Evidence for unique association signals
in SLE at the CD28–CTLA4–ICOS locus
in a family-based study

D.S. Cunninghame Graham¹, A.K. Wong¹, N.J. McHugh², J.C. Whittaker³ and Timothy J. Vyse¹,*

¹Imperial College, Molecular Genetics and Rheumatology Section, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK, ²Royal National Hospital for Rheumatic Diseases, Bath BA1 1RL, UK and ³Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Received May 24, 2006; Revised and Accepted September 14, 2006

CD28, CTLA4 (cytotoxic T lymphocyte-associated protein 4) and ICOS (inducible T cell co-stimulator) are good candidate genes for systemic lupus erythematosus (SLE) because of their role in regulating T cell activation. CTLA4 inhibits CD28-mediated T cell activation. CTLA4 is expressed on CD4⁺ and CD8⁺ activated T cells, and also B cells, but CD28 and ICOS are largely restricted to T cells. An interval encompassing the CD28–CTLA4–ICOS locus on chromosome 2q33 was linked to lupus in two genome-wide linkage scans. This large family-based association study in 532 UK SLE families represents the first high-density genetic screen of 80 SNPs at this locus. There are seven haplotype blocks across the locus. In CTLA4, the strongest signal comes from two variants, located 2.1 kb downstream from the 3'-UTR. These polymorphisms, rs231726 (SNP 43) and rs231726 (SNP 44), are in complete linkage disequilibrium (LD) (r² = 1) and are associated with SLE P = 0.0008 (GH) and P = 0.01 (family-based association test). There is also a signal in the distal 3' flanking region of CTLA4/ICOS promoter (P = 0.003). There was no confirmation of published associations for SLE in the promoter or coding region of CTLA4. These SLE risk alleles are more distal than those identified in Graves' disease and are in LD with Graves' disease protective alleles identified in both of these regions of CTLA4 (Ueda et al. 2003). These factors suggest an SLE-specific pattern of association. The functional consequences of the associated polymorphisms are likely to influence CTLA4 expression, although it is possible that genetically modulated ICOS expression is involved in SLE susceptibility.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a multi-system complex autoimmune disease of uncertain aetiology, with a wide range of clinical symptoms (OMIM 152700). The disease is characterized by the production of autoantibodies against a diverse range of nuclear and cell surface autoantigens and the formation of immune complexes that contribute to end organ damage.

The overall disease frequency is 11–250 in 100 000 (1), which is dependent on gender and racial groupings, being less common in white Caucasians. There is a ~10-fold preponderance in females of childbearing age when compared with males (1). The genetic component to SLE is borne out by familial clustering studies (sibling risk ratio, λs = 20) (2), and there is 10-fold greater concordance in monozygotic twins when compared with dizygotic twins (3). To date, there are six genome-wide linkage scans, which have identified a number of suggestive susceptibility loci (4–10), encompassing a 300 kb interval at 2q33, which encodes the genes for CD28, CTLA4 (cytotoxic T lymphocyte-associated protein 4) and ICOS (inducible T cell co-stimulator) in order. CTLA4, CD28 and ICOS are all members of the immunoglobulin superfamily of receptors. CTLA4 is expressed on both activated CD4⁺ and CD8⁺ T cells and is an inhibitor of T cell activation. CTLA4 exerts its inhibitory activity through interaction with its ligands, B7.1 (CD80) and B7.2 (CD86), both of which are expressed on antigen presenting
cells (APCs), including dendritic cells, B cells and monocytes (11). Both CD28 and ICOS are capable of stimulating the development of Th2 and Th1 cells (12), but ICOS is only expressed on T cells following activation, whereas CD28 is constitutively expressed. CD28- but not ICOS-stimulated T cells produce IL2 (13). Since this cytokine stimulates T cell proliferation, it is thought that CD28 is important in the initial stages of T cell activation, whereas ICOS plays a more important role in the maintenance of T cell activation. CD28 shares the same ligands as CTLA4, although exhibits different binding affinities for them (14), whereas ICOS has its own unique ligand, ICOSL, expressed on APC. Therefore, the overall regulation of T cell activation involves a complex interplay between the down-regulatory activity of CTLA4 and its own unique ligand, ICOSL, expressed on APC. Therefore, the overall regulation of T cell activation involves a complex interplay between the down-regulatory activity of CTLA4 and its own unique ligand, ICOSL, expressed on APC. Therefore, the overall regulation of T cell activation involves a complex.

Associations for the CD28–CTLA4–ICOS gene cluster have been studied in several different autoimmune diseases, including SLE. In lupus, most of these reports have focused on CTLA4, investigating SNPs in the promoter, coding region, 3'-UTR and 3' flanking region, but in most cases, for each variant, there is only a single paper showing association at the level of P < 0.01. For the CT60 variant in the 3' flanking region of the gene, there is a previously published association in Spanish SLE patients (P = 0.003) and for the exon 1 missense variant +49A/G, there was association in Japanese families (P = 0.003) (15). However, for the promoter variant −1722, there are two papers showing association in Korean cases (P = 0.00003) (16) and Spanish cases (P = 0.0004) (17). This overall lack of consistency in the previous studies may be due to study design and maker selection, since the published data were from small-scale case–control studies, carried out in diverse populations and utilizing a limited number of polymorphisms.

