Refined association of TSH receptor susceptibility locus to Graves’ disease in the Chinese Han population

Bing-Li Liu*, Shao-Ying Yang*, Wei Liu, Li-Qiong Xue, Xia Chen1, Chun-Ming Pan, Zhao-Hui Gu, Ming Zhan, Xiao-Mei Zhang2, Jun Liang3, Guan-Qi Gao4, Wen-Hua Du4, Guo-Yue Yuan5, Ru Ying6, Shuang-Xia Zhao and Huai-Dong Song

State Key Laboratory of Medical Genomics, Shanghai Institute of Endocrinology and Metabolism, Molecular Medicine Center, Ruijin Hospital Affiliated to Shanghai Jiaotong University (SJTU) School of Medicine, Shanghai 200025, China, 1Department of Geratology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu 210029, China, 2Department of Endocrinology, The First Hospital Affiliated to Bengbu Medical College, 287 Changhuai Road, Bengbu, Anhui Province 233004, China, 3Department of Endocrinology, The Central Hospital of Xuzhou Affiliated to Xuzhou Medical College, Xuzhou, Jiangsu Province 221009, China, 4Department of Endocrinology, The People’s Hospital of Linyi, 27 Liberation Road, Linyi, Shandong Province 276003, China, 5Department of Endocrinology, The Hospital Affiliated to Jiangsu University, Zhenjiang, Jiangsu Province 212001, China and 6Department of Endocrinology, Anhui Provincial Hospital, 17 Lu Jiang Road, Hefei, Anhui Province 230001, China

*(B-L Liu and S-Y Yang contributed equally to this work)

Correspondence should be addressed to H-D Song or S-X Zhao
Emails huaidong_s1966@163.com or zhaozhao1215@gmail.com

Abstract

Background: Convincing evidence has demonstrated the association of TSH receptor (TSHR) with Graves’ disease (GD) in the Chinese Han population.

Objective: The aim of this study was to identify the causal variants for GD in the region encompassing TSHR by a refining association study.

Design and methods: GD patients (1536) and 1516 sex-matched controls were recruited in the first stage, and an additional 3832 GD patients and 3426 sex-matched controls were recruited in the replication stage. Genotyping was performed using Illumina Human660-Quad BeadChips or TaqMan single nucleotide polymorphism (SNP) Genotyping Assays and the Fluidigm EP1 platform.

Results: When the results of regression analysis for 74 genotyped SNPs and 922 imputed SNPs in the first-stage cohort were combined, rs179243 and rs3783949 were the probable susceptibility SNPs associated with GD in TSHR. Eleven SNPs, including rs179243 and rs3783949, were selected to further refine the association in the replication study. Finally, rs12101261 and rs179243 were confirmed as independent GD susceptibility variants in the replication and combined populations. Further, we also found that the rate of persistent TSHR autoantibody positivity (pTRAb +) was significantly higher in the GD patients with the susceptible genotypes rs12101261 or rs179243 than in the GD patients carrying the protective genotypes, after the GD patients had been treated for more than 1 year.

Conclusions: These findings indicate that rs12101261 and rs179243 are the possible causal SNPs for GD susceptibility in the TSHR gene and could serve as genetic markers to predict the outcome of pTRAb + in GD patients.
Introduction

Graves’ disease (GD) is a common autoimmune disease with a prevalence of ~1–2% in the Chinese population and is characterised by presentation with thyroid-stimulating hormone receptor (TSHR) autoantibodies (TRAbs), which stimulate TSHR expression and induce hyperthyroidism. TSHR is a major thyroid-specific autoantigen and plays a key role in the regulation of thyroid function in GD. Previous studies have clearly demonstrated that GD is triggered by a combination of environmental and genetic factors. The important role of genetic factors in the pathogenesis of GD has been well established based on observation of familial clustering and evidence derived from twin studies (1, 2). In the past few years, studies using the strategy of candidate genes or genome-wide association studies (GWAS) have identified a few susceptibility loci for GD. The loci supported by solid evidence include HLA, CTLA4, TSHR, SCGB3A2, PTPN22 and FCRL3 (3, 4, 5, 6, 7), as well as seven new loci identified by our study group, RNASET2, GDCG4p14, GPR174-ITM2A, CJQTNF6-RAC2, SLAMF6, ABO and TG (8, 9). RNASET2 and GDCG4p14 have been confirmed in the Caucasian GD cohorts (10, 11).

