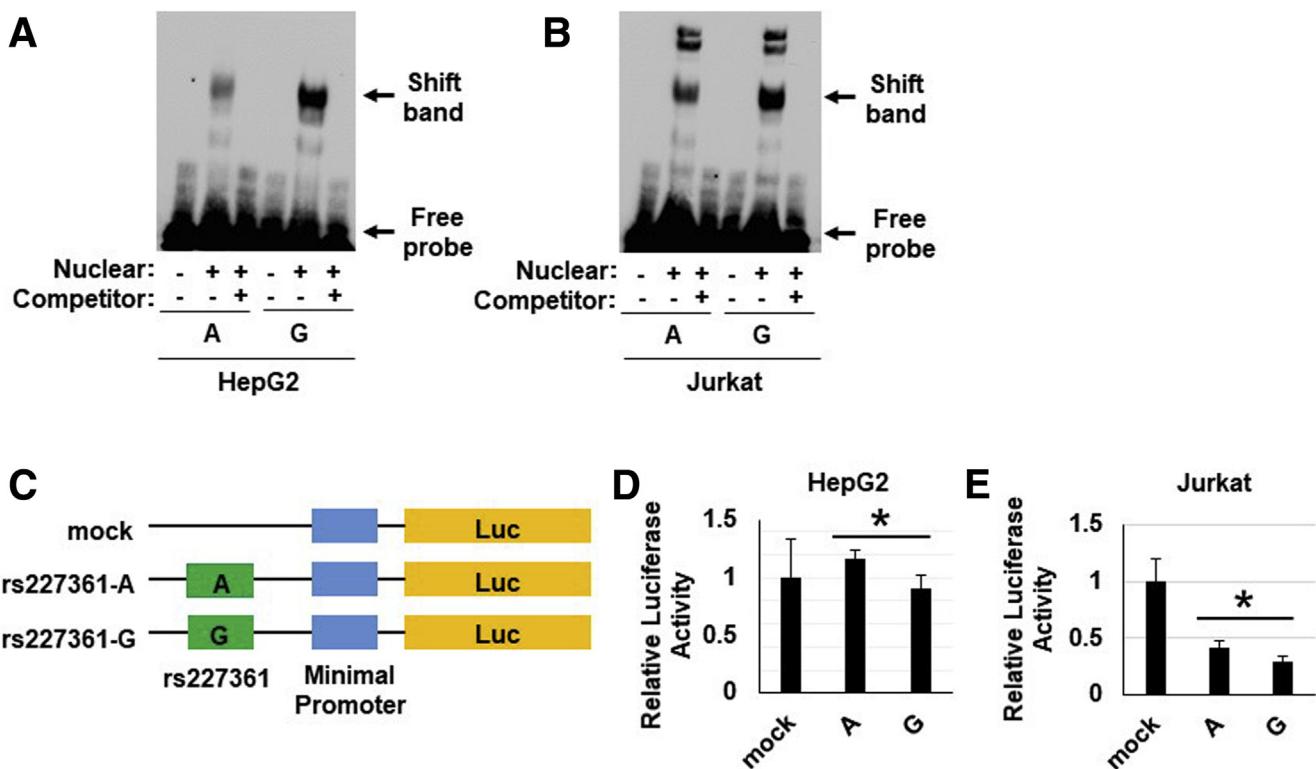
Figure 7. In vitro functional analysis of *MANBA* rs227361. (*A* and *B*) EMSA and competitor assays of rs227361 using biotin-labeled probes corresponding to the A and G alleles, and nuclear extracts of (*A*) HepG2 and (*B*) Jurkat. There was a difference in mobility shift between the 2 alleles. (*C*) Outline of reporter plasmid constructs. PCR fragments of intron 10 of *MANBA*, containing rs227361, were subcloned into pGL4.23. (*D* and *E*) Transcription was measured by cellular luciferase activity 24 hours after transfection of (*D*) HepG2 cells and (*E*) Jurkat cells. The luciferase activity of cells transfected with the PBC susceptibility allele (G allele) of rs227361 was decreased compared with those transfected with the A allele. Three independent experiments with triplicate measurements were performed for each assay, and data represent means ± SD. *$P < .05$ (Student *t* test).

is a primary functional variant in the *MANBA* locus, such that the rs227361 SNP alters RXRα binding to, and regulation of, the *MANBA* gene.

Co-segregation of the Effects of rs17032850 and rs227361 to the Disease Susceptibility of Top-Hit SNPs in This Locus

Finally, to assess the effects of these 2 putative primary functional variants on the disease susceptibility of top-hit SNPs, multiple regression analysis was performed. When rs17032850 and rs227361 were incorporated into a regression model (Homozygote of risk alleles = 1, other genotypes = 0), these SNPs entirely explained the association signal of rs230534 (TT = 2, CT = 1, CC = 0; $P = 7.55 \times 10^{-193}$). In addition, the risk genotypes of rs17032850 and rs227361 co-segregated with the rs230534-T allele (Figure 12). These results indicated that most of the effects of rs230534 were covered by rs17032850 and rs227361.

Discussion

In the present study, rs17032850, located in intron 11 of *NFKB1*, and rs227361, located in intron 10 of *MANBA*, were identified as the putative primary functional variants for disease susceptibility to PBC. The disease susceptibility

allele of rs17032850 was shown to disrupt a LEF-1 binding site and was associated with the endogenous expression levels of *NFKB1*. The disease-protective allele of rs227361 on the other hand, disrupted a RXRα binding site and was associated with significantly higher *MANBA* endogenous expression than in individuals without this allele.

