Fine mapping MHC associations in Graves’ disease and its clinical subtypes in Han Chinese

Xun Chu,1,2,3 Minjun Yang,3 Zhen-Ju Song,4 Yan Dong,1,2 Chong Li,3 Min Shen,3 Yong-Qiang Zhu,3 Huai-Dong Song,5 Sai-Juan Chen,5 Zhu Chen,5 Wei Huang3

ABSTRACT

Background The classical human leukocyte antigen (HLA) genes were the most important genetic determinant for Graves’ disease (GD). The aim of the study was to fine map causal variants of the HLA genes.

Methods We applied imputation with a Pan-Asian HLA reference panel to thoroughly investigate the major histocompatibility complex (MHC) associations with GD down to the amino acid level of classical HLA genes in 1468 patients with GD and 1490 controls of Han Chinese.

Results The strongest finding across the HLA genes was the association with HLA-DPB1 position 205 (P<2.8×10−12). HLA-DPA1*02:02 was the strongest association among the classical HLA alleles, which was in perfect linkage disequilibrium with HLA-DPβ1 residue Met11 (OR=1.90, P=1.76×10−11). Applying stepwise conditional analysis, we identified amino acid position 205 in HLA-DPB1, position 66 and 99 in HLA-B and position 28 in HLA-DRB1 explain majority of the MHC association to GD risk. We further evaluated risk of two clinical subtypes of GD, namely persistent thyroid stimulating hormone receptor antibody -positive (pTRAb−) group and ‘non-persistent TRAb positive’ (pTRAb−) group after antithyroid drug therapy. We found that HLA-B residues Lys66-Arg69-Val76 could drive pTRAb− GD risk alone, while HLA-DPB1 position 205, HLA-B position 69 and 199 and HLA-DRB1 position 28 drive pTRAb+ GD risk. The risk heterogeneity between pTRAb+ and pTRAb− GD might be driven by HLA-DPα1 Met11.

Conclusions Four amino acid positions could account for the associations of MHC with GD in Han Chinese. These distinct HLA association patterns indicated the two subtypes have distinct molecular mechanisms of pathogenesis.

INTRODUCTION

Graves’ disease (GD) is a common autoimmune disease, and the major disease autoantigen, thyroid-stimulating hormone receptor (TSHR) autoantibody (TRAb), was found to underlie its pathogenesis. Conservative therapy with antithyroid drug (ATD) is proven to be effective in achieving euthyroidism in Graves’ patients; however, 20%–50% of patients relapse after therapy withdrawal.1,2 TRAb-positive patients are more prone to relapses than TRAb-negative patients at the time of ATD withdrawal, and it was found that these two subtypes have different genetic architecture. The genetic contribution of GD was estimated to be 79%, and it was found that the human major histocompatibility complex (MHC), human leukocyte antigen (HLA), was the most predominant genetic factor.3,4 However, the high degree of linkage disequilibrium (LD) and genetic diversity challenges the determination of the causal functional variants in this region.

In Caucasian population, strong association with HLA-DR3 (DRB1*03:01) has frequently been reported, which makes the DR3 haplotype as the predisposing factor to GD.4,5 The 74th amino acids located in the peptide-binding groove of the HLA-β1 chain was found to determine the association of HLA-DR3 with GD susceptibility in Caucasians.6

In addition, the strongest association for GD in Caucasian has been reported with the HLA-class I allele, HLA-C*07.7 Recently, our understanding of the landscape of MHC association was further extended by studies in Asian population.5,10–12 Chen et al.12 performed direct and comprehensive genotyping of six classical HLA loci to four-digit resolution in Han Chinese population in Taiwan. Our previous genome-wide association study (GWAS) in Han Chinese population from China revealed three SNPs explained for most of association of MHC with GD.5 More recently, Okada et al.13 applied HLA imputation to GWAS data for GD in Japanese and found that amino acid polymorphisms of multiple class I and class II HLA genes independently contribute to disease risk. The strongest impact was observed in HLA-DPB1 in all of the above three Asian studies.11

Recent studies on HLA associations in GD and other diseases suggest that amino acid variants were the major driver conferring risk for diseases rather than the classical alleles.11,13 In this study, we conducted a fine-mapping study assessing the GD associations with SNP, HLA amino acid variants, two-digit HLA alleles and four-digit HLA alleles simultaneously in Han Chinese populations by the imputation approach in our previous GWAS data.5 Then, we explored the association patterns of HLA variants that distinguished the risks of different clinical subtypes.

