

ORIGINAL ARTICLE

# Association studies of *SEPS1* gene polymorphisms with Hashimoto's thyroiditis in Han Chinese

Miao Li<sup>1</sup>, Bailing Liu<sup>2</sup>, Lu Li<sup>3</sup>, Chen Zhang<sup>4</sup> and Qi Zhou<sup>1</sup>

Although the connection between *SEPS1* gene variants and Hashimoto's thyroiditis (HT) has been established in Europeans, the relationship between the *SEPS1* gene and HT remains unclear in Han Chinese. Here we aimed to investigate the potential association between *SEPS1* variants and HT in the Han population. In addition, the effects of *SEPS1* haplotypes on the susceptibility of the development of immune-mediated diseases with an inflammatory component will also be evaluated. Seven single-nucleotide polymorphisms (SNPs) with minor allele frequency  $\geq 0.05$  were genotyped in 1013 HT patients and 2998 healthy controls from genetically independent Han Chinese individuals. We identified that the rs28665122 SNP was significantly associated with HT, both in the female group (allelic  $P=0.002644$  and genotypic  $P=0.010326$ ) and the combined data set (allelic  $P=0.000518$  and genotypic  $P=0.002731$ ). Further analyses based on haplotypes indicated that a two-SNP haplotype (rs2009895–rs28665122) was significantly associated with HT (global  $P=0.0036$ ), which was also observed in females (global  $P=0.0162$ ) but not in males. Our findings provide further supporting evidence that confirms the results of previous studies, which suggested potential roles of the *SEPS1* gene in the pathogenesis and etiology of HT.

*Journal of Human Genetics* (2015) 60, 427–433; doi:10.1038/jhg.2015.42; published online 28 May 2015

## INTRODUCTION

Hashimoto's thyroiditis (HT) is a chronic autoimmune disease and the incidence of HT has increased worldwide and is estimated to be 0.3–1.5 cases per 1000 people per year.<sup>1</sup> The disease is characterized by abnormalities of thyroid function as a result of the production of thyroid anti-antibodies (thyroid peroxidase and antithyroglobulin) and generation of thyroid antigen-specific lymphocytes, which in turn lead to gradual destruction of the gland by infiltrating the thyroid gland with immune effector cells.<sup>2</sup> To date, previous studies have confirmed HT to be a multifactorial disorder, meaning that both genetic and environmental predisposing factors are involved in the pathogenesis of HT.<sup>3–5</sup> In addition, a recent study has indicated that females, compared with males, have a much higher risk of autoimmune thyroid disease.<sup>6</sup>

As the most common type of chronic autoimmune thyroid disease,<sup>3,7</sup> variations in several candidate genetic loci have been identified as contributing to one's risk of developing an HT phenotype by previous linkage and association studies.<sup>8,9</sup> Among these candidate genes, the *SEPS1* gene has gained much attention. This gene encodes a member of the selenoprotein family, Selenoprotein S, which is widely expressed in many tissues. The organization of *SEPS1* is not complex for encoding a 189-amino acid protein and only contains 7 exons spanning a 7-kb region located on human chromosome 15 at 15q26.3. The *SEPS1* gene was first suggested to be related to the stress response process including immune and inflammatory processes through the

study of Curran *et al.*<sup>10</sup> In addition, several diseases, including inflammatory disorders such as insulin-dependent diabetes mellitus, Alzheimer's disease and celiac disease, have been shown to be linked to the region.<sup>11–14</sup> Therefore, in the sense of function and position for involvement in inflammation-related disorders, it is reasonable for the *SEPS1* gene to be considered as a potential candidate gene.

Recent studies have provided supportive evidence for the potential association between rs28665122 in the *SEPS1* gene and a wide spectrum of disease,<sup>15–20</sup> including HT.<sup>21</sup> However, a previous study by Santos *et al.*<sup>21</sup> on the positive association of *SEPS1* with HT was from a Portuguese population. Although an association study has been proved to be a powerful tool in unraveling the genetic mechanisms of several complex disorders,<sup>22–25</sup> it is clear that an association study has not explained most of the underlying genetic risks for many diseases. Identifying individual candidate genes/variants on disease risks remains important. Despite evidence of a strongly significant association in Europeans, the underlying biological mechanisms are largely unknown and the genetic loci that contribute to the disorder remain unknown. Therefore, it is still necessary to explore the potential association between *SEPS1* polymorphisms and HT in additional genetically independent populations.

To confirm the association between the *SEPS1* gene and HT, further studies in a larger sample size may be required. In the present study, we conducted the first large genetic association study of the *SEPS1* gene containing a single-nucleotide polymorphism (SNP) of

<sup>1</sup>Department of Ultrasound, Second Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, China; <sup>2</sup>Department of Ultrasound, Children's Hospital of Xi'an, Xi'an, China; <sup>3</sup>College of Medicine and Forensics, Xi'an Jiaotong University, Xi'an, China and <sup>4</sup>The First Department of Orthopedics, Second Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, Xi'an, China  
Correspondence: Dr Q Zhou, Department of Ultrasound, Second Affiliated Hospital, School of Medicine, Xi'an Jiaotong University, 157 Xiwu Road, Xi'an 710061, China.  
E-mail: zhouqixtu@163.com

Received 11 December 2014; revised 16 March 2015; accepted 23 March 2015; published online 28 May 2015

rs28665122, which is significantly associated with HT in Europeans, to determine whether the *SEPS1* gene is associated with HT in the Han Chinese population.

