

## ORIGINAL ARTICLE

# Association of Systemic Lupus Erythematosus with *C8orf13*–*BLK* and *ITGAM*–*ITGAX*

Geoffrey Hom, Ph.D., Robert R. Graham, Ph.D., Barmak Modrek, Ph.D., Kimberly E. Taylor, Ph.D., M.P.H., Ward Ortmann, B.S., Sophie Garnier, Ph.D., Annette T. Lee, Ph.D., Sharon A. Chung, M.D., Ricardo C. Ferreira, B.S., P.V. Krishna Pant, Ph.D., Dennis G. Ballinger, Ph.D., Roman Kosoy, Ph.D., F. Yesim Demirci, M.D., M. Ilyas Kamboh, Ph.D., Amy H. Kao, M.D., M.P.H., Chao Tian, B.S., Iva Gunnarsson, M.D., Ph.D., Anders A. Bengtsson, M.D., Ph.D., Solbritt Rantapää-Dahlqvist, M.D., Ph.D., Michelle Petri, M.D., Susan Manzi, M.D., M.P.H., Michael F. Seldin, M.D., Ph.D., Lars Rönnblom, M.D., Ph.D., Ann-Christine Syvänen, Ph.D., Lindsey A. Criswell, M.D., M.P.H., Peter K. Gregersen, M.D., and Timothy W. Behrens, M.D.

## ABSTRACT

**BACKGROUND**

Systemic lupus erythematosus (SLE) is a clinically heterogeneous disease in which the risk of disease is influenced by complex genetic and environmental contributions. Alleles of HLA-DRB1, *IRF5*, and *STAT4* are established susceptibility genes; there is strong evidence for the existence of additional risk loci.

**METHODS**

We genotyped more than 500,000 single-nucleotide polymorphisms (SNPs) in DNA samples from 1311 case subjects with SLE and 1783 control subjects; all subjects were North Americans of European descent. Genotypes from 1557 additional control subjects were obtained from public data repositories. We measured the association between the SNPs and SLE after applying strict quality-control filters to reduce technical artifacts and to correct for the presence of population stratification. Replication of the top loci was performed in 793 case subjects and 857 control subjects from Sweden.

**RESULTS**

Genetic variation in the region upstream from the transcription initiation site of the gene encoding B lymphoid tyrosine kinase (*BLK*) and *C8orf13* (chromosome 8p23.1) was associated with disease risk in both the U.S. and Swedish case-control series (rs13277113; odds ratio, 1.39;  $P=1\times 10^{-10}$ ) and also with altered levels of messenger RNA in B-cell lines. In addition, variants on chromosome 16p11.22, near the genes encoding integrin alpha M (*ITGAM*, or *CD11b*) and integrin alpha X (*ITGAX*), were associated with SLE in the combined sample (rs11574637; odds ratio, 1.33;  $P=3\times 10^{-11}$ ).

**CONCLUSIONS**

We identified and then confirmed through replication two new genetic loci for SLE: a promoter-region allele associated with reduced expression of *BLK* and increased expression of *C8orf13* and variants in the *ITGAM*–*ITGAX* region.

The authors' affiliations are listed in the Appendix. Address reprint requests to Dr. Behrens at Genentech, 1 DNA Way, 45-3B, South San Francisco, CA 94080, or at behrens.timothy@gene.com.

Drs. Hom and Graham contributed equally to this article.

This article (10.1056/NEJMoa0707865) was published at www.nejm.org on January 20, 2008.

N Engl J Med 2008;358:900-9.  
Copyright © 2008 Massachusetts Medical Society.

**S**YSTEMIC LUPUS ERYTHEMATOSUS (SLE) is a chronic autoimmune disease with strong genetic and environmental components.<sup>1-3</sup> Autoantibodies play an important role in the pathogenesis of SLE, and the diverse clinical manifestations of the disease are caused by the deposition of antibody-containing immune complexes in blood vessels, leading to inflammation in the kidney, brain, and skin. Direct pathogenic effects of the autoantibodies contribute to hemolytic anemia and thrombocytopenia.

During the past 20 years, many linkage and candidate-gene studies have been performed to identify genetic factors contributing to a susceptibility to SLE. For example, haplotypes carrying the HLA class II alleles DRB1\*0301 and DRB1\*1501 are clearly associated with SLE and the presence of antibodies to nuclear autoantigens.<sup>4-6</sup> More recently, variants of the genes encoding interferon regulatory factor 5 (*IRF5*) and signal transducer and activator of transcription 4 (*STAT4*) have been discovered to be risk factors for SLE.<sup>7-10</sup> The identification of *IRF5* and *STAT4* as SLE risk genes supports the hypothesis that the type I interferon pathway is central to disease pathogenesis.<sup>11-14</sup>

In this report, we describe the results of a genomewide scan of samples from North American subjects of European descent (1311 case subjects with SLE and 3340 control subjects) and a replication analysis of Swedish case-control subjects. We identify two novel genetic loci — *C8orf13-BLK* and *ITGAM-ITGAX* — that contribute to the risk of SLE.

## METHODS

### SUBJECTS

Case subjects included 1435 patients with SLE for whom DNA samples were obtained from the following collections: 338 from the Autoimmune Biomarkers Collaborative Network, a repository funded by the National Institute of Arthritis and Musculoskeletal and Skin Diseases<sup>15</sup>; 141 from the Multiple Autoimmune Disease Genetics Consortium<sup>16</sup>; 613 from the University of California, San Francisco (UCSF), Lupus Genetics Project<sup>10,17</sup>; 335 from the University of Pittsburgh Medical Center<sup>18</sup>; and 8 from the Feinstein Institute for Medical Research. The European descent of all case subjects with SLE was determined by self-report with the use of a multiple-choice questionnaire. (In general, subjects needed to have at least

three grandparents of European origin.) The diagnosis of SLE (fulfillment of four or more of the criteria of the American College of Rheumatology [ACR]<sup>19</sup>) was confirmed in all cases by review of medical records (for 94% of case subjects) or by written documentation of criteria by the treating rheumatologist (6%). Clinical data were reviewed and tabulated at each institution. The counts and percentages for each of the 11 ACR classification criteria for SLE<sup>19</sup> are listed in Table 1 of the Supplementary Appendix (available with the full text of this article at [www.nejm.org](http://www.nejm.org)).