METHODS

Study population and clinical characteristics

All samples were recruited from Han Chinese population through collaboration with multiple hospitals in China. All the enrolled subjects provided written informed consent for participation in the study approved by the local institutional review board. In
total, 1536 patients with GD and 1516 sex-matched controls were recruited for genotyping in GWAS. Demographic information was shown in online supplementary table S1. Patients were divided into different subsets by levels of TRAb after ATD and with or without ophthalmopathy. Plasma levels of TRAb in all patients with GD who had been treated with ATD for ≥1 year were remeasured by quantitative ELISA (RSR Limited, UK). GD patients with TRAb levels ≥1.5 U/L were defined as ‘persistent TRAb positive’ (pTRAb+) and those with TRAb levels <1.5 U/L were defined as ‘non-persistent TRAb positive’ (pTRAb−). Of the 1536 patients with GD, 997 patients belong to the pTRAb+ group and 410 patients belong to the pTRAb− group. There were 602 patients diagnosed with GD orbitopathy and 866 patients without orbitopathy.

Genotyping was performed using Illumina Human660-Quad BeadChips. The details of quality control (QC) steps have been described in previous study. Briefly, after stringent QC, the genotypes of 486 049 SNPs in 1468 patients with GD and 1490 controls were kept for association analysis with an overall call rate of 98%, a minor allele frequency over 1% and the p values for Hardy-Weinberg equilibrium test in the controls over 10−6. Principal component analysis (PCA) and multidimensional scaling analysis provided minimal evidence for population stratification in the current sample collection (L pca=1.02; online supplementary figure 1).

**HLA imputation**

We extracted 2676 SNP genotypes from −25 to −35 Mbp at chromosome 6 (Hg19, build 37) harbouring the extended MHC region from the GWAS data after QC. Imputation of two-digit and four-digit classical HLA alleles and amino acid polymorphisms of the HLA genes along with the SNPs that were not genotyped in the GWAS was performed using the SNP2HLA tool (https://www.broadinstitute.org/mpg/sn2hla/) and the Pan-Asian reference panel. We applied postimputation QC criteria of minor allele frequency of ≥1% and PLINK INFO of ≥0.8 for the association analysis.

**Association analysis**

In the 5264 variants past quality control, a total of 84 amino acid residues for the eight classical HLA genes (HLA-A, HLA-B, HLA-C, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1 and HLA-DRB1). Association analysis revealed that 1381 markers met the study-wide significance (p≤9.35×10−6; online supplementary table S2).

**Associations of classical HLA genes with GD**

We first investigated the association-specific amino acid residues and classical alleles of each classical HLA gene with GD susceptibility (online supplementary table S2). The most significant amino acid residue at HLA-A was Arg151 (OR=0.38, p binary =1.71×10−13; online supplementary figure 2A), and

