

PLD4 is a genetic determinant to systemic lupus erythematosus and involved in murine autoimmune-phenotypes.

Shuji Akizuki¹, Kazuyoshi Ishigaki², Yuta Kochi³, Sze-Ming Law¹, Keitaro Matsuo^{4,5}, Koichiro Ohmura¹, Akari Suzuki³, Manabu Nakayama⁶, Yusuke Iizuka⁷, Haruhiko Koseki⁷, Osamu Ohara⁸, Jun Hirata⁹⁻¹¹, Yoichiro Kamatani^{2,12}, Fumihiko Matsuda¹², Takayuki Sumida¹³, Kazuhiko Yamamoto³, Yukinori Okada^{2,9,14}, Tsuneyo Mimori¹, Chikashi Terao^{1,2,15,16}

1. Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan.
2. Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan.
3. Laboratory for Autoimmune Diseases, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan.
4. Division of Cancer Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya 464-8681, Japan
5. Department of Epidemiology, Nagoya University Graduate School of Medicine, Nagoya, 466-8560, Japan
6. Department of Frontier Research and Development, Kazusa DNA Research Institute, Chiba 292-0818, Japan.
7. Laboratory for Developmental Genetics, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan.
8. Laboratory for Integrative Genomics, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan.
9. Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita 565-0871, Japan.
10. Pharmaceutical Discovery Research Laboratories, TEIJIN PHARMA LIMITED, Hino 191-8512, Japan.
11. Department of Human Genetics and Disease Diversity, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo 113-8510, Japan.
12. Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan
13. Department of Internal Medicine, Faculty of Medicine, University of Tsukuba, Tsukuba,

305-8575, Japan

14. Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita 565-0871, Japan.

15. Clinical Research Center, Shizuoka General Hospital, Shizuoka 420-0881, Japan.

16. Department of Applied Genetics, The School of Pharmaceutical Sciences, University of Shizuoka, Shizuoka 422-8526, Japan.

Correspondence should be addressed to: Chikashi Terao

E-mail: a0001101@kuhp.kyoto-u.ac.jp

Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan

Word count:4271

Key words: systemic lupus erythematosus, genetic study, autoimmunity, mouse study

Abstract

Objectives: Systemic lupus erythematosus (SLE) is an autoimmune disease that is characterized by autoantibody production and widespread inflammation damaging many organs. Previous genome-wide association studies (GWAS) have revealed over 80 genetic determinants of SLE, but they collectively explain a fraction of the heritability, and only a few were proven in vivo for the involvement in SLE. We conducted a meta-analysis of SLE GWAS in the Japanese population, followed by functional analyses of a susceptibility gene with use of mutant mice.

Methods: We conducted a meta-analysis of two GWAS comprising a total of 1363 cases and 5536 controls using the 1000 Genome Project data as an imputation reference. Enrichment analyses for functional annotations were conducted. We examined Phospholipase D4 (Pld4) mutant mice to assess functional involvement of a genetic determinant.

Results: We found a total of 14 significant loci, which included rs2582511 in *AHNAK2/PLD4* recently reported in a Chinese study and a novel locus of rs143181706 in *MAMLD1* ($p=7.9\times10^{-11}$ and 3.7×10^{-8} , respectively). PLD4 risk allele was associated with anti-dsDNA antibody production. Enrichment analysis of genetic signals revealed involvement of a wide range of immune-related cells and pathways. Pld4 mutant mice revealed remarkably low body weight. The mice demonstrated autoimmune phenotypes compatible with SLE, including splenomegaly and lymphadenopathy, expansion of B cells and hyper secretion of BAFF and production of autoantibodies especially anti-nuclear antibody and anti-dsDNA antibody.

Conclusions: We found a novel susceptibility gene to SLE. Pld4 mutant mice revealed autoimmune phenotypes suggesting functional involvement of PLD4 with the basics of SLE.

Introduction

Systemic lupus erythematosus (SLE) is an autoantibody-mediated autoimmune disease characterized by systemic organ involvement[1]. A variety of autoantibodies including anti-nuclear antibody (ANA), anti-DNA antibody, anti-Ro antibody and anti-RNP antibody frequently accompany SLE[1]. More than 80% of the patients with SLE are young females at the onset of this disease[1]. Severe organ damages include renal failure and central nervous system (CNS) lupus, both of which are sometimes fatal. Serologically, type I interferon, especially interferon alpha signature is assumed to play critical roles in SLE[2, 3]. Elevated levels of B-cell-activating factor of the TNF family (BAFF), cytokine activating B cell expansion and functional proliferation are features of SLE[4]. Belimumab, a monoclonal antibody targeting BAFF was approved to be the first molecular target agent for SLE[4].

Although treatments are available for this autoimmune disorder, environmental and genetic components are associated with the onset of SLE and more research is required to develop additional comprehensive therapies. A previous study analyzing 104 monozygotic twins reported 24% concordance for SLE, much higher than that in RA[5]. Genetic components are estimated to explain about 44 to 66% of the onset of SLE[6]. HLA is the strongest genetic component among the susceptibility loci[7]. Previous genetic studies, mainly genome-wide association studies (GWAS) have identified a total of more than 80 susceptibility loci to SLE[7-12], some of which were ethnic-specific[7]. However, these loci collectively explain a mere ~24% of heritability and there are a number of yet-to-be-identified genetic loci[9]. Furthermore, it is a challenge to narrow down a single causative gene and variant out of the candidate susceptibility loci. In spite of identifying over 80 susceptibility loci to SLE, a few genes or variants were functionally examined for their associations with SLE pathophysiology[13].

Expansion of subjects in a single population is an effective approach to identify susceptibility loci which is population-specific or difficult to be found through trans-ethnic meta-analysis due to different linkage disequilibrium (LD) structure or allele frequencies. In fact, previous Asian-specific approach seemed to be successful in uncovering genetic determinants for SLE[9]. While so far there have been numerous GWAS studies conducted in European and Asian populations, there is only one GWAS from the Japanese population in which a total of six non-HLA loci were identified, five of which were shared by Europeans[7]. To expand our knowledge of genetic background of SLE and its functional relevance, here we conducted a meta-analysis of the previous

Japanese GWAS and a *de novo* GWAS. In addition, mutant mouse of *Pld4*, a susceptibility gene to SLE was examined to get an insight of SLE pathophysiology.

Materials and Methods

Genetic associations studies of SLE

Subjects

A total of 1,363 cases and 5,536 controls were recruited for this study. A detailed breakdown of the samples is shown in Supplementary Table 1. We conducted a *de novo* GWAS composed of 474 case subjects and 2162 control subjects as the 1st set. We utilized the previous Japanese GWAS data as the 2nd set composed of 889 cases and 3374 controls. All of the case subjects fulfilled the revised ACR criteria in 1997[14]. Written informed consent was obtained from all of the participants. This study was approved by the ethical committee in each institution.

Quality control

Criteria of quality control before imputation was decided for each set, the details of which are described in Supplementary Table 2. Kinship was estimated in both studies by plink software. PCA was carried out in each data set. Outliers of East Asian (EAS) cluster (Supplementary Figure 1) and samples showing high degree of relatedness with other subjects were excluded. Subjects showing success rate less than 0.98 and 0.99 were excluded in sets 1 and 2, respectively.

Imputation

The two GWAS data sets were imputed onto 1000 Genome project East Asian population panel p3v5 by use of minimach3 software after phasing by shapeit2 software. After imputation, variants with Rsq less than 0.5 in either of the two studies were filtered out. We also constrained single nucleotide polymorphisms (SNPs) with frequency more than 1% in controls to avoid false positives. As a result, a total of 7,110,910 variants remained for the subsequent analyses.

LD score regression

LD score regression analysis was conducted with the use of statistics of the meta-analysis using ldsc software[15]. We assessed intercept before and after LD score regression to estimate polygenic effects on SLE susceptibility.

