Systemic lupus erythematosus (SLE) is a complex autoimmune disease. Genome-wide linkage studies implicated a region containing the adhesion molecule P-Selectin. This family-based study revealed two regions of association within P-Selectin. The strongest signal, from a 21.4-kb risk haplotype, stretched from the promoter into the first two consensus repeat (CR) regions \( (P = 8 \times 10^{-4}) \), with a second association from a 14.6-kb protective haplotype covering CR 2–9 \( (P = 0.0198) \). The risk haplotype is tagged by the rare C allele of rs3753306, which disrupts the binding site of the trans-activating transcription factor HNF-1. One other variant (rs3917657) replicated in 306 US SLE families and was significant in a joint UK-SLE data set after permutation. The protective haplotype is tagged by rs6133 (a non-synonymous variant in CR8 \( P = 10^{-3} \)), replicated in independent pseudo case-control analysis and was significant by meta-analysis \( (P = 4.37 \times 10^{-4}) \). A third associated variant on the risk haplotype (rs3917657) replicated in 306 US SLE families and was significant in a joint UK-SLE data set after permutation. The protective haplotype is tagged by rs6133 (a non-synonymous variant in CR8 \( P = 9.00 \times 10^{-4} \)), which also shows association in the pseudo case-control analysis \( (P = 1.09 \times 10^{-3}) \) and may contribute to another signal in P-Selectin. We propose that polymorphism in the upstream region may reduce expression of P-Selectin, the mechanism by which this promotes autoimmunity is unknown, although it may reduce the production of regulatory T cells. 

Keywords: P-Selectin; SLE; association; promoter; imputation

Introduction

Systemic lupus erythematosus (SLE) is a generalized autoimmune disease that behaves as a complex genetic trait. A number of groups have published genome-wide and targeted linkage analyses.\(^1\)\(^-\)\(^5\) There is considerable heterogeneity within these linkage studies, no doubt reflecting small study sizes, genetic heterogeneity and clinical heterogeneity in SLE. The most consistently mapped lupus susceptibility loci reside in the following regions: 1q23, 1q25–31, 1q41–42, 2q35–37, 4p16–15, 4p11–21 (MHC) and 16q12, all of which have been corroborated at least once in an independent cohort.\(^6\)\(^-\)\(^7\) The 1q22–25 region, which contains SELP (P-Selectin), has been identified as part of two genome-wide linkage scans in humans.\(^3\)\(^-\)\(^4\) The gene itself, alternately known as P-Selectin, CD62 or Granulocyte membrane protein (GMP140), is over 50kb long and contains 17 exons,\(^8\) most of which encode discrete domain structures, as described in the legend to Figure 1. P-Selectin is a member of the selectin gene family, which includes L-Selectin and E-Selectin. We have already studied these two genes and found no evidence of association with SLE, although two variants in the EGF domain were associated with the levels of soluble L-Selectin.\(^9\) All three selectin genes are located in a gene cluster on chromosome 1. The selectin molecules begin with an N-terminal, ‘lectin’ domain followed by an ‘EGF’ domain, nine tandem consensus repeats similar to those in complement-binding proteins, a transmembrane domain and a cytoplasmic tail. SELP is longer than the other selectin genes because it contains nine consensus repeats (CRs) compared with six in E-Selectin and two in L-Selectin.\(^10\)

We were interested in the role of P-Selectin as a candidate gene for lupus because, in addition to the linkage data, it also makes a good biological candidate for a susceptibility gene. P-Selectin is primarily expressed on the cell membrane of platelets and endothelial cells after cellular activation by molecules such as histamine or thrombin, which cause release of the protein from α-granules (platelets) or Weibel–Palade bodies
(endothelial cells). The protein functions as an adhesion molecule for a variety of leukocytes, including neutrophils, monocytes, T cells, eosinophils, basophils, platelets and some malignant cells, thereby bringing to sites of inflammation and into contact with a range of cytokines and chemokines expressed by endothelial cells. There are two receptors for P-Selectin, the major leukocyte receptor, P-selectin glycoprotein ligand 1 (PSGL1), which is the high affinity receptor on both myeloid cells and stimulated T lymphocytes, and CD24, which is expressed on a number of cells, including B-lineage cells and T lymphocytes. Furthermore, the 57 V allele of CD24 has also been recently described as a susceptibility allele for lupus in a Spanish population and for both multiple sclerosis and rheumatoid arthritis.

In addition to the genetic and biological reasons for investigating the potential role for P-Selectin, there is evidence from an animal model to support a role in autoimmune glomerulonephritis. In this model, P-Selectin-deficient mice show several defects in leukocyte function, including almost absent leukocyte rolling in mesenteric venules and delayed recruitment of neutrophils to the peritoneal cavity after laboratory-induced inflammation using intra-peritoneal thioglycollate.

The work presented in this paper describes a comprehensive family-based association study to look for potential susceptibility alleles, in a collection of UK and US (Minnesota) SLE families.