## MATERIALS AND METHODS

### Patients and controls

We implemented a case–control design for our study. A total of 4011 individuals, ranging from 21 to 68 years old, were randomly recruited for our study. We enrolled Han Chinese individuals with no migration history within the immediate three generations from Shaanxi Province. To avoid socio-economic status being a potential confounding factor, all of the participants had a similar socio-economic level. Living situation, annual income, basic education and subject characteristics served as the primary socio-economic indicators for this study. A face-to-face interview with a standard questionnaire was conducted.

A total of 1013 HT patients (791 females/222 males, range 23–66 years, mean age = 35.13 years, s.d. = 11.53) were diagnosed and recruited at the Second Affiliated Hospital of Xi'an Jiaotong University (Xi'an, China). These patients were enrolled between November 2011 and August 2014, and patients were selected based on an enlarged thyroid, characteristic ultrasound signs (hypoechoogenicity and non-homogeneous texture) and a high level of either anti-thyroid peroxidase (normal range in the method used: 0–34 IU ml<sup>-1</sup>) or anti-thyroglobulin (normal range in the method used: 0–115 IU ml<sup>-1</sup>), with or without clinical and biochemical hypothyroidism. The Roche Diagnostics COBAS 6000 E601 Module Immunochemistry Analyzer (Roche, Basel, Switzerland) was applied for thyroid function tests. For suspected cases, the diagnoses were conducted by fine needle aspiration cytology (FNAC). To control potential confounding factors, participants with a reported history of thyroid cancer and/or previous thyroid surgery were excluded from the study. All patients underwent thyroid hormone replacement therapy.

We recruited 2998 unrelated healthy controls (2345 females/653 males, range 21–68 years, mean age = 33.54 years, s.d. = 12.51) from the Health Check-up Center of the same hospital, who did not have any personal or family history of thyroid disease. It is well known that especially in females, ~10% of euthyroid individuals show an elevated level of thyroid-specific autoantibody and they have the possibility of having autoimmune thyroiditis. Thus, the healthy controls collected in our study had normal thyroid functions and no ultrasound changes were found in the thyroid, and they were negative for thyroid autoantibodies (within the normal ranges mentioned above). Exclusion criteria also included the existence of any comorbid cardiac, autoimmune, infectious, musculoskeletal or malignant disease, or a recent history of operation or trauma. This study was conducted in accordance with the ethical guidelines of the Declaration of Helsinki (version 2002) and was approved by the Xi'an Jiaotong University Ethics Committee. All participants completed written informed consent forms. Baseline characteristics and primary clinical information for all subjects were obtained and are summarized in Table 1.

### SNP selection and genotyping

We have chosen our candidate SNPs through the dbSNP and HapMap databases. As an initial screen of the most common SNPs in the Han Chinese population, a minor allele frequency  $\geq 0.05$  was used as the cut-off value. We searched for all SNPs with minor allele frequency  $\geq 0.05$  between 5 kb upstream and 5 kb downstream of the *SEPS1* gene in the CHB databases of HapMap and dbSNP. Based on the criteria, we finally selected seven SNPs (rs4965373, rs78817529, rs4965814, rs2306534, rs2009895, rs28665122 and rs34713741) within the 17-kb region containing the *SEPS1* gene.

Peripheral venous blood samples were collected in plain tubes and were stored at –20 °C. Genomic DNA was isolated from peripheral blood leukocytes according to the manufacturer's protocol (Genomic DNA kit, Axygen Scientific, Union City, CA, USA) and DNA samples were stored at –20 °C for SNP analysis. All SNPs were genotyped using high-throughput, matrix-assisted laser desorption ionization–time-of-flight mass spectrometry. The resulting spectra were processed using Typer Analyzer software (Sequenom, San Diego, CA, USA) for final data acquisition.

**Table 1** Baseline characteristics and clinical parameters of patients and controls

	Patients (n = 1013)	Controls (n = 2998)
Gender (female/male)	791/222	2345/653
Range of age (years)	23–66	21–68
Mean age (years)	35.13 ± 11.53	33.54 ± 12.51
<i>Thyroid size</i>		
Normal	58 (5.73%)	2998 (100%)
I Degree	123 (12.14%)	None
II Degree	769 (75.91%)	None
III Degree	63 (6.22%)	None
<i>Family history</i>		
Positive	160 (15.79%)	None
Negative	853 (84.21%)	2998 (100%)
<i>Clinical parameters</i>		
FT <sub>3</sub> (pmol l <sup>-1</sup> )	4.12 ± 0.75	3.34 ± 0.32
FT <sub>4</sub> (pmol l <sup>-1</sup> )	15.33 ± 2.42	13.43 ± 1.30
TSH (mIU l <sup>-1</sup> )	4.84 ± 0.65	1.64 ± 0.93
Anti-TPO (IU ml <sup>-1</sup> )	495.44 ± 271.68	11.18 ± 7.95
Anti-Tg (IU ml <sup>-1</sup> )	446.58 ± 263.37	29.47 ± 22.48

Abbreviations: FT, free T, Tg, thyroglobulin; TPO, thyroid peroxidase; TSH, thyroid stimulating hormone.

Data are expressed as mean ± s.d. The Roche Diagnostics COBAS 6000 E601 Module Immunochemistry Analyzer was applied for thyroid function tests. Normal reference ranges of each parameter are as follows: FT<sub>3</sub> (normal reference range: 3.1–6.8 pmol l<sup>-1</sup>), FT<sub>4</sub> (normal reference range: 12–22 pmol l<sup>-1</sup>), TSH (normal reference range: 0.27–4.2 IU ml<sup>-1</sup>); anti-TPO (normal reference range: 0–34 IU ml<sup>-1</sup>); anti-Tg (normal reference range: 0–115 IU ml<sup>-1</sup>).

