An Intronic HCP5 Variant Is Associated With Age of Onset and Susceptibility to Graves Disease in UK and Polish Cohorts

Laura Claire Lane,1,2,3 Aleksander Kuś,4 Tomasz Bednarczuk,4 Artur Bossowski,5 Jacek Daroszewski,6 Beata Jurecka-Lubieniecka,7 Heather Jane Cordell,8 Simon Henry Schofield Pearce,1,2 Timothy Cheetham1,3 and Anna Louise Mitchell1,2

1Translational and Clinical Research Institute, Newcastle University, Central Parkway, Newcastle-upon-Tyne, NE1 3BZ, UK; 2Endocrine Unit, Royal Victoria Infirmary, Queen Victoria Road, Newcastle-upon-Tyne, NE1 4LP, UK; 3Department of Paediatric Endocrinology, The Great North Children's Hospital, Queen Victoria Road, Newcastle-upon-Tyne, NE1 4LP, UK; 4Department of Internal Medicine and Endocrinology, Medical University of Warsaw, 02-091 Warsaw, Poland; 5Department of Pediatrics, Endocrinology and Diabetes with a Cardiology Unit, Medical University of Białystok, 15-089 Białystok, Poland; 6Department of Endocrinology, Diabetes and Isotope Therapy, Wroclaw Medical University, 50-367 Wroclaw, Poland; 7Department of Nuclear Medicine and Endocrine Oncology, Maria Sklodowska-Curie Institute - Oncology Center, Giwice Branch, 44-102 Giwice, Poland; and 8Population Health Sciences Institute, Newcastle University, Baddiley-Clark Building, Newcastle-upon-Tyne, NE2 4AX, UK

ORCiD number: 0000-0002-6630-2701 (L. Laura Claire).

Context: The genetic background of young-onset Graves disease (GD) remains largely unknown. An intronic variant in human leukocyte antigen (HLA) complex P5 (HCP5) has previously been associated with GD susceptibility and age of onset in a cohort of Polish patients.

Objective: We aimed to investigate the association of the HCP5 variant rs3094228 with GD susceptibility and age of onset in a UK cohort and conduct a meta-analysis of UK and Polish data.

Design and Participants: rs3094228 was genotyped in 469 UK patients with GD using Taqman chemistry. Genotype frequencies were compared with genotypic data available from the Wellcome Trust case-control consortium using logistic regression analysis. To determine whether rs3094228 is independently associated with age of GD onset, the HLA DRB1*0301 tagging variant, rs535777, was also genotyped.

Results: The C allele of rs3094228 was overrepresented in the UK GD cohort compared with controls (P<0.0001; odds ratio 1.76; 95% confidence interval, 1.46-2.13). This association was more marked in young-onset GD (<30 years) (P<0.0001). The meta-analysis of UK and Polish data supported the association of the C allele with GD susceptibility and age of onset (P=5.63×10^-8). Haplotype analysis demonstrated that rs3094228 is associated with age of GD onset (P=2.39×10^-6) independent of linkage disequilibrium with HLA DRB1*0301.

Conclusion: The rs3094228 HCP5 polymorphism is independently associated with GD susceptibility and age of onset in a UK GD cohort. Our findings indicate a potential role of long noncoding ribonucleic acids, including HCP5, in GD pathogenesis, particularly in the younger population. (J Clin Endocrinol Metab 105: 1–8, 2020)

Key Words: thyroid, autoimmune, polymorphism, genotyping, Graves disease, meta-analysis

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA
© Endocrine Society 2020. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Received 12 December 2019. Accepted 1 June 2020.
First Published Online 5 June 2020.
Corrected and Typeset 18 July 2020.

Graves disease (GD) is the most common cause of hyperthyroidism in young children and adolescents; however, it is still relatively rare, with a reported incidence of between 0.1 and 3 per 100,000 (1) compared with Caucasian European adults where the incidence is reported as 20 to 25 cases per 100,000 (2, 3). Longitudinal studies have reported an increasing incidence of hyperthyroidism in both adult and pediatric populations (4-6). Similar to other autoimmune conditions, there is a clear female preponderance, with GD affecting up to 3% of women and 0.5% of men, with a peak incidence occurring between 30 and 50 years of age (2).