Results
Selection of markers for analysis
Our genotyping strategy was to type the haplotype-tagging single nucleotide polymorphisms (SNPs) across SELP and also to include the synonymous/non-synonymous coding variants in SELP, with minor allele frequencies above 0.05 in Caucasian populations. We therefore typed a total of 52 SNPs across the SELP locus (Figure 1a). The haplotype block structure across SELP is constructed from 263 European Caucasian parent-proband trios from the UK study cohort. There is a single haplotype block across the gene. The haplotypes are numbered on the left of each haplotype in brackets from 1–8 with the haplotype frequencies shown to the right of each haplotype. Only haplotypes with a frequency of >2.5% are shown. The SNP numbers across the top of the haplotypes correspond to those in the gene diagram. The haplotype bounded by a solid red line in block 1 shows over-transmission in the UK SLE families and the haplotype inside the dashed blue box in block 2 shows under-transmission in the UK SLE families. The P-values for the associated haplotypes are shown below their respective haplotype blocks. A full colour version of this figure is available at the Genes and Immunity journal online.

Figure 1 Genomic organization and pattern of association across P-Selectin in UK SLE families. (a) Human P-Selectin consists of 17 exons. The translated exons are illustrated as black boxes and the 5' and 3' UTRs as white boxes. The protein domains are shown as: E2 (S1)—signaling domain 1; E3 (S2/Lec)—signaling domain 2/lectin-like domain; E4 (EGF)—epidermal growth factor-like domain; E5-13 (CR1-9)—complement repeat regions; E14 (TM)—transmembrane domain; E15 (C1)—cytoplasmic domain 1 and E16 (C2)—cytoplasmic domain 2. In the diagram, the SNPs used in the analysis have been recoded as numbers from 1 to 52: SNP 1 (rs17523783), SNP 2 (A1969G)—identified by sequencing of 60 UK SLE cases, SNP 3 (rs3753306), SNP 4 (rs3917657), SNP 6 (rs3917657), SNP 8 (rs3917862), SNP 10 (rs3917676), SNP 11 (rs3917682), SNP 12 (rs3917683), SNP 13 (rs3917691), SNP 14 (rs6125), SNP 15 (rs6131), SNP 17 (rs2076704), SNP 18 (rs3917751), SNP 19 (rs2423378), SNP 20 (rs7690594), SNP 21 (rs3917792), SNP 22 (rs2298903), SNP 23 (rs6127), SNP 24 (rs6132), SNP 25 (rs6133), SNP 26 (rs6136), SNP 27 (rs3917836), SNP 28 (rs3917848), SNP 29 (rs3753305), SNP 30 (rs6028) and SNP 31 (rs6703462) as listed in Table 4. The markers, which failed the quality control are place-marked in the sequence by gray lines. (b) The haplotype block structure across SELP is constructed from 263 European Caucasian parent-proband trios from the UK study cohort. There is a single haplotype block across the gene. The haplotypes are numbered on the left of each haplotype in brackets from 1–8 with the haplotype frequencies shown to the right of each haplotype. Only haplotypes with a frequency of >2.5% are shown. The SNP numbers across the top of the haplotypes correspond to those in the gene diagram. The haplotype bounded by a solid red line in block 1 shows over-transmission in the UK SLE families and the haplotype inside the dashed blue box in block 2 shows under-transmission in the UK SLE families. The P-values for the associated haplotypes are shown below their respective haplotype blocks. A full colour version of this figure is available at the Genes and Immunity journal online.
### Table 1: Analysis of Single SNPs across the P-Selectin by GENEHUNTER

| SNP | SNP Name | Location | MA | UK SLE Trios (n = 263) | MN SLE Trios (n = 306) |
|-----|----------|----------|----|------------------------|------------------------|
|     |          |          |    | T:U | χ² | P-value | OR | 95% CI |        | T:U | χ² | P-value | OR | 95% CI |
| 1   | rs17523783 | Promoter | T  | 69.69 | 0.0 | 1.0 | 1.0 | 0.716–1.40 |        | 68.56 | 1.16 | 0.281 | 1.21 | 0.830–1.73 |
| 2   | A-1969G    | Promoter | A  | 89.102 | 0.885 | 0.347 | 0.873 | 0.657–1.16 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 3   | rs375306   | Promoter | C  | 57.37 | 4.26 | 0.0391 | 1.54 | 1.02–2.33 |        | 68.56 | 1.16 | 0.281 | 1.21 | 0.830–1.73 |
| 4   | rs3917651  | Promoter | A  | 25.34 | 1.37 | 0.241 | 0.735 | 0.439–1.23 |        | 68.56 | 1.16 | 0.281 | 1.21 | 0.830–1.73 |
| 5   | rs3917657  | II       | T  | 20.40 | 6.67 | 9.80 × 10⁻¹⁰ | 0.500 | 0.292–0.855 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 6   | rs3917862  | I        | C  | 13.25 | 3.79 | 0.0516 | 0.520 | 0.266–1.02 |        | 68.56 | 1.16 | 0.281 | 1.21 | 0.830–1.73 |
| 7   | rs3917676  | II       | G  | 24.17 | 1.12 | 0.274 | 0.400 | 0.0776–2.06 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 8   | rs3917682  | II       | G  | 114.15 | 4.91 | 0.0267 | 0.760 | 0.596–0.970 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 9   | rs3917683  | I        | A  | 97.88 | 8.00 | 5.65 | 0.009 | 0.921–3.30 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 10  | rs3917686  | I        | C  | 97.88 | 8.00 | 5.65 | 0.009 | 0.921–3.30 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 11  | rs3917687  | I        | C  | 12.67 | 3.19 | 0.201 | 0.707 | 0.450–1.17 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 12  | rs3917691  | I        | T  | 97.88 | 8.00 | 5.65 | 0.009 | 0.921–3.30 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 13  | rs3917692  | I        | C  | 97.88 | 8.00 | 5.65 | 0.009 | 0.921–3.30 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |
| 14  | rs3917693  | I        | A  | 97.88 | 8.00 | 5.65 | 0.009 | 0.921–3.30 |        | 32.51 | 4.35 | 0.037 | 0.627 | 0.403–0.976 |