### Statistical analysis and power analysis

We used GENEPOP v4.0 (François Rousset, <http://kimura.univ-montp2.fr/~rousset/Genepop.htm>) to conduct the Hardy–Weinberg equilibrium test for each SNP in both the case and control groups. Allelic and genotypic association tests were performed using the genetic analysis software CLUMP v2.4 (Dave Curtis, <http://www.davecurtis.net/software/dcurtis/clump24.zip>). With 10 000 empirical Monte Carlo simulations implemented by this software, multiple comparison problems can be solved.<sup>26</sup> Haploview v4.2 (Broad Institute of MIT and Harvard, <http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/haploview>) was used to investigate the linkage disequilibrium (LD) structure of the candidate markers. To further examine the LD structure around the significant SNP, we used HapMap data as a reference data set to generate a regional LD plot. Haplotype frequency was calculated using GENECOUNTING v2.2 software.<sup>27,28</sup> All haplotypes including rare ones (frequency <0.01) were selected for a significant test through a likelihood ratio test, followed by permutation testing.<sup>28,29</sup> Curtis and Xu<sup>30</sup> reported that minor differences in haplotype frequency estimates can produce very large differences in heterogeneity test statistics. Thus, we also implemented score tests that account for the linkage phase ambiguity for the haplotype association.<sup>31</sup> The score tests were originally derived from generalized linear models and have been used for global tests of association and haplotype-specific tests. Permutation was implemented to obtain the maximum haplotype-specific score statistics and its *P*-values. Haplo Stats package v1.6.11 (Sinnwell JP and Schaid DJ, <http://www.mayo.edu/research/labs/statistical-genetics-epidemiology/software>) was used for these analyses.<sup>31</sup> We performed stratification analysis to investigate the potential effects of gender on the single SNP and haplotype association analyses. The global *P*-value from the haplotype analysis was calculated based on the differences of haplotypic frequency distribution among patients and controls. A *P*-value of 0.05 was chosen as a threshold for significance. In addition, in the study design stage, we calculated the statistical power using PGA v2.0 (Menashe *et al.*<sup>32</sup>). Our results showed that the power for SNP and haplotype association

**Table 2 Allele and genotype frequency of single SNP association analysis**

SNP markers/bp	Allele freq. %		P-value <sup>a</sup>	Genotype freq. %			P-value <sup>a</sup>	HWE	OR <sup>b</sup> 95% CI	
rs4965373	G	A		GG	GA	AA				
HT	101,272	66.73	33.27	0.133770	43.93	45.61	10.46	0.274095	0.387	1.09
CTR	,190	68.53	31.47		46.83	43.40	9.77		0.739	0.98–1.21
rs78817529	T	C		TT	TC	CC				
HT	101,272	73.10	26.90	0.129732	53.41	39.39	7.21	0.294993	0.961	1.09
CTR	,915	74.80	25.20		56.20	37.19	6.60		0.461	0.97–1.22
rs4965814	A	G		AA	AG	GG				
HT	101,273	58.59	41.41	0.111705	33.56	50.05	16.39	0.161255	0.317	1.09
CTR	,712	60.59	39.41		36.89	47.4	15.71		0.681	0.98–1.20
rs2306534	T	C		TT	TC	CC				
HT	101,275	83.86	16.14	0.170821	70.48	26.75	2.76	0.395849	0.708	1.10
CTR	,488	85.12	14.88		72.62	25.02	2.37		0.502	0.96–1.26
rs2009895	C	T		CC	CT	TT				
HT	101,276	51.68	48.32	0.581945	26.06	51.23	22.7	0.419665	0.411	1.03
CTR	,024	52.38	47.62		27.89	49.00	23.12		0.330	0.93–1.14
rs28665122	G	A		GG	GA	AA				
HT	101,277	83.37	16.63	<b>0.000518</b>	69.99	26.75	3.26	<b>0.002731</b>	0.260	1.28
CTR	,522	86.49	13.51		75.08	22.82	2.10		0.195	1.11–1.47
rs34713741	C	T		CC	CT	TT				
HT	101,277	76.26	23.74	0.249071	57.45	37.61	4.94	0.177088	0.218	1.07
CTR	,671	77.50	22.50		60.31	34.39	5.30		0.448	0.95–1.21

Abbreviations: CI, confidence interval; CTR, control; HT, Hashimoto's thyroiditis; HWE, Hardy–Weinberg equilibrium; OR, odds ratio; SNP, single-nucleotide polymorphism. Significant *P*-values are in italic bold.  
<sup>a</sup>*P*-value of the normal  $\chi^2$ -statistics from Monte Carlo stimulation using CLUMP (T2).  
<sup>b</sup>OR refers to risk allele OR.

tests were 95% and 87%, respectively, at the false-positive rate of 0.05 and a presumed odds ratio of 2.