There have been two published meta-analyses in SLE for a non-synonymous A/G Thr/Ala −17 change at position +49 in the CTLA4 leader peptide (+49AG). The earlier report (18), which used fixed effects meta-analysis in eight published studies, showed a significantly increased risk for SLE associated with the G allele of +49AG [OR is 1.23 (95% CI =1.08–1.41); P = 0.002]. However, a later report by Lee et al. (19), which included a further six studies in the meta-analysis and separated the analysis of the Caucasian and Asian studies, demonstrated no significant association in the Caucasian population for the +49A/G variant nor for two promoter polymorphisms, −1722 C/T (DIL92) and −318 C/T (DIL95), or a 106 bp allele of a dinucleotide repeat in the 3'-UTR. However, in the Asian population, there was association for the +49 G allele [OR is 1.246 (95% CI =1.057–1.469); P = 0.009] (19).

For the other genes in the CD28–CTLA4–ICOS cluster, there has been a single report investigating the genetic contribution of CD28 to SLE, studying variant in intron 3 of the gene, which showed no association.

In autoimmune diseases other than SLE, the main focus of association studies at the CD28–CTLA4–ICOS gene cluster has also been on CTLA4. There has been convincing data implicating CTLA4 in autoimmunity in a study of Graves' disease and confirmed in type I diabetes (T1D) and autoimmune thyroid disease (AITD), albeit with a weaker effect (20). This report showed that the strongest association signal arose from the CT60 polymorphism the 3' flanking region of CTLA4, although there was also association with the +49A/G variant in exon 1 of CTLA4. In rheumatoid arthritis (RA) association studies, there is also a significant association reported for the CT60 variant in North American/Swedish samples (P = 0.001) (21) and a suggestive association in Chinese Han RA cases (P = 0.007) (22). However, most of the other studies for these other autoimmune diseases are either negative or have P-values >0.01, including a further study in RA (23), one in Japanese AITD (24) and in coeliac disease (CD) (25,26), Addison’s disease (27), juvenile idiopathic arthritis (23) and atopic dermatitis (28). There have been additional investigations carried out for the +49AG polymorphism, but these other studies showed either no significant association, as for Addison’s disease (29,30), or no consistent pattern of association, as in T1D (31,32), CD (26,33) and Graves’ disease (34–36). Meta-analyses for SNPs in CTLA4 reported no association for +49A/G either in multiple sclerosis (37) or in RA (38). In CD, a meta-analysis was not performed on variants located within CTLA4, but on two microsatellites located upstream of CD28. The results showed that by pooling seven studies, there was significant association for two microsatellites located upstream of CD28, D2S2214 (P = 0.0014) and D2S116 (P = 0.0006) (39).

The linkage studies implicating the genomic interval containing CD28–CTLA4–ICOS as an important susceptibility locus for lupus were not limited to humans. In mice, the Bsx1 interval on mouse chromosome 1, which is linked to lupus in BxsB mice (40) is orthologous to human chromosome 2q33 that contains the CD28–CTLA4–ICOS gene cluster. Murine knockouts of all three genes in the CD28–CTLA4–ICOS gene cluster support a role for these genes in T cell activation in mice. Cita4-deficient mice develop a spontaneous T cell lymphoproliferation 5–6 days after birth. These T cells then infiltrate many organs, leading to multi-organ failure within 3–4 weeks (41). However, lymphoproliferation is reduced in the T cells from both Cd28-deficient (42,43) and Icos-deficient mice (44,45).

In this paper, we have undertaken a dense screen across the CD28–CTLA4–ICOS locus in UK SLE nuclear families. This family-based design provides good evidence for a unique association signal with CTLA4, which extends to ICOS.

RESULTS

Selection of informative markers across the CD28–CTLA4–ICOS locus

The 80 markers selected initially, as described in Materials and Methods, were genotyped in 332 UK SLE trios (Fig. 1A). Fourteen of these SNPs did not generate viable assays for the Sequenom platform or were monomorphic in the genotyping for SLE trios. The criteria for marker exclusion in this initial genotyping excluded five markers with <90% genotyping efficiency, five markers with >1% sporadic Mendelian errors and three further markers with a HWE in parental samples of P < 0.05. Consequently, a total of 53 markers were used for additional genotyping in 142 EC single parent
Figure 1. Genomic organization, haplotype architecture and haplotype-TDT across the human \(CD28-CTLA4-ICOS\) locus. (A) The genomic organization of the human \(CD28-CTLA4-ICOS\). The exons are marked with black boxes. It is not possible to visualize the 5'-UTRs by white boxes because of their small size. The 5'-untranslated region in \(CD28\) is 222 bp, in \(CTLA4\) is 155 bp and in \(ICOS\) is 35 bp. The 3'-UTRs, shown as white boxes, in \(CD28\) is 2921 bp, in \(CTLA4\) is 1148 bp and in \(ICOS\) is 1978 bp. The variants studied are numbered 1–80 under the gene diagram. The promoter of \(CD28\), the \(CD28-CTLA4\) intergenic region and the \(CTLA4-ICOS\) intergenic region are shown as dropped-down segments of the gene cluster. (B) The haplotype block structure across \(CD28-CTLA4-ICOS\), as constructed using Haploview, from parents in the 332 European parent-proband trios. Fifty-three markers were included in this analysis, with these SNP numbers shown across the top of the haplotypes, retaining the original numbering scheme shown in (A). The asterisks above the haplotypes indicate haplotype-tagging SNPs. There are seven haplotype blocks across the locus, labelled 1–7. The frequency of each haplotype is identified by a number to the right of the haplotype. Only common haplotypes at a frequency >3% are shown. The inter-block \(D^2\) is shown below the gap between adjacent haplotype blocks. The under-transmitted haplotypes in blocks 1 and 6 are outlined by a discontinuous box. The over-transmitted haplotype in block 4 is outlined by a continuous box. (C) A graph of \(-\log_{10}(P)\) against marker number for all 53 variants analysed across the \(CD28-CTLA4-ICOS\) gene cluster. The \(P\)-values were derived from GENEHUNTER-TDT analysis in 332 UK SLE parental-proband trios. The positions of the SNPs identified by the conditional analysis (Table 4) are marked with a white circle. The haplotype blocks are represented by horizontal bars and identified by numbers across the top of the graph. (D) Detailed description of the haplotype blocks across the locus. For each haplotype block, the length and separation of the haplotype blocks are shown, together with the location in the gene cluster and SNPs included in each block.
families, so that a total of 474 UK Caucasian SLE families were available for analysis.