Abbreviations: CI, confidence interval; MA, minor allele for each variant; OR, odds ratio; MN SLE, Minnesota systemic lupus erythematosus; SNP, single nucleotide polymorphism; T:U, transmission: untransmission ratio of the minor allele; major allele.

TDT analysis by GENEHUNTER for each SNP tested in the 263 UK SLE families, and 306 US Minnesota families showing the location of each analyzed SNP within the gene. Those SNPs marked with an asterisk are either non-synonymous (rs6131—Asn/Ser, rs6133—Leu/Val and rs6136—Pro/Thr) or synonymous (rs6132—Asn/Asn). 'SNP analyzed', gives the order of analyzed SNPs. For each variant analyzed SNP within the gene. Those SNPs marked with an asterisk are either non-synonymous (rs6131—Asn/Ser, rs6133—Leu/Val and rs6136—Pro/Thr) or synonymous (rs6132—Asn/Asn). 'SNP , single risk haplotype, haplotype 5, which is tagged by the under-transmitted T allele of SNP 36 (rs6133). The breakdown in LD between blocks 1 and 2, as shown by an inter-block D^2 score of 0.21 (Figure 1b), provides further indication that there may be multiple independent association signals within P-Selectin.

### Haplotype structure across SELP in UK SLE families

We defined the boundaries of the haplotype blocks across P-Selectin using Haplovie (Figure 1b). There are three haplotype blocks across the gene: block 1 stretching 21.4 kb across the promoter, the signaling domains and into the first three CR regions; block 2, which covers 14.6 kb, stretching over most of the CR and block 3, which is an 11.2 kb region including the transmembrane domain and the 3' flanking region. Block 1 contains a single risk haplotype, haplotype 5, which is tagged by the C allele of SNP 3 (rs391687). Block 2 contains an under-transmitted haplotype, haplotype 11 (P = 0.0198). This protective haplotype is uniquely tagged by the under-transmitted T allele of SNP 36 (rs6133).
Table 2  Permutation analysis of single SNPs across P-Selectin in a UK and US (Minnesota) SLE families

| rs       | Study        | (SNP 3) | (SNP 6) | (SNP 21) | (SNP 33) | (SNP 34) | (SNP 36) | (SNP 41) |
|----------|--------------|---------|---------|---------|---------|---------|---------|---------|
| T:U      | UK           | 57:37   | 20:40   | 48:62   | 117:106 | 50:71   | 82:79   | 63:62   |
|          | MN           | 68:56   | 32:51   | 93:127  | 141:139 | 80:76   | 82:79   | 63:62   |
|          | Joint        | 125:93  | 52:91   | 141:189 | 258:245 | 130:147 | 110:138 | 86:87   |
| OR (CI)  | UK           | 1.5 (1.0–2.3) | 0.50 (0.29–0.86) | 0.77 (0.53–1.1) | 1.1 (0.85–1.4) | 0.70 (0.49–1.0) | 0.48 (0.30–0.74) | 0.92 (0.52–1.6) |
|          | MN           | 1.2 (0.83–1.7) | 0.63 (0.40–0.98) | 0.73 (0.56–0.96) | 1.0 (0.80–1.3) | 1.1 (0.77–1.4) | 1.0 (0.76–1.4) | 1.0 (0.72–1.4) |
|          | Joint        | 1.3 (1.03–1.76) | 0.57 (0.41–0.80) | 0.75 (0.60–0.93) | 1.1 (0.88–0.70) | 0.88 (0.70–1.1) | 0.78 (0.62–1.0) | 0.99 (0.73–1.3) |
| P (P<0.001) | UK          | 0.039 (0.68) | 0.0099 (0.264) | 0.18 (0.99) | 0.46 (1.0) | 0.056 (0.79) | 9.04 x 10^-4 (0.025) | 0.77 (1.0) |
|          | MN           | 0.28 (0.86) | 0.037 (0.12) | 0.022 (0.12) | 0.91 (1.0) | 0.75 (1.0) | 0.82 (0.99) | 0.93 (1.0) |
|          | Joint        | 0.029 (0.18) | **0.0016 (0.0072)** | 0.014 (0.054) | 0.61 (0.99) | 0.30 (0.90) | 0.081 (0.38) | 0.97 (1.0) |

Abbreviations: CI, confidence interval; MN, Minnesota; OR, odds ratio; SLE, systemic lupus erythematosus; SNP, single nucleotide polymorphism; T:U, transmission: untransmission ratio of the minor allele: major allele.