## RESULTS

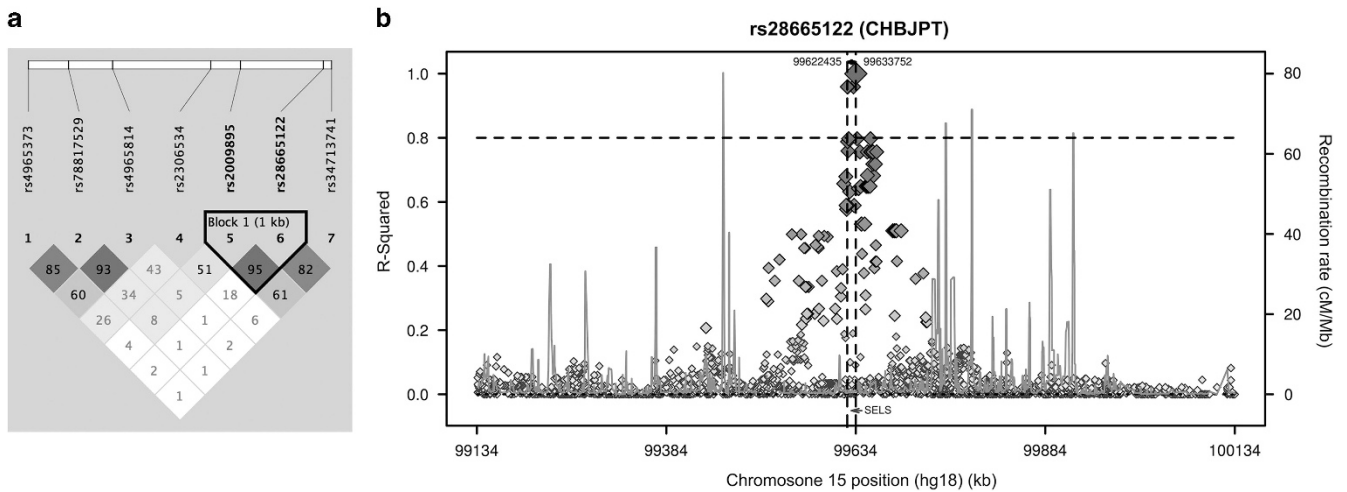
Genotype data from 7 SNPs of the 17-kb genomic region containing the *SEPS1* gene from 1013 HT cases and 2998 controls were released. Allelic and genotypic frequencies of all SNPs in the case and control groups, as well as the Hardy–Weinberg equilibrium test, are shown in Table 2. All SNPs were highly polymorphic in the case and control groups, and both groups were in Hardy–Weinberg equilibrium. Significant allelic association of rs28665122 ( $P=0.000518$ ; odds ratio = 1.28; 95% confidence interval 1.11–1.47) was identified in combined analysis. Genotypic association analysis confirmed this result with a significant *P*-value of 0.002731. There were no significant differences in the other six SNPs in both allelic and genotypic analyses (Table 2). To perform haplotype-based association analyses, we examined the LD structure within the genotype data of seven SNPs and identified a haplotype block. An LD pattern is suggested in Figure 1a, which presents the results of the LD tests between the pairs of seven SNP markers across the *SEPS1* gene. According to Figure 1a, we observed a two-SNP LD structure, which was also confirmed in a regional LD structure plot shown in Figure 1b. Several strongly correlated common SNPs with rs28665122 were identified in the plot. Haplotype association analyses were followed to test this LD block. A significant *P*-value (global  $P=0.0036$ ) was obtained through association test of the haplotype consisting of two SNPs (rs2009895 and rs28665122; Table 3). Three haplotypes, the HAP2 (TG), the HAP3 (TA) and the HAP4 (CA), were significantly associated with HT, especially the HAP4, of which the frequency increased threefold in the case group ( $P=0.027$ ), while HAP1 (CG) showed no significance ( $P=0.452$ ; Table 3). The haplotype frequencies estimated by GENECOUNTING (Jinghua Zhao, <http://legacy.mrc-epid.cam.ac.uk/~jinghua.zhao/>

software.htm) and HaploStats (Sinnwell JP and Schaid DJ, <http://www.mayo.edu/research/labs/statistical-genetics-genetic-epidemiology/software>) were very similar. Thus, the potential bias reported by Curtis and Xu<sup>30</sup> that minor differences in haplotype frequency estimates could produce very large differences in heterogeneity test statistics may not affect our analysis.

To investigate the potential effects of gender in our association analyses, we stratified our analyses by males and females. We found that the rs28665122 SNP showed both allelic ( $P=0.002644$ ) and genotypic ( $P=0.010326$ ) association with HT in females, whereas it only showed marginal significance in males (Table 4). Similar to single SNP analysis, haplotypic analysis in females showed a significant association (global  $P=0.0162$ , Table 5). The HAP3 (TA) showed significant differences for its frequency in the female case and control groups, and an almost twofold increase was observed in female cases with a *P*-value of 0.004. We observed a protective effect implicated by HAP2 (TG) ( $P=0.024$ ; Table 5) and this may be a result of the higher frequencies in female controls. Both haplotypic association patterns failed to be identified in males.

## DISCUSSION

HT is a complex disease and chronic inflammation has been considered as an important factor in the etiopathogenesis, while a wide range of environmental factors can trigger HT and cause disease progression. As one of the novel candidate genes in association studies of diseases related to chronic inflammation, the *SEPS1* gene has been the focus of genetic dissection in the past several years. A previous study has shown that the *SEPS1* gene is a significant risk factor for preeclampsia in Norwegian populations,<sup>19</sup> and another study based on a Finnish cohort has also indicated the potential relationship between the *SEPS1* gene and coronary heart disease or ischemic stroke.<sup>20</sup> However, a German study failed to identify association signals between



**Figure 1** (a) Estimation of LD between each pair of seven SNPs genotyped in *SEPS1* gene in Han Chinese population. LD structure ( $D'$ ) between marker pairs was indicated by the shaded matrices. (b) Regional LD plot of rs28665122 based on the HapMap CHB+JPT data. The  $r^2$  of other SNPs versus rs28665122 was indicated by diamond shape. Gene regions were indicated at the bottom of the plot.

