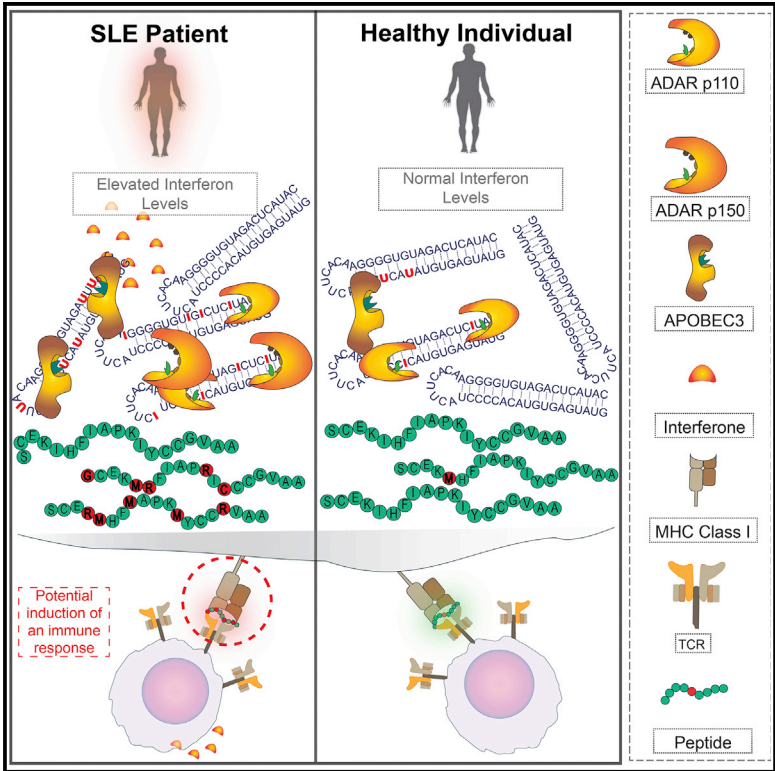# Cell Reports

# Increased RNA Editing May Provide a Source for Autoantigens in Systemic Lupus Erythematosus

## Graphical Abstract



## Authors

Shalom Hillel Roth, Miri Danan-Gotthold, Meirav Ben-Izhak, Gideon Rechavi, Cyrille J. Cohen, Yoram Louzoun, Erez Y. Levanon

## In Brief

Roth et al. show that SLE patients have elevated RNA editing, a process that modifies an RNA sequence from the sequence encoded in the genome. This is manifested as both increased numbers and various altered coding sequences, which may promote autoimmune progression by increasing autoantigenic load.

## Highlights

- SLE patients have elevated RNA editing in their blood

- In SLE, elevated editing at coding regions can potentially generate autoantigens

- Elevated RNA editing may promote SLE progression by increasing autoantigen load

CellPress

# Increased RNA Editing May Provide a Source for Autoantigens in Systemic Lupus Erythematosus

Shalom Hillel Roth,[1,6] Miri Danan-Gotthold,[1,6] Meirav Ben-Izhak,[2,3] Gideon Rechavi,[4] Cyrille J. Cohen,[5] Yoram Louzoun,[2,3] and Erez Y. Levanon[1,7,*]

[1]The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat Gan 5290002, Israel
[2]Department of Mathematics, Bar-Ilan University, Ramat Gan 5290002, Israel
[3]The Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan 5290002, Israel
[4]Cancer Research Center and the Wohl Institute for Translational Medicine, Chaim Sheba Medical Center, Tel Hashomer 52621, Israel
[5]Laboratory of Tumor Immunology and Immunotherapy, The Mina and Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat Gan 5290002, Israel
[6]These authors contributed equally
[7]Lead Contact
Correspondence: erez.levanon@biu.ac.il
https://doi.org/10.1016/j.celrep.2018.03.036

## SUMMARY

RNA-editing mechanisms, which induce nucleotide substitution in the RNA, increase transcript and protein diversities. Editing dysregulation has been shown to lead to grave outcomes, and transcriptome-wide aberrant RNA editing has been found in tumors. However, little is known about the involvement of editing in other diseases. Systemic lupus erythematosus (SLE) is a multisystemic autoimmune disease characterized by a loss of tolerance for autoantigens from various tissues and the production of multiple autoantibodies. Here, we show that blood samples from individuals with SLE have abnormally high levels of RNA editing, some of which affect proteins and potentially generate novel autoantigens. We suggest that elevated RNA editing, either by ADARs or APOBECs, may be involved in the pathophysiology of SLE, as well as in other autoimmune diseases, by generating or increasing the autoantigen load, a key requisite for the progression of autoimmunity.

## INTRODUCTION

Adenosine-to-inosine (A-to-I) editing is a prevalent form of post-transcriptional modification (Bazak et al., 2014a; Ramaswami and Li, 2014), catalyzed by the adenosine deaminase acting on RNA (ADAR) enzymes (Bass, 2002; Nishikura, 2010; Savva et al., 2012). Because inosine is interpreted by most cellular machineries as a guanosine (G), there are several possible outcomes to this deamination (Speyer et al., 1962). Hence, editing can alter splice consensus elements, microRNA seeds, and binding sites, as well as protein coding sequences (Burns et al., 1997; Ekdahl et al., 2012; Higuchi et al., 2000; Rueter et al., 1999). The reaction is catalyzed in humans by the ADAR1 (ADAR) and ADAR2 (ADARB1) enzymes. ADAR2 is confined to

the nucleus, whereas ADAR1 has two isoforms—the p110 isoform that is localized in the nucleus and the p150 interferon (IFN)-inducible isoform that shuttles between the nucleus and the cytoplasm (Eckmann et al., 2001; George and Samuel, 1999).

A-to-I RNA editing is a constitutive and finely tuned process that is involved with several systems in the body and requires a delicate balance of activity. On the one hand, insufficient editing has detrimental and even lethal effects, partly due to the accumulation of endogenous double-stranded RNAs (dsRNAs), which triggers an innate immune response (Liddicoat et al., 2015a; Mannion et al., 2014; Pestal et al., 2015). ADAR1 editing deficiency is involved in autoimmune disorders such as Aicardi-Goutières syndrome (AGS) and dyschromatosis symmetrica hereditaria (DSH) (Liu et al., 2006; Rice et al., 2012). Both show some phenotypic similarities to systemic lupus erythematosus (SLE), and AGS patients, like SLE patients, are characterized by elevated type I IFN production (Crow and Manel, 2015; Lee-Kirsch et al., 2014; Rice et al., 2012).

On the other hand, excessive ADAR activity may be harmful. Studies show that elevated levels of ADAR1 usually correlate with higher editing activity, manifested both as higher rates of editing and as editing of non-typical editing sites (Han et al., 2015; Paz-Yaacov et al., 2015). Elevated levels of editing have been shown to have significant effects on the transcriptomic diversity of cancers and, in several cases, even correlate with patient survival (Han et al., 2015; Paz-Yaacov et al., 2015).

The other type of RNA editing in humans is the deamination of cytidine to uridine (C-to-U) by several members of the APOBEC protein family of cytidine deaminases, mainly APOBEC1 and APOBEC3A (Sharma et al., 2015; Smith et al., 2012). Similar to ADAR1, APOBEC3 paralogs are upregulated by IFN (Peng et al., 2006). Overexpression of APOBEC3A was shown to affect thousands of sites that have the potential to recode more than a thousand proteins (Sharma et al., 2017).