---

**Table 1** Associations of HLA amino acid residues with Graves’ disease risk in Han Chinese

<table>
<thead>
<tr>
<th>Genes</th>
<th>Effective residue</th>
<th>PLINK INFO</th>
<th>ERF (%)</th>
<th>Cases (n=1468)</th>
<th>Controls (n=1490)</th>
<th>Europeans*</th>
<th>OR (95% CI)</th>
<th>P values</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-A</td>
<td>Arg151</td>
<td>0.99</td>
<td>19.6</td>
<td>29.6</td>
<td>11.5</td>
<td>0.58</td>
<td>(0.51 to 0.65)</td>
<td>1.71×10−18</td>
</tr>
<tr>
<td>HLA-B</td>
<td>Lys66-Arg69-Val76</td>
<td>0.95</td>
<td>14.1</td>
<td>6.5</td>
<td>–</td>
<td>2.38</td>
<td>(2.00 to 2.87)</td>
<td>5.81×10−21</td>
</tr>
<tr>
<td>HLA-C</td>
<td>Tyr116</td>
<td>1.03</td>
<td>37.9</td>
<td>27.8</td>
<td>10.9</td>
<td>1.59</td>
<td>(1.41 to 1.76)</td>
<td>1.15×10−15</td>
</tr>
<tr>
<td>HLA-DPA1</td>
<td>Met1</td>
<td>0.92</td>
<td>59.5</td>
<td>44.8</td>
<td>3.5</td>
<td>1.90</td>
<td>(1.70 to 2.12)</td>
<td>1.76×10−11</td>
</tr>
<tr>
<td>HLA-DPB1</td>
<td>Leu35</td>
<td>0.93</td>
<td>50.9</td>
<td>36.6</td>
<td>9.4</td>
<td>1.80</td>
<td>(1.72 to 2.15)</td>
<td>8.73×10−10</td>
</tr>
<tr>
<td>HLA-DQA1</td>
<td>Lys47-His52-Leu54</td>
<td>1.04</td>
<td>7.0</td>
<td>15.2</td>
<td>12.9</td>
<td>0.42</td>
<td>(0.36 to 0.51)</td>
<td>1.93×10−12</td>
</tr>
<tr>
<td>HLA-DQB1</td>
<td>Ala-10</td>
<td>0.98</td>
<td>86.9</td>
<td>78.9</td>
<td>82.4</td>
<td>1.78</td>
<td>(1.55 to 2.04)</td>
<td>1.18×10−15</td>
</tr>
<tr>
<td>HLA-DRB1</td>
<td>Asp57</td>
<td>0.96</td>
<td>72.7</td>
<td>63.6</td>
<td>54.7</td>
<td>1.53</td>
<td>(1.37 to 1.71)</td>
<td>1.79×10−13</td>
</tr>
<tr>
<td></td>
<td>Gly11-Tyr13-Lys14-Gln25-Leu30-Gln74</td>
<td>1.04</td>
<td>7.1</td>
<td>15.2</td>
<td>10.3</td>
<td>0.42 (0.36 to 0.51)</td>
<td>1.93×10−21</td>
<td></td>
</tr>
</tbody>
</table>

*Erf, effective residue frequency; HLA, human leucocyte antigen.
HLA-A*02:07 was the top classical HLA-A allele tagged by Cys99 (OR=2.06, $P_{\text{binary}}=2.07 \times 10^{-12}$; table 1, figure 1). The most significantly associated residue of HLA-B gene is Arg69 (OR=2.38, $P_{\text{binary}}=5.81 \times 10^{-21}$), which was in perfect LD with Lys66 and Val76 ($r^2=1$; table 1, online supplementary figure 2B). The most associated HLA-B classical allele was HLA-B*46:01 (OR=2.38, $P_{\text{binary}}=8.78 \times 10^{-21}$; table 2), which could be tagged by residues Arg69, Lys66 and Val76. Positions 66, 69 and 76 of the HLA-B chain are all located on the α-helix wall of peptide binding pocket (figure 1). The top association signal at HLA-C was residue Tyr116 (OR=1.59, $P_{\text{binary}}=1.15 \times 10^{-15}$) and HLA-C*01:02 was the most associated HLA-C classical allele (OR=1.84, $P_{\text{binary}}=5.40 \times 10^{-15}$; table 2, online supplementary figure 2C).

HLA-DPα1 residue Met11 is the peak binary signal of HLA-DPA1 as well as the most significant residues the MHC region (OR=1.76×10^{-31}, $P_{\text{binary}}=1.90$; online supplementary figure 2D), which tagged HLA-DPA1*02:02 (tables 1 and 2). HLA-DPα1 position 11 is located within the β-sheet floor of peptide binding pocket (figure 1). The top residue association signal at HLA-DPB1 corresponded to Leu35 (OR=1.80, $P_{\text{binary}}=8.73 \times 10^{-22}$; online supplementary figure 2E). HLA-DPB1 Glu55 was in strong LD with Leu35 ($r^2=0.96$) and had a similar association (OR=1.91, $P_{\text{binary}}=1.09 \times 10^{-26}$). HLA-DPB1 Leu35 was the second significant residues across the MHC region, which exhibits moderate LD with HLA-DPα1 residue Met11 ($r^2=0.61$). The association of HLA-DPB1 Leu35 was still significant, though not meet study-wide significance, when conditioning on HLA-DPα1 Met11 ($P=3.58 \times 10^{-5}$). Similarly, HLA-DPα1 Met11 was significantly associated with GD when conditioning on HLA-DPB1 Leu35 ($P=2.13 \times 10^{-5}$).