A comparison of the permutation analyses for single SNPs across SELP in 263 UK SLE trios (UK), 306 US SLE trios (MN) and a joint dataset (joint). Each SNP was analyzed independently in each dataset and the transmission:untransmission ratios (T:U) are shown, together with the odds ratio and 95% confidence intervals OR (CI) and the P-value after 10,000 permutations P(P<0.001). The SNP in bold, rs391657 retains significance after permutation analysis.

Joint analysis in UK and US (Minnesota) SLE families

We wanted to carry out a joint UK–US analysis across P-Selectin. Tests for stratification were carried out on the assumption that the two populations were indeed homogenous across P-Selectin: when employing Pearson’s χ²-test, we did not reject this assumption with respect to the transmission ratios (P > 0.001) (Supplementary Table 2) and using the Breslow–Day test, we did not reject the assumption with respect to the odds ratios (P > 0.001) (Supplementary Table 3). All of these factors give us the confidence to do a robust joint analysis between the two populations. The results of this joint analysis, presented in Table 2, show a stronger association for SNP 3 (rs3753306), SNP 6 (rs3917657) and SNP 21 (rs6131) in the joint UK–US data set, compared with either of the two independent UK or US data sets.

Furthermore, the association with SNP 6 remains significant after permutation analysis with 10,000 permutations (P = 7.20 x 10^-6) and a significant association for SNP 21 (rs6131) (P = 0.053), which also lies on the risk haplotype. It is interesting that when we repeated the imputation on taking both the 263 cases from the trios and the additional 270 independent cases and then re-did the pseudo case-control association study against chromosomal position. The P-values for the variants with the strongest association are presented in Table 3. These data reveal independent replications of the association seen by TDT analysis in our UK families for two variants within the bounds of the upstream risk haplotype: SNP 12 (rs3917683) (P-value = 0.0270) and SNP 13 (rs3917687) (P-value = 0.0330), with a third borderline association for SNP 21 (rs6131) (P-value = 0.053), which also lies on the risk haplotype. It is interesting that when we repeated the imputation on taking both the 263 cases from the trios and the additional 270 independent cases and then re-did the pseudo case-control analysis, there was highly significant association for SNP 13 rs3918687 (P = 2.52 x 10^-4) and a significant association for non-synonymous variant SNP 36 (rs6133) (P = 1.09 x 10^-5). The imputed SNPs in the independent cases data were not significant, whereas several SNPs showed marginal levels of significance in the combined data set (Figure 2b).

Meta-analysis

To maximize the numbers of samples for the association analysis, we combined the P-values for SNPs on the UK risk haplotype, which showed association (SNP 13—rs3917687 and SNP 12—rs3917683) or a trend for association (SNP 21—rs6131) in the pseudo case-control analysis and/or showed association in either the 263 UK parental-proband trios (rs3917687) or the 306 Minnesota trios (rs6131) and in the imputation analysis. These results revealed a highly significant P-value for SNP 13 (P = 4.37 x 10^-5) and a significant association for SNP 21 (P = 9.56 x 10^-3).

Discussion

P-Selectin is located in the SLE linkage region on human chromosome 1 (1q23). Linkage studies in humans and in animal models of lupus have suggested an important role for the gene in SLE. The current family-based analysis gives us the confidence to do a robust joint analysis across the two populations. The results of this joint analysis, presented in Table 2, show a stronger association for SNP 3 (rs3753306), SNP 6 (rs3917657) and SNP 21 (rs6131) in the joint UK–US data set, compared with either of the two independent UK or US data sets. The risk allelics of all the three variants are carried on the risk haplotype defined in the UK SLE families (Figure 1b). Furthermore, the association with SNP 6 remains significant after permutation analysis with 10,000 permutations (P = 7.20 x 10^-6). Haplotype–TDT analysis in the joint data set adds further evidence to the over-transmission of a SNP 3(C)-SNP 6(C)-SNP 21(G) haplotype (P = 0.0239). These are allelics all carried on the more densely mapped UK SLE risk haplotype. An under-transmitted T-T-A haplotype was also seen in the combined data set. This haplotype: SNP 3(T)-SNP 6(T)-SNP 21(A) haplotype (P = 3.10 x 10^-5) is carried on UK SLE haplotypes 6 and 7, both of which show a trend for under-transmission (P < 0.05). These data serve to confirm the importance of the promoter region of P-Selectin in the association signal.

Replication using pseudo case-control analysis with imputed genotypes

To extend our analysis, we used imputation to test un-genotyped variants within the chromosomal ranges of our existing UK genotyping. We then sought to replicate the associations observed in our UK trio collection by carrying out an imputation-based ‘pseudo case-control’ analysis with a set of 270 cases independent from the 263 UK SLE families used for our TDT association study.
The association testing of this candidate gene has provided significant evidence for the association to SLE in two independent populations. Specifically, we are confident that we have discovered a risk haplotype stretching across the promoter of P-Selectin, and we have additional evidence to support a second signal in the CR region of the gene.