SLE is a multisystemic heterogenic autoimmune disease. It is characterized by autoantibodies to a variety of autoantigens derived from various tissues, as well as antinuclear antibodies (Lipsky, 2001; Tsokos, 2011). The causes and pathogenesis of the disease are not fully understood. SLE patients usually have

elevated levels of circulating type I IFN and increased expression of both type I and type II IFN-stimulated genes (ISGs) (Chiche et al., 2014; Crow and Manel, 2015; Hagberg and Rönnblom, 2015). Thus, inhibitors of IFN and downstream ISGs are attractive and promising therapeutic targets for treatment (Furie et al., 2017; Hagberg and Rönnblom, 2015; Kirou and Gkrouzman, 2013; Niewold, 2008). Moreover, high artificially induced levels of IFN can cause SLE-like symptoms (Ho et al., 2008; Niewold, 2008). Altered levels of ISGs and RNA modifications have been marked as potential contributors to SLE pathogenesis in several previous studies (Graham and Utz, 2005; Hueber et al., 2004). This includes increased rates of alternative splicing (Ng et al., 2004) and indications of aberrant RNA editing in a few genes, as well as elevated expression of ADAR1 in lymphocytes (Laxminarayana et al., 2002, 2007; Orlowski et al., 2008). However, these early studies focused only on small subset of editing targets.

Here, by analyzing wide-scale, high-throughput sequencing data (Hung et al., 2015), we show that SLE patients have elevated global RNA-editing levels. In addition, we use computational predictions to demonstrate that this excessive editing can potentially generate autoantigens, which may then be presented on major histocompatibility complex (MHC) molecules and induce an immune response. Editing in coding regions, which alters the amino acid sequence (recoding), may prove immunogenic. Because proteasomal degradation is presumably unaffected by RNA editing, an edited form of a protein could be degraded and eventually be presented by the MHC system. We previously showed that medullary thymic epithelial cells (mTECs), which mediate the presentation of self-epitopes in the thymus, routinely express many edited forms of proteins and thereby forestall the induction of a dangerous immune response against edited peptide-derived autoantigens (Danan-Gotthold et al., 2016). However, possibly due to incomplete presentation and/or confinement of the process to selected edited versions, not all edited forms are presented by the mTECs. As a consequence, T cells specific for edited epitopes may escape negative selection and react to cells that present an editing-derived neo-autoantigen (Danan-Gotthold et al., 2016). Although the main activity of editing is to inhibit the innate immune response to endogenous dsRNA (Liddicoat et al., 2015b; Mannion et al., 2014; Pestal et al., 2015), our results suggest that RNA editing can trigger an immune response directed against editing-originated neo-autoantigens, exacerbating autoimmunity. We therefore suggest it is involved in the etiology and pathogenesis of SLE and possibly other autoimmune diseases.

## RESULTS

### Enhanced A-to-I Editing in SLE Patients
Many studies of editing in various diseases concentrated on previously detected editing sites. Here, we used RNA sequencing (RNA-seq) data to evaluate changes in global levels of editing, without prior knowledge of specific sites. To compare the RNA-editing levels in SLE patients with those in healthy individuals, we first used two complementary approaches to estimate the global editing rate using RNA-seq samples from Hung et al. (2015) (99 from SLE patients and 18 from healthy individuals)

(Table S1): the Alu editing index (AEI), which measures the global rate of editing in Alu repeats (Bazak et al., 2014b), and the global rates of clustered editing sites in each sample, according to the normalized number of hyper-edited (HE) sites (see Experimental Procedures) (Porath et al., 2014). Both types of analysis revealed that global editing was significantly elevated in the blood of SLE patients compared to that of controls (Wilcoxon p value = 4.96e−6 and 8.27e−7 for AEI and HE analyses, respectively). Because higher editing levels may be associated with higher levels of ISGs, we grouped the patients according to the original dataset of interferon signature metric (ISM) division, which measures the presence of the IFN-inducible genes' expression signature. As expected, ISM-high SLE patients had significantly higher editing levels than controls (Wilcoxon p value = 4.16e−7 and 2.48e−8 for AEI and HE analyses, respectively) (Figures 1A and 1B) and significantly higher editing levels than ISM-low patients (Wilcoxon p value = 1.87e−2 and 1.77e−5 for AEI and HE analyses, respectively) (Figures 1A and 1B). However, even ISM-low patients had significantly higher editing levels than controls (Wilcoxon p value = 3.28e−2 and 3.93e−2 for AEI and HE analyses, respectively) (Figures 1A and 1B). Overall, the editing signal was clean. A-to-G mismatches comprised most mismatches (82.79%), and the ADAR deamination motif (5′ neighbor preference A = U > C > G) (Eggington et al., 2011; Riedmann et al., 2008) was observable (Figures 1C and 1D), indicating that most of the A-to-G substitutions detected in the HE analysis resulted from ADAR editing. Similar results were obtained from another, but smaller, independent dataset from Rai et al. (2016) (Figure S1).

As a complement approach, we investigated the level of RNA editing at known sites that have been identified and characterized by previous studies (see Experimental Procedures). Among the sites at which editing levels were found to differ significantly between the controls and the ISM-high patients, most had increased editing levels. This was also the general trend in sites without significant changes (Figure 1E). These results further support the observation that RNA editing is elevated in the blood of ISM-high patients.

To determine whether ADAR1, which is itself an ISG, is responsible for the increased editing levels, we examined its expression levels. As expected, its expression was significantly higher in ISM-high patients compared to controls (2.24-fold change, Wilcoxon p value = 3.77e−6) (Figure 1F). However, in contrast to the results of the global RNA-editing levels, no significant differences were observable between ISM-low patients and controls (Wilcoxon p value = 0.54). This may indicate that ADAR1 expression is regulated directly by IFN, while RNA-editing levels remain relatively high in these patients even without IFN induction. In contrast, the expression of ADAR2 was lower in ISM-high patients compared to controls (0.75-fold change, Wilcoxon p value = 0.035) (Figure 1G). Altogether, our results show significantly elevated ADAR1 expression and editing activity in ISM-high patients.

### Enhanced C-to-U Editing in SLE Patients
Proteins from the additional deaminase family, the APOBEC, were also significantly upregulated in ISM-high patients compared to controls (Figure 2A; Table S2). This may lead to
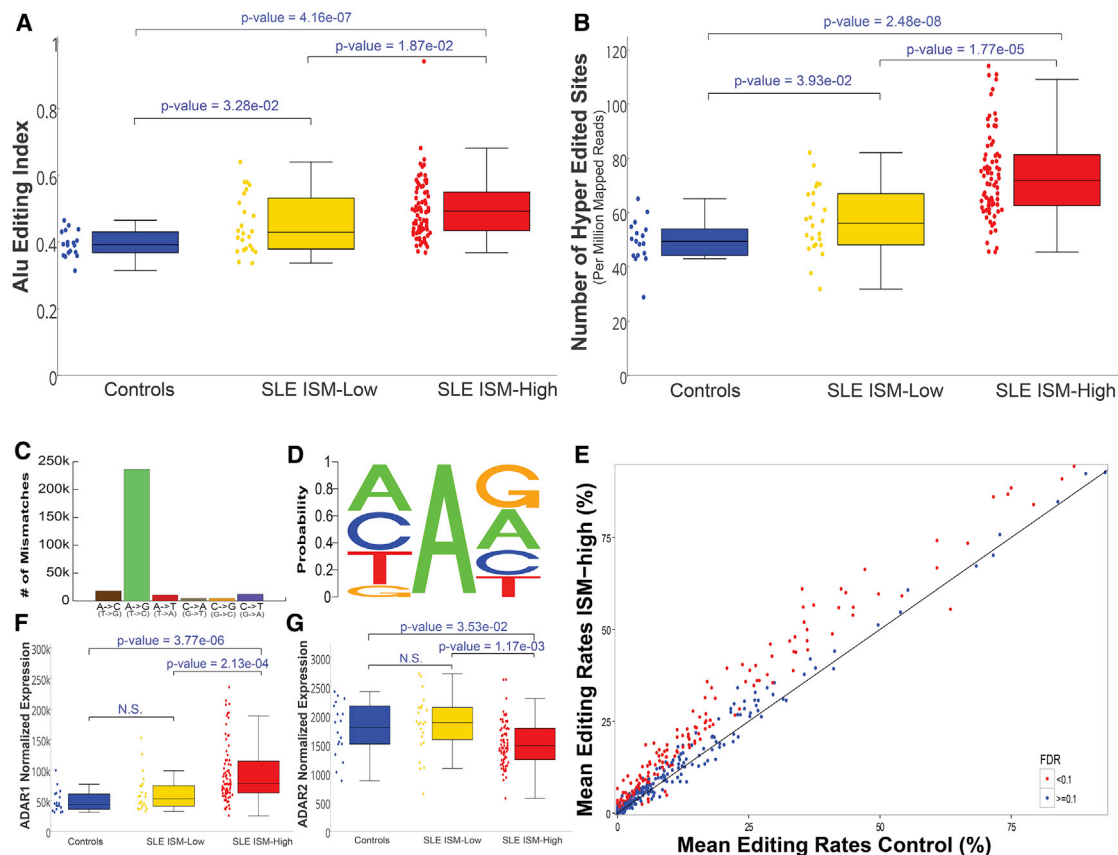
**Figure 1. A-to-I Editing Is Significantly Increased in SLE Patients**