**Table 3** Haplotypes frequency and association analysis

Haplotype ID	Haplotype <i>rs2009895</i> – <i>rs28665122</i>	Genecounting (frequency %)				Haplo stats (frequency %)			
		HT	CTR	<i>P</i> -value <sup>a</sup>	Global <i>p</i> <sup>b</sup>	HT	CTR	<i>P</i> -value <sup>a</sup>	Global <i>p</i> <sup>b</sup>
HAP1	CG	51.2	52.2	0.452	<b>0.0036</b>	51.2	52.2	0.474	<b>0.0035</b>
HAP2	TG	32.2	34.3	<b>0.019</b>		32.2	34.3	<b>0.016</b>	
HAP3	TA	16.1	13.3	<b>0.001</b>		16.1	13.3	<b>0.001</b>	
HAP4	CA	0.6	0.2	<b>0.027</b>		0.6	0.2	< <b>0.001</b>	

Abbreviations: HT, Hashimoto's thyroiditis; CTR, control; SNP, single-nucleotide polymorphism.

Significant *P*-values are in italic bold.

<sup>a</sup>Based on 10 000 permutations.

<sup>b</sup>Based on comparison of frequency distribution of all haplotypes for the combination of SNPs.

the *SEPS1* gene and inflammatory bowel disease susceptibility. In that study, the polymorphisms of the *SEPS1* gene did not contribute to either a certain disease phenotype or increased tumor necrosis factor- $\alpha$  levels in inflammatory bowel disease patients.<sup>33</sup> Another study in Italian and German populations also failed to confirm any significant association of the *SEPS1* gene with stroke in young patients and controls.<sup>34</sup>

The purpose of our study was to investigate the potential association between seven common SNPs in the *SEPS1* gene and HT in the Han Chinese population. To the best of our knowledge, we have carried out the first genetic association study of the *SEPS1* gene in Han Chinese individuals. To examine whether common risk variants exist in different populations or not, we compared our study with that of Santos *et al.*<sup>21</sup> In the two studies, a significant association signal was found between rs28665122 and HT, and the A allele was considered to be the risk allele. However, we also observed some differences between the two studies. The odds ratio of rs28665122 was 1.28 in our data set, compared with 2.22 in that reported by Santos *et al.*<sup>21</sup> The A allele frequency of rs28665122 in the cases of HT was much lower in our study than in the study of Santos *et al.*<sup>21</sup> (0.1613 in our samples and 0.2651 in the Santos *et al.*<sup>21</sup> samples). In addition, we also found a sex-specific association of rs28665122 that did not exist in the previous report.<sup>21</sup> Indeed, it is not enough to come to a conclusion simply from the analysis of some SNPs.<sup>35</sup> In our study, two specific haplotypes,

HAP3 (TA) and HAP4 (CA), have been demonstrated to be two main risk factors for HT. However, the frequency of HAP4 (CA) was too low to conclude that it was associated with HT. In addition, we found that the association signal of rs28665122 was stronger than that of the haplotypes. Considering that rs2009895 was not associated with HT, it is reasonable to assume that the association in the haplotypes may be caused by rs28665122. One previous study has reported the gender-specific effect of the *SEPS1* polymorphism with coronary heart disease samples. The authors of that paper proposed that the gender differences might be a result of the differences in disease etiology or in hormonal milieu for men and women.<sup>20</sup> Although the coronary heart disease involved in their study is different from ours, similar results in both studies indicated the gender-specific effect in the *SEPS1* gene. In fact, it has been reported that gender is believed to be related to HT in several epidemiological studies.<sup>6</sup> However, our data should be interpreted with caution, because confirmation in a larger sample that includes more male samples is necessary to properly detect whether possible gender-specific interactions exist or not.

In spite of some similarities in general association patterns between these studies, the possible reasons for these discrepancies mentioned above might be differences in ethnicity, genetic heterogeneity, type of polymorphism evaluated, sex and age of participants. However, our results are unlikely to be an artifact. First, significant signals were confirmed by both single SNP and haplotype-based analyses. Second,