Population differences in allele frequencies
A comparison was made of the parental allele frequencies for each marker in each of the three major racial sub-populations in the UK SLE study cohort: European Caucasian (EC), Indian-Pakistani (IP) and Afro-Caribbean (AF). Numerous markers showed differences in parental allele frequencies markers between the EC and AF parental samples (Supplementary Material, Table S1). These markers were located right across the locus and not focused in a particular sub-region. A further four markers were monomorphic in the AF parental samples. Furthermore, the haplotype architecture in the IP samples was different from that in the EC trios, so that the major haplotype in EC samples (haplotype 1 in Fig. 2) was present at 39.3%, but increased to 46% in the IP samples (data not shown). Consequently, association analyses were only carried out using EC samples.

Haplotype structure of the CD28–CTLA4–ICOS locus
Haplotypes across the CD28–CTLA4–ICOS locus were constructed from 332 UK EC parental-proband trios with Haploview, using the algorithm defined by Gabriel et al., with the criteria as in Materials and Methods. There is a discrete haplotype block across CD28, which extends from SNP 6, located 1.6 kb upstream of the coding sequence, to SNP 21 at 59 kb downstream of the gene (Fig. 1B). The inter-block D' score between CD28 and CTLA4 was 0.47, indicating a break-down in (LD) between the two genes. In CTLA4, there are five main haplotype blocks, which fall within a region of strong LD, stretching from the promoter into the 3' flanking region of the gene and into the promoter of ICOS (Fig. 1C). There is a separate LD block covering the 3' end of ICOS, since there are two haplotype blocks across the gene, with an inter-block D' score of 0.25 for the two haplotype blocks across the gene.

Haplotypes were also constructed across the locus, from CEU (EC CEPH) genotype data downloaded from the HAPMAP database. To perform a comparison in haplotype structure between the HAPMAP samples and the EC SLE trios, haplotypes were constructed in each data set using data from 16 SNPs genotyped in both data sets. The HAPMAP samples showed two haplotype blocks: one covering CD28 and a second covering CTLA4, with inter-block D' score between CD28 and CTLA4 of 0.5. It was not possible to include haplotype structure across ICOS in this comparison, since there were two markers in common between our SLE samples and the HAPMAP data. There are some limitations to this approach since the haplotype structure constructed in the UK SLE trios was much more extensive than the HAPMAP data, being constructed with more samples (996 versus 90 samples) and more markers (53 versus 16).

Overall association across CD28–CTLA4–ICOS cluster
Given the density of SNPs across the locus, it was necessary to pinpoint the key regions of association across the locus. It was impossible to conduct a haplotype-transmission disequilibrium test (TDT) analysis for all SNPs included in the study across the entire locus because of the computational demand of such an analysis. This also proved impossible by only taking the tagging SNPs from the study. Instead, an overall haplotype-TDT analysis was performed by taking the tagging SNPs from each haplotype block 1–7 separately, using TRANSMIT in the full UK data set of 474 UK SLE families. The tagging SNPs selected were those shown in Figure 1A–C. The results are presented in Table 1, with the P-values given as an overall test of each block, in each case pooling rare haplotypes (<5%) to avoid loss of power due to excessive degrees of freedom. The results of both these analyses show that the strongest effect across the locus is in the 3' flanking region of CTLA4 for haplotype block 4 (P = 0.06) and for block 6, with a P-value of 0.04. We now consider further analysis of both these regions via tests of individual SNPs and by considering the transmission of the haplotypes described in Figure 1A–C.

Single SNP TDT association analysis
TDT analysis for each individual SNP across the CD28–CTLA4–ICOS locus was performed using GENEHUNTER in the 332 EC parental-proband trios. The results are shown in Table 2. In the 3' flanking region of CTLA4, there were two variants showing significant association to SLE (P < 0.001). These are SNP 43 (P = 0.0004) and SNP 63 (P = 0.0004). In addition, there are 13 other variants distributed throughout the locus, exhibiting weak association to SLE (P-values of <0.05). The strongest of these associations are in the proximal 3' flanking region of CTLA4 between SNPs 42 and SNP 63. In CD28, there is possible association with SLE and the T allele of SNP 6 (P = 0.005) situated in the promoter. In addition, the data presented in Table 2 reveals that there are several variants in the ICOS promoter SNPs 68, 69, 70, 73 and 74, which all show suggestive association to lupus.