Compared cohorts are healthy controls, SLE patients with low levels of ISG (ISM-low, which corresponds to a more quiescent disease), and SLE patients with high levels of ISG (ISM-high), depicted in blue, yellow, and red, respectively. The comparisons were evaluated using the Wilcoxon test.

(A and B) Global levels of editing were assessed by determining (A) the editing in Alu elements (AEI) and (B) the number of hyper-edited (HE) sites (normalized by the number of mapped reads) per sample.

(C and D) Cleanness of the HE signal was assessed by (C) the number A-to-G mismatches compared to all types of mismatches and (D) the neighboring nucleotides frequencies at the sites detected by the HE analysis (which matched the ADAR deamination motif).

(E) Editing levels were assessed using the rates at pre-known sites. The means of the editing rates of ISM-high patients were plotted against controls per site. Sites with significant changes (FDR < 0.1) are depicted in red.

(F and G) Expression levels of (F) ADAR1 and (G) ADAR2 suggest that ADAR1 is responsible for the alterations in editing levels.

the accumulation of C-to-U mismatches, which can serve as a potential source for neo-autoantigens. A dataset of putative C-to-U editing sites in monocytes and macrophages has been published (Sharma et al., 2015). Of the sites, 252 were detected in the examined dataset (96.5% of the expressed sites), and most of them (75%) were edited in more than 20% of the samples. Out of these, 26 had significantly elevated editing rates in ISM-high patients (false discovery rate [FDR] < 0.1), and a similar global trend was observable (Figure 2B). Because the editing probably occurs only in specific cell types (e.g., monocytes and macrophages), very low editing rates were observed in the total cell population (90% of the sites had a mean rate of 0%–2%). To further assess the C-to-U editing levels in SLE patients and healthy individuals, we measured the global rates of HE C-to-U clusters in each sample (see Experimental Procedures). As was the case for the A-to-I editing, ISM-high patients had significantly more clustered C-to-U sites per sample (Wil-

coxon p value = 4.87e−3) (Figure 2C), and their numbers correlated well with APOBEC3A levels (r = 0.70) (Figure 2D) indicating that this is the main deaminase driving this elevation.

**Elevated Levels of Recoding Events in SLE Patients**

Only a small portion of A-to-I RNA editing results in the recoding of proteins. To identify recoding sites associated with SLE, we systematically searched for differentially edited sites with a non-synonymous outcome. Because the current approaches for *de novo* detection of recoding sites without a matched DNA sequences from the same individual perform poorly, we limited the analyses to high-credibility sites found within the HE regions or to the previously verified ones.

We first examined putative recoding sites detected by the HE analysis, in which the potentially editable adenosine is found on the transcribed strand. Excluding known SNPs and sites within highly polymorphic genomic regions, we identified 624 putative
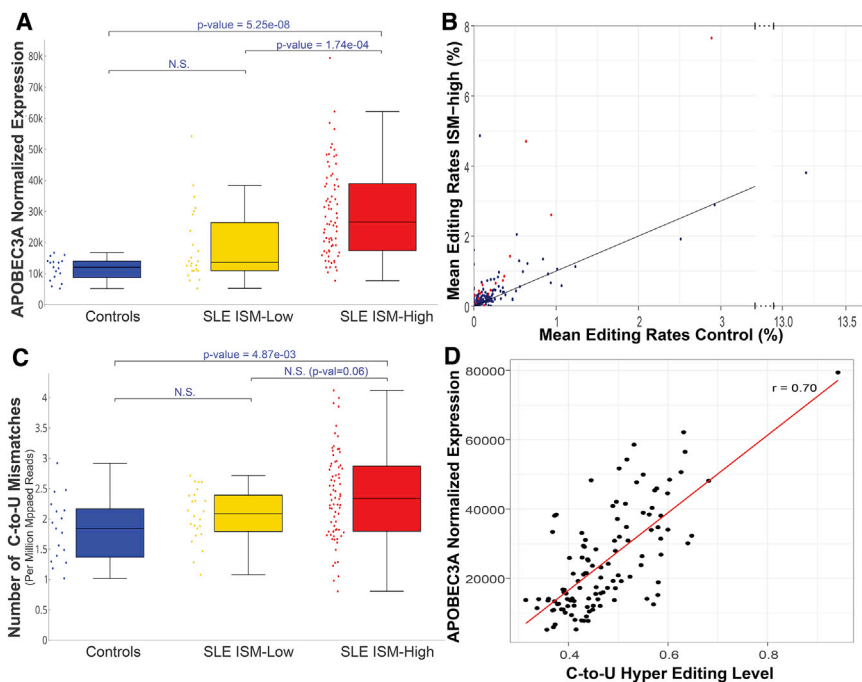
**Figure 2. Elevated C-to-U Editing in SLE Patients**

ISM-high patients, ISM-low patients, and controls are depicted in red, yellow, and blue, respectively. The paired comparisons were evaluated using the Wilcoxon test.

(A) Expression levels of APOBEC3A in ISM-high patients, ISM-low patients, and controls.

(B and C) Elevated C-to-U editing in SLE patients was assessed by (B) the mean C-to-U mismatch rates per site at pre-known sites of ISM-high patients, plotted against the controls (sites with significant changes [FDR < 0.1] are depicted in red) and (C) the number of HE C-to-U sites (normalized by the number of mapped reads) per sample.

(D) Correlation between APOBEC3A expression levels and normalized number of HE C-to-U sites.

## RNA Editing Results in Potential MHC Class I Epitopes

To investigate any putative immunogenicity caused by elevated editing in the patients, we evaluated the potential of MHC class I epitopes to be derived from peptides originating from edited transcripts, using histocompatibility leukocyte antigen (HLA) allele predictions for each sample (Szolek et al., 2014). We measured the editing rates of recoding editing events that generate peptides with a predicted high affinity for MHC class I molecules (see Experimental Procedures). SLE patients had significantly higher rates at such sites than controls (Wilcoxon p value = 1.59e−14 and 3.56e−5 for ISM-high and ISM-low patients, respectively), and ISM-high patients had higher rates than ISM-low patients (Wilcoxon p value = 1.65e−3). We focused on a subset of these sites for which editing enhances the binding affinity of the peptides (see Experimental Procedures). ISM-high patients had significantly higher rates at these sites than controls (Wilcoxon p value = 9.13e−7) (Figure 3D) and ISM-low patients (Wilcoxon p value = 1.86e−3). It is evident that these patients have higher editing rates, even at a fraction of the recoding sites, whose resulting proteins have higher affinities for the HLA alleles.