GD is characterized by the presence of thyroid receptor autoantibodies (TRAbs) that stimulate the cell-surface thyroid-stimulating hormone (TSH) receptor, directly resulting in excessive, autonomous thyroid hormone secretion. The clinical features and prognosis of GD is highly variable depending on age of disease onset, with the remission rate following a course of antithyroid medication being as low as 25% in the pediatric population compared with 50% to 60% in adults (7, 8).

GD results from a complex interaction between genetic and environmental factors involving variants in multiple susceptibility genes, each exerting modest individual effects. Family and twin studies over the past 50 years have demonstrated that up to 80% of an individual’s predisposition to GD is attributable to genetic factors (9, 10). However, only around 20% of this genetic load has been accounted for by the currently associated genomic variants (11).

Genomic polymorphisms associated with susceptibility to GD are primarily found at immune-regulatory loci, such as MHC (12), CTLA-4 (13), PTPN22 (14), and CD40 (15). A stronger genetic association is suspected in the younger population who have had less exposure to environmental factors. Several of the known susceptibility loci are also associated with a younger age of disease onset, including those at CTLA-4 (16), human leukocyte (HLA)-DRB1 (17), and PTPN22 (18), with the most strongly associated variants located at the major histocompatibility complex (MHC) locus (19, 20). Determining genetic variants associated with GD can provide mechanistic insight by highlighting pathogenic functional pathways, particularly by studying the younger population where genetics may be the dominant factor (19).

This study aimed to investigate the association of the HLA complex P5 (HCP5) gene in GD susceptibility and age of onset in a UK cohort. The association between HCP5 variants and thyroid autoimmunity was first demonstrated in a multicenter population-based genome-wide association study conducted by Medici et al for serum levels of thyroid peroxidase antibodies (21). The first study showing an association of HCP5 with susceptibility to GD (rs3094228, $P = 1.6 \times 10^{-12}$; odds ratio [OR] = 1.88), was performed in a single center by Kus et al (22). A subsequent multicenter study with a relatively large pediatric GD cohort demonstrated the HCP5 variant, rs3094228, as a risk locus for young-onset GD (YOGD) (23).

We have studied the same HCP5 polymorphism in a UK GD cohort and performed a meta-analysis of data from the UK and Polish patient cohorts.

Materials and Methods

Participants

A total of 469 patients were included in the UK cohort, including 118 patients with YOGD (aged <30 years) and 351 patients with unrelated later-onset GD (LOGD) (aged ≥30 years). The YOGD cohort included 18 (15%) male and 100 (85%) female (GD onset aged 3-29 years; median 22 years, mean 20.8 years) and the LOGD cohort included 55 (16%) male and 296 (84%) female (GD onset aged 30-92 years; median 47 years, mean 48.2 years).

The patients providing these samples were of Caucasian European background and had attended outpatient endocrinology at the Royal Victoria Infirmary or the Great North Children’s Hospital, Newcastle-upon-Tyne, UK. Each participant with GD was diagnosed by the following criteria: fully suppressed serum TSH with serum free thyroxine and/or free triiodothyronine above the reference range and the existence of detectable TSH receptor antibody (TRAb; ≥1.8 mU/L; Brahms Kryptor).

Genotype data from 5377 control samples from the Wellcome Trust case-control consortium (WTCCC2) database were used for comparison. Informed, written consent was obtained from all participants. This study was carried out with approval of the Leeds East (Ref. 05/Q1206/144) and Berkshire Valley ethics committees (Ref. 04/12/015).

HCP5 genotyping

The HCP5 variant rs3094228 was genotyped in genomic deoxyribonucleic acid extracted from venous blood using TaqMan chemistry as per the manufacturer's instructions (assay C_2995657_10) and run on the QuantStudio 7 Flex Real-Time PCR (polymerase chain reaction) System (Applied Biosystems). Twenty percent of the samples were genotyped in duplicate to ensure assay fidelity. The overall genotyping call rate was 99.8%.

