Published in final edited form as:

Ann Rheum Dis. 2016 November ; 75(11): 2007–2013. doi:10.1136/annrheumdis-2015-208441.

Decreased *SMG7* expression associates with lupus-risk variants and elevated antinuclear antibody production

Yun Deng^{1,33}, Jian Zhao^{1,33}, Daisuke Sakurai¹, Andrea L. Sestak², Vadim Osadchiy¹, Carl D. Langefeld³, Kenneth M. Kaufman^{4,5}, Jennifer A. Kelly⁶, Judith A. James^{6,7,8}, Michelle A. Petri⁹, Sang-Cheol Bae¹⁰, Marta E. Alarcón-Riquelme^{11,11,6,11} for the BIOLUPUS on behalf of GENLES networks, Graciela S Alarcón¹², Juan-Manuel Anaya¹³, Lindsey A. Criswell¹⁴, Barry I Freedman¹⁵, Diane L. Kamen¹⁶, Gary S. Gilkeson¹⁶, Chaim O. Jacob¹⁷, Joan T Merrill¹⁸, Patrick M. Gaffney⁶, Kathy Moser Sivils^{6,7}, Timothy B Niewold¹⁹, Rosalind Ramsey-Goldman²⁰, John D Reveille²¹, R Hal Scofield^{6,8,22}, Anne M Stevens^{23,24}, Susan A Boackle^{25,26}, Luis M Vilá²⁷, Woong Sohn II¹⁰, Seung Lee¹⁰, Deh-Ming Chang²⁸, Yeong Wook Song²⁹, Timothy J. Vyse³⁰, John B. Harley^{4,5}, Elizabeth E. Brown^{12,31}, Jeffrey C. Edberg¹², Robert P. Kimberly¹², Rita M. Cantor³², Bevra H. Hahn¹, Jennifer M. Grossman¹, and Betty P. Tsao^{1,*}

¹Division of Rheumatology, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

²Department of Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

³Department of Biostatistical Sciences and Center for Public Health Genomics, Wake Forest School of Medicine, Winston-Salem, NC, USA

⁴Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

⁵US Department of Veterans Affairs Medical Center, Cincinnati, OH, USA

⁶Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA

⁷Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

⁸Department of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

⁹Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA

*Corresponding Author: Betty P. Tsao, PhD., Division of Rheumatology, Department of Medicine, David Geffen School of Medicine, University of California Los Angeles, 900 Veteran Ave, Warren Hall Room 14-224, Los Angeles, CA 90095-1670, bttsao@mednet.ucla.edu, Phone: 310-825-8906, Fax: 310-825-6903.

³³These authors contributed equally to this work.

[¶]Memberships of the consortia are available in the Acknowledgments.

Current address: Betty P. Tsao, PhD, Clinical Science Bldg, MUSC, 96 Jonathan Lucas St – Suite 936C, Charleston, SC 29425

Competing Interests

None declared.

- ¹⁰Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea
- ¹¹Pfizer-Universidad de Granada-Junta de Andalucía Center for Genomics and Oncological Research, Granada, Spain
- ¹²Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA
- ¹³Center for Autoimmune Diseases Research (CREA), Universidad del Rosario, Bogotá, Colombia
- ¹⁴Rosalind Russell/Ephraim P. Engleman Rheumatology Research Center, University of California San Francisco, San Francisco, CA, USA
- ¹⁵Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA
- ¹⁶Division of Rheumatology, Medical University of South Carolina, Charleston, SC, USA
- ¹⁷Department of Medicine, University of Southern California, Los Angeles, CA, USA
- ¹⁸Clinical Pharmacology, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA
- ¹⁹Division of Rheumatology and Department of Immunology, Mayo Clinic, Rochester, MN, USA
- ²⁰Division of Rheumatology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
- ²¹Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, TX, USA
- ²²US Department of Veterans Affairs Medical Center, Oklahoma City, OK, USA
- ²³Division of Rheumatology, Department of Pediatrics, University of Washington, Seattle, WA, USA
- ²⁴Center for Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, WA, USA
- ²⁵Division of Rheumatology, University of Colorado School of Medicine, Aurora, CO, USA
- ²⁶US Department of Veterans Affairs Medical Center, Denver, CO, USA
- ²⁷Division of Rheumatology, Department of Medicine, University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico
- ²⁸Taipei Veterans General Hospital, Taipei City, Taiwan
- ²⁹Department of Molecular Medicine and Biopharmaceutical Sciences, Graduate School of Convergence Science and Technology, and College of Medicine, Medical Research Center, Seoul National University, Seoul, Korea
- ³⁰Division of Genetics and Molecular Medicine and Immunology, King's College London, London, UK
- ³¹Department of Pathology, University of Alabama at Birmingham, Birmingham, AL, USA
- ³²Department of Human Genetics, University of California Los Angeles, Los Angeles, CA, USA

Abstract

Objectives—Following up the SLE GWAS identification of *NMNAT2* at rs2022013, we fine-mapped its 150kb flanking regions containing *NMNAT2* and *SMG7* in a 15,292 case-control multi-ancestry population and tested functions of identified variants.

Methods—We performed genotyping using custom array, imputation by IMPUTE 2.1.2, and allele specific functions using qRT-PCR and luciferase reporter transfections. SLE PBMCs were cultured with siRNAs to measure antinuclear antibody (ANA) and cyto/chemokine levels in supernatants using ELISA.