We also assessed the number of possible epitopes from each peptide, both before and after editing (see Experimental Procedures), which was mostly not statistically different, suggesting that the edited forms of the peptides have the same potential to be derived to epitopes as the non-edited forms.

## DISCUSSION

The results of this study suggest the possibility of a connection between autoimmunity and excessive RNA editing. We surmise that the latter may facilitate the generation of autoantigens in peripheral tissues. Because these autoantigens might not necessarily be expressed in the thymus, reactive T cells may escape the negative selection and recognize these recoded proteins as non-self. The elevated global editing activity in SLE patients, one of the manifestations of the inflammatory condition, results
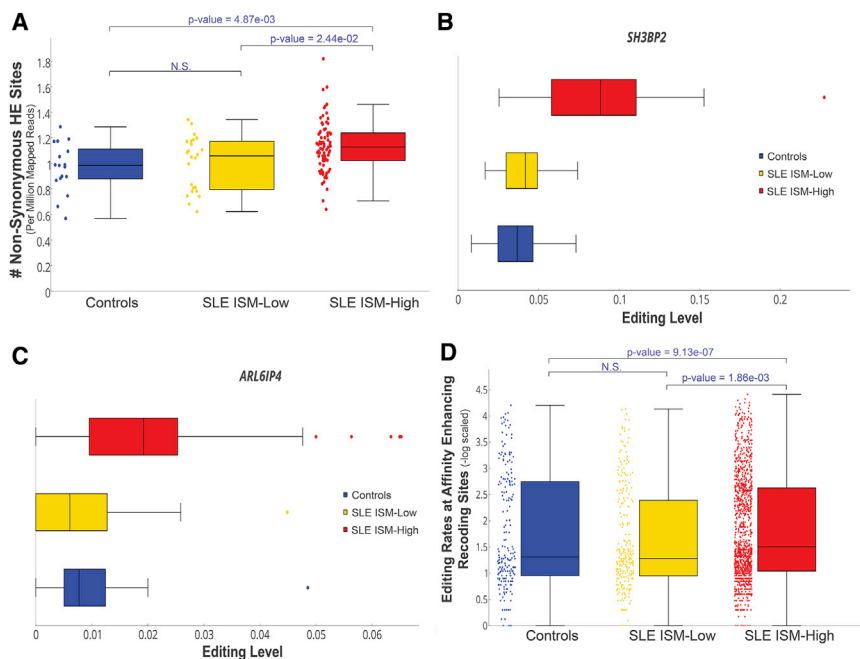
recoding sites. As expected, most of these sites (95%) were edited in SLE samples, and their prevalence was higher than in controls (an average of 1.100 sites per million mapped reads in SLE patients versus 0.967 in controls, Wilcoxon p value = 0.018) (Figure 3A). The obtained signal was noisy (Figure S2A), indicating that a portion of these sites are not *bona fide* ADAR targets. Nevertheless, the neighbor preferences obtained fairly agree with the ADAR motif (Figure S2B), supporting the contribution of genuine editing by ADAR. Many of these sites were stochastically edited (545 of 624 sites were edited in fewer than five samples). We therefore compared the relative number of samples that expressed the edited version of the transcript between the groups. Two such sites, in the genes *IFITM1* and *ODF3B*, were found to be edited in a higher percentage of ISM-high samples compared to controls (Fisher's exact test, FDR < 0.05), probably because of elevated expression levels in SLE patients (FDR < 1e−5). This elevated expression results in increased detectability by our analysis (and presumably in higher numbers of the recoded versions of peptide), which should also reflect higher visibility of them to the immune system. Several edited sites were found within the *SRSF4* and *VIM* genes, which are known targets for autoantibodies of SLE (Sherer et al., 2004).

We also examined the editing rates at recoding sites identified in other studies (see Experimental Procedures). Two sites were found to have higher editing levels in ISM-high patients (Wilcoxon, FDR < 0.05) (Figures 3B and 3C), with no significant expression difference in SLE patients and controls in the genes SH3BP2 and ARL6IP4 (at chr4:2,835,556 and chr12:123,466,262, respectively).

These results support the hypothesis that the high global editing levels in SLE patients can give rise to a variety and higher levels of the edited versions of proteins.

**Figure 3. A-to-I Editing in Non-synonymous Sites Is Significantly Elevated, Resulting in the Potential Generation of Epitopes**

ISM-high patients, ISM-low patients, and controls are depicted in red, yellow, and blue, respectively. The paired comparisons were evaluated using the Wilcoxon test.

(A) Normalized number of recoding sites within HE regions (normalized by the number of mapped reads) per sample.

(B and C) Editing levels of (B) SH3BP2 (at chr4:2,835,556) and (C) ARL6IP4 (SRp25, at chr12:123,466,262).

(D) Normalized number of sites per sample creating edited peptides that have superior HLA affinity over the non-edited forms.

in increased variety and higher levels of edited forms of proteins. Similarly, conditions that lead to editing alterations can result in new recoding events (Daniel et al., 2012). These have the potential to be processed into autoepitopes that may then be subsequently presented on the MHC molecules, thus stimulating an autoimmune response (Figure 4). Our results therefore enrich knowledge about the recently discovered role of ADAR1 and RNA editing in regulating the innate immune system and support the connection between imbalance of RNA editing and immune dysfunction.

We hypothesize that the elevated editing may be involved in positive feedback, aggravating autoimmunity (Figure 4). Inflammatory cytokines produced by an immune response triggered by elevated editing levels may maintain or even further increase the editing levels by stimulating the IFN-induced deaminases, resulting in the production of still more, potentially immunogenic, editing-recoded peptides. In addition, the edited epitope not only may prove immunogenic but also may initiate a process of epitope spreading (Lehmann et al., 1992; Sercarz, 1998; Vanderlugt and Miller, 2002). Moreover, SLE has several characteristics that may promote the immunogenicity of RNA editing, such as the upregulation of epitope presentation pathway by IFNs and the accumulation of dead cell debris (Boehm et al., 1997; Lipsky, 2001; Rusinova et al., 2013; Schroder et al., 2004).

Although the results presented here are promising and demonstrate the potential of editing to create neo-autoantigens, there are several limitations to this type of research, leading to incomplete detection of the editing-derived proteome and assessment of its immunogenicity. First, from an analytical perspective, while the HE analysis enables the discovery of recoding editing sites, the HE reads comprise only a minority (~1%) of the total reads, and presumably a small fraction of the total editing. Moreover, the recoding sites detected here have a relatively low signal-to-noise ratio, so the credibility of

each site is not high enough, even though most of them are genuine and altogether the sites are credible. Second, from a biological perspective, several limitations arise. The RNA was extracted mainly from living cells; thus, it is possible that editing sites generating particularly immunogenic peptides may be underrepresented in the data because of their elimination by a fiercer immune reaction against such cells. Another potential drawback is introduced by the RNA-seq data used for the analysis being derived from whole blood, which contains heterogeneous cell types, which may dampen the signal and add noise. Thus, it is probable that the number of recoding sites detected here is an underestimation of the actual number of recoding editing sites in SLE patients.

In summary, we have shown significantly elevated RNA editing in SLE patients and revealed its potential to give rise to neo-autoantigens, implying a role for RNA editing in the etiology and progress of SLE. These findings provide another link between RNA editing and autoimmune diseases.