Clinical information

Presence of autoantibodies in SLE, namely, anti-DNA antibody, anti-SSA antibody, anti-SSB antibody, and anti-U1RNP antibody were obtained from clinical charts. Since all of the subjects in set 1 were derived from a single institution, clinical information was intensively analyzed in set 1 to avoid data noises arising from different measurement methods.

Functional annotation and enrichment

Variants with significant associations were functionally annotated with the use of Haploreg[16] to evaluate enrichment of enhancer histone marks in significant variants or variants in LD with them.

LD estimation

Linkage disequilibrium between variants were estimated by plink software[17].

Heritability estimate and partitioning heritability

Heritability explained by a single variant or a set of significant variants was calculated by liability threshold model assuming prevalence of SLE as 0.05%[1] the details of which are described elsewhere[18]. Partitioning of heritability into cell groups or detailed cell types was conducted by ldsc software[15].

Pathway analysis

Pathway analysis for the evaluation of important molecular networks based on genetic association results was conducted by using PASCAL software[19]. We adopted the method to calculate gene scores by taking sum of all variants, not restricted to top variants, in gene regions into account.

Statistical analysis for genetic studies.

Logistic regression analysis was performed for this study, and the covariates for each GWAS are shown in Supplementary Table 2. Statistical analysis was conducted by plink1.9 software or mach2dat software for sets 1 and 2, respectively. Meta-analysis was conducted with the use of the inverse-variance method assuming fixed effects by R statistical software. As for the X chromosome, we separately conducted association studies for males and females and combined the results in each set. Significant level was set as 5.0×10^{-8} for GWAS and 0.05 for clinical phenotypes and mice study. To confirm

a significant association between a clinical phenotype and genotype, we further conducted a permutation test with 10,000 permutations by shuffling genotypes to take covariance structure among autoantibody positivity into account and obtained null distribution of the smallest p-values.

In-vivo study of *Pld4*

Mice and cells

BALB/c *Pld4 thss/thss* mutants were purchased from The Jackson Laboratory (JAX stock #012624) [20]. All animal studies used 3 to 6 female mice that are 12-15 weeks of age, repeated at least two times, unless otherwise indicated. All mice were maintained in a specific pathogen free condition at Kyoto University, Japan. All experiments were approved by the Animal Care and Use Committee of the Institute for Kyoto University and were performed in accordance with the institutional guidelines. For cell culture experiment, splenic CD19⁺ cells were isolated using mouse CD19 micro beads combined with MACS Column according to the manufacture's protocol (Miltenyi Biotec). The purity of CD19⁺ cells were typically 85-90%.

Confirmation of genotypes

DNA was extracted from tail or ear of mice and subjected to either Sanger sequencing or polymerase chain reaction followed by restriction digest by HpyCH4III enzyme. Details are described in Supplementary Methods.

Analysis of body weight and organs

Body weight was measured every week from birth for both mutant and heterozygous mice. Weights of fresh organs were compared between the two genotypes at 1 month after birth.

Antibodies, cell culture and flow cytometric analysis

Monoclonal antibodies used for the current study are described in Supplementary Methods. For B cell proliferation assay, the purified CD19⁺ cells were labelled with 5 μ M CFSE for 5 minutes at room temperature, then washed three times with PBS containing 5% FCS and cultured at a density of 1x10⁶ cells/mL in RPMI 1640 supplemented with 10% FCS, 50 mM β -mercaptoethanol, 2 mM glutamine, 100 U/mL penicillin and 100 mg/mL streptomycin for 72 hours, in the presence or absence of goat anti-mouse IgM (Southern Biotec), LPS (Sigma-Aldrich), CpG ODN 2395 (Miltenyi Biotec) at the

indicated concentrations. The dividing cells were determined as the cells with reduced intensity of CFSE labelling measured by flow cytometry. All stained cells were analyzed by FACSCalibur (BD Biosciences).

Histology and immunohistochemistry

For histology, every organ was fixed in 10% formalin, embedded in paraffin, and the sections were stained with hematoxylin and eosin (H&E). For immunohistochemistry of the kidney, the organs were snap frozen in isopentane with dry ice, mounted in OCT compound, and stored at -80°C. Cryostat sections at 10µm were cut and mounted on gelatin-coated histological slides. Sections were air dried, and fixed with ice-cold acetone for 20min. After blocking for non-specific staining with PBS containing 2% FCS, sections were incubated with FITC-conjugated anti-mouse IgG antibody for 60min at 4°C, and rinsed with PBS. For immunohistochemistry of the spleen, snap frozen sections were fixed with ice-cold acetone for 5 minutes, washed with PBS, blocked with Blocking One (Nacalai Tesque, Japan) for 30 minutes, and then incubated with indicated antibodies for 60 minutes or streptavidin for 30 minutes at room temperature. All images were obtained by fluorescence microscope FSX100 (Olympus). All slides were prepared by the Center for Anatomical, Pathological and Forensic Medical Research, Kyoto University Graduate School of Medicine.

ELISA and CBA assay

We quantified serum levels of BAFF by Mouse BAFF/Blys Quantikine ELISA kit according to the manufacturer's protocol (R&D Systems, U.S.). For determining concentration of serum immunoglobulin in mice, we used Mouse Immunoglobulin Isotyping Kit according to the manufacturer's protocol (BD Biosciences).

Anti-nuclear antibody (ANA) and anti-dsDNA antibody detection.

We determined serum ANA by using ImmuGlo™ ANA Hep-2 substrate according to the manufacturer's protocol. For detection of anti-nuclear antibody toward mouse antigen, 5×10^5 mouse embryonic fibroblast cells were plated on 24 flat-bottom well, culture dish, and cultured overnight. Attached cells were fixed with 0.5% formaldehyde for 5 min, then permeated with 0.01% triton-X. After washing the cells, serially diluted sera were applied, followed by FITC-conjugated anti-mouse IgG secondary antibody. Titer of ANA was defined as the serum dilution at which fluorescence was no longer visible. For anti-dsDNA antibody, Mouse Anti-dsDNA ELISA KIT (Shibayagi, Japan) was used following the manufacturer's instructions.

Gene expression analysis

We extracted total RNA from kidney and splenocyte using RNeasy mini kit (QIAGEN, Germany). A total of 1ug of the RNA was reverse-transcribed into cDNA with High Capacity cDNA Reverse Transcription Kits (Applied Biosystems, U.S.). For the gene expression analysis of kidney, control glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and target cDNAs were amplified using SYBR green method. Details of primers are described in Supplementary Methods. For splenocytes, mRNA expression of IL-6, IL-12b, TNFa, IL-1b were obtained using TaqMan Gene Expression Assay (Applied Biosystems), according to the manufacturer's protocol. We conducted real-time PCRs with the use of the ABI 7500 system (Applied Biosystems).

Statistical analysis for mice studies

SPSS software was used for analyses (IBM, U.S.). Welch's t test or Fisher's exact test was carried out for statistical analysis using Prism version 6 software (GraphPad, U.S.). P values less than 0.05 were considered significant.

Results

A total of 1,363 cases and 5,536 controls in two GWAS were imputed for genome-wide SNPs using 1000G p3v5 panel as reference. We conducted logistic regression analysis and combined the results by inverse-variance method to assess the overall significance.

As a result, there was no evidence for confounding bias leading to deviation of chi-square statistics from expectation across our data set ($\lambda=1.08$, Fig 1A). The HLA region showed the strongest signal (rs796780915, $p=1.3 \times 10^{-29}$). We identified a total of 13 non-HLA loci satisfying the GWAS significant level (Table 1 and Figure 1B). *AHNAK2/PLD4* on chromosome 14, a susceptibility locus very recently reported in a Chinese study[21], and *MAMLD1* on chromosome X, a novel susceptibility locus, were identified (rs2582511, $p=7.9 \times 10^{-11}$ and rs143181706, $p=3.7 \times 10^{-8}$, respectively, Table 1 and Figure 1B). Both susceptibility SNPs showed comparable effect sizes across the two studies (Table 1). Conditioning on the top SNP in each region, we did not identify independent signals ($p>0.08$). Association with SLE was previously reported for the remaining 12 loci. We found two loci (*BMP2K*, and *PTGER4*) showing suggestive associations ($p \leq 1.0 \times 10^{-6}$) which were not reported previously (Table 1).