TSHR, on chromosome 14q31, was previously well established as a susceptibility locus for GD (12, 13, 14, 15, 16, 17, 18, 19, 20). However, the results of studies investigating the causal variants for GD in this region have been less consistent (21, 22); single nucleotide polymorphisms (SNPs) in intron 7 have been suggested to be associated with GD in Japanese patients (21), but intron 1 SNPs have been found in association with GD in the Caucasian UK population (22, 23). Further, early studies using small samples claimed that three nonsynonymous SNPs, D36H, P52T and D727E, in TSHR have been confirmed in the Caucasian GD cohorts (10, 11). We have provided convincing evidence for the association of the SNPs in intron 1 of TSHR with GD in the Chinese Han population (8, 9). In this study, we performed further fine mapping, based on our GWAS data, to identify the probable susceptibility variants for GD in TSHR. It is well-known that a positive TRAb test before discontinuation of antithyroid drugs (ATD) represents the best predictor of GD relapse (24, 25), suggesting that persistent TRAb positivity (pTRAb+) could be a marker of poor clinical outcome. In our previous report, we identified a group of SNPs in intron 1 of TSHR that are specifically associated with GD in the pTRAb-positive subgroup, but not in those without pTRAb+ (8). In this study, the association of susceptibility alleles in TSHR with pTRAb+ was further investigated in GD patients who had been treated with ATD for more than 1 year.

In this study, we identified two independent GD-susceptibility SNPs (rs12101261 and rs179243) by refining the association in the TSHR. The GD patients with the risk genotypes rs12101261 or rs179243 had significantly higher rates of pTRAb+ than those with the protective genotypes, after treatment for more than 1 year. This indicates that rs12101261 and rs179243 may serve as genetic markers to predict the outcome of TRAb.

Subjects and methods

Samples and clinical characteristics

All participants were recruited from the Chinese Han population through collaboration with multiple hospitals in China. All subjects provided informed consent using protocols approved by the local institutional review board. GD patients (1536) and 1516 sex-matched controls were recruited in the first stage, and an additional 3994 GD patients and 3510 sex-matched controls were recruited in the replication stage (8, 9). Diagnosis of GD was based on documented clinical and biochemical evidence of GD, diffuse goiter, and the presence of at least one of the following: positive TRAb test, diffusely increased 131I uptake in the thyroid gland, or exophthalmos (8, 9). Graves’ eye disease was classified into six grades according to the guidelines of the American Thyroid Association. Goiters were divided into three grades according to the Common Criteria. Treatment of GD includes ATD and radiiodine (radioactive iodine, 131I). The detailed clinical information for GD patients is shown in Supplementary Table 1, see section on supplementary data given at the end of this article. The levels of sensitive TSH, free thyroxine 3 and free triiodothyronine 4 were measured using chemiluminescence immunoassay. All individuals classified as GD were interviewed and examined by experienced clinicians. The investigation was approved by the local ethics committee.

Genotyping and quality control

In the GWAS stage, genotyping was performed using Illumina Human660-Quad BeadChips (8). Genotype clustering was conducted using Illumina BeadStudio 3.3 Software (8, 9). Also, the SNPs with Hardy–Weinberg equilibrium P≤10−6, missing call rate ≥0.05, or minor
allele frequency (MAF) ≤ 0.01 were removed from the association study (8, 9).

In the replication stage, in addition to the eight SNPs genotyped in our previous GWAS study (8), an additional three SNPs (rs179243, rs2284720 and rs3783949) were genotyped using TaqMan SNP Genotyping Assays and the Fluidigm EP1 platform. All SNPs were genotyped with a call rate of >95% for further association analysis.

**Imputation**

Genotype imputation was performed using IMPUTE version 2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html/), based on the 1000 Genomes Project (June 2011) as reference. The association analysis of the imputed SNPs was carried out using ProbABEL (26). Quality filtering was performed on SNPs before analysis to ensure robust association tests. The SNPs with Hardy–Weinberg equilibrium $P \leq 10^{-6}$, missing call rate ≥ 0.05, or MAF ≤ 0.01 were discarded.

