nature genetics

Genetic variants near *TNFAIP3* on 6q23 are associated with systemic lupus erythematosus

Robert R Graham[1,11,12], Chris Cotsapas[1,2,12], Leela Davies[1], Rachel Hackett[1], Christopher J Lessard[3,4], Joanlise M Leon[5], Noel P Burtt[1], Candace Guiducci[1], Melissa Parkin[1], Casey Gates[1], Robert M Plenge[1], Timothy W Behrens[6], Joan E Wither[7], John D Rioux[8], Paul R Fortin[9], Deborah Cunninghame Graham[10], Andrew K Wong[10], Timothy J Vyse[10], Mark J Daly[1,2], David Altshuler[1], Kathy L Moser[4] & Patrick M Gaffney[4]

Systemic lupus erythematosus (SLE) is an autoimmune disease influenced by genetic and environmental factors. We carried out a genome-wide association scan and replication study and found an association between SLE and a variant in *TNFAIP3* (rs5029939, meta-analysis $P = 2.89 \times 10^{-12}$, OR = 2.29). We also found evidence of two independent signals near *TNFAIP3* associated with SLE, including one previously associated with rheumatoid arthritis (RA). These results establish that variants near *TNFAIP3* contribute to differential risk of SLE and RA.

Systemic lupus erythematosus (SLE, MIM152700) is an autoimmune disease characterized by dysregulated interferon responses and loss of self-tolerance to cellular antigens. Autoantibody production leads to immune complex formation, resulting in local and systemic inflammation and organ failure. The prevalence of SLE is estimated to be between 40 and 400 cases per 100,000 individuals, with higher prevalence rates and more severe complications occurring in persons of Hispanic or African ancestry[1]. SLE afflicts women at a rate nine times that of men and most often presents during the years between menarche and menopause[2]. The etiology of SLE is complex and poorly defined, requiring interplay between genetic predisposition and environmental triggers.

To identify genetic variants contributing to SLE, we tested association of 311,238 successfully genotyped SNPs in 431 unrelated SLE cases and 2,155 controls (see **Supplementary Methods** online). We compared the distribution of the observed *P* values to the expected distribution under the null hypothesis, and deviation of the tail of the observed *P* value distribution was evident, suggesting the presence of significant genetic effects (**Supplementary Fig. 1** online). Three regions met a pre-specified threshold for genome-wide significance ($P < 5 \times 10^{-8}$): the previously defined HLA and *IRF5-TNPO3* regions, and a previously unreported SLE locus at chromosome 6q23 within the *TNFAIP3* gene (**Fig. 1a**).

TNFAIP3, also known as A20, functions as key regulator of NF-κB signaling through ubiquitin modification of adaptor proteins RIP and TRAF6 downstream of TNFα and Toll-like receptors, respectively[3,4]. We observed evidence for association with a variant in *TNFAIP3* (rs5029939, GWAS $P = 2.55 \times 10^{-8}$; **Fig. 1b**) and two flanking SNPs (rs10499197, GWAS $P = 2.11 \times 10^{-6}$; rs7749323, GWAS $P = 9.63 \times 10^{-7}$) in strong LD with rs5029939 ($r^2 > 0.95$) (**Fig. 1b**). A SNP located ~185 kb upstream of *TNFAIP3* recently reported to be associated with risk of RA[5,6] (rs6920220) showed modest association in the SLE GWAS dataset (GWAS $P = 0.01$; **Fig. 1b**).

We tested 134 loci (excluding HLA and *IRF5*) for association with SLE in 740 independent trios (**Supplementary Table 1** online) and calculated meta-analysis *P* values to determine the overall evidence for association from the GWAS and trio replication datasets. Association with rs5029939 in *TNFAIP3* replicated in the trios (trio $P = 2.47 \times 10^{-5}$), which together with the GWAS data produced a convincing meta-analysis $P = 2.89 \times 10^{-12}$ (**Table 1**). Variants in *STAT4* and *BLK* regions also showed association with SLE in both the GWAS and trio datasets and together produced meta-analysis *P* values exceeding genome-wide significance (rs3821236 in *STAT4*, meta $P = 8.49 \times 10^{-11}$; rs2618476 in *BLK*, meta $P = 1.7 \times 10^{-8}$; **Table 1**). These two loci as well as *ITGAM* have recently been identified in two other SLE GWA scans[7–9]. In our study, a variant in *ITGAM* also showed convincing evidence of replication (rs11150610, meta $P = 1.72 \times 10^{-6}$; **Supplementary Table 1**).