It is obvious these two amino acids represent independent associations. The most significant four-digit allele at HLA-DPB1 was HLA-DPB1*05:01 (OR=1.72, $P_{\text{binary}}=1.73 \times 10^{-26}$), which was

---

**Figure 1** Three-dimensional ribbon models of HLA amino acid positions associated with GD risk. The protein structures of HLA-A, HLA-B, HLA-C, HLA-DP, HLA-DQ and HLA-DR are based on Protein Data Bank (PDB) entries 2XP2, 2BVP, 4N76, 3LQZ, 4Z7W and 3PDO, respectively. GD associated amino acid positions identified by the association analysis are highlighted. This figure was prepared using UCSF Chimera.

| Table 2 Associations of classical HLA alleles with Graves’ disease susceptibility in Han Chinese |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| HLA allele                     | PLINK INFO      | Cases (n=1468)  | Controls (n=1490) | Caucasian populations | OR (95% CI) | $P$ values   |
|--------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| HLA-A*02:07                    | 0.90            | 9.7             | 4.9             | 0.007           | 2.10 (1.70 to 2.59) | $2.07 \times 10^{-12}$ |
| HLA-B*46:01                    | 0.94            | 14.1            | 6.5             | 0.010           | 2.38 (1.99 to 2.86) | $8.78 \times 10^{-21}$ |
| HLA-C*01:02                    | 1.02            | 18.4            | 10.9            | 3.432           | 1.83 (1.57 to 2.12) | $5.40 \times 10^{-15}$ |
| HLA-DPA1*02:02                 | 0.92            | 59.5            | 44.8            | 6.324           | 1.90 (1.70 to 2.12) | $1.76 \times 10^{-15}$ |
| HLA-DPB1*05:01                 | 0.92            | 44.0            | 31.3            | 1.927           | 1.90 (1.69 to 2.14) | $1.73 \times 10^{-15}$ |
| HLA-DQA1*02:03                 | 1.04            | 7.0             | 15.2            | 12.658          | 0.43 (0.36 to 0.51) | $1.93 \times 10^{-14}$ |
| HLA-DQB1*02:01                 | 1.00            | 10.9            | 17.8            | 15.437          | 0.57 (0.49 to 0.66) | $2.31 \times 10^{-13}$ |
| HLA-DRB1*07:01                 | 1.04            | 7.1             | 15.3            | 13.337          | 0.43 (0.36 to 0.51) | $2.49 \times 10^{-13}$ |

*The allele frequency of Caucasian populations was calculated from the average frequencies of all samples of Caucasian ethnic origin from database: http://www.allelefrequencies.net. HLA, human leucocyte antigen.
found as the most associated HLA marker in our previous GWAS study and other studies.\textsuperscript{5, 11, 12} HLA-DPB1*05:01 was in strong LD with Leu35 (r\textsuperscript{2}=0.76). The association of HLA-DPB1*05:01 was not significant when conditioning on HLA-DP\textbeta{}1 Leu35 (p=0.09), whereas HLA-DPB1 Leu35 was significantly associated with GD when conditioning on HLA-DP\textbeta{}1*05:01 (p=7.12×10\textsuperscript{-6}).

The top residue association signal at HLA-DQA1 was three residues in completely perfect LD (Lys47, His52 and Leu54; OR=0.42, p=1.93×10\textsuperscript{-6} for Val205 and OR=0.40, p=1.80×10\textsuperscript{-7} for Met205) at the amino acid position 205 of HLA-DP\textbeta{}1, followed by HLA-DP\textalpha{}1 position 205 (Pomnibus=2.48×10\textsuperscript{-6}). The most associated association across all variants corresponded to the HLA-DP\textbeta{}1*02:01 (table 1). Ala-10 was the strongest associated residues of HLA-DQB1 (OR=1.78, P\textsubscript{binary}=1.18×10\textsuperscript{-15}; online supplementary figure 2G) and the most significant four-digit allele at HLA-DQB1 was HLA-DQB1*02:01 (OR=0.57, p=2.31×10\textsuperscript{-13}). Interestingly, HLA-DQB1 Asp57 is the most protective allele for type 1 diabetes, and it is interesting that it was found to increase the risk of GD (OR=1.53, p=1.79×10\textsuperscript{-15}).\textsuperscript{20}

Six HLA-DRB1 residues (Gly11, Tyr13, Lys14, Gln25, Leu30 and Gln74; OR=0.42, p=1.93×10\textsuperscript{-21}) uniquely encoded by HLA-DRB1*07:01 (OR=0.43, p=2.49×10\textsuperscript{-22}) had the strongest associations with GD among the HLA-DRB1 alleles and amino acid residues (tables 1 and 2; online supplementary figure 2H).