**Table 4** Gender-specific allele and genotype association analysis in female and male

SNP markers	Allele freq. %		P-value <sup>a</sup>	Genotype freq. %			P-value <sup>a</sup>	HWE P-value	OR <sup>b</sup> 95% CI
<i>rs4965373</i>	G	A		GG	GA	AA			
F									
HT	63.21	36.79	0.115831	38.94	48.55	12.52	0.243521	0.218	1.10
CTR	65.39	34.61		42.3	46.18	11.51		0.323	0.98–1.24
M									
HT	78.83	21.17	0.691542	61.26	35.14	3.60	0.874043	0.433	1.06
CTR	79.71	20.29		63.09	33.23	3.68		0.485	0.81–1.38
<i>rs78817529</i>	T	C		TT	TC	CC			
F									
HT	73.83	26.17	0.157260	54.61	38.43	6.95	0.366401	0.879	1.10
CTR	75.61	24.39		57.40	36.42	6.18		0.540	0.96–1.25
M									
HT	70.50	29.50	0.550244	49.10	42.79	8.11	0.759458	0.669	1.08
CTR	71.98	28.02		51.91	40.12	7.96		0.890	0.85–1.36
<i>rs4965814</i>	A	G		AA	AG	GG			
F									
HT	59.54	40.46	0.109496	35.02	49.05	15.93	0.178212	0.610	1.10
CTR	61.81	38.19		38.72	46.18	15.10		0.293	0.98–1.24
M									
HT	55.63	44.37	0.790366	28.83	53.60	17.57	0.877825	0.201	1.03
CTR	56.36	43.64		30.47	51.76	17.76		0.182	0.83–1.28
<i>rs2306534</i>	T	C		TT	TC	CC			
F									
HT	84.26	15.74	0.163784	71.05	26.42	2.53	0.371516	0.914	1.12
CTR	85.69	14.31		73.60	24.18	2.22		0.500	0.96–1.31
M									
HT	83.11	16.89	0.958784	68.92	28.38	2.70	0.956944	0.873	1.01
CTR	83.00	17.00		69.07	27.87	3.06		0.754	0.76–1.34
<i>rs2009895</i>	C	T		CC	CT	TT			
F									
HT	51.14	48.86	0.653021	25.54	51.20	23.26	0.519974	0.490	1.03
CTR	51.79	48.21		27.29	49.00	23.71		0.363	0.92–1.15
M									
HT	54.05	45.95	0.865473	28.38	51.35	20.27	0.829422	0.614	1.02
CTR	54.52	45.48		30.02	49.00	20.98		0.762	0.82–1.27
<i>rs28665122</i>	G	A		GG	GA	AA			
F									
HT	82.24	17.76	<b>0.002644</b>	68.14	28.19	3.67	<b>0.010326</b>	0.325	1.26
CTR	85.39	14.61		73.09	24.61	2.30		0.510	1.08–1.47
M									
HT	87.39	12.61	0.069117	76.58	21.62	1.80	0.064458	0.775	1.36
CTR	90.43	9.57		82.24	16.39	1.38		0.172	0.98–1.91
<i>rs34713741</i>	C	T		CC	CT	TT			
F									
HT	76.74	23.26	0.389650	58.41	36.66	4.93	0.453961	0.449	1.06
CTR	77.78	22.22		60.68	34.20	5.12		0.612	0.93–1.22
M									
HT	75.00	25.00	0.567945	54.50	40.99	4.50	0.236823	0.165	1.08
CTR	76.34	23.66		58.81	35.07	6.13		0.455	0.84–1.38

Abbreviations: CI, confidence interval; CTR, control; HT, Hashimoto's thyroiditis; HWE, Hardy–Weinberg equilibrium; OR, odds ratio; SNP, single-nucleotide polymorphism.

Significant P-values are in italic bold.

<sup>a</sup>P-value of the normal  $\chi^2$ -statistics from Monte Carlo stimulation using CLUMP (T2).

<sup>b</sup>OR refers to risk allele OR.



**Table 5 Gender-specific haplotypes frequency and association analysis in female and male**

Haplotype rs2009895–rs28665122		Female		Male		P-value <sup>a</sup>		Global <sup>b</sup> P-value
		Case	Control	Case	Control	Female	Male	
HAP1	CG	50.8	51.6	51.9	54.1	0.057	0.647	Female
HAP2	TG	31.5	33.8	34.5	36.3	<b>0.024</b>	0.408	<b>0.0162</b>
HAP3	TA	17.4	14.4	11.4	9.2	<b>0.004</b>	0.156	Male
HAP4	CA	0.4	0.2	1.2	0.4	0.087	0.145	0.254

Abbreviation: SNP, single-nucleotide polymorphism.

<sup>a</sup>Based on 10 000 permutations and significant P-values are in italic bold.

<sup>b</sup>Based on comparison of frequency distribution of all haplotypes for the combination of SNPs.

for Han Chinese, matching of the geographic population and genetic background could help us to avoid a potential population stratification problem.<sup>36</sup> Thus, population stratification is an unlikely explanation, because we recruited all subjects from the same geographical population. Furthermore, similar conclusions were reached in two ethnic populations (Han Chinese and European) that reaffirmed the association observed from those studies. Notably, two potential limitations of our study should be kept in mind. One is that our study results need to be replicated and validated in other ethnic groups, because SNPs and haplotype structures can vary among different ethnic groups, and statistically significant results may occur by chance. The other is that we could not conclusively rule out some other potential confounding factors, such as smoking and alcohol use, and the possibility that there are some other genes that may affect the onset of HT. As is known, *SEPS1* polymorphism can affect the circulating levels of pro-inflammatory cytokines. Furthermore, the human biological mechanism of inflammation is very complex and hundreds of genes may be involved in the process of inflammation. Therefore, a comprehensive analysis with all of these related genes may be significant in unraveling the effects of these polymorphisms in the pathogenesis of HT.

In the present study, our results suggested that the A allele may serve as a risk allele for HT. However, the molecular mechanism by which the *SEPS1* gene has a role in the pathogenesis of HT remains unknown and needs to be elucidated. Previous studies have shown that SNP rs28665122 is associated with circulating levels of the inflammatory cytokines interleukin-1, interleukin-6 and tumor necrosis factor- $\alpha$ , and its minor allele was related to a decrease in *SEPS1* gene expression. Impairment of *SEPS1* is furthermore associated with increased interleukin-6 and tumor necrosis factor- $\alpha$  secretion in macrophages.<sup>10</sup> The GG genotype of the SNP rs28665122 was also shown to be significantly correlated with thioredoxin reductase activity and with serum selenium; furthermore, the A allele induced lower selenoprotein S activity, while the G allele of this SNP seemed to be necessary to complete enough promoter activity in a stressful environment.<sup>37</sup> This finding was consistent with the functional studies of *SEPS1* polymorphisms mentioned above. Therefore, it is reasonable to hypothesize that A allele carriers may have a higher risk of developing HT, given the above-mentioned role of *SEPS1* in inflammation. Another observation was that the significant SNP identified in our study is located in the promoter region of the *SEPS1* gene, which indicated that this region may be a harbor for functionally relevant variants that may be of interest for future studies.