A total of 3583 samples from control subjects were examined in the association analyses. As part of this project, 1861 control samples from the New York Cancer Project collection<sup>20</sup> were selected and then genotyped on the basis of self-described ancestral origin, sex, and age. In addition, genotype data from 1722 control samples (all self-described North Americans of European descent) were obtained from the publicly available iControlDB database ([www.illumina.com/pages.ilmn?ID=231](http://www.illumina.com/pages.ilmn?ID=231)).

To test for replication, we genotyped DNA from an independent collection of samples from 793 Swedish patients with SLE (all of whom fulfilled four or more of the classification criteria for SLE, as defined by the ACR) and 857 healthy Swedish control subjects. The patients were from rheumatology clinics at the Lund, Uppsala, Karolinska (Solna), and Umeå University Hospitals.<sup>7</sup> The institutional review board at each collaborating center approved the study, and all subjects gave informed written consent.

### GENOTYPING

A total of 1861 samples from control subjects from the New York Cancer Project were genotyped on the Illumina HumanHap550 Genotyping BeadChip<sup>21</sup> at the Feinstein Institute. A total of 1465 samples (464 from case subjects and 1001 from control subjects) were genotyped on the HumanHap550v1 chip, and 1875 samples (1015 from case subjects and 860 from control subjects) were genotyped on the HumanHap550v3 chip. Genotype data from 1452 of the control samples were submitted to iControlDB and made publicly available in 2007. An additional, independent set of 1722 samples that were genotyped on the HumanHap550 BeadChip was obtained from studies 66 and 67 of the iControlDB. Samples from case subjects were genotyped at the Feinstein Institute in serial phases; series 1 consisted of the 479 case

subjects from the Autoimmune Biomarkers Collaborative Network and Multiple Autoimmune Disease Genetics Consortium, series 2 consisted of the 613 case subjects from the UCSF Lupus Genetics Project, and series 3 consisted of the 387 case subjects from the University of Pittsburgh Medical Center and the Feinstein Institute. The 545,080 single-nucleotide polymorphisms (SNPs) present on both HumanHap550 versions 1 and 3 were advanced into the analysis. Case and control samples with average call rates of less than 80% across the chip underwent repeated genotyping.

In the Swedish series, the SNPs rs11574637 (*ITGAM-ITGAX*) and rs13277113 (*C8orf13-BLK*) were genotyped with the use of homogeneous single-base extension assays with fluorescence polarization detection at the SNP Technology Platform in Uppsala ([www.genotyping.se](http://www.genotyping.se)) and reagents (PerkinElmer).<sup>22</sup> The genotype call rate in the samples was 96%, and the reproducibility was 100%, on the basis of duplicate assays of 4.6% of the genotypes. Samples from a three-generation pedigree from the Centre d'Etude du Polymorphisme Humain (CEPH) with 20 members were genotyped in parallel with the study samples, and no deviation from mendelian inheritance was observed for either of the SNPs.

#### DATA-QUALITY FILTERS

Details on the data-quality filters used in the study, tests for association and for heterogeneity among the three case-control studies, and removal of population outliers with the use of EIGENSTRAT software<sup>23</sup> are available in the Supplementary Appendix. Series 1 consisted of 411 case subjects and 1047 control subjects, series 2 consisted of 595 case subjects and 1516 control subjects, and series 3 consisted of 305 case subjects and 777 control subjects. Overall, women accounted for 93% of case subjects and 62% of control subjects. No significant differences in allele frequencies were noted between men and women.

A total of 3323 SNPs were removed because more than 2% of data were missing in at least one series or missing data were unequally distributed between case and control subjects (differential missingness,  $P < 1 \times 10^{-3}$ ). Another 13 SNPs in the pseudoautosomal region of chromosome X showed no significant association and were excluded from further analysis. Filtering of samples and markers was conducted with the use of

analytical modules within the PLINK software program.<sup>24</sup> For each series, a total of 502,033 SNPs were advanced into downstream analyses.

#### GENE-EXPRESSION ANALYSIS

We examined gene expression in B-cell lines transformed by the Epstein-Barr virus (EBV) from 210 unrelated, healthy HapMap subjects (GENEVAR project, [www.sanger.ac.uk/humgen/genevar/](http://www.sanger.ac.uk/humgen/genevar/))<sup>25</sup> and an independent set of 400 EBV-transformed B cells.<sup>26</sup> Additional details regarding these analyses are available in the Supplementary Appendix.

#### STATISTICAL ANALYSIS

The association between all SNPs and a susceptibility to SLE was calculated with the use of two-by-two contingency tables. A genomic control inflation factor ( $\lambda_{gc}$ ) was then calculated for each sample series.<sup>27</sup> The genomic control inflation factor is a metric based on the median chi-square that reflects whether the bulk of the distribution conforms to the null hypothesis ( $\lambda_{gc} = 1.0$ ). A  $\lambda_{gc}$  value of more than 1 indicates an elevation of the average chi-square association statistic owing to systemic technical artifacts or the presence of population stratification. The 50 loci with the strongest associations are listed in Table 2 of the Supplementary Appendix, and the corrected summary statistics of population stratification for all SNPs passing quality-control filters from each series and the combined association statistics are available from dbGAP (accession number phs000122.v1.p1).

## RESULTS

#### GENOMEWIDE ASSOCIATION ANALYSIS

A total of 502,033 SNPs on the Illumina chips passed quality-control filters and were tested for association with SLE in a staged fashion with the use of three case-control series (Table 1). A combined association statistic was calculated by the addition of the z scores converted from the EIGENSTRAT-corrected chi-square test statistic,<sup>23</sup> weighted for series size and adjusted for the residual  $\lambda_{gc}$  of each series (see the Supplementary Appendix).