Variants in the *TNFAIP3* region have recently been reported to independently influence risk of RA[5,6]. The minor allele of rs6920220 is associated with increased risk of RA (MAF cases > controls), whereas the minor allele of rs10499194 confers protection (MAF controls > cases) (**Supplementary Fig. 2** online). To evaluate the locus in more detail in SLE, we genotyped a panel of eight SNPs that tag the RA and SLE alleles in our complete set of trio families ($n = 991$). We also included a putative causal missense variant (rs2230926) in exon 3 of *TNFAIP3* (**Supplementary Table 2** online). We found that

[1]Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [2]Center for Human Genetic Research, Mass General Hospital, 185 Cambridge Street, and the Department of Medicine, Harvard Medical School, Boston, Massachusetts 02114, USA. [3]Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73104, USA. [4]Arthritis and Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104, USA. [5]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota 55455, USA. [6]Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. [7]Arthritis Centre of Excellence; Division of Genetics and Development, Toronto Western Hospital Research Institute, University Health Network; Departments of Medicine and Immunology, University of Toronto, Toronto, Ontario, Canada. [8]Université de Montréal and the Montreal Heart Institute Research Center, Montreal, Quebec, Canada. [9]University of Toronto Lupus Clinic, Centre for Prognosis Studies in the Rheumatic Diseases, Toronto Western Hospital, University Health Network; Department of Medicine, University of Toronto, Toronto, Ontario, Canada. [10]Molecular Genetics and Rheumatology Section, Imperial College Faculty of Medicine, Hammersmith Hospital, London, UK. [11]Present address: Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA. [12]These authors contributed equally to this work. Correspondence should be addressed to P.M.G. (gaffneyp@lupus.omrf.org).

Received 3 March; accepted 16 June; published online 1 August 2008; doi:10.1038/ng.200