These associations in parental-proband trios were studied in the entire cohort using pedigree disequilibrium test (PDT), TRANSMIT and family-based association test (FBAT), all of which allowed the inclusion of 142 single parent families in which one or more unaffected sibling was available. These extended analyses in the full UK data set of 474 families suggested that the strongest associations in the locus arise from the proximal CTLA4 3' flanking region. The results of the PDT analysis are also shown in Table 2. As with the analysis in just parental-proband trios, the strongest associations in the full UK data set including single parent families were identified between SNPs 42 and 63. The strength of the association for SNP 43, as determined by P-value, is P = 0.004 (PDT and TRANSMIT) and P = 0.005 (FBAT) and for SNP 63 is P = 0.004 (PDT and FBAT) and P = 0.020 (TRANSMIT). In CD28, there was overtransmission of the rare T allele of SNP 6, located in the promoter, with P-value of 0.018 by PDT and of 0.006 with PBAT. A second variant in CD28, SNP 13, showed a stronger association with PDT (P = 0.004) than with GENEHUNTER (Table 2). The variant showing the strongest association in the ICOS promoter (SNP 74) also showed suggestive association by PDT (P = 0.02), FBAT (P = 0.025) and TRANSMIT (P = 0.045).
Conditional analysis of individual variants

Individual SNP tests of all 53 markers included in the analysis of the CD28–CTLA4–ICOS gene cluster were conducted using UNPHASED. The most significant SNPs were then selected and tests adjusted for these SNPs performed. The results in Table 3 show that in the 474 UK SLE families, four SNPs showed significance with $P$-values < 0.005, all of which had shown association in the previous analyses (Table 2). All four variants are located in haplotype blocks 4–6 in the 3′ flanking region of CTLA4 or in the promoter of ICOS. Following correction for multiple testing using the permutation test option, there was a global $P$-value for significance of 0.008. We next tested each of these SNPs conditional on each of the others to investigate the evidence for multiple signals at this locus. The pattern of results is not easy to interpret, but it is clear that the strongest evidence for an effect lies in the CTLA4 3′ flanking region.

Haplotype-TDT analysis

GENEHUNTER TDT was performed on the haplotypes constructed across the CD28–CTLA4–ICOS locus in the 332 parental-proband trios for the associated haplotypes shown in Figure 1B and C. The results are tabulated in Table 4. In CTLA4, there is a significant over-transmission of a CG haplotype in the 3′ flanking region of the gene to cases between two variants that are in LD ($r^2 = 1$). These are SNP 43 and SNP 44 located by GENEHUNTER ($P = 0.0008$) and by FBAT ($P = 0.012$) 2.1 kb downstream of the 3′-UTR. There is a secondary signal in the distal 3′ flanking region of CTLA4. This latter effect arises from a protective TGTGGTT haplotype in block 6, which also includes several variants in the ICOS promoter ($P = 0.003$). In CD28, there is an under-transmitted GGTTGAAT haplotype ($P = 0.001$), which stretches from SNP 6 in the promoter of CD28 to SNP 21 in the 3′ flanking region of the gene.

DISCUSSION

Comparison of associated alleles in SLE and T1D families

This is the first large, family-based study to look at the genetic contribution of the entire CD28–CTLA4–ICOS gene cluster in SLE. We have shown good evidence of association to SLE from allelic variants located in the proximal 3′ flanking region of CTLA4, with secondary signal on a haplotype stretching from the distal 3′ flanking region of CTLA4 into the promoter of ICOS. Both these regions have been previously reported to contain alleles associated to T1D, Graves’ disease and AITD in a comprehensive screen across the CD28–CTLA4–ICOS locus (20). However, there are clear differences in the pattern of association between the two studies, both in the identity of the SNPs showing the strongest association and also in the risk haplotypes involved.

Figure 1C shows the location of the four haplotype blocks across CTLA4. However, the inter-block $D'$ score between these four blocks is >0.95, indicating that there is good LD right across the gene, from the promoter to the proximal 3′ flanking region of the gene, despite variation in LD between SNPs used in the construction of the major haplotype blocks.
Table 2. Analysis of single SNPs across the CD28–CTLA4–CD28 locus