There were some limitations to this study. First, although the transcription binding sites are believed to be critical components underlying disease susceptibility, important regulatory variants in autoimmune disease are only within transcription binding sites at 10% to 20%.[35] Second, the RegulomeDB is likely out of date, given that this database was last updated in 2014. Third, the PICS algorithm, which is an LD-based method for predicting SNPs that are most likely to be the causal variants responsible for associations with gene expression levels, indicated that rs227361 showed the highest PICS score in the *MANBA* locus, but did not indicate this for rs17032850 in the *NFKB1* locus. Fourth, in the logistic regression analysis used in the present study, most of the effects of rs17032850 and rs227361 were covered by rs230534, but this coverage was not complete. Therefore, some other primary functional variants presumably exist. However, results in this study indicate that the identified PBC susceptibility alleles of *NFKB1* and *MANBA* could be the functional variants that alter the

A
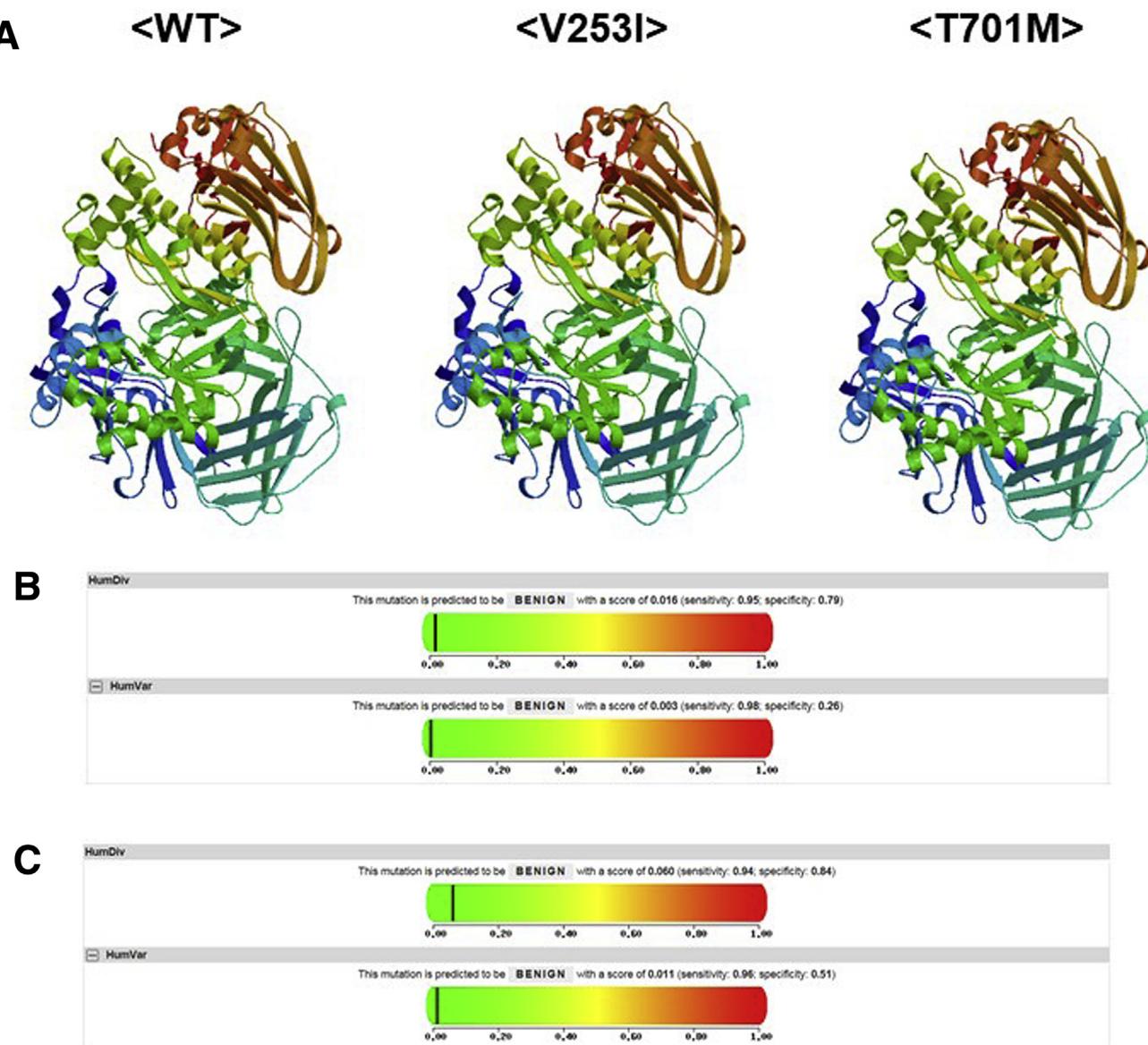
Figure 8. Prediction of the effect of nonsynonymous substitutions on MANBA protein functionality. (*A–C*) By both (*A*) SWISS model and (*B* and *C*) Polyphen2, the minor alleles of the indicated nonsynonymous substitutions were predicted to have no damaging effect on MANBA protein folding or function.

transcription of target genes by the reduction of the effect by these limitations using in vitro functional evaluation analysis.

In the present study, individuals with the PBC-susceptible allele (G allele) of rs227361, in the intron of *MANBA*, showed significantly lower endogenous levels of *MANBA* expression. To date, the contribution of β-mannosidase to the pathogenesis of digestive disorders has not been reported, and the relative level of *MANBA* expression in PBC patients and healthy controls has not been assessed. However, the phenotype of human β-mannosidosis, which is caused by the deficiency of β-mannosidase in lysosomes, may provide us with clues to the potential role of β-mannosidase in the onset of digestive disorders. Lysosomes in β-mannosidosis patients reportedly

suffer from lysosomal membrane permeabilization,[41] and this type of lysosomal damage can induce defects in autophagy.[42] Dysfunction of autophagy is well known to play a role in the pathogenesis of digestive disorders.[43] Additional studies are warranted to improve our understanding of the relationship between PBC pathogenesis and β-mannosidase.