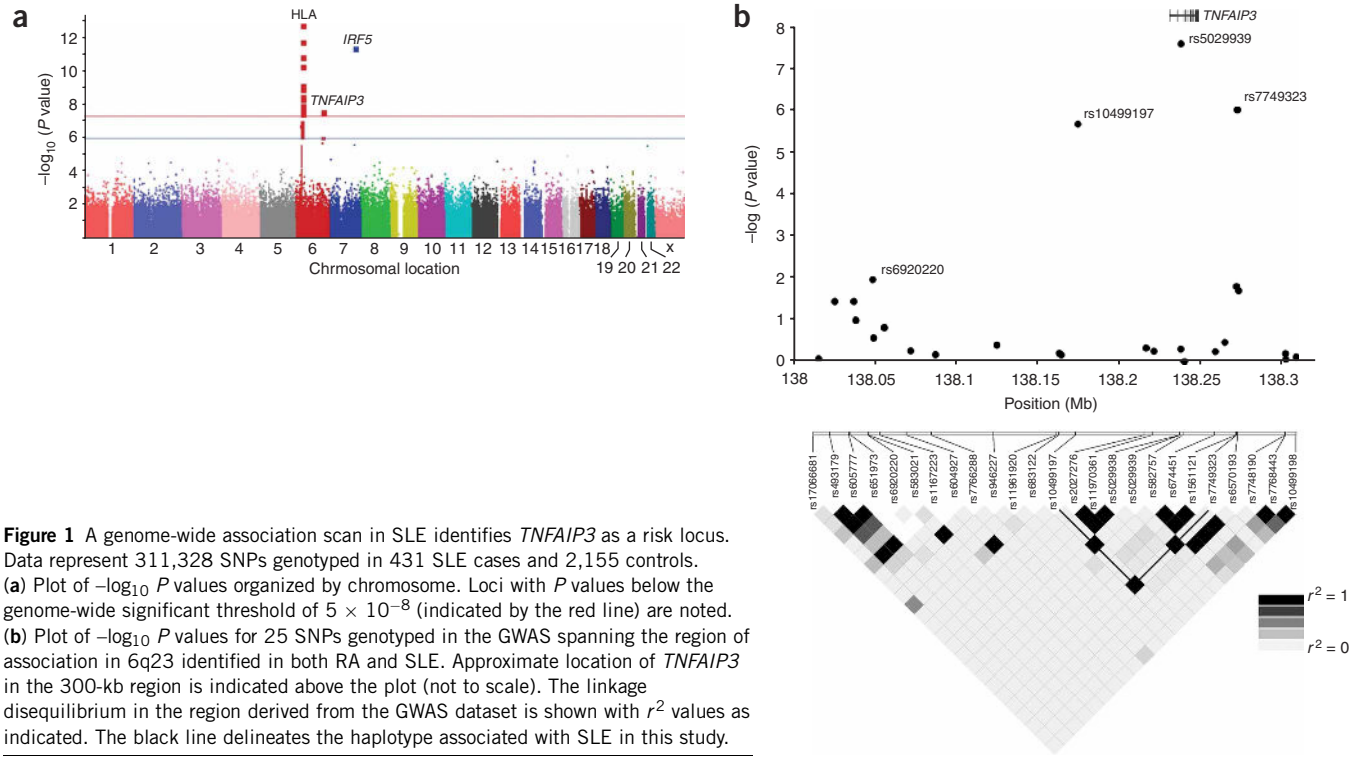

Figure 1 A genome-wide association scan in SLE identifies *TNFAIP3* as a risk locus. Data represent 311,328 SNPs genotyped in 431 SLE cases and 2,155 controls. (**a**) Plot of $-\log_{10}$ *P* values organized by chromosome. Loci with *P* values below the genome-wide significant threshold of 5×10^{-8} (indicated by the red line) are noted. (**b**) Plot of $-\log_{10}$ *P* values for 25 SNPs genotyped in the GWAS spanning the region of association in 6q23 identified in both RA and SLE. Approximate location of *TNFAIP3* in the 300-kb region is indicated above the plot (not to scale). The linkage disequilibrium in the region derived from the GWAS dataset is shown with r^2 values as indicated. The black line delineates the haplotype associated with SLE in this study.

the minor allele of rs6920220 that contributes to RA risk was also associated with susceptibility to SLE ($P = 8.92 \times 10^{-5}$); however, no evidence for association was seen for the rs10499194 RA protective variant ($P = 0.72$) (**Supplementary Table 2** and **Supplementary Fig. 2**).

We then examined the haplotypic association for the five most common haplotypes (frequency >1%) formed by the associated markers rs6920220, rs10499197, rs5029939, rs2230926 and rs7749323 in the SLE pedigrees (~224-kb span; **Supplementary Fig. 3** online). Haplotype 2, present on 19.3% of the chromosomes tested, is associated with SLE and carries only the minor allele of rs6920220 (haplotypic $P = 0.0064$). Haplotype 4, present on 1.1% of chromosomes tested, is also associated with SLE and carries the minor alleles of rs10499197, rs5029939, rs2230926 and rs7749323 but not rs6920220 (haplotypic $P = 3.67 \times 10^{-5}$). These results suggested the possibility of two independent genetic effects, one marked by the minor allele of rs6920220 (haplotype 2) and one marked by the minor alleles of rs10499197, rs5029939, rs2230926 and rs7749323 (haplotype 4). Haplotype 3, also associated with SLE, carries the minor alleles of all five markers (haplotypic $P = 1 \times 10^{-4}$).

We performed conditional association analyses to test for independent effects at this locus. The epidemiological and genetic similarities between SLE and RA suggest that not only the same genes, but also identical variants, might affect risk of each disease. We therefore sought to test the contribution in SLE of variants previously shown to modulate risk of RA. Using logistic regression, we conditioned on the signal we identified in *TNFAIP3* to determine whether association persisted at the RA-associated markers (**Supplementary Table 3** online). Conditioning on either rs7749323 or rs10499197, we observed residual association at the RA risk marker rs6920220 ($P = 0.0024$ conditional on rs10499197; $P = 0.0017$ conditional on rs7749323). The converse analysis also supports independent signals, as adjustment for the effect at rs6920220 did not eliminate signal at any of the four markers in the region surrounding *TNFAIP3* (rs10499197