**Statistical analysis**

GD patients (1536) and 1516 sex-matched controls were recruited in the first stage, and 1468 cases and 1490 controls remained after stringent quality control. An additional 3994 GD cases and 3510 sex-matched controls were recruited for the replication study. We successfully genotyped 11 SNPs in 3832 GD cases and 3426 controls in the replication stage (8). We analysed the SNPs from the GWAS and replication stages using the Cochran–Armitage trend-test in PLINK. In the combined stage, Cochran–Mantel–Haenszel stratification analysis was used to examine the associations. Forward stepwise logistic regression analysis was carried out in GWAS and the combined stage using SPSS PASW Statistics 18 packages. Conditional logistic regression analysis was performed using PLINK packages.

In stage 2, the 11 SNPs were selected using Haploview Software (http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haplovie/haplovie) and a pairwise-tagging approach. Haplotype association analyses, estimation of pairwise linkage disequilibrium (LD) and tagSNP selection were calculated using Haploview 4.1.

**Sample size and power calculations**

Sample size calculations were performed using QUANTO (version 1.2.4). Based on our GWAS data, we found that a minimum of 1343 (rs2284722) and a maximum of 3540 (rs4903964) sample pairs were needed to detect a significant difference in the 11 selected SNPs on a genome-wide significance level for each variant with allele frequencies ranging from 0.01 to 0.5 and with at least 80% power (Supplementary Table 2, see section on supplementary data given at the end of this article). In the current study, we therefore used a total of ~5000 sample pairs in the first and replication studies combined.

**Measurement of TRAb levels**

TRAb levels of GD patients who had been treated with ATD for more than 1 year were measured by quantitative ELISA (RSR Limited, Cardiff, UK) in our laboratory. GD patients with TRAb levels ≥ 1.5 U/l were defined as ‘persistently TRAb-positive’ and those with TRAb levels < 1.5 U/l were defined as ‘TRAb-negative’. Interactions between SNPs of TSHR in persistently TRAb-positive or TRAb-negative GD patients were analysed using the case-only $\chi^2$ test in the SPSS PASW Statistics 18 package.

The association between the genotypes or alleles at rs12101261 and rs179243 and the different grades of thyroid goiter in GD patients was also analysed using the $\chi^2$ test in the SPSS PASW Statistics 18 package.

**Results**

**Regression analysis of GWAS-stage and imputation data**

First, a 665-kb LD block harbouring the strongest TSHR association and separated by two recombination hotspots (recombination rate > 60 cM/Mb) located in the up- or down-streams of the LD block in the Chinese population was chosen from our GWAS data (8) for further refinement of the association (Fig. 1a). After quality control, 131 SNPs in this region, which encompasses C14orf145/TSHR, remained for the next analysis in 1468 GD patients and 1490 control subjects using the trend-test in PLINK (Supplementary Table 3, see section on supplementary data given at the end of this article). Of these SNPs, 31 showed a significance level of $P < 0.001$ and 22 showed a level of $P < 0.0001$.

To identify the independent susceptibility variants associated with GD in the region encompassing C14orf145/TSHR, the 31 SNPs with $P < 0.001$ were analysed in the GWAS data using forward stepwise regression analysis. Interestingly, two of the 31 SNPs were independently associated with GD risk in the GWAS cohorts ($P_{\text{forward}} = 1.94 \times 10^{-6}$ for rs2284720, $P_{\text{forward}} = 1.23 \times 10^{-2}$. 


for rs3783949). Moreover, rs2284720 improved 23 of the 30 SNP models in two-locus conditional logistic regression models (Fig. 2a). Interestingly, the seven SNP models, which were not improved by rs2284720, were improved by rs3783949 (Fig. 2b). Further, the region could not be perfectly explained by separately conditioning on either rs2284720 or rs3783949; however, after conditioning on the two SNPs, no other SNPs in this region were found to be

Figure 1
Regional plots of 14q31 susceptibility loci associated with Graves’ disease (GD) in the first stage. Values of −log10 are plotted against chromosome position, which is based on the 1000 Genomes Pilot 1. Each diamond represents one SNP, and the most strongly associated SNP is indicated by a large red diamond. The colour of each SNP block reflects its $r^2$ with the most strongly associated SNP within each association locus, ranging from red to white. Estimated recombination rates (based on the combined HapMap Chinese sample (CHB) and HapMap Japanese sample (JPT) from the HapMap project) are plotted in cyan to reflect the local LD structure around the associated SNPs. (a) Association results of genotyped SNPs on 14q31 with GD in the first stage. The region plotted contains 131 SNPs covering a 665-kb region that encompassed C14orf145/TSHR in a GD case-control cohort consisting of 1468 GD cases and 1490 controls. (b) The association of genotyped and imputed SNPs in TSHR with GD is magnified in panel b.
Figure 2