We further found evidence of genetic similarities between the current and previous studies. Our results contained 39 top SNPs previously reported as SLE susceptibility loci and not in Table 1 (we excluded susceptibility loci from Asian studies containing part of our results). Twenty six of the 39 markers showed p-values less than 0.05 and 37 out of the 39 shared risk alleles (Supplementary Table 3), indicating strong genetic overlap across populations.

rs2582511 is a synonymous variant of *AHNAK2* which encodes a nucleoprotein involved with calcium signaling and its role in carcinogenesis has been previously suggested[22]. The LD block containing rs2582511 spanned 400kb and contained *AHNAK2* and *PLD4* which encodes a member of the phospholipase family without phospholipase D activity[23] (Figure 2A).

rs143181706 is in an intronic region of *MAMLD1*. *MAMLD1* is the only gene in the LD block containing rs143181706 (Figure 2B), and is associated with genital development including hypospadias and gonadal dysgenesis[24].

We further analyzed whether rs2582511 and rs143181706 are associated with clinical phenotypes of SLE. We found that rs2582511 risk allele was associated with increased positivity of anti-dsDNA antibody (nominal $p=0.0073$, Table 2). We also found that the risk alleles in the two SNPs showed positive associations ($OR>1$) with all of the auto-antibodies analyzed (permutation $p=0.006$, Table 2), suggesting PLD4 works on B cell hyperactivity.

When we calculated heritability explained by the two SNPs based on liability-scale threshold model, rs2582511 and rs143181706 explained 0.37 and 0.26% of heritability, respectively. A total of 5.7% of heritability was brought about by the fourteen significant SNPs in the current study.

We found that significant non-HLA SNPs and their linked SNPs showed enrichment of enhancer histone marks in disease-relevant cell types including B cells and NK cells by Haploreg (Supplementary Table 4).

In order to assess polygenic effects on SLE susceptibility, we conducted LD score regression. Since we found a strong correlation between LD scores and chi-squared values (Supplementary Figure 2) and decreased intercept by LD score regression (from

1.086 to 1.04), yet-to-be-determined polygenic effects seem to underlie this disease. We then analyzed whether heritability enrichment considering polygenic effects was observed in specific cell groups. As a result, we found that hematopoietic cell groups showed strong enrichment of heritability ($p=1.2 \times 10^{-8}$, Supplementary Figure 3). We further took advantage of LD scores of specific cell types and found strong enrichment of histone marks, especially H3K4me1, of immune-relevant cells including CD56+, CD3+, CD4+ memory and CD8+ memory cells (Supplementary Table 5 and Supplementary Figure 4).

In order to assess important molecular pathways supported by the genetic findings, we conducted pathway analysis using PASCAL software[19] which takes non-significant signals into account. As a result, in addition to interferon (IFN) signaling pathway, other immune-related pathways showed significant associations (Supplementary Table 6). It is interesting that some immune-related molecules, including IL12 and IL10 were pinpointed.

We further dig into the *AHNAK2/PLD4* region. Previous Japanese cell-specific eQTL study did not pinpoint one of the two genes[25]. We decided to focus on *PLD4* gene for functional analysis rather than *AHNAK2* because of the following reasons. Firstly, *PLD2*, a family member gene of *PLD4*, is a susceptibility gene to SLE in the European population[11]. Secondly, *PLD4* regulates kidney fibrosis in human and mice[26], which is a severe complication of SLE. Thirdly, *Pld4* is dominantly expressed in immune-related organs and cells including B cells, monocyte and dendritic cells[27]. Fourthly, the previous GWAS meta-analysis of rheumatoid arthritis (RA) in the Japanese population showed a SNP in *PLD4* as a top signal in this locus[28].

Therefore, we have analyzed the function of PLD4 using Pld4 mutant mouse which was first identified and maintained in Jackson Laboratory. The mutant mice carry a single nucleotide transition from G to T at position 114,001,632 (NCBI build 37) which is a nonsense mutation, resulting in the introduction of a premature stop codon at residue 46 of the 503 amino acid protein (Figure 3A).

We compared mutant homozygous mice (mutant mice, m/m) with mutant heterozygous mice (w/m) in the characterization analysis. As Jackson Laboratory reported[20], mutant mice have revealed decreased weight in comparison to heterozygous mice which became evident soon after birth (Figure 3B). Adult mutant mice showed marked

splenomegaly and lymphadenopathy (Figure 3C), inversely proportional to their body size, while the other major non-lymphoid organs were in proportion (Supplementary Figure 5). Regarding the number of splenocytes, there was approximately 1.5 - 2 fold more in mutant mice compared to heterozygotes (Figure 3D). Microscopic histological analyses revealed remarkable spontaneous formation of germinal center in spleen and lymphoid organs in mutant mice (Figure 3E, arrows, Supplementary Figure 6), suggesting the involvement of *Pld4* in immunological functions especially in the context of B cell activities. All experiments were performed using female mice on the assumption that female was more prone to develop autoimmunity. However, small body and splenomegaly were observed in male homozygous mutant mice as well.

In order to characterize the lymphoid enlargement, we analyzed the immune cell profile within lymphoid organs. Mutant mice bear more T and B cells in the spleen, but cell fraction was comparable (Figure 4A). The T cells of mutant mice displayed an activated phenotype, shown as elevated percentage of CD69+ cells (Figure 4A, right panel). Furthermore, the cell fractions of marginal zone B cells, macrophages and plasmacytoid dendritic cells which is important for type I interferon secretion, were increased in mutant mice (Figure 4B). These results collectively suggest that *Pld4* mutant mouse has altered the homeostasis of lymphoid organs. Histologically, although most organs of 12-15 weeks old mice showed normal appearance in Hematoxylin-Eosin (HE) staining (Supplementary Figure 7), some mutant homozygotes showed scattered infiltration of inflammatory cells within the liver (Supplementary Figure 7).

Since expansion and activation of B cells were suggested in mutant mice, we analyzed B cells in detail. The peripheral B cell counts in the mutants were proportional to their small body at neonatal period (1 week of age). After 3 weeks of age, the cell numbers rapidly increased, exceeding the number observed in control mouse (Supplementary Figure 8A). However, we did not find much difference in the percentage of proliferation of B cells in vitro between the two genotypic groups upon anti-IgM, LPS, CpG ODN stimuli (Supplementary Figure 8B). We found escalated percentage of germinal center B cells and plasma cells (CD138+) in the mutant mice (Figure 4C). We quantified serum levels of B cells-related molecules and identified hypergammaglobulinemia and marked elevation of BAFF in mutant mice (Figure 4C). We did not find class-specific increase in immunoglobulin (Figure 4C), suggesting an expansion of plasma cells irrespective of class-switch, and possibly, somatic hypermutation. We also found increase of anti-nuclear antibody (ANA) and anti-DNA antibody in mutant mice (Figure 4D).

Since splenomegaly, lymphadenopathy, hypergammaglobulinemia and the increase of ANA and anti-dsDNA antibody suggest B-cell autoimmunity, we assessed SLE-related phenotypes in the mutant mice. We surveyed immune deposition in kidneys in the two genotypic groups. We found PAS-positive deposition in the glomeruli of the mutant mice, but not in the heterozygous mice (Figure 4E). Immuno-histochemical staining revealed deposition of IgG and C3 in the glomeruli of the mutant mice (Figure 4E). In addition, we found upregulation in the gene expression of proinflammatory cytokines (*Il6*, *Il12b*) in the spleen, and interferon signature in the kidney of mutant mice (*Isg15*, *OAS2*, *Irf4* and *Ifna*) (Supplementary Figure 9, Figure 4E lower panel). On the contrary, we did not find remarkable difference of urine protein adjusted for creatinine between the mutant mice and heterozygous mice (Figure 4E) and no significant differences were observed regarding lifespan of mutant and heterozygous mice (data not shown).