HLA genotyping

The HLA DRB1*0301 tagging variant, rs535777, was genotyped in the UK cohort as above (assay C_26546461_30). The overall genotyping call rate was 99.6%.

In the Polish cohort, the HLA-DRB1 polymorphism was genotyped using the low-resolution single specific primer-polymerase chain reaction (SSP-PCR) method with use of the Dynal All Set SSP DR Kit or the HLA-Ready Gene DR Kit, as previously described (22).
Statistical analysis

Statistical association analysis was performed using PLINK (24) and SPSS version 25 (25). All the control sample genotypes were in Hardy-Weinberg equilibrium ($P > 0.4$). Study data were compared with WTCCC2 control data using logistic regression with sex as a covariate. A subgroup association analysis was performed comparing young-onset (aged <30 years) GD to older-onset (aged ≥ 30 years) GD. A meta-analysis, using the Review Manager (RevMan) Version 5.0 program (Nordic Cochrane Centre, Copenhagen, Denmark (26)) was then undertaken, using a random effects model to calculate ORs, 95% confidence intervals (CI) and 2-sided $P$ values. The impact of heterogeneity between the cohorts was estimated using an I² index. Kaplan-Meier plots and log-rank tests were applied to determine whether genotype was significantly associated with age of GD onset. Logistic regression and haplotype analysis (UNPHASED 3.1.7 (27)) was performed to determine the independent association of $rs3094228$ with age of GD onset.

Results

GD susceptibility

The minor C allele and the CC genotype at $rs3094228$ are associated with susceptibility to GD. The frequency of the C allele was significantly increased in the GD cohort as a whole (303/938; 32%) compared with WTCCC2 controls (2118/10,754; 20%; $P = 5.08 \times 10^{-9}$; OR 1.76 [95% CI, 1.46-2.13]). There was also a significant increase in the CC genotype in the GD group (39/469; 8%) compared with WTCCC2 controls (219/5377; 4%; $P = 2.89 \times 10^{-18}$) (Table 1).

GD age of onset

Although an increased frequency of the C allele was present in both the YOGD (99/236; 42%) and LOGD groups (204/702; 29%) compared with controls (2118/10,754; 20%), the difference was more significant in the YOGD group ($P = 1.70 \times 10^{-10}$; OR 2.73 [95% CI, 2.00-3.71]) compared with the LOGD cohort ($P = 0.0008$; OR 1.49 [95% CI, 1.18-1.88]). In addition, a significant increase in the frequency of the C allele (99/236; 42% vs 204/702; 29%; $P_{\text{allele}}=0.00025$; OR 1.76 [95% CI, 1.3-2.4]) and CC genotype (16/118; 13% vs 23/351; 7%; $P_{\text{genotype}}=0.00059$) was observed in the YOGD group when compared with the LOGD cohort (Table 2). This suggests that the C allele and CC genotype have a stronger association with susceptibility to GD at a younger age.

Meta-analysis

A meta-analysis was undertaken using additional genotype data provided by Kuś et al (22) from a study that examined genetic risk loci including $rs3094228$ in a Polish GD cohort. The YOGD Polish cohort (aged <30 years) included 66 (19%) male and 280 (81%) female (GD onset aged 3-29 years; median 18 years, Table 1. Logistic regression analysis with sex as a covariate. Genotype and allele frequencies at $rs3094228$ in the GD UK cohort and healthy Controls (WTCCC2), subdivided into age of disease onset (aged < 30 years / ≥ 30 years)