A comparison of the observed P values for the meta-analysis with the P values for a null distribution is shown in Figure 1A. Significant deviation from the null distribution was observed at the tail of the distribution, which may indicate the presence of true positive associations. A strong as-

**Table 1. Summary of Samples from 1311 Case Subjects and 3340 Control Subjects in the Genomewide Association Study.\***

Series	No. of Case Subjects	Study	No. of Control Subjects†	Genomic Control Inflation Factor‡
1	411	Autoimmune Biomarkers Collaborative Network and Multiple Autoimmune Disease Genetics Consortium	1047	1.015
2	595	UCSF Lupus Genetics Project	1516	1.038
3	305	University of Pittsburgh Medical Center	777	1.030

\* UCSF denotes University of California, San Francisco.

† A total of 1783 samples from the New York Cancer Project were genotyped (559 in series 1, 802 in series 2, and 422 in series 3). Genotyping data from an additional 1557 control subjects were obtained from the iControlDB (study numbers 66 and 67), including 488 in series 1, 714 in series 2, and 355 in series 3.

‡ The genomic control inflation factor for each data set was calculated after the use of quality-control filters and correction for population stratification.

sociation with SLE was noted for three established risk loci. In the HLA class II region, rs2187668 is a near perfect predictor of the DRB1\*0301 allele<sup>28</sup> and was the variant most strongly associated with SLE in the combined analysis ( $P=3\times 10^{-21}$ ). An additional 157 SNPs in the HLA region, many of which are correlated with the DRB1\*0301 allele, had observed  $P$  values of less than  $5\times 10^{-7}$  (Fig. 1B). A strong association was observed with variants linked to the well-validated risk haplotype of *IRF5* (e.g., rs10488631;  $P=2\times 10^{-11}$ ).<sup>7-9</sup> In addition, an association with *STAT4* was observed (rs7574865,  $P=9\times 10^{-14}$ ). An association between *STAT4* variants and both SLE and rheumatoid arthritis was reported recently.<sup>10</sup> Our SLE data set overlaps with that of the earlier study<sup>10</sup> and includes an additional 341 case subjects and 2905 control subjects who were not included in the previous analysis. After removing variants in HLA, *IRF5*, and *STAT4* from the expected versus observed chi-square analysis, a deviation of  $P$  values from the null distribution remained, suggesting the existence of novel SLE loci (Fig. 1A). Multiple SNPs near the B lymphoid tyrosine kinase (*BLK*) gene and in a region that contains the integrin alpha M (*ITGAM*) and integrin alpha X (*ITGAX*) genes were highly associated with SLE in the combined analysis (Fig. 1B). Neither of these genes or regions has previously been implicated in SLE susceptibility.

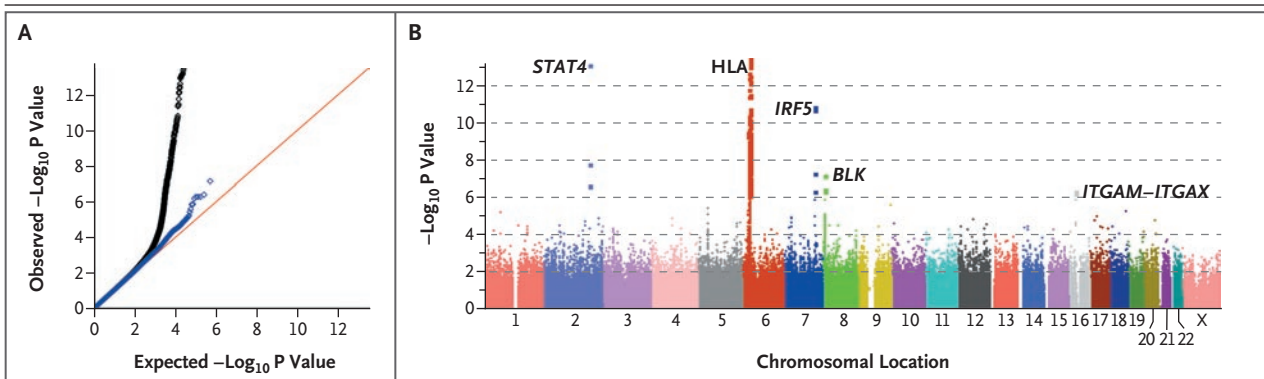
#### C8orf13–BLK

Several variants on the short arm of chromosome 8 (8p23.1) were associated with SLE (Fig. 2 and Table 2, and Table 5 of the Supplementary Appendix). The A allele of rs13277113 was highly en-

riched in the sample from U.S. case subjects as compared with controls ( $P=8\times 10^{-8}$ ; combined odds ratio, 1.39; 95% confidence interval [CI], 1.26 to 1.54). To confirm this initial observation, an independent collection of 793 samples from SLE case subjects and 857 matched controls from Sweden was typed for rs13277113, and a convincing association between the minor A allele and SLE was also observed ( $P=4\times 10^{-4}$ ; odds ratio, 1.33; 95% CI, 1.13 to 1.55) (Table 2). A combined analysis of rs13277113 with the use of both the U.S. and Swedish samples showed  $P=1\times 10^{-10}$ , which meets the rigorous criterion of  $P<5\times 10^{-8}$  for the significance of a genomewide association.<sup>31</sup>

The SNP rs13277113 maps to the interval between two genes transcribed in opposite directions: *BLK*, a tyrosine kinase in the src family that signals downstream of the B-cell receptor, and *C8orf13*, a ubiquitously expressed gene of unknown function (Fig. 2). No known coding-region variants of *BLK* or *C8orf13* are in linkage disequilibrium with rs13277113.

Common genetic variation has been shown to correlate with levels of *cis* gene expression.<sup>25,26,32</sup> Using a gene-expression data set generated from transformed B-cell lines of 210 unrelated HapMap samples,<sup>25</sup> we observed that the risk A allele of rs13277113 was associated with lower levels of messenger RNA (mRNA) expression of *BLK* (Fig. 2B). Homozygotes for the A allele had a level of expression that was approximately 50% of that of homozygotes for the G allele, and A/G heterozygotes had intermediate levels. The expression of the *C8orf13* gene also correlated with the risk haplotype, but in the opposite direction. The A allele of rs13277113 was associated with



**Figure 1.** Identification of Five Major Loci Associated with Systemic Lupus Erythematosus in a Genomewide Association Study.