Collectively, these findings in mice indicate *Pld4* involvement in contributing to autoimmune phenotypes, especially in the production of antinuclear antibody, and immune complex mediated-tissue injury.

Discussion

In the current study, we conducted a meta-analysis of GWAS for SLE in the Japanese population and identified a novel susceptibility gene and another susceptibility gene which was very recently reported by a Chinese group. Involvement of *PLD4* in B-cell autoimmunity was supported by analysis using *Pld4* mutant mice.

To the best of our knowledge, the association of *PLD4* and *MAMLD1* with SLE in the current study had not been reported in Europeans or in Asian meta-analysis before late 2017. This suggests that heterogeneity exists in the same population group and that there are potential undetermined signals which are specific to a population. Apparently, trans-ethnic meta-analysis is a promising approach to find novel signals by leveraging increased sample sizes but our results suggest the usefulness of GWAS in a single and relative uniform population. Taking advantage of the difference in LD structure and allele frequencies between populations would be a promising approach to find novel susceptibility loci.

Our analysis also suggests many unknown variants with small effect sizes associated with SLE. Increase of sample sizes in a single population, Asian populations and trans-

ethnic populations would lead to identification of novel signals. Our results also identified two suggestive loci as candidates of future replication studies.

The current analyses demonstrated promising pathways and immune cells which would serve as therapeutic targets. Designing clinical and basic experiments based on the list would be an efficient approach for the development of novel treatment for SLE.

While we prefer *PLD4* as a better candidate, *AHNAK2*, the adjacent gene in the same LD block, showed upregulated expression in murine lupus nephritis[21]. There is also a possibility of a trans-eQTL effect of a variant beyond this LD block. Thus, further investigations are needed to pinpoint a mechanism of the association in this locus.

Mammalian PLD is known to be a major player in the regulation of actin cytoskeleton, vesicle trafficking for secretion and endocytosis, and receptor signaling. Proteomics data suggest that human PLD4 is expressed in plasma, monocyte, B-lymphocyte and natural killer cell. Within the cells, PLD4 seems to localize mostly to the ER and Golgi, as well as the nucleus. We previously reported that PLD4 is also associated with systemic sclerosis (SSc)[29]. Since previous studies have shown overlapping susceptibility genes across multiple autoimmune disorders, PLD4 may have associations with a wide range of autoimmune diseases. rs2841280, a missense variant in PLD4 was suggested as a candidate causative variant of SSc[29]. This SNP is in LD with rs2582511 (top marker in this study) and showed a significant association in the current study ($r^2=0.71$, $p=3.3 \times 10^{-9}$). These findings suggest that functional alteration of PLD4 due to conformational change with or without change of PLD4 expression would lead to immunological abnormalities. A previous study using *Pld4*KO mice showed that *Pld4* was expressed mainly in the marginal zone of spleen[23]. The same group also showed that *Pld4* is implicated in the proliferation and phagocytosis function of microglia in CNS[30]. Collectively, these findings strongly indicate functional involvement of PLD4 in the development of autoimmunity through controlling functions of immune cells.

Based on the significant intra-case association between *PLD4* risk allele and the production of anti-dsDNA antibody, the risk allele seems to contribute to SLE pathophysiology through activating B cells. In spite of the lack of significant associations, the trend of the positive congruence between other antibodies and the PLD4 risk allele may support the hypothesis.

The *Pld4* mutant mice revealed remarkably low birth weight and autoimmune phenotypes compatible with SLE, including splenomegaly, lymphadenopathy, expansion of B cells, hypersecretion of BAFF and production of autoantibodies especially anti-nuclear antibody and anti-dsDNA antibody. In spite of abnormal findings of renal tissues in the mutant mice, they did not show proteinuria and short lives, suggesting that another genetic hit might be necessary to lead to lupus nephritis and renal phenotypes compatible with SLE.

Premature stop codon of *Pld4* in the mutant would lead to truncated PLD4 protein. Since all the cells in the mutant mice harbor this gene mutation, our study could not clarify whether the gene mutation has direct effect on specific immune cell types, or has resulted in indirect phenotypes due to mutations of non-hematopoietic cells. Despite the above limitation, taking into consideration that peripheral B cell subpopulations, but not B cell precursors within the bone marrow (data not shown) were altered in mutant mice, *Pld4* seems to affect peripheral maturation, or survival of B cells. Lineage-specific ablation of *Pld4* or induction of the mutation would clarify cellular designation to phenotypes in mutant mice.

We have not elucidated the detailed mechanism into how the truncated Pld4 protein or lack of normal Pld4 itself could lead to low body weight and immune dysfunction especially for acquired immunity. The latest report revealed Pld4-deficient mice demonstrated exaggerated TLR9 response and spontaneous inflammation[31]. This dysfunction of immune system may at least partly explain phenotypes found in the current study.

Mammalian *MAMLD1* encodes a mastermind-like domain containing protein. Human X-linked hypospadias type 2 is known to be caused by a loss-of-function mutation in this gene[24]. Due to lack of evidence of eQTL association and amino acid alteration of rs143181706 on *MAMLD1*, we could not infer how the variant and/or gene would contribute to SLE susceptibility. Since *MAMLD1* is associated with sex development, *MAMLD1* association might be related to high frequency of females in lupus patients. The positive associations between *MAMLD1* risk allele and autoantibody production may suggest *MAMLD1* function on B cells. Further in-vivo analysis would clarify these points.

Taken together, we identified a novel susceptibility locus to SLE and conducted in-vivo functional study of PLD4. These genes and encoding proteins may serve as candidates

of cellular and molecular targets in SLE. Further increase of sample sizes in the study subjects for GWAS would uncover genetic susceptibility loci to SLE. It would be interesting to develop mutant mice carrying multiple mutations of SLE-susceptibility genes.

Acknowledgments and affiliations: We would like to thank all the medical and technical staffs to help us to collect DNA samples. This study was supported by JSPS KAKENHI Grant Number JP16H06251 and 16K09891, Nagao Memorial Fund and THE KATO MEMORIAL TRUST FOR NAMBYO RESEARCH.

COMPETING INTEREST: none declared

Author contributions:

Wrote the paper:S.A, S.M.L, C.T

Data analysis:S.A, K.I, C.T

Performed the mouse experimental work:S.A, Y.Kochi, S.M.L, A.S, M.N, Y.I, H.K, O.O, J.H, C.T

Conceived and designed the study:C.T

Substantial contribution to acquired samples and creation of data in 1st GWAS:K.M, K.O, F.M, Y.O, T.M, C.T

Substantial contribution to acquired samples and creation of data in 2nd GWAS:K.I, Y.Kochi, Y.Kamatani, T.S, K.Y, Y.O

All authors revised and approved the manuscript to be published.

Reference

1. Terao C, Yamada R, Mimori T, et al. A nationwide study of SLE in Japanese identified subgroups of patients with clear signs patterns and associations between signs and age or sex. *Lupus* 2014;**23**:1435-1442.
2. Bengtsson AA, Ronnblom L. Role of interferons in SLE. *Best Pract Res Clin Rheumatol* 2017;**31**:415-428.
3. Chasset F, Arnaud L. Targeting interferons and their pathways in systemic lupus erythematosus. *Autoimmunity reviews* 2018;**17**:44-52.
4. Vincent FB, Morand EF, Schneider P, et al. The BAFF/APRIL system in SLE

pathogenesis. *Nature reviews Rheumatology* 2014;**10**:365-373.