Results—We confirmed association at *NMNAT2* in European American (EA) and Amerindian/Hispanic ancestries, and identified independent signal at *SMG7* tagged by rs2702178 in EA only ($P=2.4\times 10^{-8}$, OR=1.23 [95%CI=1.14–1.32]). In complete linkage disequilibrium with rs2702178, rs2275675 in the promoter region robustly associated with *SMG7* mRNA levels in multiple expression quantitative trait locus datasets. Its risk allele was dose-dependently associated with decreased *SMG7* mRNA levels in PBMCs of 86 SLE patients and 119 controls ($P=1.1\times 10^{-3}$ and 6.8×10^{-8} , respectively) and conferred reduced transcription activity in transfected HEK-293 and Raji cells ($P=0.0035$ and 0.0037 , respectively). As a critical component in the nonsense-mediated mRNA decay pathway, *SMG7* could regulate autoantigens including RNP and Sm. We showed *SMG7* mRNA levels in PBMCs correlated inversely with ANA titers of SLE patients ($r=-0.31$, $P=0.01$), and *SMG7* knockdown increased levels of ANA IgG and CCL19 in SLE PBMCs ($P=2.0\times 10^{-5}$ and 2.0×10^{-4} , respectively).

Conclusions—We confirmed *NMNAT2* and identified independent *SMG7* association with SLE. The inverse relationship between levels of the risk-allele associated *SMG7* mRNAs and ANA suggested the novel contribution of mRNA surveillance pathway to SLE pathogenesis.

Keywords

Systemic lupus erythematosus; Nonsense mediated mRNA decay; Antinuclear antibody; Genetic association study

INTRODUCTION

Recent genome-wide association studies (GWAS) have identified more than 50 susceptibility loci for systemic lupus erythematosus (SLE) [1]. At 1q25, rs2022013 located in the first intron of *NMNAT2* (encoding nicotinamide mononucleotide adenylyltransferase 2) was associated with SLE in GWAS ($P=1.1\times 10^{-7}$, OR=0.85) and in a large-scale replication study ($P=1.5\times 10^{-3}$, OR=0.92) using European-derived subjects [2, 3], suggesting *NMNAT2* as a SLE risk locus. *NMNAT2*, mainly expressed in the brain, is a central enzyme in the nicotinamide adenine dinucleotide biosynthetic pathway that is known to delay axon degeneration [4, 5]. Given no apparent clues for its involvement in immune dysregulation, we fine mapped *NMNAT2* and its neighboring genes in four major populations, confirmed *NMNAT2* association with SLE in European American (EA) and Amerindian/Hispanic (HS) ancestries and identified independent association at *SMG7* in EA only.

SMG7, located approximately 50 kb 5' of *NMNAT2*, encodes a component essential for nonsense-mediated mRNA decay (NMD) that controls mRNA quality and regulates gene expression [6]. We detected the SLE-risk alleles associated with decreased *SMG7* mRNA levels in peripheral blood mononuclear cells (PBMCs) of both SLE patients and healthy controls, and an inverse correlation between antinuclear autoantibody (ANA) titers and *SMG7* mRNA levels in SLE patients. *SMG7* reduction was associated with increased levels of ANA and chemokine (C-C motif) ligand 19 (CCL19) in SLE PBMCs, suggesting that decreased *SMG7* expression impacts the NMD pathway, contributing to autoantibody production in SLE.

METHODS

Subjects

DNA from individuals included in the multicenter Large Lupus Association Study 2 was processed at Oklahoma Medical Research Foundation (OMRF) that had site-specific and overall approval of Institutional Review Board. All patients met the classification criteria for SLE [7].

Genotyping and quality control

Genotyping was performed using Illumina custom array containing 35 tag SNPs of the *NMNAT2-SMG7* region (selected according to HapMap datasets (release#24, 2008) and 347 ancestry informative markers (AIMs). SNP genotypes that met the quality control criteria [8] were included for association tests. Subjects with missing genotype rates >10%, shared identity by descent >0.4, gender mismatch, or genetic outliers (estimated using principal components analysis and ADMIXMAP [9, 10]) were removed, resulting in 15,292 unrelated subjects with clean data.

Imputation

SNP and INDEL genotypes from the 1000 Genomes (version 3, Phase 1) were used for imputation (IMPUTE 2.1.2; [11]). Genotypes with information scores >0.9 and minor allele frequency (MAF)>0.01 were analyzed.

Quantitative real-time PCR (qRT-PCR)

Total RNAs purified from PBMCs were reverse-transcribed into cDNA to measure levels of *SMG7*, *SMG7-AS1* and housekeeping gene *RPLP0* by TaqMan assays (*SMG7*: Hs00539224_m1, *SMG7-AS1*: Hs01069005_m1, *RPLP0*: Hs99999902_m1; Life Technologies). Their relative expressions were calculated by 2^{-C_t} and Log₁₀ transformed.

Autoantibody profiles of 68 SLE patients (including ANA, anti-dsDNA or antibodies to extractable nuclear antigens) were measured in Clinical Laboratory at the time of blood drawn.

Plasmid construction and luciferase reporter assay

DNA sequences (0.3kb) surrounding rs2275675 or rs10911339 were PCR amplified using genomic DNA from homozygous subjects for the C-allele (rs2275675) and the T-allele

(rs10911339) using following primers: rs2275675, 5'-GGGGTACCGTAGAAAGAAAAGCAGAAC-3' (forward) and 5'-GAAGATCTGAGACCTGCACCAATAAG-3' (reverse); rs10911339, 5'-GGGGTACCGGTATGGGTGCCTAGC-3' (forward) and 5'-GAAGATCTCCAGGTGTGCAGACTTC-3' (reverse). The PCR products were digested using restriction enzymes, inserted into pGL3-promoter luciferase vector (Promega), and mutated to generate the respective other allele using Site-Directed Mutagenesis Kit (Stratagene). All construct inserts were sequenced for orientation and authenticity.