Data represent 502,033 variants of single-nucleotide polymorphisms (SNPs) that were genotyped in three series of DNA samples from 1311 case subjects and 3340 control subjects. Panel A shows a quantile–quantile plot of the observed P value distribution, as compared with the expected null P value distribution. The black diamonds represent all P values, and the blue diamonds represent P values after the exclusion of variants in the HLA region, *IRF5*, and *STAT4*. Panel B shows the  $-\log_{10}$  P values from the combined analysis, according to chromosome. Not shown in Panel B are an additional 34 variants in the HLA region with  $P < 1 \times 10^{-13}$ .

higher expression of *C8orf13* in the transformed lines, whereas the G allele was significantly associated with lower expression (Fig. 2C). Again, A/G heterozygotes showed intermediate levels of expression. The expression of a number of control mRNAs (e.g., beta-actin and glyceraldehyde-3-phosphate dehydrogenase [GAPDH]) did not vary in the cell lines on the basis of genotype at rs13277113 (Table 3 of the Supplementary Appendix). Consistent allelic differences in *BLK* expression with statistical significance were observed in all HapMap populations, except the Yoruba population, in which the risk allele is less frequent (Table 4 of the Supplementary Appendix). These results were confirmed by analysis of another expression data set obtained from an independent sample of 400 transformed B-cell lines.<sup>26</sup> In this data set, a marker in linkage disequilibrium with rs13277113 (rs4840568,  $r^2 = 0.77$ ) was associated with both decreased expression of *BLK* ( $P = 9 \times 10^{-27}$ , probe 206255\_at) and increased expression of *C8orf13* ( $P = 5 \times 10^{-35}$ , probe 226614\_s\_at), with effect sizes similar to those observed in the HapMap data set.

Multiple conserved sites of transcription-factor binding, including motifs for interferon regulatory factor 1 (IRF1), peroxisome proliferator-activated receptor gamma (PPARG), and an interferon-stimulated response element, are located in the 5' region of *BLK* and *C8orf13*. However, neither rs13277113 nor variants in linkage disequilibrium ( $r^2 > 0.5$ ) altered known sites of transcription-factor binding or other known functional nucleic acid

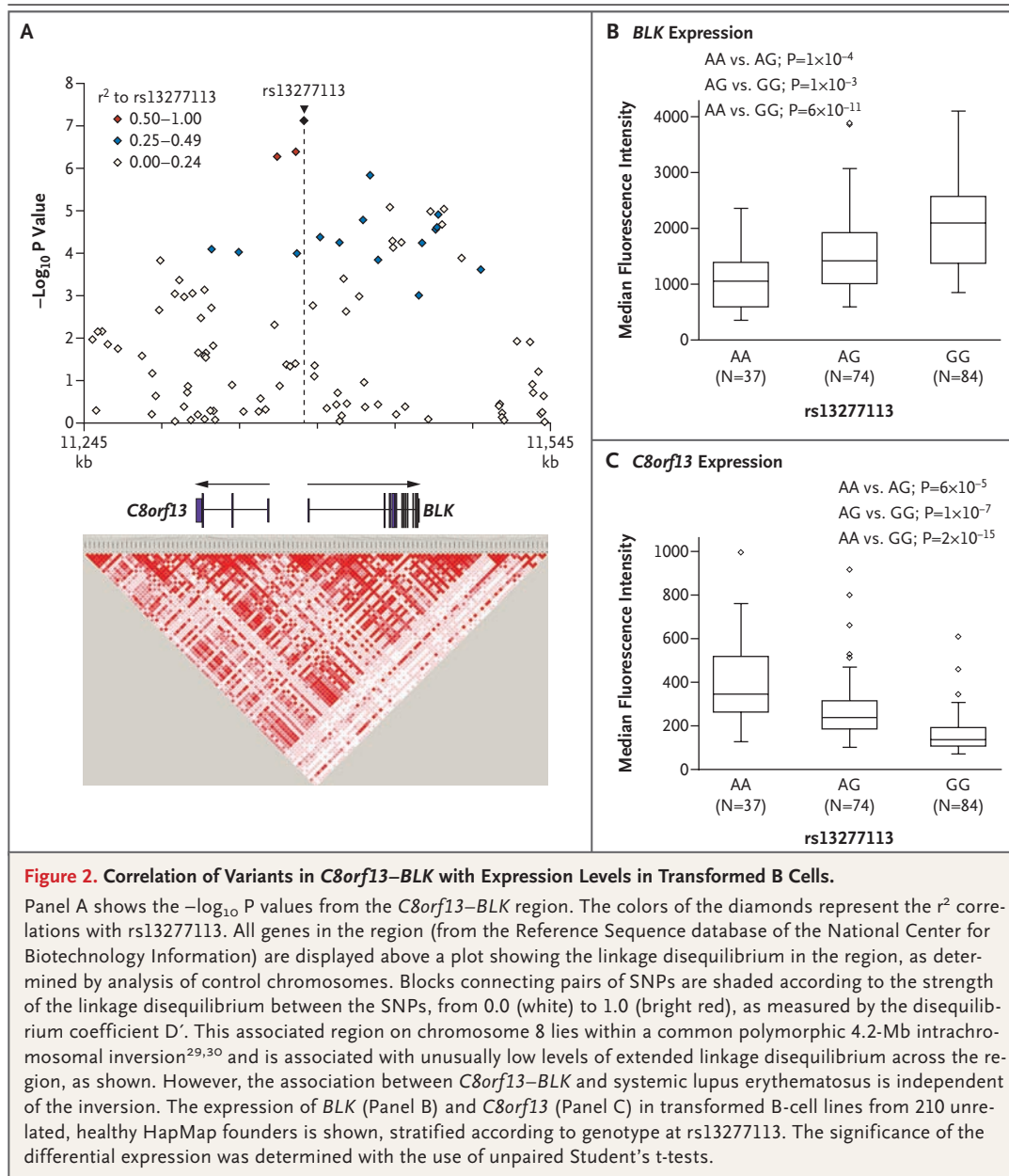
motifs. We conclude that rs13277113, or a variation that is strongly associated with rs13277113, alters the level of mRNA expression of *BLK* and *C8orf13*.