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Minor allele</th>
<th>MAF</th>
<th>Total families in analysis</th>
<th>GENEHUNTER TDT analysis (332 UK SLE parent-affected trios)</th>
<th>PDT analysis (474 UK SLE families)</th>
<th>Over/under</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T:U</td>
<td>Z-score</td>
<td>P-value</td>
</tr>
<tr>
<td>5</td>
<td>CD28</td>
<td>T</td>
<td>0.095</td>
<td>404</td>
<td>49:31</td>
<td>3.95</td>
<td>0.047</td>
</tr>
<tr>
<td>6</td>
<td>CD28</td>
<td>T</td>
<td>0.239</td>
<td>374</td>
<td>88:57</td>
<td>8.01</td>
<td>0.005</td>
</tr>
<tr>
<td>8</td>
<td>CD28</td>
<td>T</td>
<td>0.361</td>
<td>509</td>
<td>175:143</td>
<td>3.14</td>
<td>0.076</td>
</tr>
<tr>
<td>13</td>
<td>CD28</td>
<td>T</td>
<td>0.035</td>
<td>472</td>
<td>14:26</td>
<td>3.6</td>
<td>0.058</td>
</tr>
<tr>
<td>23</td>
<td>CD28–CTLA4</td>
<td>C</td>
<td>0.493</td>
<td>341</td>
<td>190:149</td>
<td>4.31</td>
<td>0.038</td>
</tr>
<tr>
<td>25</td>
<td>CTLA4</td>
<td>C</td>
<td>0.434</td>
<td>340</td>
<td>180:148</td>
<td>2.79</td>
<td>0.095</td>
</tr>
<tr>
<td>40</td>
<td>CTLA4</td>
<td>G</td>
<td>0.363</td>
<td>326</td>
<td>129:164</td>
<td>4.26</td>
<td>0.039</td>
</tr>
<tr>
<td>42</td>
<td>CTLA4</td>
<td>A</td>
<td>0.319</td>
<td>345</td>
<td>118:167</td>
<td>7.78</td>
<td>0.005</td>
</tr>
<tr>
<td>43</td>
<td>CTLA4</td>
<td>T</td>
<td>0.309</td>
<td>472</td>
<td>92:148</td>
<td>12.54</td>
<td>0.0004</td>
</tr>
<tr>
<td>44</td>
<td>CTLA4</td>
<td>A</td>
<td>0.322</td>
<td>495</td>
<td>116:158</td>
<td>5.9</td>
<td>0.015</td>
</tr>
<tr>
<td>46</td>
<td>CTLA4</td>
<td>T</td>
<td>0.482</td>
<td>273</td>
<td>151:121</td>
<td>3.79</td>
<td>0.051</td>
</tr>
<tr>
<td>50</td>
<td>CTLA4</td>
<td>C</td>
<td>0.480</td>
<td>239</td>
<td>140:102</td>
<td>6.54</td>
<td>0.011</td>
</tr>
<tr>
<td>56</td>
<td>CTLA4</td>
<td>C</td>
<td>0.316</td>
<td>499</td>
<td>107:157</td>
<td>8.35</td>
<td>0.004</td>
</tr>
<tr>
<td>61</td>
<td>CTLA4</td>
<td>B</td>
<td>0.488</td>
<td>343</td>
<td>193:150</td>
<td>4.71</td>
<td>0.030</td>
</tr>
<tr>
<td>63</td>
<td>CTLA4</td>
<td>C</td>
<td>0.306</td>
<td>465</td>
<td>67:114</td>
<td>12.57</td>
<td>0.0004</td>
</tr>
<tr>
<td>68</td>
<td>ICOS</td>
<td>promoter</td>
<td>0.311</td>
<td>496</td>
<td>105:151</td>
<td>7.22</td>
<td>0.007</td>
</tr>
<tr>
<td>69</td>
<td>ICOS</td>
<td>promoter</td>
<td>0.470</td>
<td>488</td>
<td>136:188</td>
<td>7.12</td>
<td>0.008</td>
</tr>
<tr>
<td>70</td>
<td>ICOS</td>
<td>promoter</td>
<td>0.471</td>
<td>496</td>
<td>143:188</td>
<td>5.1</td>
<td>0.024</td>
</tr>
<tr>
<td>73</td>
<td>ICOS</td>
<td>promoter</td>
<td>0.472</td>
<td>419</td>
<td>111:147</td>
<td>3.99</td>
<td>0.046</td>
</tr>
<tr>
<td>74</td>
<td>ICOS</td>
<td>promoter</td>
<td>0.478</td>
<td>429</td>
<td>70:116</td>
<td>9.19</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Each SNP across the CD28–CTLA4–CD28 gene cluster was tested individually and the variants showing association with a P-value of <0.05 are detailed here. For both GENEHUNTER-TDT and PDT, all P-values were quoted with one degree of freedom. For PDT, the AvePDT option was used, which does not give extra weight in the analysis to larger families. The minor allele frequency (MAF) is quoted for each polymorphism, and the column marked ‘T:U’ defines the ratio of the numbers of informative families showing transmission of the quoted minor allele to those families not showing transmission for the GH-TDT analysis. The column marked ‘families in analysis’ represents the total number of families for PDT, which includes single parents, since GENEHUNTER only analyses trios. The column marked ‘over/under’ indicates whether the association was an over- or under-transmission of the stated minor allele.

Table 3. Conditional analysis of individual markers across the CD28–CTLA4–ICOS gene cluster

<table>
<thead>
<tr>
<th>Marker</th>
<th>Location</th>
<th>Haplotype block</th>
<th>Conditioning SNP</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>CTLA4 3’ flank</td>
<td>4</td>
<td>43 56 63 74</td>
<td>0.0003</td>
</tr>
<tr>
<td>56</td>
<td>CTLA4 3’ flank</td>
<td>5</td>
<td>43 56 63 74</td>
<td>0.0002</td>
</tr>
<tr>
<td>63</td>
<td>CTLA4 3’ flank</td>
<td>6</td>
<td>43 56 63 74</td>
<td>0.0013</td>
</tr>
<tr>
<td>74</td>
<td>ICOS promoter</td>
<td>6</td>
<td>43 56 63 74</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

All 53 polymorphisms across the gene cluster were analysed by UNPHASED. The variants listed show the strongest association (P < 0.002), following analysis by UNPHASED. The analysis was then repeated by conditioning each variant on each of the other associated variants in turn. The results are presented as P-values with one degree of freedom.

across the locus, as determined by the $r^2$ figures. Consequently, it is possible to force the haplotypes across CTLA4 into a single block in order to facilitate the comparison of the pattern of association in SLE with the published data from Graves’ disease.