Results of conditional logistic regression analysis for SNPs with $P < 1 \times 10^{-3}$ in 14q31 in the first stage. (a and b) Two-locus conditional logistic regression for rs2284720 or rs3783949 with the other 29 genotyped SNPs with $P < 1 \times 10^{-3}$. The $P$ values of the other 29 SNPs conditioning on rs2284720 or rs3783949 are shown in blue diamonds. The $P$ values of rs2284720 or rs3783949 after conditioning on the 29 SNPs are shown in red squares. (c) After conditioning on rs2284720 and rs3783949 at the same time, the $P$ values of other SNPs in the Graves’ disease susceptibility locus on 14q31 are shown in panel c. None of the 29 SNPs in this susceptibility locus improved the model with rs2284720 and rs3783949 at the level of $P < 0.05$. Each diamond represents one SNP. The colour of each SNP reflects its $r^2$ with the most strongly associated SNP, rs2284720, ranging from red to white. Estimated recombination rates (based on the combined CHB and JPT samples from the HapMap project) are plotted in cyan to reflect the local LD structure around the associated SNPs. (d, e, and f) The results of conditional logistic regression analysis for 390 genotyped and imputed SNPs with $P < 1 \times 10^{-3}$ in the GWAS samples in TSHR. (d and e) Two-locus conditional logistic regression for rs179243 or rs3783949 with SNPs other than themselves. The influences of the other SNPs on the model conditioned by rs179243 or rs3783949 are presented in blue diamonds. The $P$ values of rs179243 or rs3783949 after conditioning on the other SNPs are shown in red squares. (f) After conditioning on rs179243 and rs3783949, no other SNPs in TSHR improved the model. The colour of each SNP reflects its $r^2$ with the most strongly associated SNP, rs179243, changing from red to white.
independently associated with GD (Fig. 2c). Notably, rs2284720 and rs3783949 were not in a LD block \( (r^2 = 0.15, \text{Supplementary Figure 1}, \text{see section on supplementary data given at the end of this article}) \) and were located in intron 1 of TSHR. These results indicate that the SNPs rs2284720 and rs3783949 in the TSHR susceptibility locus were independently associated with GD in the Chinese Han population.

Furthermore, 996 SNPs in a 443-kb region containing C14orf145/TSHR, representing an average distance of about 445 bp per SNP, were imputed based on our GWAS data using the IMPUTE 2 Software. Among the 996 SNPs, a cluster of SNPs that included rs179245 \( (P = 9.49 \times 10^{-11}), \) rs179243 \( (P = 1.01 \times 10^{-15}), \) and rs3783943 \( (P = 1.14 \times 10^{-10}) \) located in intron 1 of TSHR showed the most significant association with GD (Supplementary Table 4, see section on supplementary data given at the end of this article and Fig. 1b). Meanwhile, the SNPs in the coding region of TSHR, PS2T \( (rs80491931) \) and D727E \( (rs8068036) \), which were associated with GD in previous studies using small samples \( (12, 13, 14, 15, 16, 17, 18, 19, 20) \), were not associated with GD in our study (Supplementary Table 5, see section on supplementary data given at the end of this article).

To further refine the association of candidate SNPs remaining in the 443-kb region harbouring C14orf145/TSHR, 390 SNPs with \( P < 1 \times 10^{-5} \) were further analysed by two-locus logistic regression. Among the 390 SNPs, 42 SNPs containing rs2284720 and rs72693090 as the best markers for the TSHR region were individually tested in the logistic regression models, and only 18–50 SNPs improved these models with a cut-off of \( P < 0.05 \) (Supplementary Table 6, see section on supplementary data given at the end of this article). Next, we tested a regression model, taking each one of 390 SNPs in turn, and adding each test locus. Interestingly, the majority of the markers (more than 320 SNPs in this region) were improved by the addition of each of 42 SNPs in TSHR (Supplementary Table 6). With regard to the 42 SNPs in TSHR intron 1, rs179243 probably plays the most important role in the susceptibility to GD because only 18 SNPs in this region improved the model, with rs179243 as the most strongly associated marker (Supplementary Table 6 and Fig. 2d). Interestingly, these 42 SNPs as the best SNP predisposing to GD were perfectly linked \( (r^2 > 0.80) \) (Supplementary Figure 2a, see section on supplementary data given at the end of this article).