HEK-293 (human embryonic kidney cell line) and Raji B cells were purchased from ATCC. HEK-293 cells were seeded on 24-well plates at 2×10^5 cells/well and transiently transfected using Lipofectamine 2000 (Life Technologies) with 1 μ g rs2275675 vector (C or T), rs10911339 vector (C or T) or empty vector (pGL3-promoter) and 100 ng pRL-SV40 vector (Renilla luciferase) as an internal control. Raji cells were seeded on 24-well plates at 2×10^6 cells/well and electroporated with 5 μ g luciferase constructs and 100 ng Renilla control vector using nucleofector (Amaxa). Luciferase activity in cell lysates was measured after 24 hours using a dual luciferase reporter assay (Promega).

RNA Interference

Accell siRNAs targeting *SMG7* (E-021305), *GAPDH* (as positive control; D-001930) or with non-targeting sequences (as negative control; D-001910) were synthesized by Dharmacon [12]. PBMCs from 13 SLE patients were isolated by Ficoll-Hypaque, resuspended in Accell delivery media plus 3 μ M siRNA, distributed to 96-well plates at 2×10^5 cells/well, and divided into silence groups for *SMG7*, *GAPDH*, non-targeting and medium only mock controls. Cells were incubated at 37°C with 5% CO₂ for 5 days, supernatants were collected and cells were harvested to measure specific inhibition of *SMG7* and *GAPDH* by qRT-PCR.

Enzyme-linked immunosorbent assay (ELISA)

ANA IgG, CCL19, chemokine (C-X-C motif) ligand 10 (CXCL10), interleukin 6 (IL-6), IL-17, B-cell activating factor (BAFF) and interferon α (IFN- α) levels in supernatants were measured by ELISA kits.

Statistical analysis

Allelic and conditional haplotype-based association tests in each ancestry were performed under a logistic regression model adjusted for gender and the first three principal components estimated using AIMs. The Bonferroni corrected *P*-value threshold was adjusted to $P < 1 \times 10^{-3}$ based on the maximum number of tests across all populations (49 independent variants with $r^2 < 0.8$). For trans-ancestral meta-analysis, a fixed effect model was applied if Cochran's Q statistic showed no heterogeneity among odds ratios ($P > 0.05$); otherwise, a random effect model was used. Association analyses were performed using PLINK v1.07 [13]. Pairwise linkage disequilibrium (LD) values between SNPs were calculated using Haploview 4.2 [14]. For comparing between two groups, Student's *t*-test was used if the variance was normally distributed; whereas Mann-Whitney U test was used

if the variance was not normally distributed. Correlation between groups was evaluated using Pearson test. A P value < 0.05 was considered statistically significant.

RESULTS

We genotyped and imputed genetic variants (SNPs/INDELs) within a 308 kb region containing genes *LAMC2*, *NMNAT2*, *SMG7-AS1* and *SMG7*. After applying quality control measures, 35 genotyped and 85 to 278 imputed SNPs (varying among different ancestries) were assessed for association with SLE in 15,292 unrelated case-control participants from four ancestral groups: EA (3,438 SLEs vs. 3,417 controls), African Americans (AA: 1,679 vs. 1,934), Asians (AS: 1,265 vs. 1,260) and HS (1,492 vs. 807) (Figure 1).

Confirmation of *NMNAT2* association with SLE in EA and HS

In the largest EA dataset, 135 SNPs spanning the entire *NMNAT2* were significantly associated with SLE after Bonferroni correction ($P < 1.0 \times 10^{-3}$) for multiple comparisons (Figure 1B, Table S1). We confirmed the previous GWAS identified rs2022013 (MAF 38.1% in cases vs. 42.4% in controls, $P = 3.9 \times 10^{-7}$, OR=0.83 [95%CI=0.77–0.89]), and detected rs12146097 in *NMNAT2* (NM_015039) intron 1 that exhibited the strongest signal exceeding the GWAS significance level (16.7% vs. 12.9%, $P = 1.5 \times 10^{-10}$, OR=1.38 [95%CI=1.25–1.53]).

In HS, significant associations with SLE were localized within *NMNAT2* intron 1 after Bonferroni correction that rs536586 showed the strongest signal (44.2% vs. 39.2%, $P = 2.7 \times 10^{-4}$, OR=1.26 [95%CI=1.11–1.43]) (Figure 1B, Table S2), but no association at rs2022013 ($P = 0.11$).

In AA and AS, we observed only weak associations at SNPs in the *NMNAT2* region ($P < 0.05$) and none reached the Bonferroni-corrected significance level (Figure 1B, Table S3–S4).

Comparing across EA and HS, 8 SNPs ($P < 1.0 \times 10^{-3}$) within a ~51 kb interval in *NMNAT2* intron 1 showed consistent association with SLE, of which 6 SNPs (rs564146, rs681054, rs664422, rs502870, rs548292 and rs12146097) showed $P_{meta} < 5 \times 10^{-8}$ in trans-ancestral meta-analyses (Figure 1C, Table S5). Except rs12146097 (group 1), the other 5 SNPs (group 2) were in strong LD (r^2 0.97, $D' = 1.0$) and with rs2022013 (r^2 0.70, $D' = 0.96$) in both EA and HS. Conditioning on group 1 or 2 SNPs, *NMNAT2* associations were eliminated or reduced to baseline in HS, but retained residual signals in EA (Table S1–S2). We confirmed *NMNAT2* association with SLE in EA and HS.