5. Terao C, Ikari K, Nakayamada S, et al. A twin study of rheumatoid arthritis in the Japanese population. *Mod Rheumatol* 2016;**26**:685-689.
6. Kuo CF, Grainge MJ, Valdes AM, et al. Familial Aggregation of Systemic Lupus Erythematosus and Coaggregation of Autoimmune Diseases in Affected Families. *JAMA Intern Med* 2015;**175**:1518-1526.
7. Okada Y, Shimane K, Kochi Y, et al. A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus erythematosus in Japanese. *PLoS genetics* 2012;**8**:e1002455.
8. Remmers EF, Plenge RM, Lee AT, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med* 2007;**357**:977-986.
9. Sun C, Molineros JE, Looger LL, et al. High-density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry. *Nat Genet* 2016;**48**:323-330.
10. Morris DL, Sheng Y, Zhang Y, et al. Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. *Nat Genet* 2016;**48**:940-946.
11. Bentham J, Morris DL, Graham DSC, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat Genet* 2015;**47**:1457-1464.
12. Yang J, Yang W, Hirankarn N, et al. ELF1 is associated with systemic lupus erythematosus in Asian populations. *Hum Mol Genet* 2011;**20**:601-607.
13. Manjarrez-Orduno N, Marasco E, Chung SA, et al. CSK regulatory polymorphism is associated with systemic lupus erythematosus and influences B-cell signaling and activation. *Nat Genet* 2012;**44**:1227-1230.
14. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;**40**:1725.
15. Finucane HK, Bulik-Sullivan B, Gusev A, et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* 2015;**47**:1228-1235.
16. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012;**40**:D930-934.
17. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559-575.
18. Terao C, Kawaguchi T, Dieude P, et al. Transethnic meta-analysis identifies GSDMA

- and PRDM1 as susceptibility genes to systemic sclerosis. *Ann Rheum Dis* 2017;**76**:1150-1158.
19. Lamparter D, Marbach D, Rueedi R, et al. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLoS Comput Biol* 2016;**12**:e1004714.
 20. B. H, PF. W-B, DE. B, et al. Thin hair with small size (thss) is a new recessive hair mutation on Chromosome 12. *MGI Direct Data Submission* 2011.
 21. Wen L, Zhu C, Zhu Z, et al. Exome-wide association study identifies four novel loci for systemic lupus erythematosus in Han Chinese population. *Ann Rheum Dis* 2018;**77**:417.
 22. Saito M, Fujiwara Y, Asao T, et al. The genomic and epigenomic landscape in thymic carcinoma. *Carcinogenesis* 2017;**38**:1084-1091.
 23. Yoshikawa F, Banno Y, Otani Y, et al. Phospholipase D family member 4, a transmembrane glycoprotein with no phospholipase D activity, expression in spleen and early postnatal microglia. *PLoS One* 2010;**5**:e13932.
 24. Ogata T, Sano S, Nagata E, et al. MAMLD1 and 46,XY disorders of sex development. *Semin Reprod Med* 2012;**30**:410-416.
 25. Ishigaki K, Kochi Y, Suzuki A, et al. Polygenic burdens on cell-specific pathways underlie the risk of rheumatoid arthritis. *Nat Genet* 2017;**49**:1120-1125.
 26. Trivedi P, Kumar RK, Iyer A, et al. Targeting Phospholipase D4 Attenuates Kidney Fibrosis. *J Am Soc Nephrol* 2017;**28**:3579-3589.
 27. ImmGen C. Open-source ImmGen: mononuclear phagocytes. *Nat Immunol* 2016;**17**:741.
 28. Okada Y, Terao C, Ikari K, et al. Meta-analysis identifies nine new loci associated with rheumatoid arthritis in the Japanese population. *Nat Genet* 2012;**44**:511-516.
 29. Terao C, Ohmura K, Kawaguchi Y, et al. PLD4 as a novel susceptibility gene for systemic sclerosis in a Japanese population. *Arthritis Rheum* 2013;**65**:472-480.
 30. Otani Y, Yamaguchi Y, Sato Y, et al. PLD4 is involved in phagocytosis of microglia: expression and localization changes of PLD4 are correlated with activation state of microglia. *PLoS One* 2011;**6**:e27544.
 31. Gavin AL, Huang D, Huber C, et al. PLD3 and PLD4 are single-stranded acid exonucleases that regulate endosomal nucleic-acid sensing. *Nat Immunol* 2018;**19**:942-953.

Figure legends.

Figure1. Manhattan plot of the meta-analysis.

A) Quantile-quantile plot of the meta-analysis is indicated

B) P-values are shown according to their chromosomal positions. The two loci we focused in detail are shown in red.

Figure 2. PLD4 and MAMLD1 regions identified in the current study. Detailed plots surrounding the significant markers are shown for A) PLD4 and B) MAMLD1 regions.

Figure 3. Pld4 mutant mice revealed remarkable splenomegaly and lymphadenopathy in spite of low weight

- A. Mutation in Pld4 mutant mice. Sanger sequencing results of w/w, w/m and m/m mice are indicated.
- B. Pld4 mutant mice revealed low weight compared to heterozygous mice during the observed time period. Body size of both mice in the same scale is indicated in the upper panel.
- C. Splenomegaly and lymphadenopathy in Pld4 mutant mice. Spleen, inguinal lymphnode (iLN) and mesenteric lymph node (mLN) are compared between m/w and m/m mouse at 15 weeks after birth.
- D. Increased number of splenocyte in Pld4 mutant mice. Cell numbers of splenocyte are compared between w/w, w/m and m/m mice.
- E. Spontaneous germinal center formation in the spleen of Pld4 mutant mice. Representative microscopic images (HE) of spleen of 15 weeks m/m and w/m mice (N=6). Arrow heads indicate enhanced germinal center formation.

* indicates p-values less than 0.01

Figure 4. Pld4 mutant mice demonstrate autoimmune phenotypes including production of anti-DNA antibody.

- A. Pld4 mutant mice reveal increased T cells with expression of cell surface molecule of activated phenotype.
- B. Pld4 mutant mice reveal increased follicular T-cell (FOL), marginal zone T cell (MZ), macrophage and plasmacytoid dendritic cell (pDC). cDC: conventional dendritic cell. Cell numbers of FOL, MZ (left panel), macrophages (middle panel) and pDC and cDC are compared between mutant mice and heterozygote mice.
- C. Increased B cells and hypergammaglobulinemia and serum BAFF levels in Pld4 mutant mice. B cells in the spleen were analyzed by flowcytometry to compare the numbers of germinal center B cells (CD19⁺ CD95⁺ GL-7⁺), and plasma cells

(B220dim CD138+). Serum concentration of various subclasses of immuno-globulin (14-15 weeks old age) was measured (N=5-9). Circulating BAFF levels were compared between m/w and m/m mice.

- D. Autoantibody production in Pld4 mutant mice. Higher levels of anti-dsDNA antibody (left panel) and anti-nuclear antibody (right panel) are observed in Pld4 mutant mice. MEF cells were used for ANA quantification.
- E. PAS-positive deposition and accumulated IgG and C3 were found in glomeruli in Pld4 mutant mice (upper panel). Interferon-signature genes were highly expressed in the kidney of Pld4 mutant mice (left lower panel). No proteinuria was observed in Pld4 mutant mice (right lower panel).

Mutant mice are indicated in dark closed circles or bars. Heterozygote mice are indicated in white closed circles or bars. * indicates p-values less than 0.05, ** indicates p-values less than 0.01.

Table 1. Genetic associations with SLE susceptibility found in the current study.