#### ITGAM–ITGAX

Variants within a cluster of genes encoding the integrin alpha chains on chromosome 16 were also significantly associated with SLE (Table 2 and Fig. 3). Reproducible association of the C allele of rs11574637 was observed across the three series of SLE cases ( $P = 5 \times 10^{-7}$ ; odds ratio, 1.30; 95% CI, 1.17 to 1.45). The C allele of rs11574637 showed similarly strong enrichment in the Swedish replication series ( $P = 4 \times 10^{-7}$ ; odds ratio, 1.59; 95% CI, 1.33 to 1.91) (Table 2), and combined analysis showed a combined P value of  $3 \times 10^{-11}$  (Table 2).

SNP rs11574637 is part of a block of correlated SNPs that covers approximately 150 kb and encodes several genes, including *ITGAM* and the 5' portion of *ITGAX* (Fig. 3A). Both *ITGAM* and *ITGAX* are expressed at detectable levels in EBV-transformed B cells, but rs11574637 was not significantly associated with mRNA levels of either gene (data not shown).

Of potential interest is the observation that in the North American control subjects, rs11574637 was correlated with two nonsynonymous variants of *ITGAM*. The first SNP, rs1143678 ( $r^2 = 0.85$ ), results in a Pro1146Ser substitution (association with SLE,  $P = 3 \times 10^{-5}$ ). The C allele of rs11574637 and the 1146Ser allele form a haplotype on 18.2%



of control chromosomes; the C allele is also present on a distinct haplotype that is present on approximately 2% of chromosomes from controls lacking the 1146Ser allele. The second nonsynonymous SNP, rs1143683 ( $r^2=0.45$  in HapMap CEU [CEPH Utah residents with ancestry from northern and western Europe]), results in an Ala858Val substitution and was not directly genotyped in this study. At this time it is not known whether the *ITGAM* nonsynonymous variants or additional alleles underlie the association within the *ITGAM*–*ITGAX* region.

## DISCUSSION

By studying a large number of patients with SLE (1311) and an even larger group of control subjects (3340), we had excellent power to detect associations between SLE and genetic variants. The strong associations between SLE and SNPs in the HLA region, *IRF5*, and *STAT4* serve as positive controls for the experiment and confirm that these loci confer susceptibility to SLE.

The src family tyrosine kinase *BLK* is an interesting new candidate gene for SLE. Expression

**Table 2. Association between Variants in C8orf13–BLK and ITGAM–ITGAX and Systemic Lupus Erythematosus in a Genomewide Association Study, a Replication Series, and an Analysis of the Combined Studies.**

Locus	Chromosome	SNP	Position	Minor Allele	Genomewide Association Study*					Replication Series†					Combined Studies‡		
					Allele Frequency	Odds Ratio (95% CI)	Chi-Square	P Value§	Allele Frequency	Odds Ratio (95% CI)	Chi-Square	P Value	Odds Ratio (95% CI)	P Value¶	Odds Ratio (95% CI)	P Value	
C8orf13–BLK	8p21.3	rs2736340	11,381,383	T	0.292	0.232	1.37 (1.24–1.51)	25.7	4×10 <sup>-7</sup>	Case	Control						
C8orf13–BLK	8p21.3	rs13277113	11,386,596	A	0.293	0.229	1.39 (1.26–1.54)	28.9	8×10 <sup>-8</sup>	0.305	0.248	1.33 (1.13–1.55)	12.7	4×10 <sup>-4</sup>	1.39 (1.28–1.51)	1×10 <sup>-10</sup>	
ITGAM–ITGAX	16p11.2	rs9937837	31,206,441	G	0.316	0.265	1.28 (1.16–1.41)	24.7	7×10 <sup>-7</sup>								
ITGAM–ITGAX	16p11.2	rs11574637	31,276,376	C	0.233	0.189	1.30 (1.17–1.45)	25.1	5×10 <sup>-7</sup>	0.222	0.152	1.59 (1.33–1.91)	25.6	4×10 <sup>-7</sup>	1.33 (1.22–1.46)	3×10 <sup>-11</sup>	

\* In series 1 through 3 of the genomewide association study, samples from 1311 case subjects and 3340 control subjects were genotyped for 550,000 single-nucleotide polymorphisms (SNPs).

† In the Swedish replication series, samples from 793 case subjects and 857 control subjects were genotyped.

‡ The combined genomewide study and replication study included samples from 2104 case subjects and 4197 control subjects.

§ The combined P values were calculated by the addition of the corrected z scores weighted according to series size and adjusted according to the residual genomic control inflation factor for each series.