Independent *SMG7* association with SLE in EA only

In EA, 21 SNPs in strong LD (r^2 0.98, $D' = 0.99$ termed group –3) located at ~32kb 5' of *NMNAT2* were strongly associated with SLE (2.4×10^{-8} P 4.3×10^{-7}), exhibiting the best signal at rs2702178 in *SMG7* intron 1 (40.8% vs. 36.2%, $P = 2.4 \times 10^{-8}$, OR=1.23 [95%CI=1.14–1.32]) (Figure 1B, Table S1). All group 3 *SMG7* SNPs were in low LD with group 1 & 2 *NMNAT2* SNPs (r^2 0.22, $D' = 0.80$), and genetic effects of these three groups

were independent in EA by conditional tests (Table S1). Unlike results in EA, *SMG7* SNPs exhibited modest associations with SLE (7.5×10^{-3} P 0.02) that none passed Bonferroni-corrected significance in HS (Table S2), and were not well imputed in AA and AS which might explain the lack of significant association (Table S3–S4). Taken together, we identified multiple *SMG7* SNPs showing independent association with SLE in EA only.

SLE-associated SNPs show allelic differences in *SMG7* expression

Given group 1&2 (5.5kb at *NMNAT2* intron 1) and group 3 (97kb from promoter to intron 17 of *SMG7*) SNPs located in non-coding regulatory regions, we assessed their potential functions using the UCSC Genome Browser (Figure S1). While group 1&2 SNPs showed low or modest overlap with regulatory elements, group 3 SNPs (especially rs2275675 and rs10911339, $r^2=1$) colocalized with annotations for H3K4Me3/H3K27Ac marks (promoter/enhancer activity), DNase hypersensitivity and transcription factor-binding signals, suggesting their potential effects in regulating gene expression.

We evaluated eQTL datasets [15–19] to characterize allelic differences in flanking gene expression, including *LAMC1*, *LAMC2*, *NMNAT2*, *SMG7-AS1*, *SMG7*, *NCF2*, *ARPC5*, *APOBEC4* and *RGL1*. All three group SNPs exhibited eQTL evidence for *SMG7*, but not for *LAMC1*, *LAMC2*, *NMNAT2*, *SMG7-AS1*, *APOBEC4* and *RGL1*, or weak effects on *NCF2* and *ARPC5* in diverse cell types from European-derived donors, including fibroblasts, adipocytes, lymphoblastoid cell lines, peripheral blood cells, primary T, B cells and monocytes. Compared to group 1&2 SNPs, group 3 SNPs (tagged by rs10911339 or rs2275675) had stronger eQTL effect of association between SLE-risk alleles and decreased *SMG7* levels (Table S6). No eQTLs were reported at *NMNAT2* in studies from a variety of human brain tissues [20–27]. Surprisingly, *cis* associations between SLE-risk SNPs and *SMG7* expression were also identified in brain tissues [23, 24], suggesting that modulation of expression levels is a likely functional mechanism of SLE-associated *SMG7* variants.

Dose-dependent association between the SLE-risk allele of rs2275675 and decreased *SMG7* levels

To support eQTL findings, we assessed association of rs12146097 and rs2275675 with *SMG7* mRNA levels in PBMCs from 86 SLE patients and 119 controls of European ancestry. Consistently, the SLE-risk C allele of rs2275675 was dose-dependently associated with decreased *SMG7* mRNA levels in both SLE and control subjects ($P=1.1 \times 10^{-3}$ and 6.8×10^{-8} in linear regression, respectively; Figure 2A). SLE patients and controls of the same genotype showed insignificant differences in *SMG7* mRNA levels (genotype TT: $P=0.34$, CT: $P=0.12$, CC: $P=0.87$, SLE vs. controls in *t* test). *SMG7* expression was not significantly associated with rs12146097, which might be due to fewer minor-allele carriers (MAF: 14% of rs12146097 vs. 37% of rs2275675 in Europeans; 1000 Genomes).

Expression of long non-coding RNA (lncRNA) often correlates with its *cis* protein-coding genes, which are subjected to eQTL regulation [28]. lncRNA *SMG7-AS1* (*SMG7* antisense RNA 1) is located 0.4kb upstream of *SMG7*; transcript levels of *SMG7* and *SMG7-AS1* were positively correlated in control PBMCs ($r=0.60$, $P=3.5 \times 10^{-9}$; Figure S2A). Located in

its second intron, the rs2275675 risk allele was also associated with decreased *SMG7-AS1* levels ($P=0.008$, Figure S2B).

We performed luciferase reporter assays to test the effects of rs2275675 and its linked SNP rs10911339 on transcription activity. Compared to the non-risk T-allele, transfections with the rs2275675 risk C-allele construct showed significantly lower luciferase activity in both HEK-293 and Raji cells ($P=0.0035$ and 0.0037 , respectively; Figure 2B). The rs10911339 constructs showed no significant allelic differences (Figure 2C). Consistent results from *ex vivo* and *in vitro* studies suggest that rs2275675 might best tag SLE-associated *SMG7* SNPs to affect *SMG7* expression, with the risk allele conferring decreased mRNA levels.

Decreased *SMG7* levels are associated with increased ANA titers

The NMD pathway degrades mRNAs with premature termination codons (PTCs), preventing production of truncated, deleterious proteins [29]. Given that *SMG7* participates in UPF1 dephosphorylation regulating the NMD pathway [30], we hypothesize that decreased *SMG7* levels may affect NMD efficiency, leading to accumulation of PTC-containing transcripts complexed with proteins (termed mRNA ribonucleoprotein particles, mRNPs) and autoantibody production in SLE. First, we observed a significant inverse correlation between ANA titers and *SMG7* mRNA levels in PBMCs from 68 European-derived SLE patients ($r=-0.31$, $P=0.01$; Figure 3A).