<table>
<thead>
<tr>
<th>$rs3094228$</th>
<th>WTCCC Controls (%)</th>
<th>All GD UK Cohort (%)</th>
<th>GD Aged &lt; 30 Years (%)</th>
<th>GD Aged ≥ 30 Years (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>219 (4)</td>
<td>39 (8)</td>
<td>16 (13)</td>
<td>23 (7)</td>
</tr>
<tr>
<td>CT</td>
<td>1680 (31)</td>
<td>225 (48)</td>
<td>67 (57)</td>
<td>158 (45)</td>
</tr>
<tr>
<td>TT</td>
<td>3478 (65)</td>
<td>205 (44)</td>
<td>35 (30)</td>
<td>170 (48)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>$2.89 \times 10^{-10}$</td>
<td>$1.90 \times 10^{-15}$</td>
<td>$6.97 \times 10^{-9}$</td>
<td></td>
</tr>
<tr>
<td>Allele frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>2118 (20)</td>
<td>303 (32)</td>
<td>99 (42)</td>
<td>204 (29)</td>
</tr>
<tr>
<td>T</td>
<td>8636 (80)</td>
<td>635 (68)</td>
<td>137 (58)</td>
<td>498 (71)</td>
</tr>
<tr>
<td>$P$ value (OR [95% CI])</td>
<td>$5.08 \times 10^{-9}$</td>
<td>$1.70 \times 10^{-10}$</td>
<td>0.0008</td>
<td></td>
</tr>
<tr>
<td>Abbreviations: CI, confidence interval; GD, Graves disease; OR, odds ratio; WTCCC, Wellcome Trust case-control consortium.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Association analysis of sex-matched young-onset (aged < 30 years) GD compared with later-onset (aged ≥ 30 years) GD

<table>
<thead>
<tr>
<th>$rs3094228$</th>
<th>GD Aged &lt; 30 Years (%)</th>
<th>GD Aged ≥ 30 Years (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype frequency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>16 (13)</td>
<td>23 (7)</td>
</tr>
<tr>
<td>CT</td>
<td>67 (57)</td>
<td>158 (45)</td>
</tr>
<tr>
<td>TT</td>
<td>35 (30)</td>
<td>170 (48)</td>
</tr>
<tr>
<td>$P$ value</td>
<td>0.00059</td>
<td></td>
</tr>
<tr>
<td>Allele frequency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>99 (42)</td>
<td>204 (29)</td>
</tr>
<tr>
<td>T</td>
<td>137 (58)</td>
<td>498 (71)</td>
</tr>
<tr>
<td>$P$ value (OR [95% CI])</td>
<td>0.00025</td>
<td></td>
</tr>
<tr>
<td>Abbreviations: CI, confidence interval; GD, Graves disease; OR, odds ratio.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Copyedited by: oup

Downloaded from https://academic.oup.com/jcem/article-abstract/105/9/dgaa347/5851730 by guest on 06 August 2020
mean 19 years), and the LOGD Polish cohort (aged ≥ 30 years) included 194 (22%) male and 672 (78%) female (GD onset aged 30-81 years; median 47 years, mean 48 years).

Using a random effects model, the association of the C allele with susceptibility to GD \( (P = 1.79 \times 10^{-5}; \text{OR} 1.71 [95\% \text{ CI}, 1.34-2.19]) \) and an earlier age of disease onset YOGD versus LOGD \( (P = 5.63 \times 10^{-8}; \text{OR} 1.56 [95\% \text{ CI}, 1.33-1.83]) \) was confirmed (Fig. 1A).

Using the combined UK and Polish cohorts, Kaplan-Meier plots and log-rank tests demonstrated a significant association with genotype and age of GD onset, confirming the association of the CC genotype with earlier-onset GD \( (P = 0.003) \) (Fig. 1B). In addition, a recessive model demonstrated an increased risk of earlier-onset GD in homozygotes for the minor allele compared with carriers for the common allele \( (P = 0.001) \). The median age of GD onset in those with the CC genotype was 34 years compared with 40 and 43 years with the CT and TT genotypes, respectively.

### Haplotype analysis

To determine whether the effect of rs3094228 on age of GD onset is independent of the common GD susceptibility locus, HLA DRB1*0301, a gender-adjusted logistic regression analysis on the Polish data was undertaken, including 439 patients with GD with available data on HLA DRB1*03, rs3094228, and age of GD onset. This demonstrated that the observed effect of rs3094228 on age of GD onset is independent from HLA DRB1*03 \( (P = 0.006) \).