## EXPERIMENTAL PROCEDURES

The Supplemental Experimental Procedures includes more details of the applied analyses.

### Datasets

A whole-blood (Paxgene) RNA-seq dataset of 18 healthy individuals and 99 active SLE patients (GEO: GSE72509) from Hung et al. (2015) and a smaller dataset of 12 SLE patients and 4 controls (GEO: GSE80183) from Rai et al. (2016) were downloaded. Details are provided in Table S1.

The reads were aligned to the human genome (hg19) using STAR 2.4.2 (Dobin et al., 2013).

### HE

We used a recently described pipeline (Porath et al., 2014) that enables the measurement of editing of heavily edited reads, which standard schemes fail to align correctly (Carmi et al., 2011)

### AEI

To measure the editing in *Alu* elements, we used a previously described method (Bazak et al., 2014b). Due to the magnitude of the effect and the specific occurrence of editing in *Alu* elements (such as editing in clusters), this
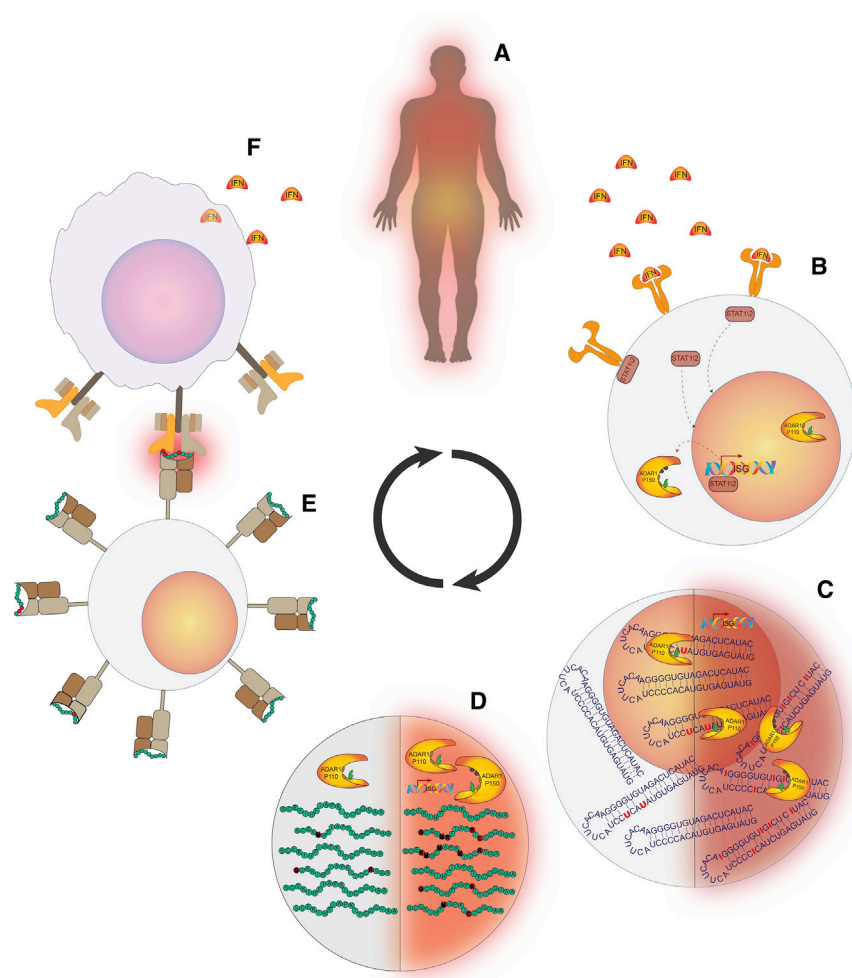
**Figure 4. Pro-inflammatory Editing Feedback**

(A) As part of an inflammatory response, interferons are released into the intercellular environment.

(B) Uptake of the interferons triggers a response via the JAK-STAT signaling pathway. Among the many genes affected, the ADAR1 p150 isoform is produced.

(C) This upregulation of ADAR1 leads to higher editing levels, due both to the increasing rates at normally edited sites and to the editing of sites that are not edited under normal conditions (left, non-inflamed cell; right, inflamed cell).

(D) Elevated editing levels lead to increased numbers and a variety of recoded proteins.

(E) Some recoded proteins might be degraded into peptides that will eventually become epitopes, resulting in presentation of recoded epitopes that are usually not found on cell surfaces.

(F) These edited epitopes might trigger an immune response (thus becoming neo-autoantigens), contributing to the inflammatory state and potentially initiating a positive feedback loop.

measurement results in a clean signal with a minimal false-positive rate (Bazak et al., 2014a).

**Known Sites**

RNA editing levels for a list of pre-known editing sites were calculated using REDIToolKnown, which is part of the REDItools package (Picardi and Pesole, 2013). The list of editing sites was compiled from previous studies (Khermesh et al., 2016; Pinto et al., 2014; Ramaswami et al., 2012).

**Expression Analysis**

The DESeq package (Anders and Huber, 2010) in R was used for analyzing differential gene expression in all control and SLE patient samples.

**Statistical Analysis**

The statistical analysis was done using R (R Project for Statistical Computing, http://www.r-project.org/). Unless otherwise specified, the statistically significant difference (5% confidence level) among two groups was tested using the Wilcoxon rank test. The tests were performed using the default parameters and in a two-sided manner. Where appropriate, p values were corrected for multiple testing.

**Recoding Editing Sites Analyses**

To provide the peptide sequence for the analyses and find recoding, HE and pre-known sites filtered for SNPs and HLA genes were annotated using wANNOVAR (Chang and Wang, 2012; Wang et al., 2010). HLA alleles for

each sample were derived using OptiType (Szolek et al., 2014). For each sample, recoding sites (detected by the HE analysis) inside peptides that have at least one window (see Supplemental Experimental Procedures) with strong affinity to the sample-predicted HLA alleles were selected according to the NetMHCPan4.0 (Jurtz et al., 2017) predictions. A recoding event was considered to have an affinity-enhancing effect if the edited version of the peptide, for all windows on average, had a higher affinity (of at least 10%). The number of epitopes per peptide chain was calculated as previously described in Agranovich et al. (2013).

## REFERENCES

Agranovich, A., Maman, Y., and Louzoun, Y. (2013). Viral proteome size and CD8+ T cell epitope density are correlated: the effect of complexity on selection. Infect. Genet. Evol. 20, 71–77.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. 11, R106.

Bass, B.L. (2002). RNA editing by adenosine deaminases that act on RNA. Annu. Rev. Biochem. 71, 817–846.

Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., Isaacs, F.J., Rechavi, G., Li, J.B., Eisenberg, E., and Levanon, E.Y. (2014a). A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. Genome Res. 24, 365–376.

Bazak, L., Levanon, E.Y., and Eisenberg, E. (2014b). Genome-wide analysis of Alu editability. Nucleic Acids Res. 42, 6876–6884.

Boehm, U., Klamp, T., Groot, M., and Howard, J.C. (1997). Cellular responses to interferon-γ. Annu. Rev. Immunol. 15, 749–795.

Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E., and Emeson, R.B. (1997). Regulation of serotonin-2C receptor G-protein coupling by RNA editing. Nature 387, 303–308.

Carmi, S., Borukhov, I., and Levanon, E.Y. (2011). Identification of widespread ultra-edited human RNAs. PLoS Genet. 7, e1002317.

Chang, X., and Wang, K. (2012). wANNOVAR: annotating genetic variants for personal genomes via the web. J. Med. Genet. 49, 433–436.