Figure 2 shows that the over-transmitted alleles in Graves’ disease, shown in grey boxes, are the G allele of SNP 25 in the promoter of CTLA4, the G allele of SNP 33 in exon 1 and the G allele of both SNPs 37 and 46 in the 3’ flanking region of the gene, together with the T allele of SNP 50, also in the 3’ flanking region. The over-transmitted alleles in SLE are the C allele of SNP 43 and the G allele of SNP 44 in block 4 and the A alleles of SNPs 69 and 70 and the C alleles of SNPs 73 and 74 in block 6. In Figure 2, these over-transmitted alleles for SLE are shown in white circles. All these over-transmitted alleles in SLE are carried on the most common haplotype, haplotype 1, which represents 39.3% of the total chromosomes. However, haplotype 1 does not carry any of the over-transmitted alleles from the study of Ueda et al., meaning that the susceptibility alleles in SLE are in LD with the protective alleles in Graves’ disease, TID orAITD. Similarly, haplotype 2, which carries all the over-transmitted alleles from the report by Ueda et al., does not carry any of the over-transmitted alleles from the SLE families. Since haplotype 2 has a frequency of 26.5%, this means that in a total of over 65% of the haplotypes, there is no LD between the susceptibility alleles in SLE compared with the organ-specific autoimmune diseases reported by Ueda et al. However, SLE is not unique in showing the association of the protective alleles from Graves’ disease. A report by Munthe-Kaas et al. (46) showed association of several Graves’ protective alleles with phenotypes of allergy and asthma.

Since the haplotypes in SLE families (Fig. 1B) are similar to those seen in the CEU samples from HAPMAP, the question arises as to whether the discrepancy is due to a difference in haplotype structure between this UK SLE study and the haplotype pattern in report of the Ueda et al. (also in UK samples). This was investigated by a comparison of the haplotype break points across the locus in the two studies. These break-points were found to be between CD28 and CTLA4 and also in the first intron of ICOS in both the SLE study and the previous report by Ueda et al. (20), which give
three haplotype blocks across the locus. In terms of defining SNPs, in both this study and the report of Ueda et al., the same variant, SNP 22 (CTAF305), defines the 22–CTLA4 haplotype break-point. Figure 2 also illustrates a ‘downstream shift’ in the location of the unique associated alleles in SLE compared with those in Graves’ disease. CT60 (SNP 37) is the variant showing maximal association in T1D (P = 6 × 10⁻⁵), Graves’ disease (P = 2.72 × 10⁻⁸) and AITD (P = 5 × 10⁻⁴). This variant is located within the first 300 bp of the 3’ flanking region of CTLA4 (20). However, in our SLE study, CT60 is not associated with disease, and furthermore, the association in SLE appears to arise from further downstream in the 3’ flanking region of CTLA4. At the 3’ end of the group of best associated SNPs in the two studies, the most distal associated variant in Graves’ disease is the T allele of SNP 50, found 6.3 kb downstream of CTLA4. In contrast, in SLE, the most distal associated allele is the C allele of SNP 74, in the ICOS promoter, located 45.6 kb downstream of SNP 50.

Therefore, in the UK population, there appears to be a clearer disease specificity in the association patterns between SLE and the three autoimmune diseases reported by Ueda et al. (20). However, this is not the first example of disease-specific susceptibility for causal alleles in the same gene within the same population. In a Dutch family-based association study of three autoimmune diseases, T1D, RA and CD, there was also evidence of such a disease specificity for association. Two of the SNPs (MH30 and CT60) showing strong association in the article by Ueda et al. (20) were carried on an over-transmitted haplotype in the Dutch T1D families, but not in the RA or CD families. However, when the CD cases were stratified for an age of onset of <10 years, there was clear association of a second haplotype not associated in either T1D or RA (47).

Consequently, it is unlikely that there are major shared pathogenic alleles in these four autoimmune diseases. This could be related to the fact that SLE has a more wide-spread pathology, whereas Graves’ disease, T1D and AITD are all organ specific. The other factor is that the current data indicate an associated haplotype in SLE, and not individual aetiological alleles. It is also possible that neither study has identified the common susceptibility allele(s) for autoimmune disease, despite the functional implications proposed by Ueda et al. The published report showed that in healthy controls, the G allele of CT60 was correlated to decreased levels of soluble sCTLA-4, a splicing variant of CTLA4. However, a more recent study by Anjos et al. (48) failed to replicate these findings. An additional effect comes from the +49A/G disease-protective A allele because healthy individuals possessing the protective +49A-CT60(A) haplotype produce higher levels of sCTLA-4 than those possessing the +49A-CT60(G) haplotype (20). However, the role of these particular polymorphisms in influencing sCTLA-4 levels is contentious (49). Given the discrepancy in associated alleles between the published data and this article, it is unclear whether the CT60 and +49A/G alleles will play a part in modulating the levels of serum sCTLA-4 in SLE. However, it is known that there are increased levels of serum sCTLA-4 in SLE (50–52), rather than the drop in the levels of serum sCTLA-4 correlated with risk alleles in T1D, which mirrors the pattern of association seen between the two diseases.

**Comparison of associated alleles for CTLA4 with previous studies in SLE**

The identification of aetiological alleles of CTLA4 with SLE has not always been helped by previous studies looking for association, which utilized a ‘targeted variant’ approach, by genotyping a few SNPs reported to be associated in other autoimmune diseases. The variants typed in these previous studies were SNP 26, 27, 31, 33 and 37 and were restricted to the region defined by haplotype block 3 in CTLA4 (Fig. 1C). One of these studies showed association for variant CT60 (53), but this was not replicated in our UK samples. Two other published reports in SLE showed over-transmission of the T allele of SNP 26 (16,17). In contrast, in this study, we have genotyped a dense map of SNPs right across the flanking region of the gene, but no significant association for any SNPs further upstream. The overall lack of agreement for individual SNPs included in previous association studies (18,19) is not surprising in the light of the data that we present, in which the major signal in SLE is not in the promoter of CTLA4, but resides in its 3’ flanking region.