Next, we tested whether siRNA-mediated knock down of *SMG7* levels in SLE PBMCs might affect autoantibody production. We selected seropositive SLE patients (ANA positive at the time of blood drawn) with inactive disease (SLE Disease Activity Index score <4 [31]) and prednisone dosage less than 15 mg/d to minimize disease activity or medication variations. We confirmed the specificity of gene inhibition after 5-day culturing (Figure S3). Consistent with our *ex vivo* findings of inverse correlation between *SMG7* levels and ANA titers, PBMCs treated with *SMG7* siRNA showed a 8.0% increase in ANA IgG levels ($P=2.0\times 10^{-5}$, compared to negative control group; Figure 3B). We also tested if *SMG7* was linked to the cyto/chemokine disturbance reported in SLE patients [32] by measuring levels of CCL19, CXCL10, IL-6, IL-17, BAFF and IFN- α in culture supernatants. Levels of CCL19, an IFN-regulated chemokine and a proposed biomarker for lupus activity [33], showed a 150% increase in *SMG7* siRNA treated group compared to negative controls ($P=2.0\times 10^{-4}$; Figure 3C). Other cyto/chemokines exhibited no differences after siRNA treatments (Figure S4). Taken together, decreased *SMG7* expression could affect levels of ANA IgG and CCL19 in SLE PBMCs, supporting a potential role for *SMG7* in autoantibody production.

DISCUSSION

Trans-ancestral fine-mapping of *NMNAT2/SMG7* confirmed previous SLE GWAS reported association within *NMNAT2* intron 1 in EA and HS, and identified multiple *SMG7* SNPs showing robust and independent genetic effect in EA. Of interest, risk alleles of SLE-associated SNPs at both *NMNAT2* and *SMG7* showed association with decreased transcript levels of *SMG7*, but not *NMNAT2* in public eQTL datasets of immune cells from European participants. Consistently, rs2275675 located in the *SMG7* promoter that tagged SLE-

associated *SMG7* SNPs was experimentally confirmed for dose-dependent association between the risk allele and low *SMG7* mRNA levels in PBMCs and reduced transcription activity in transfected cell lines. SLE patients with higher ANA titers had lower *SMG7* mRNA levels in PBMCs and inhibition of *SMG7* expression resulted in increased levels of ANA IgG and CCL19 in SLE PBMC cultures. Our data demonstrated *SMG7* as a novel SLE-risk gene and implicated the NMD pathway affected by lowered *SMG7* expression promoting the development of SLE manifestation.

The NMD pathway functions in RNA surveillance to post-transcriptionally regulate transcriptomes [34]. PTC-containing mRNAs, arising from nonsense or frameshift mutations, errors in alternative splicing, or programmed DNA rearrangements occurred in T-cell receptor and immunoglobulin genes, comprise the major group of NMD substrates [30]. If escaping NMD, the encoded truncated proteins with no/deleterious functions could affect clinical phenotypes [35]. In rare SLE cases, PTC-containing mRNAs of *DNASE1* and *TREX1* encode truncated DNASE1 with reduced enzyme activity and truncated TREX1 with altered subcellular localization, which may impair DNA clearance and lead to disease manifestations [36, 37]. In addition, NMD can modulate the expression of approximately 10% of physiological mRNAs involved in diverse cellular processes, including spliceosomal components snRNP and Sm which are lupus autoantigens [30, 38]. Depletion of *SMG7* inhibits NMD activity in HeLa cells, resulting in enhanced expression of PTC-containing transcripts [39]. We postulated that inefficient NMD pathway attributable to decreased *SMG7* expression may cause RNA-protein complex accumulation, leading to autoantibody production in SLE. Interestingly, the rs2275675 risk-allele which conferred decreased *SMG7* mRNA levels was associated with the presence of anti-dsDNA autoantibody (Table S7). Further evidence supporting our hypothesis included a) inverse correlation between low *SMG7* mRNA levels and high ANA titers in SLE patients, and b) increased ANA IgG levels in SLE PBMC cultures after *SMG7* inhibition. However, our limited sample size might preclude us to further delineate associated autoantibody specificity (Figure S5). In addition to ANA, we also found elevated CCL19 levels after *SMG7* silencing in SLE PBMCs, suggesting a potential role for *SMG7* in regulating CCL19. CCL19, a ligand for the chemokine receptor CCR7, promotes migration/interaction of B-Th cells, formation of germinal center and production of high-affinity, class-switched antibodies [40, 41]. How *SMG7* depletion affects autoantibody and CCL19 production are currently unclear.

Recently, a SLE GWAS conducted in European-derived population reported association in the *SMG7/NCF2* region [42]. Our functional data provides the first evidence supporting *SMG7* involved in SLE pathogenesis. Rs2275675, located in a regulatory region, might best tag SLE-associated *SMG7* SNPs affecting *SMG7* expression. Further analysis by the HaploReg database predicted allelic differences of rs2275675 in binding to a transcription factor THAP1 that functions in endothelial cell proliferation and proapoptotic processes [43, 44]. The exact molecular mechanism by which rs2275675 regulates *SMG7* expression is yet to be experimentally validated.

Proximity to the *NMNAT2/SMG7* region at 1q25 is *NCF2*, encoding a subunit of nicotinamide adenine dinucleotide phosphate oxidase complex involved in the reactive oxygen species generation. Genetic associations of *NCF2* with SLE include four coding

(H389Q, A202A, R395W and V297A) and three intronic (rs10911359, rs34423782 and rs34680162) variants [42, 45, 46], but none showed strong LD with SNPs at the *NMNAT2/SMG7* locus (Figure S6). Association signals at *NMNAT2/SMG7* identified by this study may not be driven by SLE-associated SNPs at *NCF2*.

Limitations of the current study include a) insufficient statistical power due to the relatively small sample size of non-EA ancestries, b) we might have not comprehensively investigated SNPs in the *NMNAT2/SMG7* region to localize causal variant(s) in SLE, c) although a trend association of rs12146097 with risk for cognitive impairment was observed in a subset of EA patients (25 vs. 284 patients with vs. without impairment assessed by [47], $P=0.07$), functional relevance of *NMNAT2* to SLE pathogenesis remains inconclusive, and d) although we observed correlation between *SMG7* and *SMG7-AS1* transcript levels, how *SMG7* expression is regulated by its local lncRNA need further investigation.