This is the first report of an association for P-Selectin in the upstream region of the gene. In our 263 UK SLE families, a single risk haplotype, haplotype 5, is responsible for the signal from the upstream region of SELP—haplotype block 1 in Figure 1b. This haplotype remains significant following permutation analysis ($P_{\text{pseudo}} = 0.0330$). There are no other haplotypes in this region, which show association. The presence of this single associated haplotype in block 1 may be because of the pattern of association for the associated SNPs within the block, namely the allele frequencies of the over-transmitted alleles. In this respect, SNP 3 (rs3753306) is unique, in that it has an over-transmitted rare allele carried on the risk haplotype. For all the remaining associated polymorphisms in this region, SNPs 6, 8, 11 and 13, it is the common allele, which is the risk allele.

The higher frequency of these major risk alleles means that they are present on multiple different haplotypes. In these situations, pinpointing the ‘etiologic haplotype’ will require a greater density of typing to allow discovery of possible recombination events. However, we believe that there are associated variants in the upstream region of P-Selectin because we have replicated the association for two alleles, which are carried on the risk haplotype (SNPs 6 and 21) in the Minnesota SLE families. Furthermore, we then went on to not only show that the association for SNP 6 is stronger in the joint UK–US data set compared with the individual populations, but that the association is still significant following permutation analysis ($P_{\text{pseudo}} = 7.20 \times 10^{-4}$) (Table 2).

To determine whether we could localize the causal alleles for association to the upstream region of P-Selectin, we carried out imputation analysis for variants across P-Selectin in an independent collection of 270 UK SLE cases. The strongest result was the independent replication of SNP 13 (rs3917687) ($P = 0.0339$) (Table 3). We confirmed that it was the same (C) allele of this variant that was both under-represented in SLE cases (Table 3) and under-transmitted in the 263 UK trios (Table 1). The
case for an association in the upstream region of \textit{P-Selectin} was further strengthened by the finding that the same variant, SNP 13, was highly significantly associated following pseudo case-control analysis, using both the cases from the 263 trios and the 270 independent UK cases (Figure 2b) \((P = 2.52 \times 10^{-5}).\) Furthermore, we have also shown that the association from SNP 13 remained highly significant following permutation analysis in our 263 UK trios \((P_{\text{per}} < 1 \times 10^{-5}).\)

Having established that variants in the upstream region of \textit{P-Selectin} may contribute to disease susceptibility, we went on to look for the potential functional significance for any of the associated SNPs on the risk haplotype. We discovered that the risk C allele of rs3753306 disrupts the binding site for the HNF-1α, a trans-activating transcription factor, expression of which is predominantly restricted to the liver and kidney. It is possible that this loss-of-function risk allele in the promoter may reduce production of \textit{P-Selectin}, thereby reducing recruitment of pro-inflammatory leukocytes. However, in a mouse model, loss of \textit{P-Selectin} in MRLfas mice by a gene-targeted deletion leads to increased glomerulonephritis.\(^25\) In another mouse model of glomerulonephritis, \textit{P-Selectin}-deficient mice show increased proteinuria, glomerular damage and mortality.\(^24\) However, although these animal models have the potential to provide a link between decreased production of \textit{P-Selectin} and increased risk of renal disease, there may be a complication in the interpretation the results. This complication was first identified by Bygrave \textit{et al.}, who discovered that the 129-derived chromosome 1 segment used for gene-targeted deletions may itself produce symptoms of autoimmunity. Therefore, the observed increase in renal disease noted in the \textit{P-Selectin-deficient} mice may be the result of the effect of a 129-derived chromosome 1 segment rather than a primary effect from \textit{P-Selectin} deficiency.\(^26\) However, in human renal disease, a reduction of \textit{P-Selectin} transcription has been reported. No \textit{P-Selectin} mRNA transcript was detected in the glomerulus of renal cases, despite an increased expression of glomerular \textit{P-Selectin}.\(^28\) The increase in protein expression may reflect the release of stored \textit{P-Selectin} from storage granules rather than increased gene expression, as \textit{P-Selectin} is constitutively produced in the secretory granules of platelets (\(\alpha\)-granules) and endothelial cells (Weibel–Palade bodies), but is not expressed on the cell surface until an appropriate stimulus is provided.\(^14\)

Given the evidence presented above, it is intriguing to hypothesize that variants in \textit{P-Selectin} may have a role in renal lupus, but we do not have sufficient power within the UK SLE cohort used for this paper to assess whether the association in \textit{P-Selectin} is stronger for UK SLE families in which the affected offspring has renal disease.

An alternative functional consequence of a reduced expression of \textit{P-Selectin} is a potential reduction in the numbers of regulatory T cells. A recent publication by Urzainqui \textit{et al.}\(^27\) reported that the engagement of \textit{P-Selectin} ligand on dendritic cells with \textit{P-Selectin} may increase the production ability of the dendritic cells to generate CD4\(^+\) CD25\(^+\) Foxp3\(^+\) regulatory T cells. This reduction in the production of regulatory T cells may in turn lead to increased incidence of autoimmune disease.