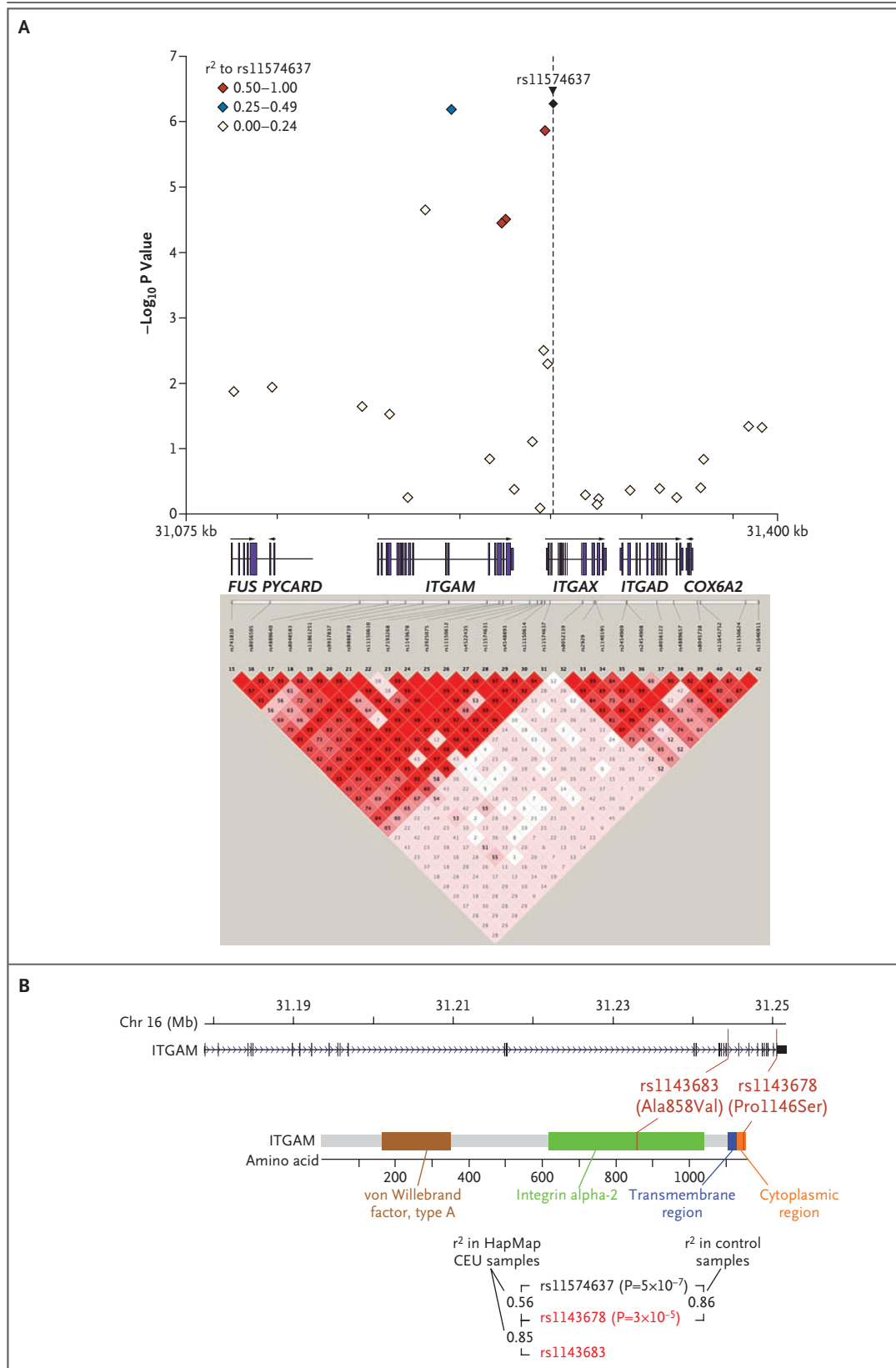
**Figure 3 (facing page). Correlation of Variants in ITGAM–ITGAX with Systemic Lupus Erythematosus.**

Panel A shows the  $-\log_{10}$  P values from the ITGAM–ITGAX region. The colors of the diamonds represent the  $r^2$  correlations with rs11574637. All genes in the region (from the Reference Sequence database of the National Center for Biotechnology Information) are displayed above a plot showing the linkage disequilibrium in the region, as determined by analysis of control chromosomes. Blocks connecting pairs of SNPs are shaded according to the strength of the linkage disequilibrium between the SNPs, from 0.0 (white) to 1.0 (bright red), as measured by the disequilibrium coefficient  $D'$ . Panel B depicts the genomic structure of ITGAM, the conserved major protein domains, and the relationship between rs11574637 and two nonsynonymous alleles of ITGAM.

of BLK is highly restricted to the B-cell lineage.<sup>33</sup> Blk expression in mice is first observed in cycling late pro-B cells, continues throughout B-cell development, and is subsequently down-regulated in plasma B cells.<sup>34</sup> Mice that are deficient in Blk have no gross phenotype,<sup>35</sup> and we are not aware of studies of the function of BLK in human B cells. It has been hypothesized that BLK is one of the tyrosine kinases that transduces signals downstream of the B-cell receptor and that it has a redundant role in mice (given the lack of a phenotype in Blk-deficient mice). There is a precedent for major species differences in the role of kinases associated with the B-cell receptor. For example, in humans, a deficiency in Bruton's tyrosine kinase (BTK) leads to X-linked agammaglobulinemia and a complete lack of B cells.<sup>36</sup> However, deficiency of Btk in mice is associated with a much milder phenotype.<sup>37</sup>

B-cell-receptor signaling is important for establishing the B-cell repertoire through induction of anergy, deletion, and receptor editing during B-cell development.<sup>38,39</sup> As our study shows, the risk allele at BLK is associated with reduced expression of BLK mRNA in transformed B-cell lines. We speculate that altered protein levels of BLK might influence tolerance mechanisms in B cells, predisposing persons to systemic autoimmunity. A similar mechanism has recently been shown for Ly108, one of the major genetic loci in the NZM2410 mouse model of lupus.<sup>40</sup> Our data also suggest an effect of the risk haplotype on expression of the ubiquitously expressed gene C8orf13, but the function of this gene is currently unknown.

We identified a second locus, ITGAM–ITGAX.



Although we cannot exclude *ITGAX* because of the strong linkage disequilibrium in the region that extends into its 5' region, we think that variants of *ITGAM* are driving the association. *ITGAM* (also known as CD11b, Mac-1, and complement receptor type 3) is a well-characterized molecule in the integrin alpha chain family that is expressed by a variety of myeloid cell types, including dendritic cells, macrophages, monocytes, and neutrophils.<sup>41-43</sup> *ITGAM* forms a heterodimer with integrin beta-2 (*ITGB2*, or CD18) and mediates adhesion between cell types in the immune system and the adhesion of myeloid cells to endothelium.<sup>44</sup> Mice that are deficient in *ITGAM* have enhanced disease progression and inflammation in several models of autoimmunity,<sup>45-47</sup> including lupus. *ITGAM* may suppress differentiation of helper T-cell type 17 (Th17),<sup>48</sup> a pathway that has been linked with induction of autoimmunity. Moreover, *ITGAM* expression has been reported to be elevated on neutrophils from patients with active SLE.<sup>49</sup>

In summary, we have identified two new susceptibility loci for SLE: *BLK-C8orf13* on chromosome 8 and *ITGAM-ITGAX* on chromosome 16. The most likely candidate genes within these two loci are *BLK* and *ITGAM*. Further dissection of these genetic loci is warranted, as are investigations to determine whether the variants contribute to the risk of other autoimmune disorders.