It was noted that 47 of the 390 SNPs in the TSHR susceptibility locus remained significantly associated with GD after conditioning on rs72693090 (Supplementary Table 6). However, the majority of models with the 47 SNPs were improved by the addition of each of another cluster of 34 SNPs in a high-LD block (Supplementary Table 7 and Figure 2b, see section on supplementary data given at the end of this article). Further, the models improved by 34 SNPs could not be improved by addition of these 42 SNPs, and the relatively weak association between the 34 SNPs and GD was detected in our data \( (P \text{ values from } 2.43 \times 10^{-5} \text{ to } 1.85 \times 10^{-6}, \text{Supplementary Table 7}) \). Among these 34 SNPs, the model with rs3783949 was improved by ten of 47 SNPs, and the models with each of the other 33 SNPs were improved by 16–32 SNPs (Supplementary Table 7 and Fig. 2e). Interestingly, neither rs179243 nor rs3783949 could perfectly explain the association of the SNPs in TSHR with GD (Fig. 2d and e). However, after conditioning on the two SNPs at the same time, no SNPs in TSHR were independently associated with GD at a cut-off of \( P < 0.05 \) (Fig. 2f). These results suggest that rs179243 and rs3783949 are probably independent susceptibility loci to GD in TSHR. The results are consistent with our conclusion based on the analysis of GWAS data, because rs179243 and rs2284720 are in a complete LD block \( (r^2 = 0.93) \) (Supplementary Figure 1). Combining the results of logistic regression analysis of the GWAS and imputation data, it can be inferred that the SNPs rs179243 and rs3783949 are the major susceptibility variants in TSHR for GD in the Chinese Han population.

**Association analysis in the replication and combined populations**

The GWAS and imputation data showed that the SNPs rs179243 and rs3783949 were the major susceptibility variants in the TSHR region. These results strongly suggested the existence of two independent susceptibility SNPs to GD in the TSHR region, rs3783949 and rs179243. Therefore, in order to further refine the association of causal SNPs and GD in the TSHR region, 11 SNPs were selected for genotyping in 3994 GD cases and 3510 controls. We successfully genotyped 11 SNPs in 3832 GD cases and 3426 controls. The maximum number of sample pairs needed in the combined stage was 3540, on the basis of the results of the association with GD in the first stage (Supplementary Table 2). Among the 11 SNPs, rs2284720 and rs2284722 were tightly linked to rs179243, and rs12101261 was tightly linked to rs3783949 \( (r^2 > 0.8, \text{Supplementary Figure 1}) \). In addition, six tagSNPs, which can tag 13 other SNPs with \( P < 0.0001 \) but were not in a high-LD block with rs179243 or rs3783949, were selected (Supplementary Table 8, see section on supplementary data given at the end of this article). Combining the genotyping results from GWAS and the replication stage, we confirmed the associations of the 11 SNPs with GD in...
SNP, single nucleotide polymorphism; OR, odds ratio for the risk allele.