**Associations of HLA amino acid positions and dependence analysis**

We also tested the influence of the polymorphic amino acid positions by means of an omnibus test (online supplementary table S2). The most associated association across all variants corresponded to the HLA-DP\textbeta{}1 position 205 (P\textsubscript{omnibus}=2.48×10\textsuperscript{-15}), followed by HLA-DP\textalpha{}1 position 11 and HLA-DP\textbeta{}1 position 35. There are three possible residues (Val205, Met205 and Deletion205) at the amino acid position 205 of HLA-DP\textbeta{}1. Among them, Met205 showed risk effects for GD (OR=1.88, P\textsubscript{binary}=5.81×10\textsuperscript{-26}), while Val205 and Deletion205 were protective (OR=0.72, P\textsubscript{binary}=1.99×10\textsuperscript{-7} for Val205 and OR=0.40, P\textsubscript{binary}=5.17×10\textsuperscript{-7} for Deletion205).

We applied stepwise conditional regression analysis to find independent HLA amino acid positions confer independent risks on GD (table 3; figure 2). When conditioning on the top associated HLA-DP\textbeta{}1 position 205, HLA-B position 66 showed the most significant independent evidence of association (conditional P\textsubscript{omnibus}=1.31×10\textsuperscript{-16}), with HLA-B position 69 demonstrating a similar level of evidence (conditional P\textsubscript{omnibus}=1.41×10\textsuperscript{-16}). Conditioning on positions HLA-DP\textbeta{}1 position 205 and HLA-B position 66 demonstrated an independent association of HLA-B position 199 (conditional P\textsubscript{omnibus}=8.65×10\textsuperscript{-16}). When conditioning on HLA-DP\textbeta{}1 position 205 and HLA-B position 66 and 199, we detected the most significant independent association at HLA-DRB1 position 28 (conditional P\textsubscript{omnibus}=1.29×10\textsuperscript{-3}).

No significant associations were observed after adjusting for the effects of HLA-DP\textbeta{}1 position 205, HLA-B position 66 and 199 and HLA-DRB1 position 28 (conditional P\textsubscript{omnibus}>9.35×10\textsuperscript{-6}), suggesting that the combination of these amino acid positions explain the majority of the HLA risk in Chinese.

To account the high LD of the HLA region, we further performed a conditional haplotype test combining HLA-DP\textbeta{}1 205 and HLA-DRB1 28 in the class II region as well as HLA-B 66 and 199 (online supplementary table S3). Although different haplotypes surpassed the study-wide significance threshold, none of them showed a significant improvement in the association observed for the associated amino acids of HLA-DP\textbeta{}1 205, HLA-DRB1 28 and HLA-B 66 and 199 independently. HLA-DRB1 28 has independent haplotype effect after we controlled for HLA-DP\textbeta{}1 205 (p=1.78×10\textsuperscript{-8}). HLA-B 199 has independent haplotype effect after we controlled for HLA-B 66 (p=3.27×10\textsuperscript{-5}). No evidence of epistatic interactions between known non-HLA risk loci\textsuperscript{2} and any of the HLA alleles described here.

**Associations of HLA genes with GD clinical subtypes**

Evaluated levels of hyrotropin receptor antibodies (TRAb) after ATD therapy are the predictors of relapse. We further evaluated associations of HLA genes with risk for pTRAb+ and pTRAb−GD, respectively.