In conclusion, our work provides further supporting evidence for association of the *SEPS1* gene with HT. Moreover, as an intriguing new insight into the pathogenesis of HT, we have also confirmed previous reports that suggest that the gene is responsible for having an important role in the etiology of HT. Given the complex patterns of

findings from association studies focusing on HT and its underlying genetic heterogeneity and mechanism of chronic inflammation, further inquiries and wider replications are required, especially within different ethnic samples.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

This research was totally supported by the National Natural Science Foundation of China (number 81171356 and number 81301562). We thank all of the donors for their assistance in accessing collections and their advice and comments during the preparation of this paper.

- Vanderpump, M., Tunbridge, W., French, J., Appleton, D., Bates, D., Clark, F. *et al.* The incidence of thyroid disorders in the community: a twenty-year follow-up of the Whickham Survey. *Clin. Endocrinol.* **43**, 55–68 (1995).
- Brix, T. H., Kyvik, K. O. & Hegedüs, L. A population-based study of chronic autoimmune hypothyroidism in danish twins. *J. Clin. Endocrinol. Metab.* **85**, 536–539 (2000).
- Tomer, Y. & Huber, A. The etiology of autoimmune thyroid disease: a story of genes and environment. *J. Autoimmun.* **32**, 231–239 (2009).
- Aust, G., Krohn, K., Morgenthaler, N., Schröder, S., Schütz, A., Edelmann, J. *et al.* Graves' disease and Hashimoto's thyroiditis in monozygotic twins: case study as well as transcriptomic and immunohistological analysis of thyroid tissues. *Eur. J. Endocrinol.* **154**, 13–20 (2006).
- Tomer, Y. Genetic dissection of familial autoimmune thyroid diseases using whole genome screening. *Autoimmun Rev.* **1**, 198–204 (2002).
- McLeod, D. S. & Cooper, D. S. The incidence and prevalence of thyroid autoimmunity. *Endocrine* **42**, 252–265 (2012).
- Brown, R. S. Autoimmune thyroid disease: unlocking a complex puzzle. *Curr. Opin. Pediatr.* **21**, 523–528 (2009).
- Ban, Y., Greenberg, D. A., Davies, T. F., Jacobson, E., Concepcion, E., Tomer, Y. 'Linkage analysis of thyroid antibody production: evidence for shared susceptibility to clinical autoimmune thyroid disease. *J. Clin. Endocrinol. Metab.* **93**, 3589–3596 (2008).
- Tomer, Y., Ban, Y., Concepcion, E., Barbesino, G., Villanueva, R., Greenberg, D. A. *et al.* Common and unique susceptibility loci in Graves and Hashimoto diseases: results of whole-genome screening in a data set of 102 multiplex families. *Am. J. Hum. Genet.* **73**, 736–747 (2003).
- Curran, J. E., Jowett, J. B., Elliott, K. S., Gao, Y., Gluschenko, K., Wang, J. *et al.* Genetic variation in selenoprotein S influences inflammatory response. *Nat. Genet.* **37**, 1234–1241 (2005).
- Zamani, M., Pociot, F., Raeymaekers, P., Nerup, J. & Cassiman, J.-J. Linkage of type 1 diabetes to 15q26 (IDDM3) in the Danish population. *Hum. Genet.* **98**, 491–496 (1996).
- Field, L. L., Tobias, R. & Magnus, T. A locus on chromosome 15q26 (IDDM3) produces susceptibility to insulin-dependent diabetes mellitus. *Nat. Genet.* **8**, 189–194 (1994).
- Blacker, D., Bertram, L., Saunders, A. J., Moscarillo, T. J., Albert, M. S. & Wiener, H. *et al.* Results of a high-resolution genome screen of 437 Alzheimer's disease families. *Hum. Mol. Genet.* **12**, 23–32 (2003).
- Susi, M., Holopainen, P., Mustalahti, K., Mäki, M. & Partanen, J. Candidate gene region 15q26 and genetic susceptibility to coeliac disease in Finnish families. *Scand. J. Gastroenterol.* **36**, 372–374 (2001).