Chiche, L., Jourde-Chiche, N., Whalen, E., Presnell, S., Gersuk, V., Dang, K., Anguiano, E., Quinn, C., Burtey, S., Berland, Y., et al. (2014). Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis Rheumatol. 66, 1583–1595.

Crow, Y.J., and Manel, N. (2015). Aicardi-Goutières syndrome and the type I interferonopathies. Nat. Rev. Immunol. 15, 429–440.

Danan-Gotthold, M., Guyon, C., Giraud, M., Levanon, E.Y., and Abramson, J. (2016). Extensive RNA editing and splicing increase immune self-representation diversity in medullary thymic epithelial cells. Genome Biol. 17, 219.

Daniel, C., Venø, M.T., Ekdahl, Y., Kjems, J., and Öhman, M. (2012). A distant cis acting intronic element induces site-selective RNA editing. Nucleic Acids Res. 40, 9876–9886.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

Eckmann, C.R., Neunteufl, A., Pfaffstetter, L., and Jantsch, M.F. (2001). The human but not the Xenopus RNA-editing enzyme ADAR1 has an atypical nuclear localization signal and displays the characteristics of a shuttling protein. Mol. Biol. Cell 12, 1911–1924.

Eggington, J.M., Greene, T., and Bass, B.L. (2011). Predicting sites of ADAR editing in double-stranded RNA. Nat. Commun. 2, 319.

Ekdahl, Y., Farahani, H.S., Behm, M., Lagergren, J., and Öhman, M. (2012). A-to-I editing of microRNAs in the mammalian brain increases during development. Genome Res. 22, 1477–1487.

Furie, R., Khamashta, M., Merrill, J.T., Werth, V.P., Kalunian, K., Brohawn, P., Illei, G.G., Drappa, J., Wang, L., and Yoo, S.; CD1013 Study Investigators (2017). Anifrolumab, an Anti-Interferon-α Receptor Monoclonal Antibody, in Moderate-to-Severe Systemic Lupus Erythematosus. Arthritis Rheumatol. 69, 376–386.

George, C.X., and Samuel, C.E. (1999). Characterization of the 5′-flanking region of the human RNA-specific adenosine deaminase ADAR1 gene and identification of an interferon-inducible ADAR1 promoter. Gene 229, 203–213.

Graham, K.L., and Utz, P.J. (2005). Sources of autoantigens in systemic lupus erythematosus. Curr. Opin. Rheumatol. 17, 513–517.

Hagberg, N., and Rönnblom, L. (2015). Systemic Lupus Erythematosus–A Disease with A Dysregulated Type I Interferon System. Scand. J. Immunol. 82, 199–207.

Han, L., Diao, L., Yu, S., Xu, X., Li, J., Zhang, R., Yang, Y., Werner, H.M.J., Eterovic, A.K., Yuan, Y., et al. (2015). The Genomic Landscape and Clinical Relevance of A-to-I RNA Editing in Human Cancers. Cancer Cell 28, 515–528.

Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P.H. (2000). Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature 406, 78–81.

Ho, V., Mclean, A., and Terry, S. (2008). Severe systemic lupus erythematosus induced by antiviral treatment for hepatitis C. J. Clin. Rheumatol. 14, 166–168.

Hueber, W., Zeng, D., Strober, S., and Utz, P.J. (2004). Interferon-alpha-inducible proteins are novel autoantigens in murine lupus. Arthritis Rheum. 50, 3239–3249.

Hung, T., Pratt, G.A., Sundararaman, B., Townsend, M.J., Chaivorapol, C., Bhangale, T., Graham, R.R., Ortmann, W., Criswell, L.A., Yeo, G.W., and Behrens, T.W. (2015). The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. Science 350, 455–459.

Jurtz, V., Paul, S., Massimo, A., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J. Immunol. 199, 3360–3368.

Khermesh, K., D'Erchia, A.M., Barak, M., Annese, A., Wachtel, C., Levanon, E.Y., Picardi, E., and Eisenberg, E. (2016). Reduced levels of protein recoding by A-to-I RNA editing in Alzheimer's disease. RNA 22, 290–302.

Kirou, K.A., and Gkrouzman, E. (2013). Anti-interferon alpha treatment in SLE. Clin. Immunol. 148, 303–312.

Laxminarayana, D., Khan, I.U., and Kammer, G. (2002). Transcript mutations of the alpha regulatory subunit of protein kinase A and up-regulation of the RNA-editing gene transcript in lupus T lymphocytes. Lancet 360, 842–849.

Laxminarayana, D., O'Rourke, K.S., Maas, S., and Olorenshaw, I. (2007). Altered editing in RNA editing adenosine deaminase ADAR2 gene transcripts of systemic lupus erythematosus T lymphocytes. Immunology 121, 359–369.

Lee-Kirsch, M.A., Wolf, C., and Günther, C. (2014). Aicardi-Goutières syndrome: a model disease for systemic autoimmunity. Clin. Exp. Immunol. 175, 17–24.

Lehmann, P.V., Forsthuber, T., Miller, A., and Sercarz, E.E. (1992). Spreading of T-cell autoimmunity to cryptic determinants of an autoantigen. Nature 358, 155–157.

Liddicoat, B.J., Piskol, R., Chalk, A.M., Ramaswami, G., Higuchi, M., Hartner, J.C., Li, J.B., Seeburg, P.H., and Walkley, C.R. (2015a). RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. Science 349, 1115–1120.

Liddicoat, B.J., Chalk, A.M., and Walkley, C.R. (2015b). ADAR1, inosine and the immune sensing system: distinguishing self from non-self. Wiley Interdiscip. Rev. RNA 7, 157–172.

Lipsky, P.E. (2001). Systemic lupus erythematosus: an autoimmune disease of B cell hyperactivity. Nat. Immunol. *2*, 764–766.

Liu, Q., Jiang, L., Liu, W.-L., Kang, X.-J., Ao, Y., Sun, M., Luo, Y., Song, Y., Lo, W.H.Y., and Zhang, X. (2006). Two novel mutations and evidence for haploin-sufficiency of the ADAR gene in dyschromatosis symmetrica hereditaria. Br. J. Dermatol. *154*, 636–642.

Mannion, N.M., Greenwood, S.M., Young, R., Cox, S., Brindle, J., Read, D., Nellåker, C., Vesely, C., Ponting, C.P., McLaughlin, P.J., et al. (2014). The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. Cell Rep. *9*, 1482–1494.

Ng, B., Yang, F., Huston, D.P., Yan, Y., Yang, Y., Xiong, Z., Peterson, L.E., Wang, H., and Yang, X.-F. (2004). Increased noncanonical splicing of autoan-tigen transcripts provides the structural basis for expression of untolerized epi-topes. J. Allergy Clin. Immunol. *114*, 1463–1470.

Niewold, T.B. (2008). Interferon alpha-induced lupus: proof of principle. J. Clin. Rheumatol. *14*, 131–132.

Nishikura, K. (2010). Functions and regulation of RNA editing by ADAR deam-inases. Annu. Rev. Biochem. *79*, 321–349.

Orlowski, R.J., O'Rourke, K.S., Olorenshaw, I., Hawkins, G.A., Maas, S., and Laxminarayana, D. (2008). Altered editing in cyclic nucleotide phosphodies-terase 8A1 gene transcripts of systemic lupus erythematosus T lymphocytes. Immunology *125*, 408–419.

Paz-Yaacov, N., Bazak, L., Buchumenski, I., Porath, H.T., Danan-Gotthold, M., Knisbacher, B.A., Eisenberg, E., and Levanon, E.Y. (2015). Elevated RNA Edit-ing Activity Is a Major Contributor to Transcriptomic Diversity in Tumors. Cell Rep. *13*, 267–276.