Table 1 Replicated loci reaching genome-wide significance in this study

			GWAS				TRIO		
Locus	Chr.	Position (Mb)	*P*	OR	Case frequency ($n = 431$)	Control frequency ($n = 2,155$)	*P*	T:U	Meta-analysis *P*
TNFAIP3									
rs5029939	6q23.3	138.237416	2.55×10^{-8}	2.28	0.071	0.031	2.47×10^{-5}	131:57	2.89×10^{-12}
STAT4									
rs3821236	2q32.3	191.728264	2.87×10^{-5}	1.49	0.264	0.194	1.35×10^{-6}	449:320	8.49×10^{-11}
BLK									
rs2618476	8p23.1	11.389950	9.58×10^{-4}	1.29	0.314	0.263	7.87×10^{-6}	490:353	1.70×10^{-8}

Position is location in Build 35 of the human genome. T:U, number of informative transmitted and nontransmitted alleles from the transmission disequilibrium test in 740 SLE trio pedigrees. Meta-analysis *P*, *P* value for the combined case-control and family-based association study.

$P = 4.51 \times 10^{-5}$, rs5029939 $P = 9.00 \times 10^{-5}$, rs2230926 $P = 0.0012$, rs7749323 $P = 1.75 \times 10^{-5}$). These results suggest that two independent variants in the *TNFAIP3* region contribute to SLE risk, although we are unable to exclude the possibility that a single variant with multiple risk alleles (for example, the *INS* VNTR in type 1 diabetes) or one causal allele partially captured by rs6920220 and rs10499197-rs7749323 underlies the signal.

The haplotype and conditional analyses also suggest that the missense SNP in exon 3 of *TNFAIP3* (rs2230926) may not be the only putative causal variant. Haplotype 5 carries the minor allele of rs2230926 but is not associated with SLE (**Supplementary Fig. 3**). In addition, the association at rs2230926 can be completely explained by the association at rs10499197 and rs7749323 (**Supplementary Table 3**). Taken together, these results suggest that an untyped variant(s) carried on the haplotype tagged by the minor alleles of rs10499197 and rs7749323 is responsible for this SLE association effect. However, additional genotyping and resequencing will be necessary to identify the causal allele(s) and definitively determine the contribution of the *TNFAIP3* locus to SLE.

This genome-wide association study identifies *TNFAIP3* as a new susceptibility locus in SLE. Through its action as a negative regulator of the NF-κB pathway, TNFAIP3 plays a key role in modulating a broad range of cellular functions, including cell activation, cytokine signaling and apoptosis[10]. Mechanistically, TNFAIP3 catalyzes the ubiquitin modification of adaptor proteins downstream of TNFR, TLR and IL1R[3,4,11]. These pro-inflammatory pathways contribute to the pathogenesis of RA and SLE.

Data for RA[5,6] and our results for SLE suggest a genetic model of three alleles with independent effects: (i) a risk allele ~185 kb upstream of *TNFAIP3* at rs6920220 for RA and SLE, (ii) a nearby protective allele at rs10499194 for RA but not apparent in SLE, and (iii) a risk haplotype for SLE marked by a haplotype comprised of the minor alleles of rs10499197 and rs7749323 that directly spans *TNFAIP3*.

In addition to our study, three other genome-wide scans in SLE have been performed[7,8,12]. New and convincing genetic associations present in two or more of these scans include variants in HLA, *IRF5*, *STAT4*, *ITGAM* and *BLK*. Other genes, including *BANK1*, *PXK* and *TNFAIP3*, achieved genome-wide significance in only one study. The sensitivity to detect association in a GWAS can be influenced by a variety of factors, including sample size, cohort demographics, study design, analytical strategies, environmental factors, phenotype definition, variation in the estimation of control allele frequencies and, in the case of *TNFAIP3*, SNP selection. Our ability to recognize the *TNFAIP3* association was facilitated, in part, because SNPs that capture the associated haplotype were well represented on the Affymetrix 5.0 array. In contrast, the other studies used genotyping arrays with SNPs that fail to capture the *TNFAIP3* risk haplotype effectively, thus making the signal virtually impossible to detect.

In summary, we have identified a previously unreported association between variants in the *TNFAIP3* region and SLE. Our data support the presence of genetic effects that are both shared and distinct from the recently described association with RA. TNFAIP3, through its role as a key regulator of NF-κB signaling, represents a compelling candidate locus for further studies of autoimmune pathogenesis.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

Support for this work was obtained from the US National Institutes of Health (grants AI063274 (P.M.G.), AR052125 (P.M.G.) and AR043247 (K.L.M.)), NIH NSRA award (5F32AR50927-RRG), Lupus Foundation of Minnesota (P.M.G., K.L.M.) and the Arthritis Foundation (K.L.M.).