In summary, we confirmed *NMNAT2* association in EA and HS, and identified *SMG7* as a novel risk gene for SLE in EA. Our data shows a link between SLE-associated variants in the *NMNAT2/SMG7* locus and decreased *SMG7* mRNA levels, and provides evidence for functional relevance of *SMG7* in production of ANA and chemokines in SLE. Given *SMG7* encoding a protein involved in the NMD pathway, our study implicates the novel contribution of this important regulatory pathway to SLE pathogenesis. Further investigation of its effects may lead to new directions in the development of therapeutic targets for SLE.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank all subjects for participation in this study, and Erika Magdangal for help with DNA preparation and organization. The BIOLUPUS network is composed of Johan Frostegård, MD, PhD (Huddinge, Sweden), Lennart Truedsson, MD, PhD (Lund, Sweden), Enrique de Ramón, MD PhD (Málaga, Spain), José M. Sabio, MD, PhD (Granada, Spain), María F. González-Escribano, PhD (Sevilla, Spain), Javier Martin, MD, PhD (Granada, Spain), Norberto Ortego-Centeno (Granada, Spain), José Luis Callejas MD (Granada, Spain), Julio Sánchez-Román, MD (Sevilla, Spain), Sandra D'Alfonso, PhD (Novara, Italy), Sergio Migliarese MD (Napoli, Italy), Gian-Domenico Sebastiani MD (Rome, Italy), Mauro Galeazzi MD (Siena, Italy), Torsten Witte, MD, PhD (Hannover, Germany), Bernard R. Lauwerys, MD, PhD (Louvain, Belgium), Eموke Endreffy, PhD (Szeged, Hungary), László Kovács, MD, PhD (Szeged, Hungary), Carlos Vasconcelos, MD, PhD (Porto, Portugal) and Berta Martins da Silva, PhD (Porto, Portugal). The members of GENLES Network are Hugo R. Scherbarth, Pilar C. Marino, Estela L. Motta, Susana Gamron, Cristina Drenkard, Emilia Menso, Alberto Allievi, Guillermo A. Tate, Jose L. Presas, Simon A. Palatnik, Marcelo Abdala, Mariela Bearzotti, Alejandro Alvarelos, Francisco Caeiro, Ana Bertoli, Sergio Paira, Susana Roverano, Cesar E. Graf, Estela Bertero, Cesar Caprarulo, Griselda Buchanan, Carolina Guillerón, Sebastian Grimaudo, Jorge Manni, Luis J. Catoggio, Enrique R. Soriano, Carlos D. Santos, Cristina Prigione, Fernando A. Ramos, Sandra M. Navarro, Guillermo A. Berbotto, Marisa Jorfen, Elisa J. Romero, Mercedes A. Garcia, Juan C Marcos, Ana I. Marcos, Carlos E. Perandones, Alicia Eimon, Sanatorio Parque and Cristina G. Battagliotti in Argentina; Eduardo Acevedo and Mariano Cucho in Perú; Ignacio García de la Torre, Mario Cardiel Ríos, José Francisco Moctezuma and Marco Maradiaga Ceceña in Mexico.

Funding

This work was supported by the US National Institutes of Health [R01AR043814 and R21AR065626 (B.P.T.), P01AI083194 (J.B.H., K.M.S., R.P.K., L.A.C., T.J.V., M.E.A.R., C.O.J., B.P.T. and P.M.G.), P01AR049084 (R.P.K., J.C.E., E.E.B., G.S.A., J.D.R., R.R.G. and M.A.P.), R01AR064820 (E.E.B, M.A.P., R.R.G., J.D.R. and L.M.V.), P30GM103510, P30AR053483, U01AI101934 and U19AI082714 (J.A.J.), R01CA141700 and RC1AR058621 (M.E.A.R.), P60AR053308 and UL1TR000004 (L.A.C.), P60AR062755 and UL1RR029882 (G.S.G. and D.L.K.), R01AR057172 (C.O.J.), R01AI063274 (P.M.G.), R01AR043274 (K.M.S.), K08AI083790,

LRPAI071651 and UL1RR024999 (T.B.N.), R01AR43727 (M.A.P.), K24AR002138, P60AR066464 and 1U54TR001018 (R.R.G.), U54RR023417 (J.D.R.), R01AR051545 and UL1RR025014 (A.M.S.), R21AI070304 (S.A.B.), R01AR042460, N01AR062277, P20RR020143 and R37AI024717 (J.B.H.), P30AR048311 and P30AR055385 (E.E.B.), R01AR033062 (R.P.K.)], the Lupus Foundation of America (B.P.T.), the Alliance for Lupus Research (B.P.T., Y.D., K.M.S., T.B.N., L.A.C., C.O.J. and S.A.B.), the Lupus Research Institute (T.B.N.), the US Department of Veterans Affairs (Merit Awards; J.B.H. and G.S.G.), the US Department of Defense (PR094002, J.B.H.), the Arthritis National Research Foundation (Eng Tan Scholar Award; J.Z. and T.B.N.), the Arthritis Foundation (A.M.S. and P.M.G.), the Korea Healthcare Technology R&D Project, Ministry for Health and Welfare, Republic of Korea (HI13C2124; S.C.B.), the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI13C1754; Y.W.S.), the European Science Foundation RNP (BIOLUPUS Research Network), the Wellcome Trust (T.J.V.), Arthritis Research UK (T.J.V.), a Kirkland Scholar Award (L.A.C.), the Wake Forest School of Medicine Center for Public Health Genomics (C.D.L.), and UCLA Clinical and Translational Science Institute (CTSI) UL1RR033176 and UL1TR000124. Some RNA samples of healthy controls used in this study were provided by the UCLA/CFAR Virology Core Lab which was supported by the NIH grant P30AI028697. The funders had no role in study design, data collection, analysis, and interpretation, writing of the report, or decision to submit the paper for publication.