Peng, G., Lei, K.J., Jin, W., Greenwell-Wild, T., and Wahl, S.M. (2006). Induc-tion of APOBEC3 family proteins, a defensive maneuver underlying interferon-induced anti-HIV-1 activity. J. Exp. Med. *203*, 41–46.

Pestal, K., Funk, C.C., Snyder, J.M., Price, N.D., Treuting, P.M., and Stetson, D.B. (2015). Isoforms of RNA-Editing Enzyme ADAR1 Independently Control Nucleic Acid Sensor MDA5-Driven Autoimmunity and Multi-organ Develop-ment. Immunity *43*, 933–944.

Picardi, E., and Pesole, G. (2013). REDItools: high-throughput RNA editing detection made easy. Bioinformatics *29*, 1813–1814.

Pinto, Y., Cohen, H.Y., and Levanon, E.Y. (2014). Mammalian conserved ADAR targets comprise only a small fragment of the human editosome. Genome Biol. *15*, R5.

Porath, H.T., Carmi, S., and Levanon, E.Y. (2014). A genome-wide map of hy-per-edited RNA reveals numerous new sites. Nat. Commun. *5*, 4726.

Rai, R., Chauhan, S.K., Singh, V.V., Rai, M., and Rai, G. (2016). RNA-seq Anal-ysis Reveals Unique Transcriptome Signatures in Systemic Lupus Erythema-tosus Patients with Distinct Autoantibody Specificities. PLoS ONE *11*, e0166312.

Ramaswami, G., and Li, J.B. (2014). RADAR: a rigorously annotated database of A-to-I RNA editing. Nucleic Acids Res. *42*, D109–D113.

Ramaswami, G., Lin, W., Piskol, R., Tan, M.H., Davis, C., and Li, J.B. (2012). Accurate identification of human Alu and non-Alu RNA editing sites. Nat. Methods *9*, 579–581.

Rice, G.I., Kasher, P.R., Forte, G.M.A., Mannion, N.M., Greenwood, S.M., Szynkiewicz, M., Dickerson, J.E., Bhaskar, S.S., Zampini, M., Briggs, T.A., et al. (2012). Mutations in ADAR1 cause Aicardi-Goutières syndrome associ-ated with a type I interferon signature. Nat. Genet. *44*, 1243–1248.

Riedmann, E.M., Schopoff, S., Hartner, J.C., and Jantsch, M.F. (2008). Spec-ificity of ADAR-mediated RNA editing in newly identified targets. RNA *14*, 1110–1118.

Rueter, S.M., Dawson, T.R., and Emeson, R.B. (1999). Regulation of alterna-tive splicing by RNA editing. Nature *399*, 75–80.

Rusinova, I., Forster, S., Yu, S., Kannan, A., Masse, M., Cumming, H., Chapman, R., and Hertzog, P.J. (2013). Interferome v2.0: an updated database of annotated interferon-regulated genes. Nucleic Acids Res. *41*, D1040–D1046.

Savva, Y.A., Rieder, L.E., and Reenan, R.A. (2012). The ADAR protein family. Genome Biol. *13*, 252.

Schroder, K., Hertzog, P.J., Ravasi, T., and Hume, D.A. (2004). Interferon-gamma: an overview of signals, mechanisms and functions. J. Leukoc. Biol. *75*, 163–189.

Sercarz, E.E. (1998). Immune focusing vs diversification and their connection to immune regulation. Immunol. Rev. *164*, 5–10.

Sharma, S., Patnaik, S.K., Taggart, R.T., Kannisto, E.D., Enriquez, S.M., Goll-nick, P., and Baysal, B.E. (2015). APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. Nat. Commun. *6*, 6881.

Sharma, S., Patnaik, S.K., Kemer, Z., and Baysal, B.E. (2017). Transient over-expression of exogenous APOBEC3A causes C-to-U RNA editing of thou-sands of genes. RNA Biol. *14*, 603–610.

Sherer, Y., Gorstein, A., Fritzler, M.J., and Shoenfeld, Y. (2004). Autoantibody explosion in systemic lupus erythematosus: more than 100 different antibodies found in SLE patients. Semin. Arthritis Rheum. *34*, 501–537.

Smith, H.C., Bennett, R.P., Kizilyer, A., McDougall, W.M., and Prohaska, K.M. (2012). Functions and regulation of the APOBEC family of proteins. Semin. Cell Dev. Biol. *23*, 258–268.

Speyer, J.F., Lengyel, P., Basilio, C., and Ochoa, S. (1962). Synthetic polynu-cleotides and the amino acid code. II. Proc. Natl. Acad. Sci. USA *48*, 63–68.

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics *30*, 3310–3316.

Tsokos, G.C. (2011). Systemic lupus erythematosus. N. Engl. J. Med. *365*, 2110–2121.

Vanderlugt, C.L., and Miller, S.D. (2002). Epitope spreading in immune-medi-ated diseases: implications for immunotherapy. Nat. Rev. Immunol. *2*, 85–95.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.

**Supplemental Information**

**Increased RNA Editing May Provide a Source**

**for Autoantigens in Systemic Lupus Erythematosus**

Shalom Hillel Roth, Miri Danan-Gotthold, Meirav Ben-Izhak, Gideon Rechavi, Cyrille J. Cohen, Yoram Louzoun, and Erez Y. Levanon