However, in this article, we do not claim to have fully defined the aetiological SNPs associated with SLE, although our data do support independent effects in the 3’ flanking region of CTLA4. The two SNPs showing maximal association...
are in SNPs 43 and 44 in the 3' flanking region of the CTLA4, in haplotype block 4. Nonetheless, there are other more distal variants, which may act in combination with these variants, located in haplotype block 6, which contains variants in the distal 3' flanking region of CTLA4/ICOS promoter. One possibility is that the two signals act on different genes in the locus, since the proximity of SNP 43 to the 3'-UTR of CTLA4 would suggest that it had an effect on CTLA4 expression, whereas the SNPs in block 6, in the ICOS promoter, are more likely to play a role in the regulation of ICOS.

Association in ICOS

We have also presented evidence of an association in the distal 3' flanking region of CTLA4 on a haplotype which includes variants in the promoter of ICOS. There are individual associations in SNPs 69, 70, 73 and 74 (Table 2). Further evidence to support the idea that the ICOS promoter is an independent effect from the CTLA4 flanking region came from Ueda et al. (20), who found a third peak of association in the ICOS promoter in their Graves' disease study. This is the first study to show association of variants upstream of ICOS with SLE, but the functional consequences and their role in disease remain to be established. ICOS may have a role in the development of germinal centres (GCs) within the B cell follicles because CXCR5+ T cells producing ICOS are enriched in the GCs. Changes in the level of ICOS may alter the production of memory B cells and the generation of affinity-matured autoreactive B cells. Furthermore, it is known that ICOS is up-regulated on the CD4+ and CD8+ T cells of SLE patients with active disease (54). It is possible that the over-transmitted promoter SNPs in block 6 act to increase the expression of ICOS.

Association in CD28

We have also found some evidence for under-transmission of haplotype in the promoter of CD28, a protein which is a major stimulator of T cell activation through interaction with CD86/CD80 at the T–B cell interface. This association is carried by a haplotype extending from the promoter into intron 1 of the gene and including the signal sequence in exon 1. However, we have defined an associated haplotype in CD28 and not identified an individual variant responsible for this genetic effect. It is possible that this represents a false-positive result. The under-transmitted rare T allele of SNP 8 is not carried on this protective GGTGGAAT haplotype, but it is sufficient to reverse a highly significant under-transmitted haplotypic association.

There is only one other published association study on CD28, in a Japanese population, which found no association with a polymorphism in intron 3, SNP 10, in a small-scale case–control study of 113 Japanese lupus patients and 200 normal controls (15). We confirm this lack of association in our large family-based UK population, although there is a difference in the minor allele frequency between the Japanese (4.4%) and UK (18.9%) populations. In spite of this lack of genetic studies for CD28 in SLE, there have been several studies indicating that aberrant function of CD28 may play a role in the aetiology of lupus. The plasma levels of sCD28 (and sCTLA-4) were increased in patients with SLE compared with controls (50).

In summary, we have confirmed that the 3' flanking region of CTLA4 is an important region for association to SLE, in addition to the previous reports in Graves' disease, T1D andAITD. However, the pattern of association is distinct from that seen in these other autoimmune diseases. The variants contributing to the association in SLE are more distal to CTLA4 than for Graves' disease and extend into the promoter of ICOS. These more distal variants may affect expression levels of ICOS and consequently the activity of follicular B cells. Thus, in SLE, there appears to be a complex and unique pattern of genetic factors that predispose to disease.

MATERIALS AND METHODS

Family collection

The large collection of SLE families in the laboratory is predominantly EC. These EC families make up 82% of the total collection, with the remaining being equally divided into IP and AF samples. All probands conformed to the ACR criteria for SLE (55), with the diagnosis being established by telephone interview, health questionnaire and details from clinical notes. Written consent was obtained from participants, including relatives. Ethical approval was obtained from Multi-Centre Research Ethics Committee. The total study cohort consisted of 532 UK EC SLE families, of which 332 were parental-proband trios and 142 were families with a single parent together with one or more unaffected offspring. However, 25 families showing consistent ex-paternity were removed from the analyses.

Genotyping methodology and marker selection

The polymorphisms were all typed by MALDI-TOF mass spectrometry (Sequenom, Hamburg, Germany), except for SNP 33, which was typed by PCR-RFLP, using TseI (56). All PCRs were carried out on DNA Engine Tetrad (MJ Research Incorporated, Waltham, MA, USA), in either a 96 well or 384 well format. The details of all amplification and extension primers are available from the authors. A total of 80 SNPs were chosen across the CD28–CTLA4–ICOS locus. Markers were selected from a series of haplotype tagging SNPs provided by John Todd (Juvenile Diabetes Research Foundation Wellcome Trust Diabetes and Inflammation Laboratory), with additional polymorphisms chosen from the SNP databases (http://www.ncbi.nlm.nih.gov/SNP/index.html) and HAPMAP (http://www.hapmap.org/). One other variant, SNP 27, was chosen because it had been typed in two previous association analyses with SLE (16,57) (Fig. 1A).

Statistical analysis

All sample genotype and phenotype data were managed by and analysis files generated with the BC/GENE and BC/CLIN software (Biocomputing Platforms Ltd, Finland). The integrity of the genotyping was checked by genotyping frequency, HWE test (58) and PEDCHECK (59). Markers were excluded if the percentage genotyping was <90%, the
HWE P-value < 0.05 and there were more than three families showing sporadic Mendelian errors.