<table>
<thead>
<tr>
<th>HLA variant</th>
<th>Case (n=1468)</th>
<th>Control (n=1490)</th>
<th>OR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-DP\textbeta{}1 amino acid position 205</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>45.1</td>
<td>32.4</td>
<td>1.88 (1.67 to 2.12)</td>
<td>1.18×10\textsuperscript{-2}</td>
</tr>
<tr>
<td>Deletion</td>
<td>3.8</td>
<td>9.1</td>
<td>0.40 (0.32 to 0.5)</td>
<td>3.22×10\textsuperscript{-15}</td>
</tr>
<tr>
<td>Val</td>
<td>51.1</td>
<td>58.5</td>
<td>(reference)</td>
<td></td>
</tr>
<tr>
<td>HLA-B amino acid position 66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>14.1</td>
<td>6.5</td>
<td>2.39 (2.00 to 2.87)</td>
<td>4.42×10\textsuperscript{-11}</td>
</tr>
<tr>
<td>Asn</td>
<td>5.0</td>
<td>6.5</td>
<td>0.76 (0.61 to 0.95)</td>
<td>6.17×10\textsuperscript{-2}</td>
</tr>
<tr>
<td>Ile</td>
<td>80.8</td>
<td>87.0</td>
<td>(reference)</td>
<td></td>
</tr>
<tr>
<td>HLA-B amino acid position 199</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>2.2</td>
<td>5.3</td>
<td>0.40 (0.30 to 0.54)</td>
<td>1.80×10\textsuperscript{-7}</td>
</tr>
<tr>
<td>Ala</td>
<td>97.8</td>
<td>94.7</td>
<td>(reference)</td>
<td></td>
</tr>
<tr>
<td>HLA-DRB1 amino acid position 28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>18.9</td>
<td>29.5</td>
<td>0.57 (0.50 to 0.64)</td>
<td>1.55×10\textsuperscript{-7}</td>
</tr>
<tr>
<td>His</td>
<td>15.3</td>
<td>11.3</td>
<td>1.49 (1.20 to 1.62)</td>
<td>2.73×10\textsuperscript{-1}</td>
</tr>
<tr>
<td>Asp</td>
<td>64.3</td>
<td>58.7</td>
<td>(reference)</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{4}Amino acid residues with frequency ≥0.005 in the controls are shown.
\textsuperscript{1}Obtained from the multivariate regression model that included position 205 in HLA-DP\textbeta{}1, position 66 and 99 in HLA-B and position 28 in HLA-DR\textbeta{}1 identified by stepwise regression analysis.
Figure 2  Regional association plots of HLA loci associated with GD in Han Chinese. Each dot represents the –log10 P of the variants, including SNPs, classical HLA alleles and amino acid polymorphisms encoded by the HLA genes. Physical positions are based on NCBI build 36 of the human genome. The green horizontal dashed line represents p=9.35×10^{-6}. (A) The top signal was from HLA-DPβ1 position 205. (B) After conditioning on HLA-DPβ1 position 205, the top signal was from HLA-B position 66. (C) After conditioning on HLA-DPβ1 position 205 and HLA-B position 66, the top signal was from HLA-B position 199. (D) After conditioning on HLA-DPβ1 position 205 and HLA-B position 66 and 199, the most significant independent association was at HLA-DRβ1 position 28. (E) No study-wide significant associations were observed after adjusting for the effects of HLA-DPβ1 position 205, HLA-B position 66 and 199 and HLA-DRβ1 position 28 (conditional $P_{\text{corr}}$<9.35×10^{-6}). GD, Graves’ disease; MHC, major histocompatibility complex.
The most strongly associated variant with pTRAb+ GD risk was observed at HLA-DPβ1 positions 205 (997 pTRAb+ patients vs 1490 controls; \( p=7.52 \times 10^{-15} \), online supplementary table S4). The effect sizes of HLA-DPβ1 position 205 tended to be greater in the pTRAb+ GD patients (997 pTRAb+ patients vs 1490 controls) than in the total of patients with GD (1468 patients vs 1490 controls). Then, we put HLA-DPβ1 position 205 in the regression model as the best locus and employed stepwise forward logistic regression to identify independent markers driving pTRAb+ GD risk. The analysis revealed that the HLA association with pTRAb+ GD could be explained by four amino acid positions, namely HLA-DPβ1 position 205, HLA-B position 66 and 199 and HLA-DRB1 position 74 with conditional \( P \) omnibus values of 7.91\( \times 10^{-22} \), 1.79\( \times 10^{-19} \), 1.51\( \times 10^{-9} \) and 3.04\( \times 10^{-7} \), respectively.

The strongest association with pTRAb+ GD risk was at HLA-B residues Lys66-Arg69-Val76 (410 pTRAb+ patients vs 1490 controls; \( OR=2.64, \) \( P_{\text{binary}}=9.73 \times 10^{-11} \)) uniquely encoded by HLA-B*46:01, which could explain the disease risk in pTRAb patients alone (online supplementary table S5).

When we directly assessed comparative risk between pTRAb+ and pTRAb− subjects, we found the lowest \( P \) value of the nominal association signal at HLA-DPα1 Met11 (\( P_{\text{binary}}=1.12 \times 10^{-7} \) for 410 pTRAb− patients vs 997 pTRAb+ patients). After conditioning on HLA-DPα1 Met11, we observed no significant association in the MHC region (conditional \( p>9.35 \times 10^{-10} \)). HLA-DPα1 Met11 increased pTRAb+ susceptibility in comparison with pTRAb− susceptibility (OR=1.50, 95%CI 1.28 to 1.77). Examining the classical alleles, we noted that HLA-DPA1*02:02 demonstrated the lowest \( P \) value for pTRAb+ versus pTRAb− association (\( P_{\text{binary}}=1.12 \times 10^{-7} \)) similar significantly with HLA-DPα1 Met11.