<table>
<thead>
<tr>
<th>SNP</th>
<th>P value</th>
<th>OR</th>
<th>95% CI</th>
<th>P value</th>
<th>OR</th>
<th>95% CI</th>
<th>P value</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs7158936</td>
<td>2.46×10^{-6}</td>
<td>1.32</td>
<td>1.18–1.49</td>
<td>1.54×10^{-2}</td>
<td>1.09</td>
<td>1.02–1.18</td>
<td>3.65×10^{-6}</td>
<td>1.16</td>
<td>1.09–1.23</td>
</tr>
<tr>
<td>rs12050151</td>
<td>2.23×10^{-8}</td>
<td>1.42</td>
<td>1.26–1.61</td>
<td>8.71×10^{-5}</td>
<td>1.17</td>
<td>1.08–1.26</td>
<td>2.02×10^{-10}</td>
<td>1.24</td>
<td>1.16–1.32</td>
</tr>
<tr>
<td>rs179243</td>
<td>3.32×10^{-10}</td>
<td>1.47</td>
<td>1.30–1.65</td>
<td>6.15×10^{-12}</td>
<td>1.30</td>
<td>1.20–1.40</td>
<td>3.30×10^{-20}</td>
<td>1.34</td>
<td>1.26–1.43</td>
</tr>
<tr>
<td>rs179247</td>
<td>1.73×10^{-6}</td>
<td>1.31</td>
<td>1.18–1.47</td>
<td>3.53×10^{-17}</td>
<td>1.36</td>
<td>1.27–1.45</td>
<td>4.48×10^{-22}</td>
<td>1.34</td>
<td>1.27–1.43</td>
</tr>
<tr>
<td>rs2284720</td>
<td>3.82×10^{-10}</td>
<td>1.45</td>
<td>1.29–1.63</td>
<td>1.78×10^{-8}</td>
<td>1.24</td>
<td>1.15–1.33</td>
<td>3.40×10^{-16}</td>
<td>1.29</td>
<td>1.22–1.38</td>
</tr>
<tr>
<td>rs2284722</td>
<td>1.03×10^{-10}</td>
<td>1.45</td>
<td>1.30–1.63</td>
<td>1.79×10^{-7}</td>
<td>1.21</td>
<td>1.13–1.30</td>
<td>2.95×10^{-15}</td>
<td>1.28</td>
<td>1.20–1.36</td>
</tr>
<tr>
<td>rs3783949</td>
<td>2.03×10^{-6}</td>
<td>1.31</td>
<td>1.18–1.47</td>
<td>3.35×10^{-16}</td>
<td>1.35</td>
<td>1.25–1.45</td>
<td>4.74×10^{-21}</td>
<td>1.34</td>
<td>1.25–1.43</td>
</tr>
<tr>
<td>rs12101261</td>
<td>6.01×10^{-5}</td>
<td>1.32</td>
<td>1.19–1.47</td>
<td>9.26×10^{-19}</td>
<td>1.37</td>
<td>1.28–1.47</td>
<td>4.67×10^{-24}</td>
<td>1.35</td>
<td>1.28–1.43</td>
</tr>
<tr>
<td>rs4903964</td>
<td>3.83×10^{-5}</td>
<td>1.25</td>
<td>1.12–1.39</td>
<td>2.34×10^{-15}</td>
<td>1.31</td>
<td>1.23–1.41</td>
<td>7.37×10^{-19}</td>
<td>1.29</td>
<td>1.22–1.37</td>
</tr>
<tr>
<td>rs2300525</td>
<td>1.67×10^{-5}</td>
<td>1.28</td>
<td>1.14–1.44</td>
<td>1.28×10^{-3}</td>
<td>1.13</td>
<td>1.05–1.21</td>
<td>4.82×10^{-7}</td>
<td>1.17</td>
<td>1.10–1.24</td>
</tr>
<tr>
<td>rs17111394</td>
<td>1.71×10^{-5}</td>
<td>1.31</td>
<td>1.16–1.48</td>
<td>8.26×10^{-6}</td>
<td>1.19</td>
<td>1.10–1.29</td>
<td>1.26×10^{-9}</td>
<td>1.23</td>
<td>1.15–1.31</td>
</tr>
</tbody>
</table>