Supported by grants (NO1-AR1-2256 and NO1-AI95386, to Dr. Gregersen) from the National Institutes of Health (NIH) for the collections of the Autoimmune Biomarkers Collaborative Network and the Multiple Autoimmune Disease Genetics Consortium; grants (R01-AR44804 and K24-AR02175) from the NIH, a Kirkland Scholar Award, and a grant from the Rosalind Russell Medical Research Center for Arthritis (all to Dr. Criswell); a grant (5-M01-RR00079) to the Moffitt Hospital, University of California, San Francisco, from the National Center for Research Resources; a grant (AR050267, to Dr. Seldin) from the NIH; a grant (AR43737) from the NIH to the Hopkins Lupus Cohort; a grant (M01-RR00052) from the NIH to the Johns Hopkins University Outpatient General Clinical Research Center; grants (R01-AR046588, to Dr. Manzi; and NHLBI-HL54900 and HL74165, to Dr. Kamboh) from the NIH; grants from the Swedish Research Council for Medicine (to Dr. Syvänen); a grant from the Knut and Alice Wallenberg Foundation to the SNP Technology Platform in Uppsala; grants from the Swedish Research Council for Medicine, the Swedish Rheumatism Association, the King Gustaf V 80-Year Foundation, and the Ulla and Roland Gustafsson Foundation (all to Dr. Rönnblom); a grant from the King Gustaf V 80-Year Foundation (to Dr. Gunnarsson); and a grant (K23-AR051044, to Dr. Kao) from the NIH.

Drs. Hom, Graham, Modrek, and Behrens and Mr. Ortman and Mr. Ferreira report being employees of Genentech and having an equity interest in the company; Drs. Pant and Ballinger, being employees of Perlegen Sciences and having an equity interest in the company; Dr. Ballinger, also having an equity interest in Genentech and Amgen; Dr. Manzi, receiving grant support from Genentech; Dr. Criswell, receiving consulting fees from Celera; and Dr. Gregersen, serving on the Abbott Scholar Award Advisory Committee and receiving honoraria from Biogen Idec, Genentech, and Roche Pharmaceuticals. No other potential conflict of interest relevant to this article was reported.

We thank Anthony Liew, Houman Khalili, and Alamelu Chandrasekaran for their assistance with genotyping; Alkes Price and Nick Patterson for their assistance with EIGENSTRAT software; and members of the Molecular Medicine group in Uppsala for their assistance with genotyping.

#### APPENDIX

The following is a list of the authors' affiliations: Genentech, South San Francisco, CA (G.H., R.R.G., B.M., W.O., R.C.F., T.W.B.); University of California at San Francisco, San Francisco (K.E.T., S.A.C., L.A.C.); Uppsala University, Uppsala, Sweden (S.G., A.-C.S., L.R.); Feinstein Institute for Medical Research—Long Island Jewish Health System, Manhasset, NY (A.T.L., P.K.G.); Perlegen Sciences, Mountain View, CA (P.V.K.P., D.G.B.); University of California at Davis, Davis (R.K., C.T., M.F.S.); University of Pittsburgh Medical Center, Pittsburgh (F.Y.D., M.I.K., A.H.K., S.M.); Karolinska Institutet—Karolinska Hospital, Stockholm (I.G.); Lund University Hospital, Lund, Sweden. (A.A.B.); Umeå University Hospital, Umeå, Sweden (S.R.-D.); and Johns Hopkins University School of Medicine, Baltimore (M.P.).

#### REFERENCES

- Hochberg MC. The epidemiology of systemic lupus erythematosus. In: Wallace DJ, Hahn BH, eds. *Dubois' lupus erythematosus*. 5th ed. Baltimore: Williams & Wilkins, 1997:49-65.
- Wakeland EK, Liu K, Graham RR, Behrens TW. Delineating the genetic basis of systemic lupus erythematosus. *Immunity* 2001;15:397-408.
- Nath SK, Kilpatrick J, Harley JB. Genetics of human systemic lupus erythematosus: the emerging picture. *Curr Opin Immunol* 2004;16:794-800.
- Goldberg MA, Arnett FC, Bias WB, Shulman LE. Histocompatibility antigens in systemic lupus erythematosus. *Arthritis Rheum* 1976;19:129-32.
- Graham RR, Ortman WA, Langefeld CD, et al. Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. *Am J Hum Genet* 2002;71:543-53.
- Graham RR, Ortman W, Rodine P, et al. Specific combinations of HLA-DR2 and DR3 class II haplotypes contribute graded risk for disease susceptibility and autoantibodies in human SLE. *Eur J Hum Genet* 2007;15:823-30.
- Sigurdsson S, Nordmark G, Göring HH, et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am J Hum Genet* 2005;76:528-37.
- Graham RR, Kozyrev SV, Baechler EC, et al. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat Genet* 2006;38:550-5.
- Graham RR, Kyogoku C, Sigurdsson S, et al. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci U S A* 2007;104:6758-63.