We also investigated association of HLA with GD orbitopathy. The comparison of HLA variation between GD patients with orbitopathy and without orbitopathy showed no significant difference.

DISCUSSION

The HLA genes were responsible for the strongest association signals for GD susceptibility.14 15 Numerous studies have tried to clarify the biological mechanism of classical HLA genes underlying disease susceptibility and found some molecular clues. In the current study, we found that combinations of amino acid polymorphisms in multiple class I and II HLA genes explained the majority of risk in the MHC region for GD in Han Chinese.

We observed the top association signal at HLA-DPβ1 position 205, which was among the most associated variants in Japanese.11 In Asians, the HLA-DPβ1 locus was constantly not covered in association analysis of HLA genes with GD previously, and it was overlooked until the report in 1992.21 Association of HLA-DPβ1*05:01 with GD was first reported in Japanese21 and later replicated in several independent sample set of Japanese and Chinese populations.10 12 22 In our previous GWAS for GD, rs2281388 conferred the top risk with susceptibility in Chinese population. Since rs2281388 was a tagging SNP that could predict HLA-DPβ1*05:01 (\( r^2=0.90 \)), we suggested the association of HLA-DPβ1*05:01 with GD susceptibility in previous report. The frequency of HLA-DPβ1*05:01 is mostly less than 5% in Caucasians, and the association of HLA-DPβ1*05:01 with GD has not been reported in Caucasians.12 22 It could be speculated that HLA-DPβ1 contributed much less to autoimmune disease in Caucasians than in Asians.

In current analysis, HLA-DPA1*02:02 showed the strongest association among the HLA classical alleles with GD in Han Chinese population, and HLA-DPB1*05:01 was the second significant HLA classical allele. Although HLA-DPA1*02:02 was not the most associated HLA alleles in recent Japanese study, it still belonged to the top association signals.11 The frequency of HLA-DPA1*02:02 is 44.8% in the current control population, while it is about 3%–6% in Caucasians.23 To our knowledge, no association of HLA-DPA1*02:02 and HLA-DPB1*05:01 with GD has been reported in Caucasians. Since only six loci are routinely typed by laboratories, that is, HLA-A, HLA-B and HLA-C for class I and HLA-DRB1, HLA-DQB1, HLA-DPB1 for class II, the HLA-DPA1 locus was constantly neglected in previous studies. Several studies that reported the association of HLA-DPB1*05:01 with GD did not cover the HLA-DPA1 locus in their study design including several recent comprehensive analysis of HLA association with GD in Chinese and Japanese.10 12 22 Fortunately, the HLA-DPA1 variation was included in the commercial genome-wide SNP arrays, and it showed association with chronic hepatitis B virus (HBV) infection and related diseases in Asians.24 25 Special attention should be paid to the attribution of HLA-DPA1 to susceptibility of autoimmune diseases in Asian populations in future studies.

We observed the peak residue signal at HLA-DPα1 Met11 in our sample set of Han Chinese, which conferred risk to GD susceptibility. The second leading signal is in HLA-DPβ1 residue Leu35, which exhibits moderate LD with HLA-DPα1 Met11 (\( r^2=0.52 \)). Conversely, HLA-DPβ1 Leu35 as the strongest association with GD in Japanese, while HLA-DPα1 Met11 was the second significant HLA classical allele.11 The association significances of HLA-DPα1 Met11 and HLA-DPβ1 Leu35 were similar, but conditional regression analysis showed that two amino acids conferred mutually independent risk to GD. The different significant level of HLA-DPα1 Met11 and HLA-DPβ1 Leu35 association with GD between Chinese and Japanese might be caused by genetic heterogeneity or stochastic nature of sample collection. Undoubtedly, both HLA-DPA and HLA-DPB contribute substantially to the genetic architecture of GD in Asians.

HLA-B harbours independent association signals for GD risk, and HLA-B positions 66, 69 and 76 showed the top associated signals among the HLA-I gene. HLA-B Arg69, Lys66 and Val76 is coded by HLA-B*46:01. In our data, HLA-B*46:01 was the most significantly associated classical allele of HLA-I genes, which had good replications in Asians.26 HLA-B*46:01 is very rare in Caucasians, and the frequencies are mostly less than 1%. It should be noticed that the strongest association for Caucasian GD has been reported with HLA-C*07,27 which showed a weak association with GD in our sample-set (\( P=0.009 \), online supplementary table S2).