					1 st set		2 nd set		Meta-analysis		
Ch				Alt freq		Alt freq					
SNP	r	position	Gene	Ref	Alt	Case/Cont	P	Case/Cont	P	P	OR (95%CI)
Significant loci											
rs2205959	1	173191371	TNFSF4	A	G	0.36/0.29	5.1x10 ⁻⁶	0.34/0.29	1.3x10 ⁻⁷	3.4x10 ⁻¹²	1.39 (1.27-1.52)
rs11889341	2	191943742	STAT4	C	T	0.4/0.3	1.3x10 ⁻⁸	0.39/0.29	6.4 x10 ⁻¹⁷	2.9x10 ⁻²⁴	1.59 (1.45-1.73)
rs796780915	6	32403397	HLA	A	AC	0.48/0.36	1.4x10 ⁻¹¹	0.53/0.41	2.9 x10 ⁻²⁰	1.3x10 ⁻²⁹	1.67 (1.52-1.82)
rs9494895	6	138234778	TNFAIP3	C	T	0.12/0.07	4.6x10 ⁻⁷	0.11/0.07	6.1 x10 ⁻⁸	6.0x10 ⁻¹⁴	1.74 (1.51-2.01)
rs11185602	7	50299077	IKZF1	A	G	0.33/0.44	9.3x10 ⁻¹⁰	0.38/0.44	2.0 x10 ⁻⁷	7.6x10 ⁻¹⁵	0.7 (0.64-0.76)
rs77009341	7	75188681	HIP1	G	C	0.04/0.03	0.0034	0.07/0.04	6.2x10 ⁻¹⁰	3.8x10 ⁻¹²	2.01 (1.65-2.44)
rs2618480	8	11354107	BLK	C	T	0.22/0.29	1.5x10 ⁻⁵	0.22/0.29	1.2 x10 ⁻⁸	1.6x10 ⁻¹²	0.7 (0.63-0.77)
			PCDH15,W								
rs11101537	10	50076325	DFY4	G	T	0.46/0.39	4.3x10 ⁻⁵	0.46/0.41	1.9x10 ⁻⁵	4.2x10 ⁻⁹	1.3 (1.19-1.42)
rs12576753	11	128304141	ETS1	C	A	0.46/0.4	0.00036	0.48/0.39	8.7x10 ⁻¹⁰	1.8x10 ⁻¹²	1.36 (1.25-1.48)
			GPR19/CDK								
rs11055009	12	12817965	N1B	T	G	0.13/0.17	0.00063	0.13/0.18	1.9x10 ⁻⁶	8.8x10 ⁻⁹	0.69 (0.61-0.79)
			AHNAK2/PL								
rs2582511	14	105416010	D4	T	C	0.30/0.38	2.1x10 ⁻⁶	0.33/0.38	5.6x10 ⁻⁶	7.9x10 ⁻¹¹	0.73 (0.66-0.8)
rs11117431	16	86015316	IRF8	A	G	0.07/0.1	0.0014	0.07/0.11	7.6x10 ⁻⁷	1.3x10 ⁻⁸	0.62 (0.52-0.73)
rs5998672	22	21966442	UBE2L3	G	A	0.57/0.46	5.2x10 ⁻⁹	0.52/0.48	0.0014	1.2x10 ⁻⁹	1.30 (1.20-1.42)

rs143181706	X	149673253	MAMLD1	C	T	0.15/0.11	6.2x10 ⁻⁵	0.15/0.12	2.7x10 ⁻⁵	3.7x10 ⁻⁸	1.50 (1.30-1.74)
Suggestive loci											
rs78382932	4	79831096	BMP2K	T	G	0.32/0.25	9.2x10 ⁻⁶	0.3/0.26	0.0032	4.4x10 ⁻⁷	1.28 (1.16-1.4)
rs79566524	5	40461731	PTGER4	G	GA	0.28/0.24	0.013	0.28/0.23	1.5x10 ⁻⁵	7.4x10 ⁻⁷	1.28 (1.16-1.41)
rs526531	6	106570056	PRDM1	G	A	0.37/0.33	0.011	0.4/0.34	2.6x10 ⁻⁵	8.6x10 ⁻⁷	1.25 (1.14-1.36)
rs3735718	8	102704965	NCALD	C	T	0.22/0.15	1.4x10 ⁻⁶	0.18/0.16	0.020	1.0x10 ⁻⁶	1.33 (1.18-1.49)
rs80022860	13	41568872	ELF1	G	A	0.29/0.25	4.5x10 ⁻³	0.3/0.24	7.2x10 ⁻⁶	9.1x10 ⁻⁸	1.29 (1.18-1.42)

Ref: reference allele, Alt: alternative allele, OR: odds ratio, CI: confidence interval, Alt freq: alternative allele frequency, Cont: control

Table 2. A significant association between rs2582511 and anti-dsDNA antibody production in intra-case analyses.

Phenotype	Number of positive subjects*	rs2582511		rs143181706	
		OR (95%CI)	P	OR (95%CI)	P
anti-dsDNA	245 (59.3%)	1.52 (1.12-2.06)	0.0073	1.05 (0.71-1.54)	0.81
anti-RNP	245 (57.6%)	1.23 (0.92-1.66)	0.16	1.12 (0.77-1.63)	0.56
anti-SSA	280 (65.1%)	1.11 (0.82-1.51)	0.49	1.07 (0.72-1.57)	0.74
anti-SSB	106 (32.0%)	1.32 (0.92-1.9)	0.14	1.13 (0.72-1.78)	0.59
		Permutation test		P	
		Smallest p-values of the eight association studies		0.0475	
		Positive associations of all of the eight association studies		0.0060	

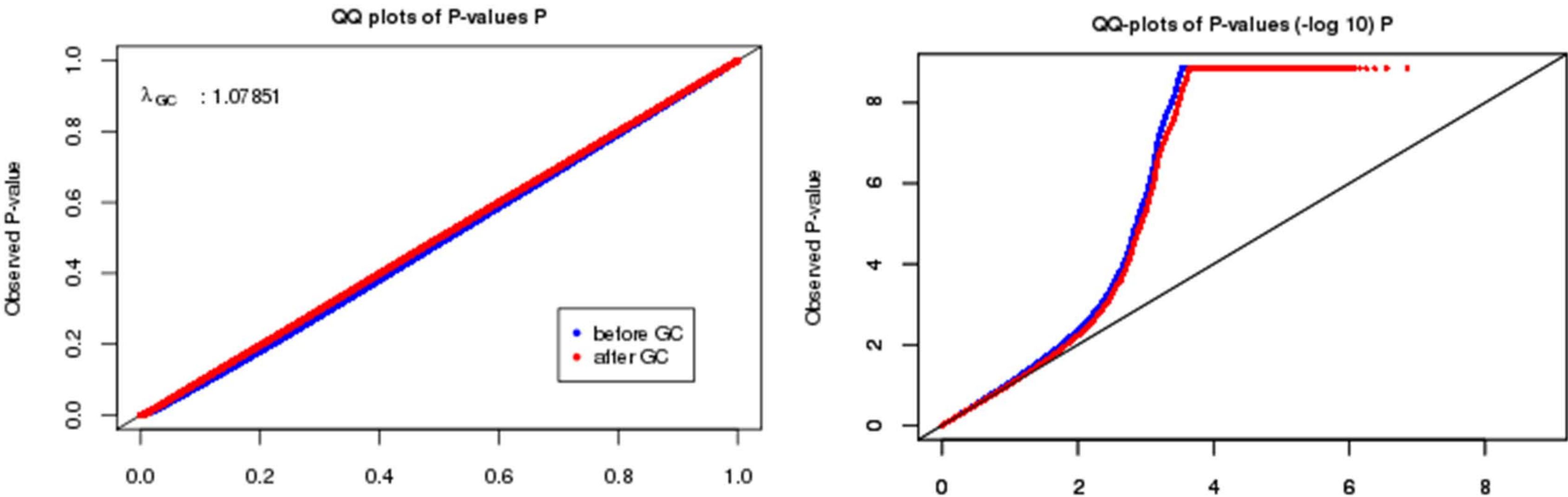
OR: odds ratio of risk allele for SLE, CI: confidence interval *subjects positive for antibodies in the set 1. Subjects with information of the autoantibodies set as a denominator.

Key Messages

- We conducted a meta-analysis of GWAS for SLE in Japanese and identified 14 significant loci including a novel locus.
- PLD4, a susceptibility gene to SLE very recently reported by a Chinese group, showed a significant association.
- The PLD4 risk variant was associated with autoantibody positivity in intra-case analyses.
- PLD4 mutant mice showed autoimmune phenotypes including splenomegaly and autoantibody production suggesting dysregulation of B cells.