To date, the following pathogenic pathway related to adaptive immune responses has been proposed in PBC: (1) release of intact PDC-E2, a well-defined, specific PBC autoantigen, from biliary epithelial cells during apoptosis; (2) endocytosis of PDC-E2 by antigen-presenting cells; (3) deficits in immune tolerance as the result of genetic and environmental factors; (4) presentation of degraded PDC-E2 peptide to autoreactive CD4[+] T cells via HLA class II, and autoreactive CD8[+] T cells via HLA class I; (5) interaction
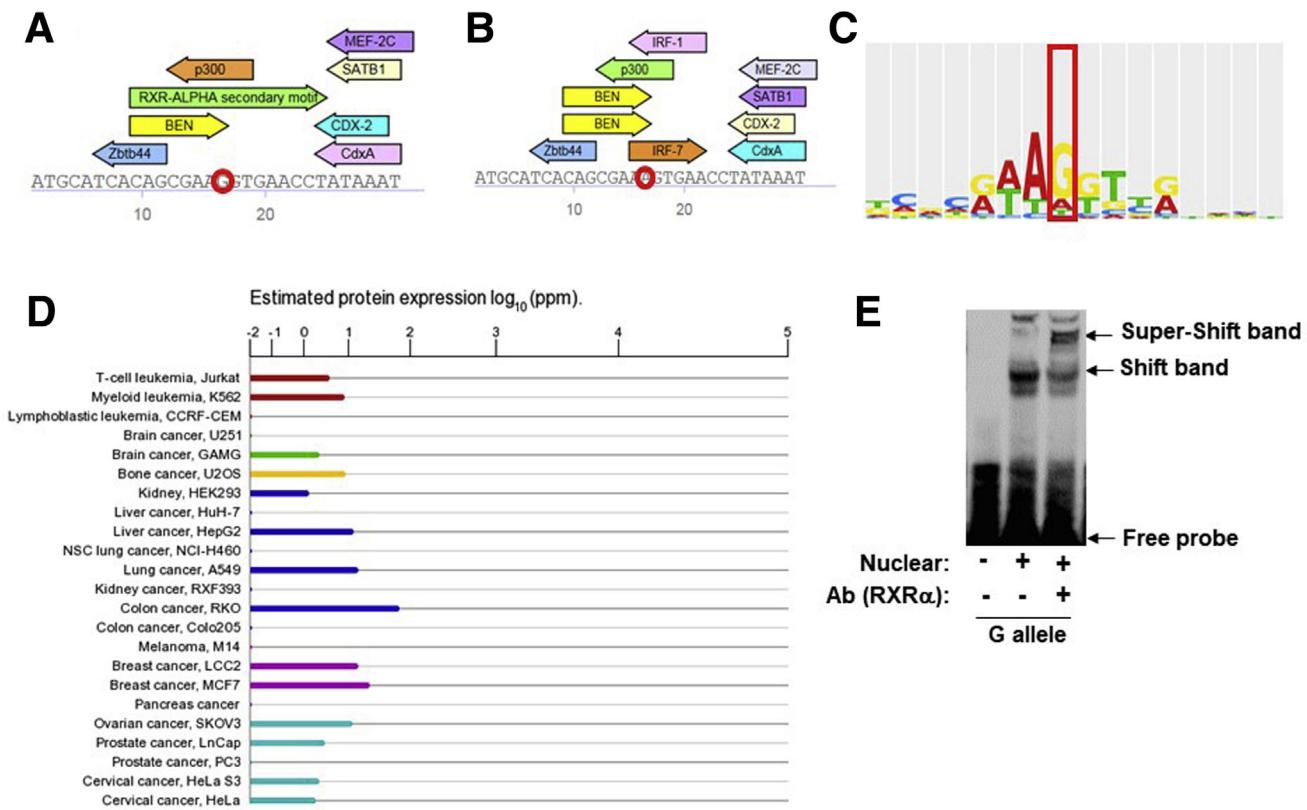
Figure 9. Identification of transcription factors targeting the G allele of rs227361. (*A* and *B*) Prediction of transcription factor binding using the TRANSFAC database. The G allele of rs227361 produces a (*A*) RXRα binding site, however, this sequence is disrupted in the (*B*) A allele. (*C*) The binding motif of RXRα, with the position of rs227361 shown as a *rectangle*. (*D*) RXRα expression in cancer cell lines. Both Jurkat and HepG2 cells showed abundant expression of *RXRA*. Data were obtained using GeneCards (http://www.genecards.org/). (*E*) Supershift was observed by incubation of HepG2 cell nuclear extracts with an anti-RXRα antibody. Three independent experiments were performed in each assay. Ab, antibody.

between autoreactive CD4[+] T cells and autoreactive B cells that produce antimitochondrial autoantibody, and (6) immune-mediated destruction of biliary epithelial cells in the small bile duct.[3] Innate immune responses mediated by natural killer cells also contribute to the pathogenesis of PBC.[4] Indeed, most of the susceptibility genes to PBC identified to date have functions related to this disease-development pathway. For example, the protein products of *HLA-DP*, *HLA-DQ*, and *HLA-DR*, reported as the strongest susceptibility genes for PBC,[44,45] are known as HLA class II molecules expressed in antigen-presenting cells. The protein product of *TNFSF15* has been reported to be involved in polarization to Th1 and Th17 cells.[46,47] The product of *IL12RB2* is a subunit of interleukin 12, which activates both innate and acquired immune responses.[48] Although the exact disease pathway in which β-mannosidase operates remains to be elucidated, the present study provides important clues in regard to their role in the pathogenesis of PBC. Crohn's disease and PBC in Asian populations have been reported to share a genetic background related to tumor necrosis factor signaling,[49,50] and *NFKB1/MANBA* has been reported as the susceptibility region for both of these diseases.[16,19] Both human genetic studies and mouse studies[51–54] have reported that defects in autophagy are

involved in the pathogenesis of Crohn's disease, and, therefore, rs227361, which was identified as the putative primary functional variant in *MANBA* in the present study, also may contribute to the dysfunction of autophagy via lysosomal damage in Crohn's disease. Although the top-hit SNP in the *NFKB1/MANBA* region in Crohn's disease (rs13126505)[19] is not found in Asian populations, the top-hit SNP in the same region in ulcerative colitis (rs3774959)[18] shows strong LD ($r^2 = 0.99$, D' = 0.99) with rs17032850, one of the putative primary functional variants in the present study. Thus, rs17032850 may operate as a functional variant in both PBC and ulcerative colitis.

GWASs have identified thousands of genetic variants associated with complex human diseases, including digestive disorders.[55] However, these genetic variants often explain only a relatively small proportion of the heritability of the disease.[56] To overcome the missing heritability problem, SNP imputation analysis using a large-scale, whole-genome sequence reference panel has been used as an efficient approach for the supplementation of GWAS data.[57] Particularly for Japanese studies, ethnicity-matched reference data have been required for appropriate SNP imputation analysis because of the higher proportion of variants that are specific to the Japanese population

*P<1.0E-04   **P<1.0E-05   ***P<1.0E-06

Figure 10. rs227361 genotypes are associated with differences in endogenous *NFKB1*, *MANBA*, *UBE2D3*, *CISD2*, and *BDH2* expressions. Comparisons between the genotype of rs227361 and endogenous expression levels of (*A*) *NFKB1*, (*B*) *MANBA*, (*C*) *UBE2D3*, (*D*) CISD2, and (*E*) *BDH2* in all tissues registered in the GTEX database. Statistical significance levels after Bonferroni multiple comparison correction: $P = .000108$ (0.05/[11 genes × 45 organs]).