This study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available from http://www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113. Permission for use of these data was obtained from the oversight committees of the respective cohorts. Control subjects from the National Institute of Mental Health Schizophrenia Genetics Initiative (NIMH-GI), data and biomaterials are being collected by the "Molecular Genetics of Schizophrenia II" (MGS-2) collaboration. The investigators and co-investigators are as follows: ENH/Northwestern University, MH059571, P.V. Gejman (Collaboration Coordinator; PI), A.R. Sanders; Emory University School of Medicine, MH59587, F. Amin (PI); Louisiana State University Health Sciences Center, New Orleans, Louisiana, MH067257, N. Buccola (PI); University of California-Irvine, MH60870, W. Byerley (PI); Washington University, St. Louis, U01, MH060879, C.R. Cloninger (PI); University of Iowa, Iowa, MH59566, R. Crowe (PI), D. Black; University of Colorado, MH059565, R. Freedman (PI); University of Pennsylvania, MH061675, D. Levinson (PI); University of Queensland, MH059588, B. Mowry (PI); Mt. Sinai School of Medicine, MH59586, J. Silverman (PI). The samples were collected by V.L. Nimgaonkar's group at the University of Pittsburgh, as part of a multi-institutional collaborative research project with J. Smoller and P. Sklar (Massachusetts General Hospital) (grant MH 63420).

We thank T. Hudson for helpful discussions, T. McKenzie for coordinating the GenES study, J.O. Claudio for ensuring ethics approval of DNA sample sharing, J. Su for assistance with management of the GenES database and all the CaNIOS investigators that have contributed to GenES. GenES is funded by the Canadian Institute of Health Research (CIHR# 62840).

We thank S. Gabriel and the entire Affymetrix genotyping team at the Broad Institute of Harvard and MIT. We thank F. Kuruvilla, J. Korn and S. McCarroll for assistance with the Birdseed genotype-calling algorithm.

We would also like to acknowledge the research assistants and coordinators that recruited the individuals in the study. Most importantly, we thank the individuals who have participated in and contributed to these studies.

AUTHOR CONTRIBUTIONS

R.R.G., C.C., C.J.L., J.M.L., D.A., K.L.M. and P.M.G. wrote the manuscript with input from coauthors. R.R.G., C.C. and R.M.P. performed the analyses under the supervision of D.A. and M.J.D. L.D., R.H., N.P.B., C. Guiducci, M.P. and C. Gates generated the genotype data. J.E.W., J.D.R., P.R.F., D.C.G., A.K.W., T.J.V. and T.W.B. assisted in the collection and characterization of SLE samples. R.R.G., C.C., D.A., M.J.D., K.L.M. and P.M.G. were responsible for the study design. D.A., K.L.M. and P.M.G. directed the project.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Helmick, C.G. *et al. Arthritis Rheum.* **58**, 15–25 (2008).
2. Kaslow, R.A. & Masi, A.T. *Arthritis Rheum.* **21**, 473–479 (1978).
3. Boone, D.L. *et al. Nat. Immunol.* **5**, 1052–1060 (2004).
4. Wertz, I.E. *et al. Nature* **430**, 694–699 (2004).
5. Plenge, R.M. *et al. Nat. Genet.* **39**, 1477–1482 (2007).
6. Thomson, W. *et al. Nat. Genet.* **39**, 1431–1433 (2007).
7. Harley, J.B. *et al. Nat. Genet.* **40**, 204–210 (2008).
8. Hom, G. *et al. N. Engl. J. Med.* **358**, 900–909 (2008).
9. Remmers, E.F. *et al. N. Engl. J. Med.* **357**, 977–986 (2007).
10. Beyaert, R., Heyninck, K. & Van Huffel, S. *Biochem. Pharmacol.* **60**, 1143–1151 (2000).
11. Heyninck, K. & Beyaert, R. *FEBS Lett.* **442**, 147–150 (1999).
12. Kozyrev, S.V. *et al. Nat. Genet.* **40**, 211–217 (2008).