References

1. Deng Y, Tsao BP. Advances in lupus genetics and epigenetics. *Curr Opin Rheumatol*. 2014; 26:482–92. [PubMed: 25010439]
2. Harley JB, Alarcon-Riquelme ME, Criswell LA, et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXXK, KIAA1542 and other loci. *Nat Genet*. 2008; 40:204–10. [PubMed: 18204446]
3. Gateva V, Sandling JK, Hom G, et al. A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. *Nat Genet*. 2009; 41:1228–33. [PubMed: 19838195]
4. Berger F, Lau C, Dahlmann M, et al. Subcellular compartmentation and differential catalytic properties of the three human nicotinamide mononucleotide adenylyltransferase isoforms. *J Biol Chem*. 2005; 280:36334–41. [PubMed: 16118205]
5. Mayer PR, Huang N, Dewey CM, et al. Expression, localization, and biochemical characterization of nicotinamide mononucleotide adenylyltransferase 2. *J Biol Chem*. 2010; 285:40387–96. [PubMed: 20943658]
6. Isken O, Maquat LE. The multiple lives of NMD factors: balancing roles in gene and genome regulation. *Nat Rev Genet*. 2008; 9:699–712. [PubMed: 18679436]
7. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum*. 1997; 40:1725. [PubMed: 9324032]
8. Sakurai D, Zhao J, Deng Y, et al. Preferential binding to Elk-1 by SLE-associated IL10 risk allele upregulates IL10 expression. *PLoS Genet*. 2013; 9:e1003870. [PubMed: 24130510]
9. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006; 38:904–9. [PubMed: 16862161]
10. Hoggart CJ, Shriver MD, Kittles RA, et al. Design and analysis of admixture mapping studies. *Am J Hum Genet*. 2004; 74:965–78. [PubMed: 15088268]
11. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009; 5:e1000529. [PubMed: 19543373]
12. Feng X, Li R, Huang J, et al. Olf1/EBF associated zinc finger protein interfered with antinuclear antibody production in patients with systemic lupus erythematosus. *Arthritis Res Ther*. 2010; 12:R59. [PubMed: 20359360]
13. 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. [PubMed: 17701901]
14. Barrett JC, Fry B, Maller J, et al. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005; 21:263–5. [PubMed: 15297300]
15. Dimas AS, Deutsch S, Stranger BE, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science*. 2009; 325:1246–50. [PubMed: 19644074]

16. Nica AC, Parts L, Glass D, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet.* 2011; 7:e1002003. [PubMed: 21304890]
17. Stranger BE, Montgomery SB, Dimas AS, et al. Patterns of cis regulatory variation in diverse human populations. *PLoS Genet.* 2012; 8:e1002639. [PubMed: 22532805]
18. Fairfax BP, Makino S, Radhakrishnan J, et al. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat Genet.* 2012; 44:502–10. [PubMed: 22446964]
19. Westra HJ, Peters MJ, Esko T, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet.* 2013; 45:1238–43. [PubMed: 24013639]
20. Myers AJ, Gibbs JR, Webster JA, et al. A survey of genetic human cortical gene expression. *Nat Genet.* 2007; 39:1494–9. [PubMed: 17982457]
21. Heinzen EL, Ge D, Cronin KD, et al. Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* 2008; 6:e1. [PubMed: 19222302]
22. Webster JA, Gibbs JR, Clarke J, et al. Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet.* 2009; 84:445–58. [PubMed: 19361613]
23. Gibbs JR, van der Brug MP, Hernandez DG, et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* 2010; 6:e1000952. [PubMed: 20485568]
24. Liu C, Cheng L, Badner JA, et al. Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol Psychiatry.* 2010; 15:779–84. [PubMed: 20351726]
25. Colantuoni C, Lipska BK, Ye T, et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature.* 2011; 478:519–23. [PubMed: 22031444]
26. Kim S, Cho H, Lee D, et al. Association between SNPs and gene expression in multiple regions of the human brain. *Transl Psychiatry.* 2012; 2:e113. [PubMed: 22832957]
27. Zou F, Chai HS, Younkin CS, et al. Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLoS Genet.* 2012; 8:e1002707. [PubMed: 22685416]
28. Popadin K, Gutierrez-Arcelus M, Dermizakis ET, et al. Genetic and epigenetic regulation of human lincRNA gene expression. *Am J Hum Genet.* 2013; 93:1015–26. [PubMed: 24268656]
29. Maquat LE. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol.* 2004; 5:89–99. [PubMed: 15040442]
30. Bhuvanagiri M, Schlitter AM, Hentze MW, et al. NMD: RNA biology meets human genetic medicine. *Biochem J.* 2010; 430:365–77. [PubMed: 20795950]
31. FitzGerald JD, Grossman JM. Validity and reliability of retrospective assessment of disease activity and flare in observational cohorts of lupus patients. *Lupus.* 1999; 8:638–44. [PubMed: 10568901]
32. Yu SL, Kuan WP, Wong CK, et al. Immunopathological roles of cytokines, chemokines, signaling molecules, and pattern-recognition receptors in systemic lupus erythematosus. *Clin Dev Immunol.* 2012; 2012:715190. [PubMed: 22312407]
33. Bauer JW, Petri M, Batliwalla FM, et al. Interferon-regulated chemokines as biomarkers of systemic lupus erythematosus disease activity: a validation study. *Arthritis Rheum.* 2009; 60:3098–107. [PubMed: 19790071]
34. Rehwinkel J, Raes J, Izaurralde E. Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors. *Trends Biochem Sci.* 2006; 31:639–46. [PubMed: 17010613]
35. Khajavi M, Inoue K, Lupski JR. Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *Eur J Hum Genet.* 2006; 14:1074–81. [PubMed: 16757948]
36. Yasutomo K, Horiuchi T, Kagami S, et al. Mutation of DNASE1 in people with systemic lupus erythematosus. *Nat Genet.* 2001; 28:313–4. [PubMed: 11479590]
37. Lee-Kirsch MA, Gong M, Chowdhury D, et al. Mutations in the gene encoding the 3′-5′ DNA exonuclease TREX1 are associated with systemic lupus erythematosus. *Nat Genet.* 2007; 39:1065–7. [PubMed: 17660818]
38. Saltzman AL, Kim YK, Pan Q, et al. Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Mol Cell Biol.* 2008; 28:4320–30. [PubMed: 18443041]