Figure 11. The effects of rs227361 on gene expression in the knock-in clones of the Jurkat cell line generated using a CRISPR/Cas9 system. (A–E) Expression levels of (A) *MANBA*, (B) NFKB1, (C) *UBE2D3*, (D) *CISD2*, and (E) *BDH2* comparing between rs227361-A/A knock-in clones and rs227361-G/G knock-in clones. (F) *MANBA* expression levels in each knock-in clone after the knock-down of RXR-α. *$P < .01$, **$P < .001$ (Mann–Whitney U test).

compared with other East Asian populations.[58] In this regard, we have constructed a large Japanese population reference panel based on whole-genome sequences (1KJPN).[32] By using this reference panel, we previously identified primary functional variants in PBC susceptibility gene loci by high-density association mapping.[16,59,60] Through the application of high-density association mapping based on SNP imputation, using ethnicity-matched reference panels, additional primary functional variants in other digestive disease susceptibility gene loci are likely to be found in future studies.

In conclusion, high-density association mapping based on SNP imputation analysis, together with in silico/in vitro functional analyses, showed the molecular mechanisms by which rs17032850 and rs227361 may contribute to the development of PBC. Altered binding of the transcription factors LEF-1 and RXRα to the enhancer regions, which includes rs17032850 and rs227361, respectively, may affect the transcription of NFKB1 and *MANBA*. These results indicate that NFkB/LEF-1 and β-mannosidase/RXRα independently have important roles in the development of PBC. In regard to PBC susceptibility genes in the Japanese population,[16] we previously identified the primary functional

variants of *TNFSF15*, *PRKCB*, and the locus on chromosome 17q12-21, and described their molecular roles.[16,59,60] Similar approaches for other susceptibility genes are needed to further clarify the molecular mechanisms of disease development.

Materials and Methods
Subjects, Guidelines, and Regulations

Information regarding the participants of this study has been described previously.[16] We obtained written informed consent from all participants, and the study was approved by the committee on research ethics and genetically modified organisms of the Graduate School of Medicine at the University of Tokyo.

Genotype Imputation

The methodology used for SNP filtering and genotype imputation using a phased reference panel of 1KJPN[32] was described in our previous study.[59]

Notably, we applied the variable interval size (depending on the density of SNPs) to the reference panel. Starting from a minimal interval size of 5 Mb, the size was extended until
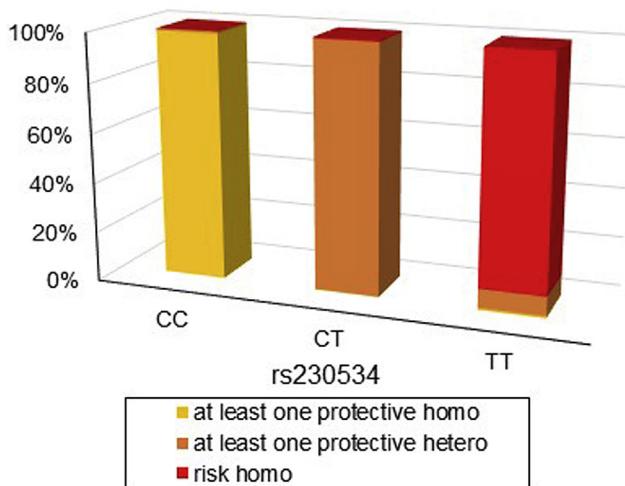
Figure 12. Co-segregation of the risk alleles of rs17032850 and rs227361 with the rs230534 T allele. The 2 putative primary functional variants co-segregated with the rs230534 T allele, which strongly associates with the susceptibility to PBC in the Japanese population. For each rs230534 genotype, the graph shows the percentages of individuals who are heterozygous or homozygous for at least 1 PBC-protective allele. homo, homozygote; hetero, heterozygote.

50 SNPs were included in the interval. The length of the buffer region was set to 1 Mb regardless of interval size.

Databases

The RegulomeDB database,[33] the UCSC genome browser,[34] LD link (https://ldlink.nci.nih.gov/),[61] and NCBI Variation viewer (https://www.ncbi.nlm.nih.gov/variation/view/) were used to evaluate the probability that candidate functional variants might influence transcription regulation, and to obtain LD information. PICS online[35] was used to obtain the probability score of SNPs that are likely to be the causal variants responsible for the association. TRANSFAC Professional[36] was used for the prediction of transcription factor binding. SWISS model, from the University of Basel,[38] and Polyphen2,[39] were used for the prediction of protein structure and the functional consequences of nonsynonymous substitutions.

Gene expression levels in each cell line were examined using data available from GeneCards at the Weizmann Institute of Science (http://www.genecards.org/).[62]

The correlation between rs17032850 or rs227361 genotypes and gene expression was examined using data from the GTEx portal database version 7.[37] P values less than .05, adjusted for multiple testing (Bonferroni correction), were regarded as statistically significant.

EMSA

EMSA was performed using the LightShift Chemiluminescent EMSA Kit (Thermo-Fisher Scientific, Waltham, MA) and biotin-labeled, double-stranded, oligonucleotide probes corresponding to each major and minor allele (Supplementary Table 4), according to the manufacturer's instructions. These oligonucleotide probes (10 fmol/μL) were incubated with a nuclear extract (2.5 μg/mL) of HepG2 or Jurkat cells (Nuclear Extract Kit; Active Region, Carlsbad, CA) for 30 minutes at 25°C.

The supershift assay was performed by incubating anti–LEF-1 antibody ab53293 (Abcam, Cambridge, UK) and anti-RXRα antibody sc-46659 (Santa Cruz Biotechnology, Santa Cruz, CA) for 24 hours at 4°C with the nuclear extracts, before subsequently incubating these complexes with the biotin-labeled probe for 30 minutes at 25°C. Each assay was performed independently 3 times.