rs179243 was the most significant signal at the stage 1 study, with the most significant signal at stage 2 study, with the most significant signal at stage 3 study, followed by rs12101261. We then used forward stepwise regression analysis to identify the main-effect susceptibility variants for GD in the 11 SNPs. After conditioning on rs179243, none of the other 10 SNPs were independently associated with GD in the replicated and combined cohorts (Table 2). Two-locus logistic regression analysis showed that neither rs179243 nor rs12101261 of these 11 SNPs could perfectly improve the models of the ten SNPs except rs179243 nor rs12101261 of these 11 SNPs could perfectly improve the models of the ten SNPs except rs179243. Two SNPs, rs12101261 and rs179243, were independently associated with GD in the replicated and combined cohorts (Table 2).
Table 2  Forward stepwise logistic regression analysis for 11 SNPs in TSHR susceptibility locus to Graves’ disease in replication and combined cohorts.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr. position</th>
<th>Association with GD</th>
<th>Replication stage (3832/3426)</th>
<th>Combined stage (5300/4916)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OR 95% CI</td>
<td>P value in final model</td>
<td>OR 95% CI</td>
</tr>
<tr>
<td>rs7158936</td>
<td>80385545</td>
<td>1.54 × 10^{-2}</td>
<td>0.41</td>
<td>3.65 × 10^{-6}</td>
</tr>
<tr>
<td>rs12050151</td>
<td>80438570</td>
<td>8.71 × 10^{-5}</td>
<td>0.5</td>
<td>2.02 × 10^{-10}</td>
</tr>
<tr>
<td>rs179243</td>
<td>80498317</td>
<td>6.15 × 10^{-12}</td>
<td>0.87</td>
<td>3.30 × 10^{-20}</td>
</tr>
<tr>
<td>rs179247</td>
<td>80502299</td>
<td>3.53 × 10^{-17}</td>
<td>0.03</td>
<td>4.48 × 10^{-22}</td>
</tr>
<tr>
<td>rs2284720</td>
<td>80512920</td>
<td>1.78 × 10^{-8}</td>
<td>0.1</td>
<td>3.40 × 10^{-16}</td>
</tr>
<tr>
<td>rs2284722</td>
<td>80514120</td>
<td>1.79 × 10^{-7}</td>
<td>0.03</td>
<td>2.95 × 10^{-15}</td>
</tr>
<tr>
<td>rs3783949</td>
<td>80518135</td>
<td>3.35 × 10^{-16}</td>
<td>0.75</td>
<td>4.74 × 10^{-21}</td>
</tr>
<tr>
<td>rs12101261</td>
<td>80520982</td>
<td>9.26 × 10^{-19}</td>
<td>1.42 × 10^{-10}</td>
<td>4.67 × 10^{-24}</td>
</tr>
<tr>
<td>rs4903964</td>
<td>80538707</td>
<td>2.34 × 10^{-15}</td>
<td>0.46</td>
<td>7.37 × 10^{-19}</td>
</tr>
<tr>
<td>rs2300525</td>
<td>80567146</td>
<td>1.28 × 10^{-3}</td>
<td>0.86</td>
<td>4.82 × 10^{-7}</td>
</tr>
<tr>
<td>rs17111394</td>
<td>80592881</td>
<td>8.26 × 10^{-6}</td>
<td>0.41</td>
<td>1.26 × 10^{-9}</td>
</tr>
</tbody>
</table>

Chr., chromosome; SNP, single nucleotide polymorphism; OR, odds ratio.

Discussion

Although TMR on chromosome 19q13 is a well-established susceptibility locus for GD, the causal variants for GD in this region have remained controversial (12, 14, 16, 18, 22). In this study, we carried out a two-stage association study for 11 SNPs, including rs12101261 and rs179243, which were the two independent GD-susceptibility variants in the Chinese Han population. In addition, the rate of pTRAb+ was significantly higher in GD patients with the susceptibility genotypes of rs12101261 or rs179243 than in those carrying the protective genotypes, after which the patients had been treated for more than 1 year. In earlier studies using small sample size, three nonmonosynonymous SNPs, D36H, P52T, and D727E, were claimed as susceptibility variants for GD in TSHR (19, 27). However, the risk genotypes or alleles at rs1210261 and rs179243 were not related to the different grades of thyroid goiter in GD patients (Supplementary Table 9 and 10).

Chrs 2 and 7 were perfectly linked to rs12101255 ($r^2$ = 1.0), and rs179243 in intron 1 of TSHR were further considered as independent GD susceptibility SNPs, rs12101261 and rs179243 in intron 1 of TSHR were associated with GD with a MAF < 1% in the Chinese Han population. Moreover, the nonmonosynonymous D36H in TSHR is a rare variant, with a MAF < 1% in the Chinese Han population, and found that rs1210261 and rs179243 are the two independent GD-susceptibility variants in the Chinese Han population. In addition, the rate of pTRAb+ was significantly higher in GD patients with the susceptibility genotypes of rs12101261 or rs179243 than in those carrying the protective genotypes after which the patients had been treated for more than 1 year. In earlier studies using small sample size, three nonmonosynonymous SNPs, D36H, P52T, and D727E, were claimed as susceptibility variants for GD in TSHR (19, 27). However, the risk genotypes or alleles at rs1210261 and rs179243 were not related to the different grades of thyroid goiter in GD patients (Supplementary Table 9 and 10).
findings indicate that the two independent GD-susceptibility SNPs rs12101261 and rs179243 in intron 1 of TSHR might be potentially involved in the development of GD when all cases are considered, but that in some patients only one of these two SNPs might be involved. Interestingly, although the risk allele A of rs179243 has MAF in both the Chinese and Caucasian populations, the frequency of risk allele T at rs12101261 was dramatically different in the two ethnic populations. The frequency of the risk allele T at rs12101261 was 64.07% in the control individuals recruited from the Chinese Han population (8), but only 28% in healthy Caucasian Europeans (22). This finding indicates that genetic factors, such as the occurrence of rs12101261, might play an important role in the mechanism underlying the difference in GD risk variants in TSHR between Chinese and Caucasian Europeans.