- 15 Olsson, M., Olsson, B., Jacobson, P., Thelle, D. S., Björkegren, J., Walley, A. *et al*. Expression of the selenoprotein S (SELS) gene in subcutaneous adipose tissue and SELS genotype are associated with metabolic risk factors. *Metabolism* **60**, 114–120 (2011).
- 16 Sutherland, A., Kim, D.-H., Relton, C., Ahn, Y.-O. & Hesketh, J. Polymorphisms in the selenoprotein S and 15-kDa selenoprotein genes are associated with altered susceptibility to colorectal cancer. *Genes Nutr.* **5**, 215–223 (2010).
- 17 Méplan, C., Hughes, D. J., Pardini, B., Naccarati, A., Soucek, P., Vodickova, L. *et al*. Genetic variants in selenoprotein genes increase risk of colorectal cancer. *Carcinogenesis* **31**, 1074–1079 (2010).
- 18 Shibata, T., Arisawa, T., Tahara, T., Ohkubo, M., Yoshioka, D., Maruyama, N. *et al*. Selenoprotein S (SEPS1) gene -105G> A promoter polymorphism influences the susceptibility to gastric cancer in the Japanese population. *BMC Gastroenterol.* **9**, 2 (2009).
- 19 Moses, E. K., Johnson, M. P., Tømmerdal, L., Forsmo, S., Curran, J. E., Abraham, L. J. *et al*. Genetic association of preeclampsia to the inflammatory response gene *SEPS1*. *Am. J. Obstet. Gynecol.* **198**, 336.e1–336.e5 (2008).
- 20 Alanne, M., Kristiansson, K., Auro, K., Silander, K., Kuulasmaa, K., Peltonen, L. *et al*. Variation in the selenoprotein S gene locus is associated with coronary heart disease and ischemic stroke in two independent Finnish cohorts. *Hum. Genet.* **122**, 355–365 (2007).
- 21 Santos, L. R., Durães, C., Mendes, A., Prazeres, H., Alvelos, M. I., Moreira, C. S. *et al*. A polymorphism in the promoter region of the selenoprotein S gene (*SEPS1*) contributes to Hashimoto's thyroiditis susceptibility. *J. Clin. Endocrinol. Metab.* **99**, E719–E723 (2014).
- 22 Guan, F., Zhang, C., Wei, S., Zhang, H., Gong, X., Feng, J. *et al*. Association of PDE4B polymorphisms and schizophrenia in Northwestern Han Chinese. *Hum. Genet.* **131**, 1047–1056 (2012).
- 23 Guan, F., Wei, S., Feng, J., Zhang, C., Xing, B., Zhang, H. *et al*. Association study of a new schizophrenia susceptibility locus of 10q24.32-33 in a Han Chinese population. *Schizophrenia Res.* **138**, 63–68 (2012).
- 24 Guan, F., Wei, S., Zhang, C., Zhang, H., Zhang, B., Xing, B. *et al*. A population-based association study of 2q32.3 and 8q21.3 loci with schizophrenia in Han Chinese. *J. Psychiatr. Res.* **47**, 712–717 (2013).
- 25 Guan, F., Zhang, B., Yan, T., Li, L., Liu, F., Li, T. *et al*. MIR137 gene and target gene CACNA1C of miR-137 contribute to schizophrenia susceptibility in Han Chinese. *Schizophrenia Res.* **152**, 97–104 (2014).
- 26 Sham, P. & Curtis, D. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann. Hum. Genet.* **59**, 97–105 (1995).
- 27 Zhao, J. H. 2LD GENECOUNTING and HAP: computer programs for linkage disequilibrium analysis. *Bioinformatics* **20**, 1325–1326 (2004).
- 28 Curtis, D., Knight, J. & Sham, P. Program report: GENECOUNTING support programs. *Ann. Hum. Genet.* **70**, 277–279 (2006).
- 29 Zhao, J. H., Lissarrague, S., Essioux, L. & Sham, P. C. GENECOUNTING: haplotype analysis with missing genotypes. *Bioinformatics* **18**, 1694–1695 (2002).
- 30 Curtis, D. & Xu, K. Minor differences in haplotype frequency estimates can produce very large differences in heterogeneity test statistics. *BMC Genet.* **8**, 38 (2007).
- 31 Schaid, D. J., Rowland, C. M., Tines, D. E., Jacobson, R. M. & Poland, G. A. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am. J. Hum. Genet.* **70**, 425–434 (2002).
- 32 Menashe, I., Rosenberg, P. S. & Chen, B. E. PGA: power calculator for case-control genetic association analyses. *BMC Genet.* **9**, 36 (2008).
- 33 Seiderer, J., Dambacher, J., Kühnlein, B., Pfennig, S., Konrad, A., Török, H. P. *et al*. The role of the selenoprotein S (SELS) gene -105G>A promoter polymorphism in inflammatory bowel disease and regulation of SELS gene expression in intestinal inflammation. *Tissue Antigens* **70**, 238–246 (2007).
- 34 Hyrenbach, S., Pezzini, A., Del Zotto, E., Giossi, A., Lichy, C., Kloss, M. *et al*. No association of the -105 promoter polymorphism of the selenoprotein S encoding gene *SEPS1* with cerebrovascular disease. *Eur. J. Neurol.* **14**, 1173–1175 (2007).
- 35 Korostishevsky, M., Kaganovich, M., Cholostoy, A., Ashkenazi, M., Ratner, Y., Dahary, D. *et al*. Is the G72/G30 locus associated with schizophrenia? Single nucleotide polymorphisms, haplotypes, and gene expression analysis. *Biol. Psychiatr.* **56**, 169–176 (2004).
- 36 Chen, J., Zheng, H., Bei, J.-X., Sun, L., Jia, W.-h., Li, T. *et al*. Genetic structure of the Han Chinese population revealed by genome-wide SNP variation. *Am. J. Hum. Genet.* **85**, 775–785 (2009).
- 37 Karunasinghe, N., Han, D. Y., Zhu, S., Yu, J., Lange, K., Duan, H. *et al*. Serum selenium and single-nucleotide polymorphisms in genes for selenoproteins: relationship to markers of oxidative stress in men from Auckland, New Zealand. *Genes Nutr.* **7**, 179–190 (2012).