Figure.1

A



B

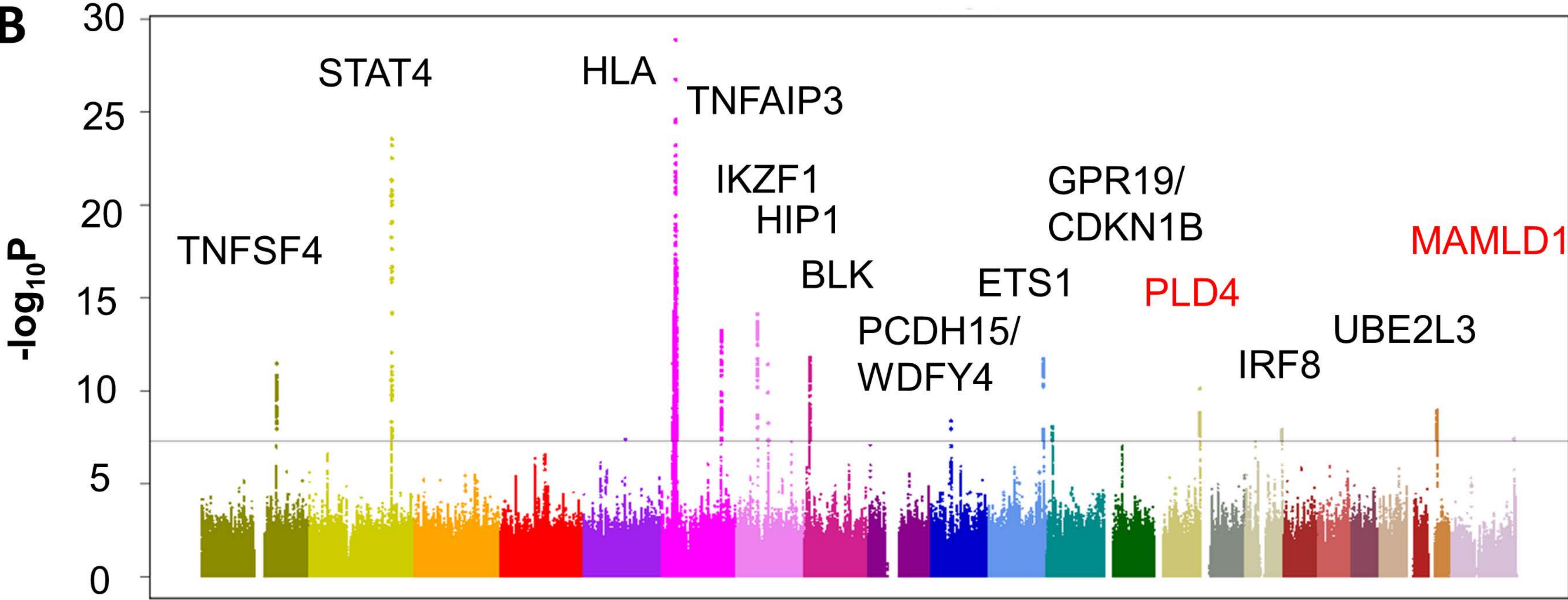


Figure.2

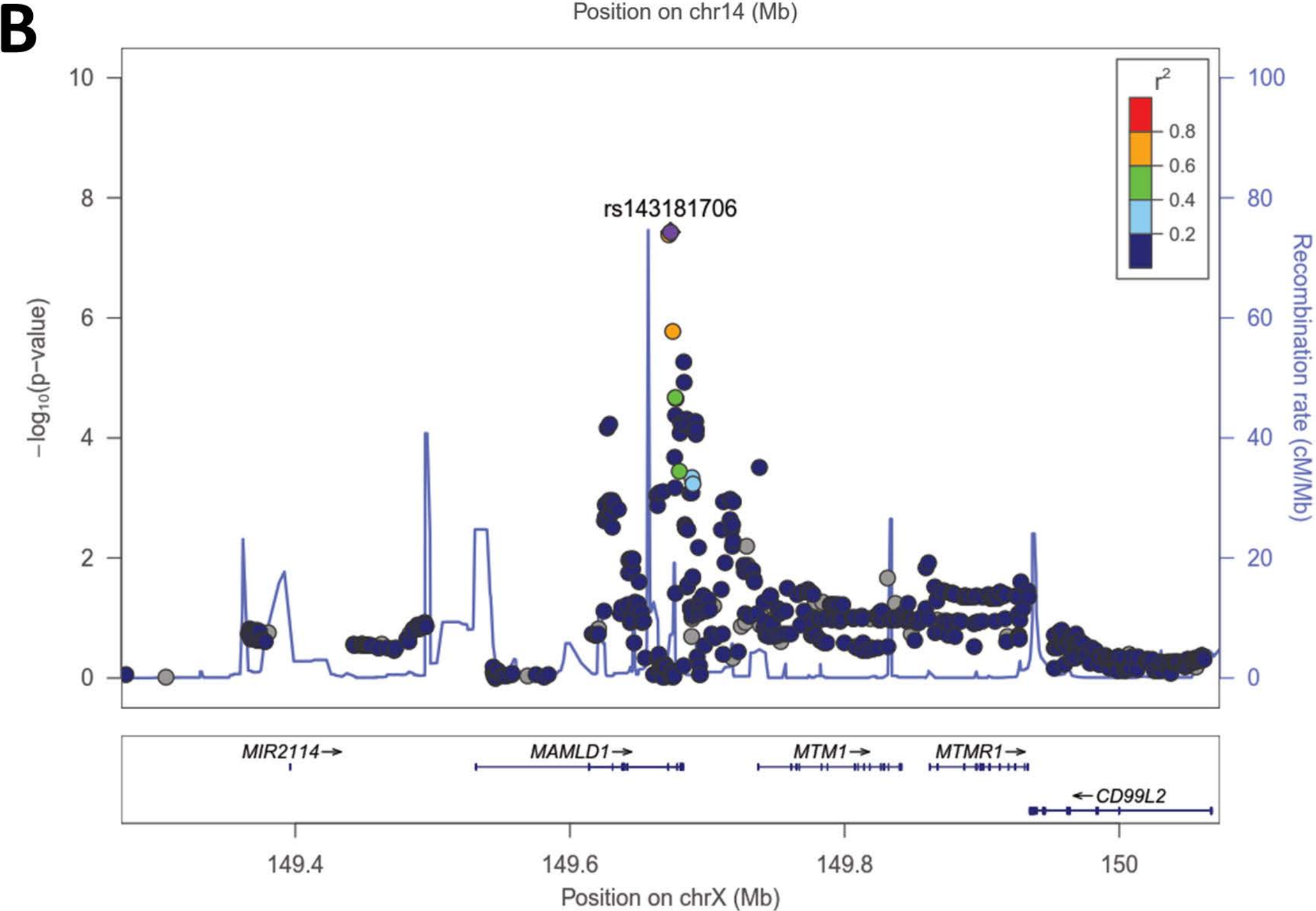
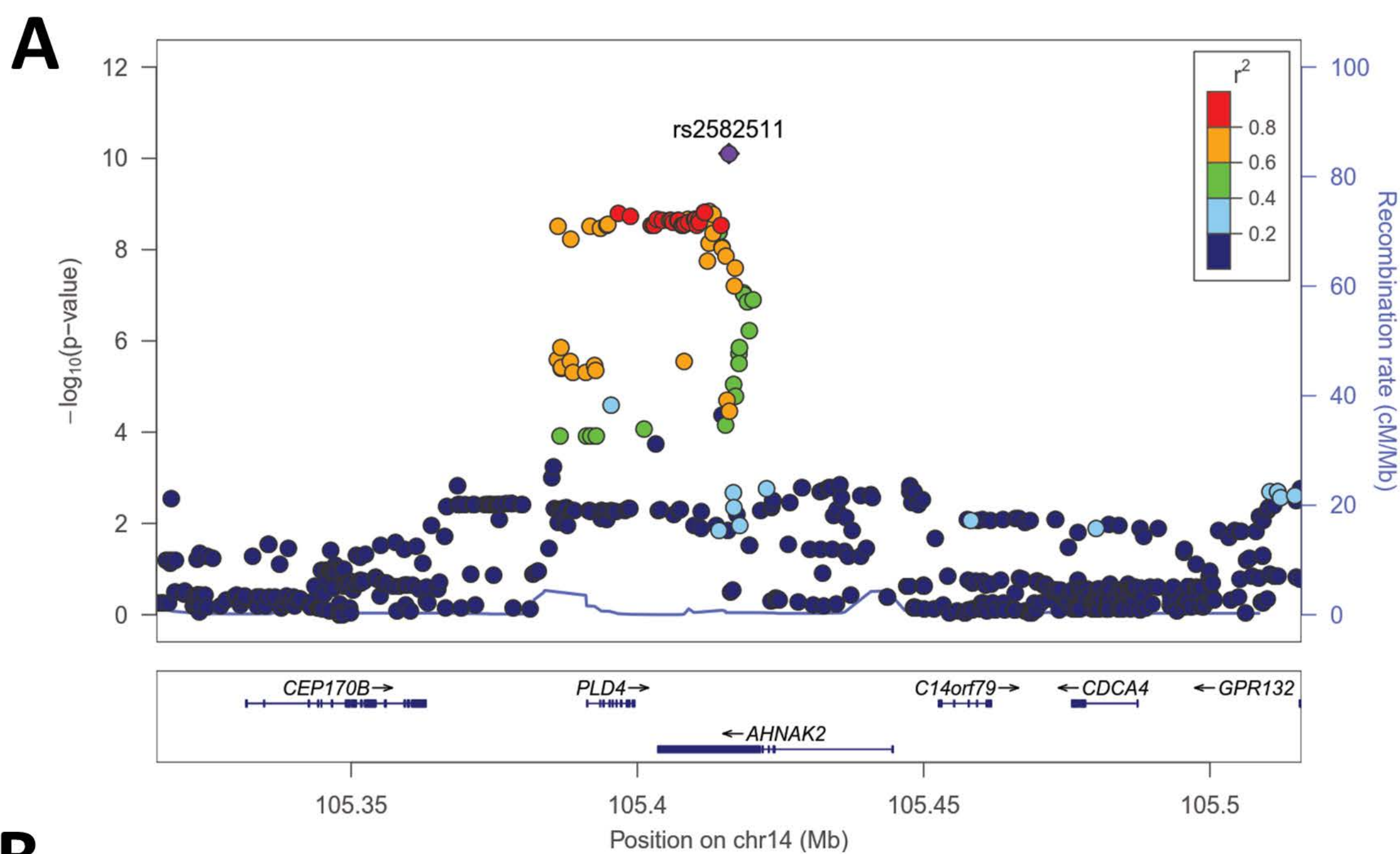