Figure 3
Results of conditional logistic regression analysis for 11 SNPs in the replication (a, b, and c) and combined (d, e, and f) cohorts. Two-locus conditional logistic regression analysis for rs179243 or rs12101261 with the other ten SNPs in replication (a and b) and combined (d and e) cohorts. The P values of the other SNPs conditioning on rs179243 or rs3783949 are presented in blue diamonds. The P values of rs179243 or rs12101261 after conditioning on the other SNPs are presented in red squares. (c and f) After simultaneous conditioning on rs179243 and rs12101261, the other SNPs in TSHR did not significantly improve the model containing rs179243 and rs12101261 in replicated (c) and combined (f) populations. The colour of each SNP reflects its $r^2$ with the most strongly associated SNP, rs12101261, changing from red to white.

Table 3  Effect of risk genotypes at rs12101261 and rs179243 on the rate of pTRAb+ in GD patients after more than 1 year of ATD treatment.

<table>
<thead>
<tr>
<th>SNP</th>
<th>rs12101261</th>
<th>rs179243</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Genotype</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TT/CT</td>
<td>CT/CC</td>
</tr>
<tr>
<td>P value</td>
<td>$6.45 \times 10^{-3}$</td>
<td>$2.94 \times 10^{-4}$</td>
</tr>
<tr>
<td>% TRAb ≥1.5</td>
<td>74%/69%</td>
<td>69%/58%</td>
</tr>
<tr>
<td>Risk genotype</td>
<td>TT</td>
<td>CT</td>
</tr>
<tr>
<td>OR</td>
<td>1.24</td>
<td>1.59</td>
</tr>
<tr>
<td>95% CI</td>
<td>1.06–1.44</td>
<td>1.23–2.04</td>
</tr>
</tbody>
</table>

pTRAb+, persistent TRAb-positivity; SNP, single nucleotide polymorphism; OR, odds ratio.
and Caucasian populations. It should be pointed out that additional international collaboration studies in distinct ethnic groups will be required to further address the effects of genetic factors on the differences in GD prevalence among distinct ethnic populations.

It is well-known that patients with GD have a higher risk of relapse after ATD treatment when TRAb is persistently positive (24, 25). Moreover, several previous reports have revealed that a positive TRAb test before discontinuation of ATD represents the best predictor of GD relapse (30, 31), suggesting that TRAb persistence could be a marker of poor clinical outcome. In our most recent GWAS, we identified a group of SNPs in intron 1 of TSHR that are specifically associated with GD in the persistent TRAb-positive subgroup, but not in those without pTRAb+ (8). In this study, we found that the frequencies of pTRAb+ were significantly higher in GD patients carrying the risk genotypes of the two independent susceptibility SNPs (rs12101261 and rs179243) than in those carrying the protective genotypes after the patients had been treated with ATD for more than 1 year. These data may provide an insight to predicting the outcome of pTRAb+ using genetic markers in GD patients.

In conclusion, the association of two independent susceptibility variants in TSHR, rs12101261 and rs179243, was refined in a large-scale study of the Chinese Han population. We also found that the rate of pTRAb+ was significantly higher in the GD patients with the susceptible rs12101261 or rs179243 genotypes than in those carrying the protective genotypes, after the patients had been treated for more than 1 year. These findings indicate that the SNPs rs12101261 and rs179243 might be used as genetic markers to predict the outcome of pTRAb+ in GD patients. However, further functional studies are needed to investigate the mechanism underlying the induction of GD onset by the two SNPs.

Acknowledgements
We thank all patients and control subjects for participating in this study.

References