However, SNP 3 (rs3753306) is not the only susceptibility allele in the upstream region of \textit{P-Selectin} because although this variant is located \(< 1\) kb upstream from the 5' UTR, it is carried on a risk haplotype stretching into the CR regions. There are several other protein domains within the bounds of this haplotype, which may contribute to disease pathogenesis. The region between SNP 3 and SNP 6 (found \(< 1\) kb downstream of the 5' UTR) contains the proximal promoter, a 249-bp section of upstream sequence, which has been shown to promote a high level of gene expression in cultured endothelial cells.\(^29\) The risk haplotype also includes the lectin and EGF domains, which are necessary for \textit{P-Selectin} to bind to its ligand PSGL.\(^29\) Changes in the affinity of the binding between \textit{P-Selectin} and its ligand will affect the recruitment of leukocytes to the sites of inflammation. Although the two SNPs we typed in these domains failed the quality control, it is probable that there are other unknown variants, which may be important in disease etiology. All these data would suggest that either SNP 3, SNP 6 and/or another variant(s) in LD with them play an important role in increasing the risk of SLE. Nevertheless, to identify the exact causal variants in \textit{P-Selectin}, it will be necessary to undertake extensive re-sequencing of the 21.4-kb risk haplotype.

Establishing a convincing association within the CR region of the gene proved to be more challenging. As there were many variants in the coding region of the gene, around the nine CR it was an attractive hypothesis to believe that the core association came from these sequences. However, as shown in Table 4, many of these variants failed to generate reliable assays or were of low minor allele frequency. This may be because of the high level of sequence similarity between the separate CRs, which affect genotyping efficiency. Although our study is the first to report an association for \textit{P-Selectin} arising from the upstream region of the gene, there has been one earlier report of an association in \textit{SELp} with lupus from an SNP in the coding region.\(^30\) The variant in question, is SNP 38 (rs3917815) a Ser/Asn non-synonymous polymorphism located in E12. We genotyped SNP 38 as part of our current study, but the variant proved to be monomorphic, and therefore was excluded from our analyses. These data agree with the genotype data on the CEU samples from the HAPMAP database. Nevertheless, in the same exon in the SNP described by Jacob, we find significant association from a Leu/Val non-synonymous variant in CR 8 (SNP 36—rs6133). The associated allele was the under-transmitted allele T (Val) allele, which tags the protective haplotype. We went on to show that the association from rs6133 remains significant after permutation analysis in our 263 UK families \((P_{\text{per}} < 1 \times 10^{-5})\) (Table 2). The consensus repeat domains are thought to be important, but not essential, in binding \textit{P-Selectin} to its ligand on leukocytes. It is possible that the over-transmitted G (Leu) allele of SNP 36 results in a stronger ligand binding, and if so, may contribute to an increase in inflammatory activity, resulting from greater recruitment of leukocytes to sites of inflammation. We also know that the same G (Leu) allele of this variant has earlier been reported as being associated with thromboembolic stroke\(^32\) and atopy.\(^33\) As in our UK SLE families, the T allele of SNP 36 (rs6133) is unique to the under-transmitted haplotype in block 2 (Figure 1b) and we observed no LD with the risk haplotype in block 1, we think that rs6133 may tag a second signal within \textit{P-Selectin}. 

\(\text{Figures 1b and 2b}\)
Materials and methods

Details of study cohorts

The collection of the UK Caucasian SLE families consisted of a total of 263 complete trios and 270 independent UK families. All probands conformed to the American College of Rheumatology criteria for SLE, with a diagnosis of SLE being established by telephone interview, health questionnaire and details from clinical notes. Written consent was obtained from all participants, including relatives. In the United Kingdom, ethical approval was obtained from Multi-Centre Research Ethics Committee. Genomic DNA from the UK samples was isolated from...
anti-coagulated whole blood by a standard phenol-
chloroform extraction.

**Non-UK family collection for corroborative genotyping**

Corroboration of the UK association study in SELP was investigated using a total of 306 US parental-affected trios, taken from the Minnesota SLE collection. These studies were approved by the Human Subject Institutional Review Boards at the University of Minnesota, and informed consent was obtained from all the patients.

**Selection of markers**

In this study a total of 52 markers were selected from across P-Selectin, taken from the public databases, dbSNP (http://www.ncbi.nlm.nih.gov/SNP/index.html) and from HAPMAP (http://www.hapmap.org/). A single SNP (A1969G) was identified by re-sequencing 60 UK probands from HAPMAP (http://www.hapmap.org/). A single marker and the Breslow–Day test in PLINK was used to compare the ratio of transmitted to untransmitted (T:U) families for each marker and the Breslow–Day test in PLINK was used to test homogeneity of odds ratios between the two populations.

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 with affected offspring. This analysis was carried out using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/).