Haplotype patterns were constructed on the basis of a method described by Gabriel et al. (60), using Haploview. This programme creates haplotypes based on the D' measure of LD (61), together with a LOD score as a measure of significance and 95% confidence intervals to state the accuracy of the P-value. r² values were used (62,63) to confirm the pairwise LD for SNPs across each gene and to refine the overall haplotypic architecture. Only markers having an MAF > 5% were included in the haplotype constructions. The haplotype block definitions were based on confidence limits for strong LD of 0.98 (upper) and 0.70 (lower), upper confidence interval maximum for strong recombination of 0.90 and at least 95% of LD of 0.98 (upper) and 0.70 (lower), upper confidence interval maximum for strong recombination of 0.90 and at least 95% of LD in informative comparisons, as described (56).

Association of alleles to SLE for both individual and multiple SNPs was tested by the TDT, which compares the observed and expected transmission of alleles from heterozygous parents to the affected offspring. The analysis in parental-proband trios was carried out using GENEHUNTER 2.0 beta (64,65). Inclusion of further data from single parent families for TDT analysis was achieved using the PDT (66), TRANSMIT v 2.5 (67) and the FBAT (68,69). PDT allows the analysis of both trios and discordant sibships, and TRANSMIT is able to infer haplotypes of the missing parents, using genotype data from siblings. The P-value for suggestive association was set at 0.05, and the threshold for significant association was set at 0.001. P-values between 0.083 and 0.05 were deemed to show a trend towards association. The UNPHASED package was used to confirm the associations from GH-TDT and to determine the interactions between associated variants.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

This work was funded by the Wellcome Trust through a Senior Fellowship awarded to T.J.V. We acknowledge the work of Paul Spencer in recruiting patients and families into the study, and we would like to thank our clinical colleagues for helping us to recruit study participants. Our thanks and appreciation is extended to all the patients and their relatives for generously donating blood samples and all the general practitioners and practice nurses for collecting them; many thanks to David Perkins for maintaining the computerized database.

Conflict of Interest statement. None declared.

REFERENCES

22. Lei, C., Dongqing, Z., Yeqing, S., Oaks, M.K., Lishan, C., Jianzhong, J.,
    et al. (2005) Expression of inducible nitric oxide synthase 1 in
    systemic lupus erythematosus: association with disease and
    in vivo clearance of immune complexes. *Arthritis Rheum.*, 43,
    2793–2800.
23. Torres, B., Aguilar, F., Franco, E., Sanchez, E., Sanchez-Roman, J.,
    Jimenez, A.J., Nunez-Roldan, A., Martin, J. and Gonzalez-Escobedo,
    co-stimulator in peripheral blood T lymphocytes in the patients with
25. Tan, E.M., Cohen, A.S., Fries, J.F., Mai, A.T., McShane, D.J., Rothfield,
    revised criteria for the classification of systemic lupus erythematosus.
26. Harper, K., Balzano, C., Rouvier, E., Mattei, M.G., Luciani, M.F. and
    Golstein, P. (1991) CTLA4 and CD28 molecules are associated with
27. Kuok, T., Sawai, Y., Gardine, C.A., Fisfalen, M.E., Alegre, M.L. and
    Blomhoff, A., Lie, B.A., Myhre, A.G., Kemp, E.H., Weetman, A.P.,
    J. Hum. Genet.*, 6, 410–413.
28. Nistico, L., Buzzetti, R., Pritchard, L.E., Van der, A.B., Giovannini,
    C., Bosi, E., Larrad, M.T., Rios, M.S., Chow, C.C., Cockram, C.S.,
    et al. (1996) The CTLA4 gene region of chromosome 2q33 is linked
29. Dihlali-Saiah, I., Schmitz, J., Harfouch-Hammoud, E., Mougenot, J.F.,
    6, 823–828.
30. Heward, J.M., Allahabadia, A., Armitage, M., Hattersley, A., Dodson, P.M.,
    et al. (1999) The development of Graves' disease and the association
    with CTLA4 polymorphisms in multiple ethnic groups. *Hum.
    Genet.*, 114, 280–287.
    and Wijmenga, C. (2004) CTLA4 +4A/G and CT60 polymorphisms in
33. Kemp, E.H., Ajjan, R.A., Husebye, E.S., Peterson, P., Uibo, R., Kumia, K.,
    et al. (1997) Insulin-dependent diabetes mellitus (IDDM) is associated
    with the cytotoxic T lymphocyte antigen-4 gene region on chromosome
34. McAdam, A.J., Greenwald, R.J., Levin, M.A., Chernova, T., Malenkovich, N.,
    Ling, V., Freeman, G.J. and Sharpe, A.H. (2001) ICOS is critical for
    between the levels of soluble cytotoxic T-lymphocyte associated
    antigen-4 (CTLA-4) and the CT-60 genotypes. *J. Autoimmun.,* 2, 8.
36. McAdam, A.J., Greenwald, R.J., Levin, M.A., Chernova, T., Malenkovich, N.,
    between the levels of soluble cytotoxic T-lymphocyte associated
    antigen-4 (CTLA-4) and the CT-60 genotypes. *J. Autoimmun.,* 2, 8.
38. McAdam, A.J., Greenwald, R.J., Levin, M.A., Chernova, T., Malenkovich, N.,
    between the levels of soluble cytotoxic T-lymphocyte associated
    antigen-4 (CTLA-4) and the CT-60 genotypes. *J. Autoimmun.,* 2, 8.
40. McAdam, A.J., Greenwald, R.J., Levin, M.A., Chernova, T., Malenkovich, N.,