39. Loh B, Jonas S, Izaurralde E. The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev.* 2013; 27:2125–38. [PubMed: 24115769]
40. Comerford I, Harata-Lee Y, Bunting MD, et al. A myriad of functions and complex regulation of the CCR7/CCL19/CCL21 chemokine axis in the adaptive immune system. *Cytokine Growth Factor Rev.* 2013; 24:269–83. [PubMed: 23587803]
41. Asquith DL, Bryce SA, Nibbs RJ. Targeting cell migration in rheumatoid arthritis. *Curr Opin Rheumatol.* 2015; 27:204–11. [PubMed: 25603038]
42. Bentham J, Morris DL, Cunninghame Graham DS, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat Genet.* Published Online First: 26 Oct 2015.
43. Cayrol C, Lacroix C, Mathe C, et al. The THAP-zinc finger protein THAP1 regulates endothelial cell proliferation through modulation of pRB/E2F cell-cycle target genes. *Blood.* 2007; 109:584–94. [PubMed: 17003378]
44. Roussigne M, Cayrol C, Clouaire T, et al. THAP1 is a nuclear proapoptotic factor that links prostate-apoptosis-response-4 (Par-4) to PML nuclear bodies. *Oncogene.* 2003; 22:2432–42. [PubMed: 12717420]
45. Jacob CO, Eisenstein M, Dinanuer MC, et al. Lupus-associated causal mutation in neutrophil cytosolic factor 2 (NCF2) brings unique insights to the structure and function of NADPH oxidase. *Proc Natl Acad Sci U S A.* 2012; 109:E59–67. [PubMed: 22203994]
46. Kim-Howard X, Sun C, Molineros JE, et al. Allelic heterogeneity in NCF2 associated with systemic lupus erythematosus (SLE) susceptibility across four ethnic populations. *Hum Mol Genet.* 2014; 23:1656–68. [PubMed: 24163247]
47. Petri M, Naqibuddin M, Carson KA, et al. Depression and cognitive impairment in newly diagnosed systemic lupus erythematosus. *J Rheumatol.* 2010; 37:2032–8. [PubMed: 20634244]

Figure 1. Association of SNPs in the *NMNAT2/SMG7* region with SLE

(A) The genomic structure of the *NMNAT2/SMG7* region and positions of genetic variants are indicated. (B) The allelic P value ($-\log_{10}P$ value) of each genetic variant with SLE is plotted against its position as a circle (genotyped) or a triangle (imputed) for European American (EA), Amerindian/Hispanic (HS), African American (AA) and Asian (AS), respectively. Genetic variants are highlighted using different colors according to their strength of linkage disequilibrium (r^2) with rs2022013 (shown as a blue diamond) in each ancestry. The dashed line represents a Bonferroni corrected $P < 1 \times 10^{-3}$. Arrows identify rs2022013 and SNPs demonstrating the most significant association signals in each ancestry. Black rectangle identifies group 3 SNPs at *SMG7* strongly associated with SLE in EA and the best-associated SNP rs2702178 is indicated. (C) Trans-ancestral meta-analysis is conducted on 8 SNPs that remain significant associations after Bonferroni correction in both EA and HS. Black rectangle identifies group 1&2 SNPs at *NMNAT2* intron 1 showing P_{meta} values exceeding the GWAS significance level. The dashed line represents the significance level of 5×10^{-8} .

Figure 2. The SLE-risk C allele of *rs2275675* associated with decreased *SMG7* expression
 (A) Association of *rs2275675* genotypes with *SMG7* mRNA levels in PBMCs from European-derived SLE patients and healthy controls, respectively. Each symbol represents an individual and horizontal lines indicate mean $\pm$ SEM values. (B&C) Allelic differences of *rs2275675* and *rs10911339* in luciferase activity. Data show the mean $\pm$ SEM results of three independent experiments.

Figure 3. Decreased *SMG7* levels associated with increased ANA

(A) *SMG7* mRNA levels in PBMCs correlated inversely with ANA titers of SLE patients. Among 68 patients, 56 were recruited at UCLA and 12 were recruited at OMRF, and their ANA titers were measured in Clinical Laboratory at each site. (B&C) Increased ANA IgG and CCL19 levels after *SMG7* silencing in SLE PBMCs. PBMCs from ANA positive SLE patients ($n=13$) were incubated for 5 days in the presence or absence of siRNA targeting *SMG7*, *GAPDH* or siRNA with a non-targeting sequence (NC), respectively. ANA IgG (B) and CCL19 (C) levels in culture supernatants were measured by ELISA, and plotted as fold change with respect to mock (culture medium only). Results are presented as mean $\pm$ SEM.