Haplotyp–TDT was carried out using Haplovie. For SNPs in each gene, haplotype patterns were generated using haplotype block definitions on the basis of the solid-spine algorithm, described on the Haplovie website, with an LD cut-off at \( D' = 0.6 \). This program constructs haplotypes on the basis of the \( D' \) measure of LD, together with a logarithm of the odds (LOD) score as measure of significance and 95% confidence intervals to state the accuracy of the \( P \)-value. The pairwise linkage LD for SNPs across each gene was confirmed by the \( r^2 \) values. Only markers having minor allele frequency of \( > 5\% \) were included in the haplotype constructions and haplotypes with a frequency of \( > 3.0\% \) were included in the LD diagrams.

**Imputation analysis**

Imputation was carried out on the cases and pseudo-controls using IMPUTE, by the method described by Marchini et al.\(^\text{39}\) We imputed all SNPs in the HAPMAP within the range of our data plus 20 kb either side—giving us 141 SNPs (110 imputed) for our case-control analysis.

The output from IMPUTE gives probabilities for each genotype, rather than point estimates. The use of probabilities allowed us to account for the uncertainty in imputation within the case-control analysis using SNPTEST.

**Pseudo case-control analysis using SNPTEST**

For each affected child in the study, we created a matched control using the untransmitted alleles from their parents. As phasing the data was possible, we could have matched each case to all haplotypes that could have been passed on from the parents. However, this would not have been possible for the imputed SNPs, therefore we only matched to genotypes made up from the untransmitted alleles.

The data were analyzed using classical methods, which returned \( P \)-values. Logistic regression additive models were fitted with SLE as the outcome.

We carried out a meta-analysis by combining the \( P \)-values from the United Kingdom and Minnesota parental-affected trios and the UK-independent pseudo case-control study using Fisher’s method.\(^\text{40}\)

**Conflict of interest**

The authors declare no conflict of interest.

**Acknowledgements**

This work was funded through a Senior Fellowship Award from the Wellcome Trust to Timothy J Vyse and we are grateful for the support from the NIHR Biomedical Research Centre funding scheme at the Imperial College. We acknowledge the work of Andrew Wong recruiting patients and families into the study and we would like to thank our clinical colleagues for helping us 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.
References

1 Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Asundi N et al. Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in African-American pedigrees. Proc Natl Acad Sci USA 1998; 95: 14869–14874.

2 Shai R, Quismorio Jr FP, Li L, Kwon OJ, Morrison J, Wallace DJ et al. Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families. Hum Mol Genet 1999; 8: 639–644.

3 Gaffney PM, Kears GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML et al. A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families. Proc Natl Acad Sci USA 1998; 95: 14875–14879.

4 Gaffney PM, Ortmann WA, Selby SA, Shark KB, Ockenden RS, Ortmann WA et al. A genome-wide screen of large-familiy linkage for susceptibility genes in systemic lupus erythematosus. Proc Natl Acad Sci USA 1999; 96: 422–429.

5 Johanneson B, Lima G, von SJ, Alarcon-Segovia D, Alarcon-Segovia D et al. The genetic basis of systemic lupus erythematosus. Semin Immunol 2001; 13: 397–408.

6 Tsao BP. The genetics of human systemic lupus erythematosus. Trends Immunol 2003; 24: 595–602.

7 Johnston GI, Bliss GA, Newman PJ, McEver RP. Structure of the human gene encoding granule membrane protein-140, a member of the selectin family of adhesion receptors for leukocytes 1. J Biol Chem 1990; 265: 21381–21385.

8 Russell AI, Cunningham GA, Chadha S, Roberton C, Russell AI et al. Deficiency of P-selectin glycoprotein ligand-1 leads to accelerated development of glomerulonephritis and increased expression of CC chemokine ligand 2 in lupus-prone mice. J Immunol 2006; 177: 8748–8756.

9 Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Ortmann WA et al. A genome-wide search of large-family linkage for susceptibility genes in systemic lupus erythematosus. Proc Natl Acad Sci USA 1999; 96: 422–429.

10 Flaumenhaft R. P-Selectin: Basic Aspects. J Thromb Thrombolysis 1998; 6: 93–102.

11 Gibson RM, Kansas GS, Tedder TF, Furie B, Furie BC. Lectin and epidermal growth factor domains of P-selectin at physiologic density are the recognition unit for leukocyte recruitment. Blood 1995; 85: 151–158.

12 Crovello CS, Furie BC, Furie B. Rapid phosphorylation and selective dephosphorylation of P-selectin accompanies platelet activation. J Biol Chem 1993; 268: 14590–14595.

13 Celi A, Pellegrini G, Lorenzet R, De Blasi A, Ready N, Furie BC et al. P-selectin induces the expression of tissue factor on monocytes. Proc Natl Acad Sci USA 1994; 91: 8767–8771.

14 Patel KD, Cuvelier SL, Wiehler S. Selectins: critical mediators of leukocyte recruitment. Semin Immunol 2002; 14: 73–81.

15 Chen M, Geng JG. P-selectin mediates adhesion of leukocytes, platelets, and cancer cells in inflammation, thrombosis, and cancer growth and metastasis. Arch Immunol Ther Exp (Warsz) 2006; 54: 75–84.

16 Veldman GM, Bean KM, Cumming DA, Eddy RL, Sait SN, Shows TB. Genomic organization and chromosomal localization of the gene encoding human P-selectin glycoprotein ligand. J Biol Chem 1995; 270: 16470–16475.