Luciferase Reporter Assay

Specific polymerase chain reaction (PCR) primers (Supplementary Table 5) were used to amplify the 11th intron of *NFKB1* and the 7th intron of *MANBA* from human genomic DNA. Amplicons then were subcloned into the luciferase reporter pGL4.23 (luc2/minP) vector (Promega, Madison, WI). pGL4.23 constructs of each allele (500 ng) and the pGL4.74 (hRluc/TK) vector (50 ng), which was used as an internal control, were transfected into Jurkat and HepG2 cells using Lipofectamine 3000 (Thermo-Fisher Scientific). For the measurement of luciferase activity, the Dual-Luciferase Reporter Assay system (Promega) was used. Differences in relative luciferase activity were compared between major and minor alleles of each SNP using the Student t test. P values less than .05 were regarded as statistically significant. Each figure shows representative data of experiments that were performed independently 3 times. The data in the figures represent the means ± SD of triplicate assays in a single experiment.

Gene Editing Using CRISPR/Cas9

The sequences of the guide RNAs (gRNAs) were designed using CRISPR search and design tools (https://apps.thermofisher.com/apps/crispr/index.html#/search), and gRNAs were synthesized using the GeneArt Precision gRNA Synthesis Kit (Thermo-Fisher Scientific) (Supplementary Table 6). gRNAs, Protein GeneArt Platinum Cas9 Nuclease (Thermo-Fisher Scientific), and donor DNA for each allele were transfected into HepG2 cells using Lipofectamine CRISPR-MAX (Thermo-Fisher Scientific).

After the single-cell cloning and the extraction of genomic DNA using a PureLink Genomic DNA Mini Kit (Thermo-Fisher Scientific), gene editing of the target site and of off-target sites (site with homologies of >75% compared with the target site) were checked by Sanger sequencing (ABI prism 3730-XL, Thermo-Fisher Scientific) using the primer sets shown in Supplementary Table 6.

Quantitative Reverse-Transcription PCR

For Jurkat clones treated either with or without transfection of RXRα small interfering RNA sc-36447 (Santa Cruz Biotechnology) and LEF-1 small interfering RNA sc-35804 (Santa Cruz Biotechnology), total RNA was extracted using a RNeasy kit (Qiagen) and first-strand complementary DNA was synthesized using a High-Capacity Complementary DNA Reverse-Transcription Kit (Thermo-Fisher Scientific).

Reverse-transcription PCR to detect each transcript was performed using primers as shown in Supplementary Table 7 and FastStart Taq polymerase (Sigma-Aldrich, St. Louis, MO). To achieve linear amplification, 21 and 27 cycles were found to be optimal in preliminary experiments for *GAPDH* and other genes, respectively. Quantitation of each transcript was performed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). These experiments were repeated 3 times with essentially identical results.