The third independent association corresponded to HLA-DRB1. HLA-DRB1 position β74 was a well-established genetic and functional candidate locus for GD risk,7 9 27 Arginine at position 74 of the HLA-DRβ1 chain conferred risk to GD susceptibility independently to HLA-DR3 in Caucasian5; however, the susceptible molecule in the current Chinese population was not arginine (\( p=0.20 \) for Arg74) but glutamic acid and leucine (\( p=1.16 \times 10^{-6} \) for Glu74; \( p=7.76 \times 10^{-6} \) for Leu74; online supplementary table S2). More importantly, the most effective molecule at position 74 in Chinese population was glutamine, which has a strong protective effect against GD (OR=0.43, \( p=3.14 \times 10^{-23} \)). Although our data suggested HLA-DRB1 position 28 as one of primary determinant of susceptibility to GD, we cannot rule out the possibility that this association is due to other variation at the HLA-DRB1 locus.

ATD is the first choice of treatment for patients with GD in China, and the relapse rates after therapy withdrawal are estimated to be 30%–50%. 

1. Evaluated levels of thyrotropin receptor antibodies (TRAb) after ATD therapy are predictors of relapse.

2. Since different genetic contributions associating with different phenotypes point to different etiopathogenesis, we illustrated that pTRAb+ GD and pTRAb− GD exhibit different genetic associations in our previous study. 

3. TSHR gene variation was associated with disease risk in pTRAb+ patients but not in pTRAb− patients. 

4. Our current data refined the HLA associations with these two clinical subtypes. Association of disease risk for pTRAb− patients with HLA genes was explained by combinations of amino acid polymorphisms in multiple class I and II HLA genes. These distinct HLA signatures underlying the two subtypes suggest that they are genetically heterogeneous and might have distinct molecular mechanisms of pathogenesis.

5. It is conceivable that viral or bacterial peptides and HLA class I molecules high-risk allele might have high affinity to HLA-class II molecule caused autoimmune attack of the thyroid gland in both pTRAb+ and pTRAb− patients with GD. 

6. After GD initiation, HLA class II molecules binding peptides derived from TSHR might influence the production of TRAb. Since the extracellular domain of human TSHR (TSHR-ECD) is shed into the circulation, TSHR-ECD is a preferentially immunogenic portion of TSHR. 

7. The complex of class-II and TSHR-ECD epitope is presented to CD4+ T cells. The activated CD4+ T cells help B cell differentiation into autoantibody producing plasma cells that produce TRAb. 

8. The TSHR-ECD epitope encoded by high-risk allele have high affinity to HLA-class II molecules high-risk alleles, and such complex probably stimulates persistent T cell response. 

9. Therefore, low-risk and high-risk HLA and TSHR alleles could affect the production of TRAb. It could be hypothesised that pTRAb+ GD patients with the risk alleles of HLA-II and/or TSHR have a high efficiency in antigen presentation causing persistent TRAb production and therefore has poor clinical outcome after ATD therapy.

10. In conclusion, our study revealed four amino acids could account for most HLA association with GD in Han Chinese. Two clinical subphenotypes corresponding to different responses after ATD therapy correlated with the inheritance of different sets of disease-risk HLA amino acids variants. Our study extends the knowledge of HLA-DPA1 association with GD association, and special attention should be paid to HLA-DPA1 variation in further study of HLA contribution to disease susceptibility in Asians.

Acknowledgements

The authors would like to thank all the clinical physicians for the sample collection in the collaborating hospitals and all participants in this research. We gratefully acknowledge Dr Zhihui Bu and Dr Meng Zhu for their valuable comments and kind help in improving the quality of the manuscript.

Contributors

XC designed research and wrote the paper, XC, MY, Z-S and CL analysed data, MS, YD, Y-Q Z, H-DS, S-JC, ZC and WH collected clinical samples and performed experiments.

Funding

This work was supported by the National Natural Science Foundation of China (31671317, 31471190, 81471840 and 31271343).

Competing interests

None declared.

Patient consent

Obtained.

Ethics approval

Ethical Committee of the Chinese National Human Genome Center at Shanghai.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data sharing statement

The SNP allele frequencies of the cases and controls could be shared with researchers by contacting corresponding author by email.

Open access

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES


Immunogenetics