This was further studied in the UK GD cohort using a known tagging single nucleotide polymorphism for HLA DRB1*0301 in the Northern and Western European (CEU) population \( (rs535777; r^2 = 0.87, D' = 0.99) \) (28). Consistent with the Polish data, a gender-adjusted
logistic regression analysis on the UK cohort demonstrated that the effect of rs3094228 on age of GD onset occurs independently of HLA DRB1*0301 (P = 0.004). A combined logistic regression analysis including the Polish and UK cohorts (906 patients) demonstrated that the interaction of HCP5 and HLA DRB1*0301 was significantly associated with age of GD onset (P = 0.046) (Table 3). Haplotype analysis with HCP5 as the test marker and HLA DRB1*0301 as the conditioning marker demonstrated that HCP5 is independently associated with age of GD onset (P = 2.39 × 10⁻⁶).

Table 3. Logistic regression analysis of rs3094228 and HLA DRB1*03, using the combined UK and Polish cohorts

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>B</th>
<th>SE</th>
<th>P Value</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3094228</td>
<td>0.330</td>
<td>0.151</td>
<td>0.028</td>
<td>1.39 (1.04-1.87)</td>
</tr>
<tr>
<td>HLA DRB1*03 / rs355777</td>
<td>-0.497</td>
<td>0.200</td>
<td>0.013</td>
<td>0.61 (0.41-0.90)</td>
</tr>
<tr>
<td>rs3094228 +</td>
<td>0.389</td>
<td>0.195</td>
<td>0.046</td>
<td>1.48 (1.01-2.17)</td>
</tr>
</tbody>
</table>

rs355777 was used as tagging variant for the HLA DRB1*0301 locus in the UK cohort.
Abbreviations: B, unstandardized regression coefficient; CI, confidence interval; OR, odds ratio; SE, standard error.

Discussion

This meta-analysis includes more than 1600 participants with GD, including more than 460 early-onset cases, from UK and Polish populations. Given the relative rarity of GD in the pediatric population, combining genotype data sets is essential to improve study power. Despite the established strong hereditary component of GD, particularly in the development of GD at a younger age, there is limited data documenting genetic risk variants specifically accounting for YOGD (19, 29, 30). This study finds a robust association between susceptibility to GD and an earlier age of disease onset with an intronic HCP5 polymorphism in a UK cohort, replicating previous findings in a Polish population (23).

The HCP5 ribonucleic acid (RNA) gene is located within the MHC class I region, centromeric of the HLA-B gene between the MICA and MICB genes, and it encodes a long noncoding RNA (lncRNA) (31). MHC genes encode cell-surface antigen-presenting proteins that are essential to mount an autoimmune response. Genome-wide association studies have identified polymorphisms at the MHC locus as risk variants for many autoimmune and inflammatory diseases, including GD in both Asian and European populations (32, 33). Furthermore, Immunochip genetic analysis has demonstrated that variants within the MHC have the strongest association with YOGD (aged <30 years) (19). The role of the MHC class II region is well established in GD, where antigen presentation of thyroid-stimulating hormone receptor to CD4+ T cells is crucial to drive the B cells to produce the pathogenic TRAb autoantibodies (34). However, genetic variants in the MHC I region (HLA-B and HLA-C) have also previously been independently associated with GD (35). Indeed, messenger RNAseq analysis of GD thyroid tissue found that HLA-C in the MHC I region was the most overexpressed gene compared with controls (36). The primary function of MHC I molecules is to present nonself antigens derived from intracellular sources, such as viruses, to CD8+ cytotoxic T cells, which then mount a cytotoxic response against the presented antigen. It has been proposed that viruses may trigger an autoimmune response through molecular mimicry where viral antigens are structurally similar to self-antigens, or bystander activation where the viral infection triggers a nonspecific activation of autoreactive cells (37).

Various HCP5 variants are demonstrated to promote susceptibility to adverse immune-related cutaneous drug reactions, such as Stevens-Johnson syndrome, as well as being associated with disease progression and viral load in untreated patients with HIV, suggesting a possible specific role for HCP5 in modifying the immune response to medications or viral infections (38, 39). Certain genetic variants within the MHC I region, such as HCP5 polymorphisms, may predispose to GD by modifying an individual’s response to an infectious agent, resulting in an excessive immune response with the potential to initiate an autoimmune reaction in genetically susceptible individuals.