References

1. Kaplan MM, Gershwin ME. Primary biliary cholangitis. N Engl J Med 2005;353:1261–1273.
2. Selmi C, Bowlus CL, Gershwin ME, Coppel RL. Primary biliary cirrhosis. Lancet 2011;377:1600–1609.
3. Gershwin ME, Mackay IR. The causes of primary biliary cirrhosis: convenient and inconvenient truths. Hepatology 2008;47:737–745.
4. Shimoda S, Hisamoto S, Harada K, Iwasaka S, Chong Y, Nakamura M, Bekki Y, Yoshizumi T, Shirabe K, Ikegami T, Maehara Y, He XS, Gershwin ME, Akashi K. Natural killer cells regulate T cell immune responses in primary biliary cirrhosis. Hepatology 2015;62:1817–1827.
5. Jones DE, Watt FE, Metcalf JV, Bassendine MF, James OF. Familial primary biliary cholangitis reassessed: a geographically based population study. J Hepatol 1999;30:402–407.
6. Selmi C, Mayo MJ, Bach N, Ishibashi H, Invernizzi P, Gish RG, Gordon SC, Wright HI, Zweiban B, Podda M, Gershwin ME. Primary biliary cholangitis in monozygotic and dizygotic twins: genetics, epigenetics, and environment. Gastroenterology 2004;127:485–492.
7. Hirschfield GM, Liu X, Xu C, Lu Y, Xie G, Lu Y, Gu X, Walker EJ, Jing K, Juran BD, Mason AL, Myers RP, Peltekian KM, Ghent CN, Coltescu C, Atkinson EJ, Heathcote EJ, Lazaridis KN, Amos CI, Siminovitch KA. Primary biliary cholangitis associated with HLA, IL12A, and IL12RB2 variants. N Engl J Med 2009; 360:2544–2555.
8. Hirschfield GM, Liu X, Han Y, Gorlov IP, Lu Y, Xu C, Lu Y, Chen W, Juran BD, Coltescu C, Mason AL, Milkiewicz P, Myers RP, Odin JA, Luketic VA, Speiciene D, Vincent C, Levy C, Gregersen PK, Zhang J, Heathcote EJ, Lazaridis KN, Amos CI, Siminovitch KA. Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cholangitis. Nat Genet 2010;42:655–657.
9. Liu X, Invernizzi P, Lu Y, Kosoy R, Lu Y, Bianchi I, Podda M, Xu C, Xie G, Macciardi F, Selmi C, Lupoli S, Shigeta R, Ransom M, Lleo A, Lee AT, Mason AL, Myers RP, Peltekian KM, Ghent CN, Bernuzzi F, Zuin M, Rosina F, Borghesio E, Floreani A, Lazzari R, Niro G, Andriulli A, Muratori L, Muratori P, Almasio PL, Andreone P, Margotti M, Brunetto M, Coco B, Alvaro D, Bragazzi MC, Marra F, Pisano A, Rigamonti C, Colombo M, Marzioni M, Benedetti A, Fabris L, Strazzabosco M, Portincasa P, Palmieri VO, Tiribelli C, Croce L, Bruno S, Rossi S, Vinci M, Prisco C, Mattalia A, Toniutto P, Picciotto A, Galli A, Ferrari C, Colombo S, Casella G, Morini L, Caporaso N, Colli A, Spinzi G, Montanari R, Gregersen PK, Heathcote EJ, Hirschfield GM, Siminovitch KA, Amos CI, Gershwin ME, Seldin MF. Genome-wide meta-analyses identify three loci associated with primary biliary cholangitis. Nat Genet 2010;42:658–660.
10. Mells GF, Floyd JA, Morley KI, Cordell HJ, Franklin CS, Shin SY, Heneghan MA, Neuberger JM, Donaldson PT, Day DB, Ducker SJ, Muriithi AW, Wheater EF, Hammond CJ, Dawwas MF, UK PBC Consortium; Wellcome Trust Case Control Consortium 3Jones DE, Peltonen L, Alexander GJ, Sandford RN, Anderson CA. Genome-wide association study identifies 12 new susceptibility loci for primary biliary cholangitis. Nat Genet 2011;43:329–332.
11. Hirschfield GM, Xie G, Lu E, Sun Y, Juran BD, Chellappa V, Coltescu C, Mason AL, Milkiewicz P, Myers RP, Odin JA, Luketic VA, Bacon B, Bodenheimer H, Liakina V, Vincent C, Levy C, Pillai S, Lazaridis KN, Amos CI, Siminovitch KA. Association of primary biliary cholangitis with variants in the CLEC16A, SOCS1, SPIB and SIAE immunomodulatory genes. Genes Immun 2012;13:328–335.
12. Liu JZ, Almarri MA, Gaffney DJ, Mells GF, Jostins L, Cordell HJ, Ducker SJ, Day DB, Heneghan MA, Neuberger JM, Donaldson PT, Bathgate AJ, Burroughs A, Davies MH, Jones DE, Alexander GJ, Barrett JC, Sandford RN, Anderson CA; UK Primary Biliary Cirrhosis (PBC) Consortium; Wellcome Trust Case Control Consortium 3. Dense fine-mapping study identifies new susceptibility loci for primary biliary cholangitis. Nat Genet 2012;44:1137–1141.
13. Juran BD, Hirschfield GM, Invernizzi P, Atkinson EJ, Li Y, Xie G, Kosoy R, Ransom M, Sun Y, Bianchi I, Schlicht EM, Lleo A, Coltescu C, Bernuzzi F, Podda M, Lammert C, Shigeta R, Chan LL, Balschun T, Marconi M, Cusi D, Heathcote EJ, Mason AL, Myers RP, Milkiewicz P, Odin JA, Luketic VA, Bacon BR, Bodenheimer HC Jr, Liakina V, Vincent C, Levy C, Franke A, Gregersen PK, Bossa F, Gershwin ME, deAndrade M, Amos CI, Italian PBC Genetics Study GroupLazaridis KN, Seldin MF, Siminovitch KA. Immunochip analyses identify a novel risk locus for primary biliary cirrhosis at 13q14, multiple independent associations at four established risk loci and epistasis between 1p31 and 7q32 risk variants. Hum Mol Genet 2012; 21:5209–5221.
14. Cordell HJ, Han Y, Mells GF, Li Y, Hirschfield GM, Greene CS, Xie G, Juran BD, Zhu D, Qian DC, Floyd JA, Morley KI, Prati D, Lleo A, Cusi D, Canadian-US PBC Consortium; Italian PBC Genetics Study Group; UK-PBC ConsortiumGershwin ME, Anderson CA, Lazaridis KN, Invernizzi P, Seldin MF, Sandford RN, Amos CI, Siminovitch KA. International genome-wide meta-analysis identifies new primary biliary cholangitis risk loci and targetable pathogenic pathways. Nat Commun 2015;6:8019.
15. Nakamura M, Nishida N, Kawashima M, Aiba Y, Tanaka A, Yasunami M, Nakamura H, Komori A, Nakamuta M, Zeniya M, Hashimoto E, Ohira H, Yamamoto K, Onji M, Kaneko S, Honda M, Yamagiwa S, Nakao K, Ichida T, Takikawa H, Seike M, Umemura T,

Ueno Y, Sakisaka S, Kikuchi K, Ebinuma H, Yamashiki N, Tamura S, Sugawara Y, Mori A, Yagi S, Shirabe K, Taketomi A, Arai K, Monoe K, Ichikawa T, Taniai M, Miyake Y, Kumagi T, Abe M, Yoshizawa K, Joshita S, Shimoda S, Honda K, Takahashi H, Hirano K, Takeyama Y, Harada K, Migita K, Ito M, Yatsuhashi H, Fukushima N, Ota H, Komatsu T, Saoshiro T, Ishida J, Kouno H, Kouno H, Yagura M, Kobayashi M, Muro T, Masaki N, Hirata K, Watanabe Y, Nakamura Y, Shimada M, Hirashima N, Komeda T, Sugi K, Koga M, Ario K, Takesaki E, Maehara Y, Uemoto S, Kokudo N, Tsubouchi H, Mizokami M, Nakanuma Y, Tokunaga K, Ishibashi H. Genome-wide association study identified TNFSF15 and POU2AF1 as susceptibility locus for primary biliary cholangitis in the Japanese population. Am J Hum Genet 2012;91:721–728.

16. Kawashima M, Hitomi Y, Aiba Y, Nishida N, Kojima K, Kawai Y, Nakamura H, Tanaka A, Zeniya M, Hashimoto E, Ohira H, Yamamoto K, Abe M, Nakao K, Yamagiwa S, Kaneko S, Honda M, Umemura T, Ichida T, Seike M, Sakisaka S, Harada M, Yokosuka O, Ueno Y, Senju M, Kanda T, Shibata H, Himoto T, Murata K, Miyake Y, Ebinuma H, Taniai M, Joshita S, Nikami T, Ota H, Kouno H, Kouno H, Nakamuta M, Fukushima N, Kohjima M, Komatsu T, Komeda T, Ohara Y, Muro T, Yamashita T, Yoshizawa K, Nakamura Y, Shimada M, Hirashima N, Sugi K, Ario K, Takesaki E, Naganuma A, Mano H, Yamashita H, Matsushita K, Yamauchi K, Makita F, Nishimura H, Furuta K, Takahashi N, Kikuchi M, Masaki N, Tanaka T, Tamura S, Mori A, Yagi S, Shirabe K, Komori A, Migita K, Ito M, Nagaoka S, Abiru S, Yatsuhashi H, Yasunami M, Shimoda S, Harada K, Egawa H, Maehara Y, Uemoto S, Kokudo N, Takikawa H, Ishibashi H, Chayama K, Mizokami M, Nagasaki M, Tokunaga K, Nakamura M. Genome-wide association study identified PRKCB as a genetic susceptibility locus for primary biliary cholangitis in a Japanese population. Hum Mol Genet 2017;26:650–659.

17. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet 2016; 48:709–717.

18. Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, Lee JC, Schumm LP, Sharma Y, Anderson CA, Essers J, Mitrovic M, Ning K, Cleynen I, Theatre E, Spain SL, Raychaudhuri S, Goyette P, Wei Z, Abraham C, Achkar JP, Ahmad T, Amininejad L, Ananthakrishnan AN, Andersen V, Andrews JM, Baidoo L, Balschun T, Bampton PA, Bitton A, Boucher G, Brand S, Büning C, Cohain A, Cichon S, D'Amato M, De Jong D, Devaney KL, Dubinsky M, Edwards C, Ellinghaus D, Ferguson LR, Franchimont D, Fransen K, Gearry R, Georges M, Gieger C, Glas J, Haritunians T, Hart A, Hawkey C, Hedl M, Hu X, Karlsen TH, Kupcinskas L, Kugathasan S, Latiano A, Laukens D, Lawrance IC, Lees CW, Louis E, Mahy G, Mansfield J, Morgan AR, Mowat C, Newman W, Palmieri O, Ponsioen CY, Potocnik U, Prescott NJ, Regueiro M, Rotter JI, Russell RK, Sanderson JD, Sans M, Satsangi J, Schreiber S, Simms LA, Sventoraityte J, Targan SR, Taylor KD, Tremelling M, Verspaget HW, De Vos M, Wijmenga C, Wilson DC, Winkelmann J, Xavier RJ, Zeissig S, Zhang B, Zhang CK, Zhao H, International IBD Genetics Consortium (IIBDGC)Silverberg MS, Annese V, Hakonarson H, Brant SR, Radford-Smith G, Mathew CG, Rioux JD, Schadt EE, Daly MJ, Franke A, Parkes M, Vermeire S, Barrett JC, Cho JH. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119–124.

19. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, Jostins L, Rice DL, Gutierrez-Achury J, Ji SG, Heap G, Nimmo ER, Edwards C, Henderson P, Mowat C, Sanderson J, Satsangi J, Simmons A, Wilson DC, Tremelling M, Hart A, Mathew CG, Newman WG, Parkes M, Lees CW, Uhlig H, Hawkey C, Prescott NJ, Ahmad T, Mansfield JC, Anderson CA, Barrett JC. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. Nat Genet 2017;49:256–261.

20. Pai S, Thomas R. Immune deficiency or hyperactivity-Nf-kb illuminates autoimmunity. J Autoimmun 2008; 31:245–251.

21. Artis D, Speirs K, Joyce K, Goldschmidt M, Caamaño J, Hunter CA, Scott P. NF-kappa B1 is required for optimal CD4+ Th1 cell development and resistance to Leishmania major. J Immunol 2003;170:1995–2003.

22. Erdman S, Fox JG, Dangler CA, Feldman D, Horwitz BH. Typhlocolitis in NF-kappa B-deficient mice. J Immunol 2001;166:1443–1447.

23. Hilliard B, Samoilova EB, Liu TS, Rostami A, Chen Y. Experimental autoimmune encephalomyelitis in NF-kappa B-deficient mice: roles of NF-kappa B in the activation and differentiation of autoreactive T cells. J Immunol 1999;163:2937–2943.

24. Das J, Chen CH, Yang L, Cohn L, Ray P, Ray A. A critical role for NF-kappa B in GATA3 expression and TH2 differentiation in allergic airway inflammation. Nat Immunol 2001;2:45–50.

25. Alkhayat AH, Kraemer SA, Leipprandt JR, Macek M, Kleijer WJ, Friderici KH. Human beta-mannosidase cDNA characterization and first identification of a mutation associated with human beta-mannosidosis. Hum Mol Genet 1998;7:75–83.

26. Cooper A, Sardharwalla IB, Roberts MM. Human beta-mannosidase deficiency. N Engl J Med 1986;315:1231.

27. Wenger DA, Sujansky E, Fennessey PV, Thompson JN. Human beta-mannosidase deficiency. N Engl J Med 1986;315:1201–1205.

28. Dorland L, Duran M, Hoefnagels FE, Breg JN, Fabery de Jonge H, Cransberg K, van Sprang FJ, van Diggelen OP. Betamannosidosis in two brothers with hearing loss. J Inherit Metab Dis 1988;11(Suppl 2):255–258.

29. Wijburg H, de Jong J, Wevers R, Bakkeren J, Trijbels F, Sengers R. Beta-mannosidosis and ethanolaminuria in a female patient. Eur J Pediatr 1992;151:311.

30. Poenaru L, Akli S, Rocchiccioli F, Eydoux P, Zamet P. Human beta-mannosidosis: a 3-year-old boy with speech impairment and emotional instability. Clin Genet 1992;41:331–334.

31. Levade T, Graber D, Flurin V, Delisle MB, Pieraggi MT, Testut MF, Carrière JP, Salvayre R. Human

beta-mannosidase deficiency associated with peripheral neuropathy. Ann Neurol 1994;35:116–119.

32. Nagasaki M, Yasuda J, Katsuoka F, Nariai N, Kojima K, Kawai Y, Yamaguchi-Kabata Y, Yokozawa J, Danjoh I, Saito S, Sato Y, Mimori T, Tsuda K, Saito R, Pan X, Nishikawa S, Ito S, Kuroki Y, Tanabe O, Fuse N, Kuriyama S, Kiyomoto H, Hozawa A, Minegishi N, Douglas Engel J, Kinoshita K, Kure S, Yaegashi N, ToMMo Japanese Reference Panel ProjectYamamoto M. Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals. Nat Commun 2015;6:8018.

33. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22:1790–1797.

34. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res 2002;12:996–1006.

35. Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, Ryan RJ, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 2015;518:337–343.

36. Wingender E, Dietze P, Karas H, Knüppel R. TRANSFAC: a database on transcription factors and their DNA binding sites. Nucleic Acids Res 1996;24:238–241.

37. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580–585.

38. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, Schwede T. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 2014; 42:W252–W258.

39. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–249.

40. Shulman AI, Mangelsdorf DJ. Retinoid x receptor heterodimers in the metabolic syndrome. N Engl J Med 2005; 353:604–615.

41. Uchino Y, Fukushige T, Yotsumoto S, Hashiguchi T, Taguchi H, Suzuki N, Konohana I, Kanzaki T. Morphological and biochemical studies of human beta-mannosidosis: identification of a novel beta-mannosidase gene mutation. Br J Dermatol 2003; 149:23–29.

42. Serrano-Puebla A, Boya P. Lysosomal membrane permeabilization in cell death: new evidence and implications for health and disease. Ann N Y Acad Sci 2016; 1371:30–44.

43. Wu DJ, Adamopoulos IE. Autophagy and autoimmunity. Clin Immunol 2017;176:55–62.

44. Yasunami M, Nakamura H, Tokunaga K, Kawashima M, Nishida N, Hitomi Y, Nakamura M. Principal contribution of *HLA-DQ* alleles, *DQB1*06:04* and *DQB1*03:01*, to

disease resistance against primary biliary cholangitis in a Japanese population. Sci Rep 2017;7:11093.

45. Invernizzi P, Ransom M, Raychaudhuri S, Kosoy R, Lleo A, Shigeta R, Franke A, Bossa F, Amos CI, Gregersen PK, Siminovitch KA, Cusi D, de Bakker PI, Podda M, Gershwin ME, Seldin MF; Italian PBC Genetics Study Group. Classical HLA-DRB1 and DPB1 alleles account for HLA associations with primary biliary cirrhosis. Genes Immun 2012;13:461–468.

46. Bamias G, Martin C 3rd, Marini M, Hoang S, Mishina M, Ross WG, Sachedina MA, Friel CM, Mize J, Bickston SJ, Pizarro TT, Wei P, Cominelli F. Expression, localization, and functional activity of TL1A, a novel Th1-polarizing cytokine in inflammatory bowel disease. J Immunol 2003;171:4868–4874.

47. Zhou M, Liu R, Su D, Feng X, Li X. TL1A increased the differentiation of peripheral Th17 in rheumatoid arthritis. Cytokine 2011;69:125–130.

48. Wolf SF, Sieburth D, Sypek J. Interleukin 12: a key modulator of immune function. Stem Cells 1994; 12:154–168.

49. Aiba Y, Yamazaki K, Nishida N, Kawashima M, Hitomi Y, Nakamura H, Komori A, Fuyuno Y, Takahashi A, Kawaguchi T, Takazoe M, Suzuki Y, Motoya S, Matsui T, Esaki M, Matsumoto T, Kubo M, Tokunaga K, Nakamura M. Disease susceptibility genes shared by primary biliary cirrhosis and Crohn's disease in the Japanese population. J Hum Genet 2015;60:525–531.

50. Sun Y, Irwanto A, Toyo-Oka L, Hong M, Liu H, Andiappan AK, Choi H, Hitomi Y, Yu G, Yu Y, Bao F, Wang C, Fu X, Yue Z, Wang H, Zhang H, Kawashima M, Kojima K, Nagasaki M, Nakamura M, Yang SK, Ye BD, Denise Y, Rotzschke O, Song K, Tokunaga K, Zhang F, Liu J. Fine-mapping analysis revealed complex pleiotropic effect and tissue-specific regulatory mechanism of TNFSF15 in primary biliary cholangitis, Crohn's disease and leprosy. Sci Rep 2016;6:31429.

51. Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, Albrecht M, Mayr G, De La Vega FM, Briggs J, Günther S, Prescott NJ, Onnie CM, Häsler R, Sipos B, Fölsch UR, Lengauer T, Platzer M, Mathew CG, Krawczak M, Schreiber S. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet 2007; 39:207–211.

52. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, Green T, Kuballa P, Barmada MM, Datta LW, Shugart YY, Griffiths AM, Targan SR, Ippoliti AF, Bernard EJ, Mei L, Nicolae DL, Regueiro M, Schumm LP, Steinhart AH, Rotter JI, Duerr RH, Cho JH, Daly MJ, Brant SR. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat Genet 2007;39:596–604.

53. Saitoh T, Fujita N, Jang MH, Uematsu S, Yang BG, Satoh T, Omori H, Noda T, Yamamoto N, Komatsu M, Tanaka K, Kawai T, Tsujimura T, Takeuchi O, Yoshimori T, Akira S. Loss of the autophagy protein Atg16L1 enhances endotoxin-induced IL-1beta production. Nature 2008;456:264–268.

54. Cadwell K, Liu JY, Brown SL, Miyoshi H, Loh J, Lennerz JK, Kishi C, Kc W, Carrero JA, Hunt S, Stone CD, Brunt EM, Xavier RJ, Sleckman BP, Li E, Mizushima N, Stappenbeck TS, Virgin HW 4th. A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. Nature 2008;456:259–263.

55. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014;42:D1001–D1006.

56. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. Nature 2009;461:747–753.

57. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. Annu Rev Genomics Hum Genet 2009;10:387–406.

58. 1000 Genomes Project ConsortiumAuton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. A global reference for human genetic variation. Nature 2015;526:68–74.

59. Hitomi Y, Kojima K, Kawashima M, Kawai Y, Nishida N, Aiba Y, Yasunami M, Nagasaki M, Nakamura M, Tokunaga K. Identification of the functional variant driving ORMDL3 and GSDMB expression in human chromosome 17q12-21 in primary biliary cholangitis. Sci Rep 2017;7:2904.

60. Hitomi Y, Kawashima M, Aiba Y, Nishida N, Matsuhashi M, Okazaki H, Nakamura M, Tokunaga K. Human primary biliary cirrhosis-susceptible allele of rs4979462 enhances TNFSF15 expression by binding NF-1. Hum Genet 2015;134:737–747.

61. Machiela MJ, Chanock SJ. LDlink a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 2015;31:3555–3557.

62. Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. Trends Genet 1997;13:163.

Correspondence

Address correspondence to: Yuki Hitomi, PhD, Department of Human Genetics, Graduate School of Medicine, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. e-mail: yhitomi-tky@umin.ac.jp; fax: (81) 3-5802-8619.