17 Moore KL, Patel KD, Bruehl RE, Li F, Johnson DA, Lichenstein JP et al. P-selectin glycoprotein ligand-1 mediates rolling of human neutrophils on P-selectin. J Cell Biol 1995; 128: 661–671.

18 Sanchez A, Abelson AK, Sabio JM, Gonzalez-Gay MA, Ortego-Centeno N, Jimenez-Alonso J et al. Association of a CD24 gene polymorphism with susceptibility to systemic lupus erythematosus. Arthritis Rheum 2007; 56: 3080–3086.

19 Zhou Q, Ramnahan K, Lin S, Robinson N, Li O, Liu X et al. CD24 is a genetic modifier for risk and progression of multiple sclerosis. Proc Natl Acad Sci USA 2003; 100: 15041–15046.

20 Sanchez E, Fernandez-Gutierrez B, Gonzalez-Gay MA, Balsa A, Garcia A, Rodriguez L et al. Investigating the role of CD24 gene polymorphisms in rheumatoid arthritis. Ann Rheum Dis 2008; 67: 1197–1198.

21 Mayadas TN, Johnson RC, Rayburn H, Hynes RO, Wagner DD. Leukocyte rolling and extravasation are severely compromised in P-selectin-deficient mice. Cell 1993; 74: 541–554.

22 Cunningham GA, Graham DS, Graham RR, Manku H, Wong AK, Whittaker JC, Gaffney PM et al. Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. Nat Genet 2008; 40: 83–89.

23 He X, Schoeb TR, Panoskaltsis-Mortari A, Zimm KR, Kesterson RA, Zhang J et al. Deficiency of P-selectin or P-selectin glycoprotein ligand-1 leads to accelerated development of glomerulonephritis and increased expression of CC chemokine ligand 2 in lupus-prone mice. J Immunol 2006; 177: 8748–8756.

24 Rosenkranz AR, Mendrick DL, Cotran RS, Mayadas TN. P-selectin deficiency exacerbates experimental glomerulonephritis: a protective role for endothelial P-selectin in inflammation. J Clin Invest 1998; 103: 649–659.

25 Bygraves AE, Rose KL, Cortese-Hernandez J, Warren J, Bigby RJ, Cook HT et al. Spontaneous autoimmune in 129 and C57BL/6 mice-implications for autoimmunity described in gene-targeted mice. PLoS Biol 2004; 2: E243.

26 Segawa C, Wada T, Takaeda M, Furutichi K, Matsuda I, Hisada Y et al. In situ expression and soluble form of P-selectin in human glomerulonephritis. Kidney Int 1997; 52: 1054–1063.

27 Uraznaiqui A, Martinez del Hoyo G, Lanana A, de la Fuente H, Barreiro O, Olazabal IM et al. Functional role of P-selectin glycoprotein ligand 1/P-selectin interaction in the generation of tolerogenic dendritic cells. J Immunol 2007; 179: 7457–7465.

28 Pan J, McEver RP. Characterization of the promoter for the human P-selectin gene. J Biol Chem 1993; 268: 22680–22686.

29 Mehta P, Patel KD, Laue TM, Erickson HP, McEver RP. Soluble monomeric P-selectin containing only the lectin and epidermal growth factor domains binds to P-selectin glycoprotein ligand-1 on leukocytes. Blood 1997; 90: 2381–2389.

30 Holenbaugh D, Bajorath J, Stenkamp R, Aruffo A. Interaction of P-selectin (CD62) and its cellular ligand: analysis of critical residues. Biochemistry 1993; 32: 2960–2966.

31 Jacob CO, Reiff A, Armstrong DL, Myones BL, Silverman E, Klein-Gitelman M et al. Identification of novel susceptibility genes in childhood-onset systemic lupus erythematosus using a uniquely designed candidate gene pathway platform. Arthritis Rheum 2007; 56: 4164–4173.

32 Zee CY, Cook NR, Cheng S, Reynolds R, Erlich HA, Lindpaintner K et al. Polymorphism in the P-selectin and interleukin-4 genes as determinants of stroke: a population-based, prospective genetic analysis. Hum Mol Genet 2004; 13: 389–396.

33 Bourgain C, Hofmann S, Nicolae R, Newman D, Steiner L, Walker K et al. Novel case-control test in a founder population identifies P-selectin as an atopy-susceptibility locus. Am J Hum Genet 2003; 73: 612–629.

34 Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982; 25: 1271–1277.

35 Storm N, rhnoer-Patel B, van den BD, Rodi CP. MALDI-TOF mass spectrometry-based SNP genotyping. Methods Mol Biol 2003; 212: 241–262.

36 Lewontin RC. The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics 1964; 49: 46–47.

37 Hill WG, Weir BS. Maximum-likelihood estimation of gene location by linkage disequilibrium. Am J Hum Genet 1994; 54: 705–714.

38 Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 1995; 29: 311–322.
Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007; 39: 906–913.

Fisher RA. Statistical Methods for Research Workers. Hafner: New York, 1954.

Supplementary Information accompanies the paper on Genes and Immunity website (http://www.nature.com/gene)