The HCP5 polymorphism rs3094228 investigated in this study has previously been associated with susceptibility to GD in a Polish cohort (22), and an association has been demonstrated between the C allele and thyroid peroxidase antibody levels in autoimmune thyroid disease (21). Other HCP5 variants have been associated with susceptibility and autoantibody production in various autoimmune disorders, including systemic lupus erythematosus (40), Sjögren syndrome (41), psoriasis, and psoriatic arthritis (42).
The MHC gene region is highly polymorphic and characterized by extended linkage disequilibrium (LD), making it challenging to determine functional variants from tagging single nucleotide polymorphisms, which highlights the importance of rigorous case-control matching in genetic association studies. LD analysis indicates that the known GD risk locus, HLA DRB1*03, and rs3099844 studied in systemic lupus erythematosus and Sjögren syndrome, are in partial LD with rs3094228 (HLA DRB1*03: $r^2 = 0.45$, $D' = 0.86$ in the Polish population (22), rs3099844: $r^2 = 0.69$, $D' = 0.93$ in the British population (28)). However, our haplotype analysis with HLA DRB1*03 demonstrates that HCP5 is independently associated with age of GD onset.

Interestingly, there is accumulating evidence suggesting that lncRNA, such as HCP5, has a crucial role in the development of autoimmunity by altering the adaptive and innate immune response through transcriptional and epigenetic regulation (32, 43-45). Studies have demonstrated that lncRNAs are associated specifically with autoimmune thyroid disease (AITD), including GD. Indeed, it has been proposed that variation and dysregulation of lncRNAs in the MHC region is highly polymorphic and characterized by extended linkage disequilibrium (LD), which highlights the importance of rigorous case-control matching in genetic association studies. LD analysis indicates that the known GD risk locus, HLA DRB1*03, and rs3099844 studied in systemic lupus erythematosus and Sjögren syndrome, are in partial LD with rs3094228 (HLA DRB1*03: $r^2 = 0.45$, $D' = 0.86$ in the Polish population (22), rs3099844: $r^2 = 0.69$, $D' = 0.93$ in the British population (28)). However, our haplotype analysis with HLA DRB1*03 demonstrates that HCP5 is independently associated with age of GD onset. The aging immune system may contribute to the phenotype differences observed between YOGD and LOGD, where involution of the thymus and reduced B- and T-lymphocyte production may alter the mechanisms driving the autoimmune response in the older population (49). Phenotypic differences may also be explained by genetic variation including those in the MHC region (HLA subtypes DB1*02, DQA1*05, and DRB1*03), which have been associated with a higher risk of relapse in GD (50).

As this study was performed in a Caucasian population, further studies are required to investigate whether a similar effect is also detected in other (non-Caucasian) populations. Further functional studies should also aim to elucidate the underlying mechanism behind the observed association.

**Conclusion**

This study has confirmed a significant association of the HCP5 polymorphism, rs3094228, with GD susceptibility and age of disease onset in a UK cohort and replicates the findings from a study of patients with GD in Poland. Adult-onset and young-onset GD share multiple common genetic risk variants, many of which remain unknown. Our findings indicate a potential role for HCP5 as a contributor to GD susceptibility, particularly in the younger population. Further research to determine the role of lncRNAs, including HCP5, in the pathogenesis of early-onset GD is now warranted.

**Acknowledgments**

**Financial Support:** This work was supported by the Medical Research Council (MRC) (Grant number MR/S001611/1) and the National Science Center, Poland (Grant number 2014/15/N/NZ5/01656).

**Additional Information**

**Correspondence and Reprint Requests:** Dr Laura C. Lane, Translational and Clinical Research Institute, Central Parkway, Newcastle upon Tyne, NE1 3BZ. E-mail: Laura.Lane@newcastle.ac.uk.

**Disclosure Summary:** I certify that neither I nor my coauthors have a conflict of interest as described above that is relevant to the subject matter or materials included in this work.

**Data Availability:** All data generated or analyzed during this study are included in this published article or in the data repositories listed in References.

**References**


44. Shirasawa S, Harada H, Furugaki K, et al. SNPs in the promoter of a B cell-specific antisense transcript, SAS-ZFAT, determine