Figure.3

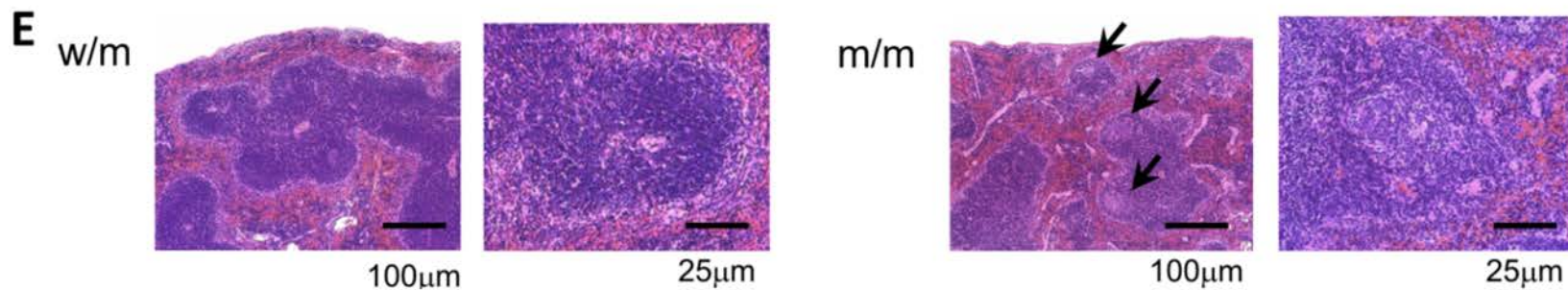
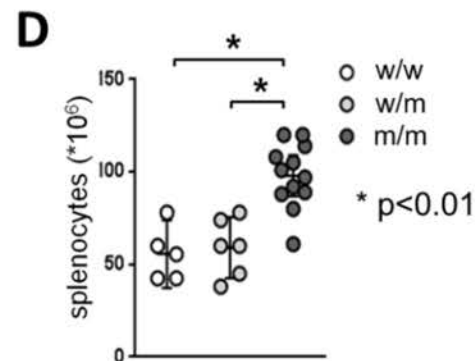
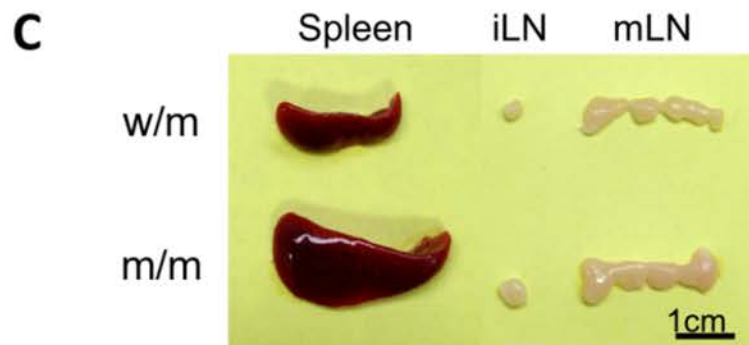
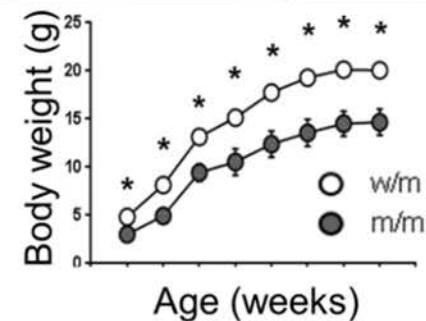
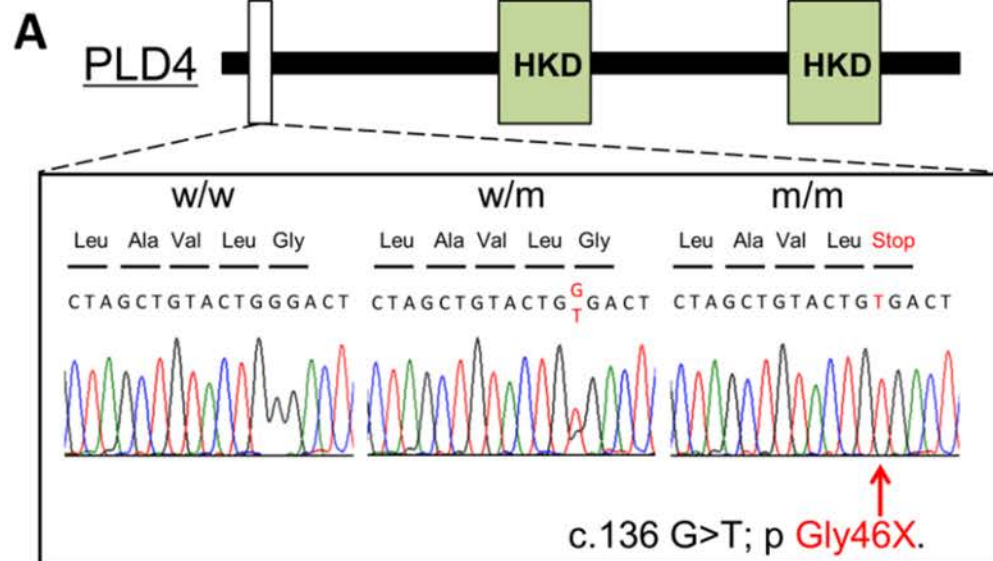


Figure.4