10. Remmers EF, Plenge RM, Lee AT, et al. *STAT4* and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med* 2007;357:977-86.
11. Rönnblom L, Alm GV. A pivotal role for the natural interferon alpha-producing cells (plasmacytoid dendritic cells) in the pathogenesis of lupus. *J Exp Med* 2001;194:F59-F63.
12. Baechler EC, Gregersen PK, Behrens TW. The emerging role of interferon in human systemic lupus erythematosus. *Curr Opin Immunol* 2004;16:801-7.
13. Banchereau J, Pascual V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity* 2006;25:383-92.
14. Miyagi T, Gil MP, Wang X, Louten J, Chu WM, Biron CA. High basal *STAT4* balanced by *STAT1* induction to control type 1 interferon effects in natural killer cells. *J Exp Med* 2007;204:2383-96.
15. Bauer JW, Baechler EC, Petri M, et al. Elevated serum levels of interferon-regulated chemokines as biomarkers for active human systemic lupus erythematosus. *PLoS Med* 2006;3(12):e491.
16. Criswell LA, Pfeiffer KA, Lum RF, et al. Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. *Am J Hum Genet* 2005;76:561-71.
17. Seligman VA, Suarez C, Lum R, et al. The Fcγ receptor IIIA-158F allele is a major risk factor for the development of lupus nephritis among Caucasians but not non-Caucasians. *Arthritis Rheum* 2001;44:618-25.
18. Demirci FY, Manzi S, Ramsey-Goldman R, et al. Association of a common interferon regulatory factor 5 (IRF5) variant with increased risk of systemic lupus erythematosus (SLE). *Ann Hum Genet* 2007;71:308-11.
19. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997;40:1725.
20. Mitchell MK, Gregersen PK, Johnson S, Parsons R, Vlahov D. The New York Cancer Project: rationale, organization, design, and baseline characteristics. *J Urban Health* 2004;81:301-10.
21. Gunderson KL, Steemers FJ, Ren H, et al. Whole-genome genotyping. *Methods Enzymol* 2006;410:359-76.
22. Hsu TM, Chen X, Duan S, Miller RD, Kwok PY. Universal SNP genotyping assay with fluorescence polarization detection. *Biotechniques* 2001;31:560, 562, 564-8.
23. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904-9.
24. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-75.
25. Stranger BE, Nica AC, Forrest MS, et al. Population genomics of human gene expression. *Nat Genet* 2007;39:1217-24.
26. Dixon AL, Liang L, Moffatt ME, et al. A genome-wide association study of global gene expression. *Nat Genet* 2007;39:1202-7.
27. Devlin B, Roeder K, Wasserman L. Genomic control for association studies: a semiparametric test to detect excess-haplotype sharing. *Biostatistics* 2000;1:369-87.
28. de Bakker PI, McVean G, Sabeti PC, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet* 2006;38:1166-72.
29. Giglio S, Broman KW, Matsumoto N, et al. Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *Am J Hum Genet* 2001;68:874-83.
30. Sugawara H, Harada N, Ida T, et al. Complex low-copy repeats associated with a common polymorphic inversion at human chromosome 8p23. *Genomics* 2003;82:238-44.
31. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;6:95-108.
32. Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. *Nature* 2005;437:1365-9.
33. Dymecki SM, Zwollo P, Zeller K, Kuhajda FP, Desiderio SV. Structure and developmental regulation of the B-lymphoid tyrosine kinase gene *blk*. *J Biol Chem* 1992;267:4815-23.
34. Wasserman R, Li YS, Hardy RR. Differential expression of the *blk* and *ret* tyrosine kinases during B lineage development is dependent on Ig rearrangement. *J Immunol* 1995;155:644-51.
35. Texido G, Su IH, Mecklenbräuker I, et al. The B-cell-specific Src-family kinase *Blk* is dispensable for B-cell development and activation. *Mol Cell Biol* 2000;20:1227-33.
36. Tsukada S, Saffran DC, Rawlings DJ, et al. Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. *Cell* 1993;72:279-90.
37. Khan WN, Alt FW, Gerstein RM, et al. Defective B cell development and function in *Btk*-deficient mice. *Immunity* 1995;3:283-99.
38. Cornall RJ, Goodnow CC. B cell antigen receptor signalling in the balance of tolerance and immunity. *Novartis Found Symp* 1998;215:21-30.
39. Nemazee D, Weigert M. Revising B cell receptors. *J Exp Med* 2000;191:1813-7.
40. Kumar KR, Li L, Yan M, et al. Regulation of B cell tolerance by the lupus susceptibility gene *Ly108*. *Science* 2006;312:1665-9.
41. Hynes RO. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* 1992;69:11-25.
42. Lu H, Smith CW, Perrard J, et al. *LFA-1* is sufficient in mediating neutrophil emigration in *Mac-1*-deficient mice. *J Clin Invest* 1997;99:1340-50.
43. Abbas AR, Baldwin D, Ma Y, et al. Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* 2005;6:319-31.
44. Dunne JL, Collins RG, Beaudet AL, Ballantyne CM, Ley K. *Mac-1*, but not *LFA-1*, uses intercellular adhesion molecule-1 to mediate slow leukocyte rolling in TNF-α-induced inflammation. *J Immunol* 2003;171:6105-11.
45. Hammerberg C, Katiyar SK, Carroll MC, Cooper KD. Activated complement component 3 (C3) is required for ultraviolet induction of immunosuppression and antigenic tolerance. *J Exp Med* 1998;187:1133-8.
46. Sohn JH, Bora PS, Suk HJ, Molina H, Kaplan HJ, Bora NS. Tolerance is dependent on complement C3 fragment iC3b binding to antigen-presenting cells. *Nat Med* 2003;9:206-12.
47. Watts GM, Beurskens FJ, Martin-Padura I, et al. Manifestations of inflammatory arthritis are critically dependent on *LFA-1*. *J Immunol* 2005;174:3668-75.
48. Ehrlichou D, Xiong Y, Xu G, Chen W, Shi Y, Zhang L. *CD11b* facilitates the development of peripheral tolerance by suppressing Th17 differentiation. *J Exp Med* 2007;204:1519-24.
49. Buyon JP, Shadick N, Berkman R, et al. Surface expression of *Gp 165/95*, the complement receptor *CR3*, as a marker of disease activity in systemic lupus erythematosus. *Clin Immunol Immunopathol* 1988;46:141-9.

Copyright © 2008 Massachusetts Medical Society.