# Supplemental Information
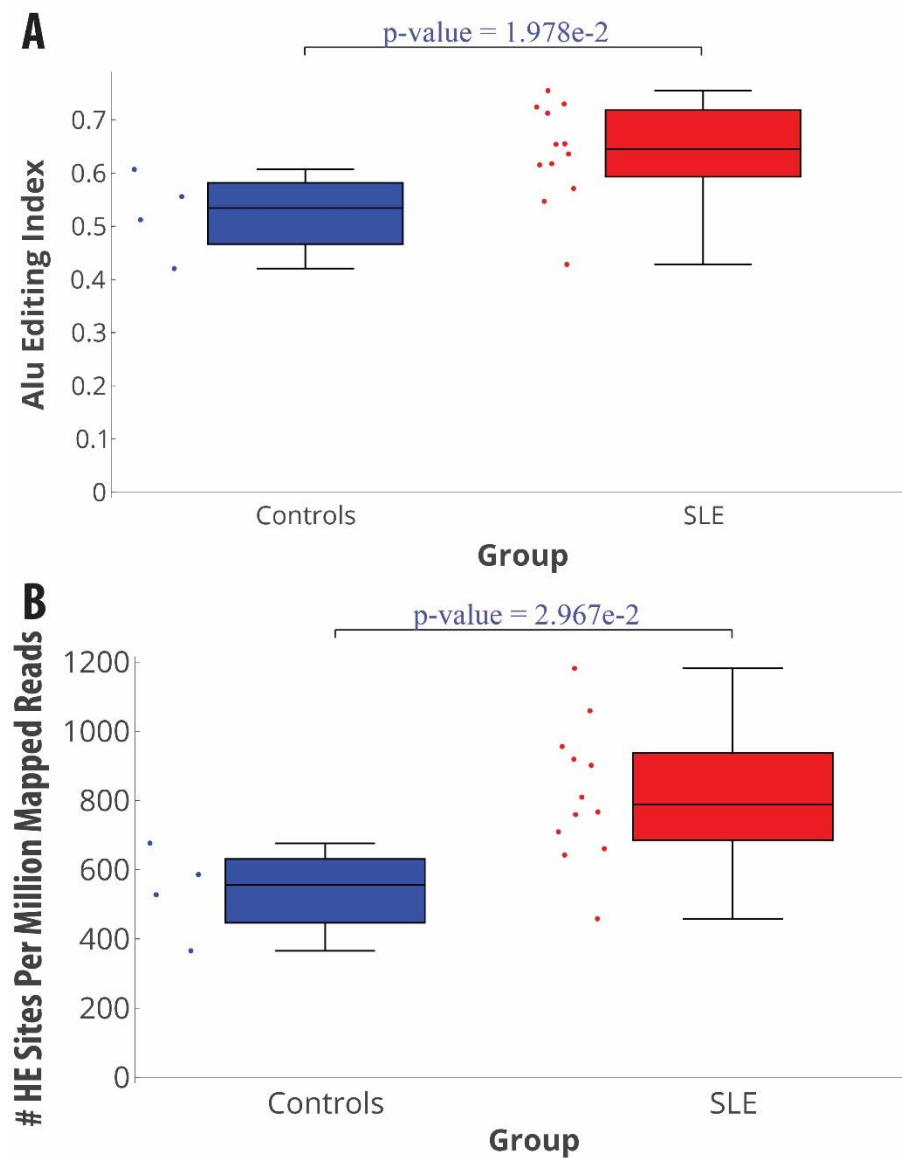
## Supplemental Figures



**Figure S1. Editing is significantly increased in SLE patients (dataset taken from (Rai et al., 2016)), Related to Figure 1.** The compared cohorts are healthy controls (blue) and SLE patients (red). Paired comparisons were evaluated using the Wilcoxon test. The global levels of editing were assessed by determining (A) the editing in *Alu* elements (*Alu* Editing Index - AEI) and (B) the number of hyper edited (HE) sites (normalized by the number of mapped reads) per sample.
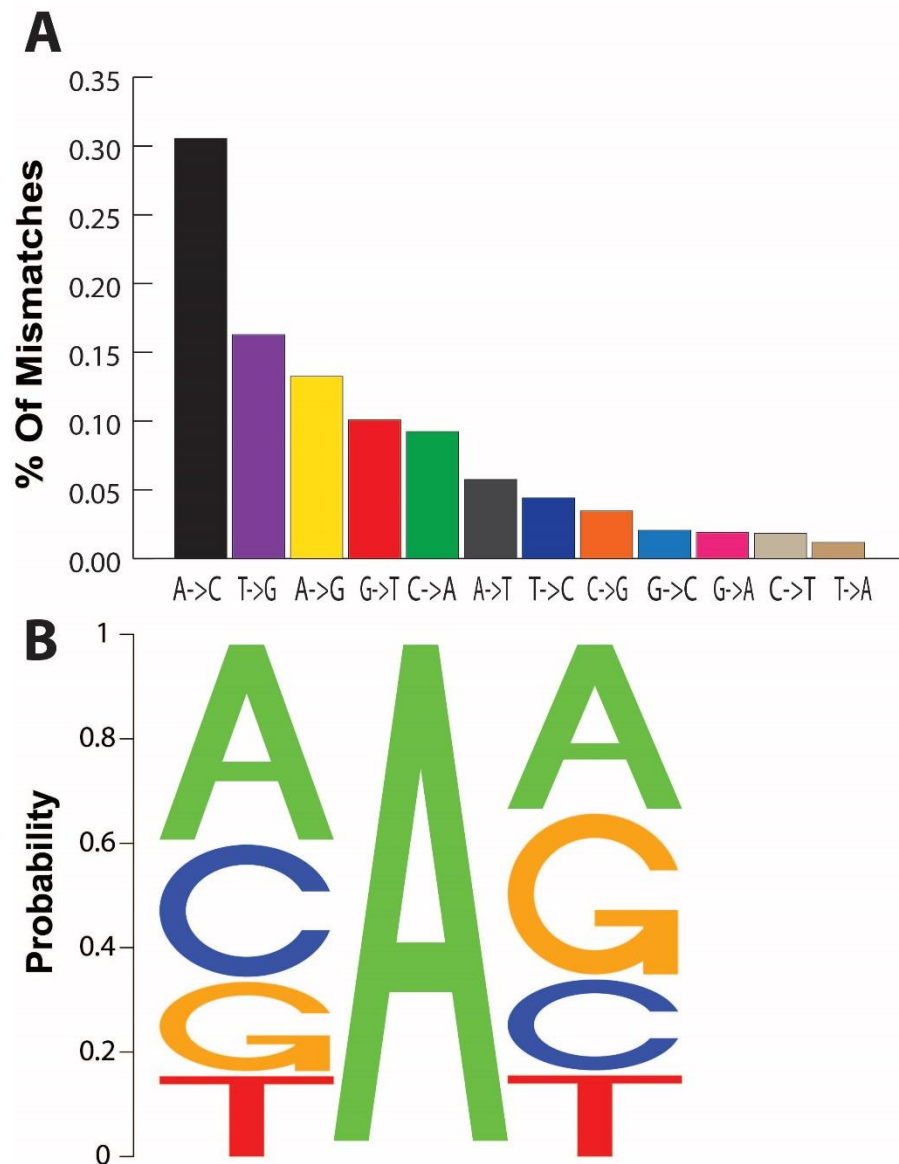
**Figure S2. The signal cleanness of recoding sites at hyper edited regions, Related to Figure 3.**
(A) Percentage of all mismatches and (B) the nucleotide frequencies of the neighboring nucleotides around the A-to-G sites

# Supplemental Tables

**Table S2. Elevated APOBEC proteins levels in SLE, Related to Figure 2.**

| Gene name | Fold Change between ISM-high patients and controls | Wilcoxon p-value |
|-----------|---------------------------------------------------|------------------|
| *APOBEC3A* | 2.58 | 5.25e-08 |
| *APOBEC3B* | 2.48 | 0.00237 |
| *APOBEC3F* | 1.80 | 1.96e-08 |
| APOBEC3G | 1.48 | 4.28e-08 |
| *APOBEC3H* | 1.84 | 0.000885 |

# Supplemental Experimental Procedures

**Reads alignment**

The reads were aligned to the human genome (hg19) using STAR 2.4.2 (Dobin et al., 2013), using alignment parameters that limit the number of mismatches to 0.1 of the mapped length and accept only uniquely aligned reads (default parameters except alignIntronMax - 1000000, alignMatesGapMax - 1000000, outFilterMismatchNmax - 999, outFilterMismatchNoverReadLmax - 0.1, outFilterMultimapNmax - 1, outReadsUnmapped - Fastx, quantMode – GeneCounts). This resulted in three output files for each sample (an alignment file, a fastq file containing all the unmapped reads, and a file containing the number of reads per gene) that were used in the next analyses.

**Hyper Editing**

RNA editing by ADAR1 usually happens in clusters. Thus, many heavily edited reads may differ from the DNA to the extent that standard schemes fail to align them correctly (Carmi et al., 2011). A recently described pipeline enables the measurement of such heavily edited reads (Porath et al., 2014). It was observed  that HE is closely correlated with ADAR1 activity (Paz-Yaacov et al., 2015; Porath et al., 2014).  In this approach, the reference genome and the reads are transformed, changing all As to Gs in order to map the reads.

We extracted HE levels for every sample (this reflects the general level of editing in the sample (Porath et al., 2014)), as well as a list of the sites at which HE was detected. As input to the pipeline we used only reads that were not mapped to the genome (typically 10%-15% of reads per sample). We used the suggested parameters and Hg19 genome, except for the fraction of A-to-G out of all mismatches, which was set to 0.8. The numbers of hyper-edited sites were normalized by the number of the mapped reads in each sample. Two samples in which we identified a very high rate of clustered A-to-C substitutions (more than 70% of all clustered substitutions found in the sample) were excluded from this analysis and from all further calculations using sites detected by it, since A-to-C clustered substitutions were previously shown to result from an Illumina machine artifact (Dohm et al., 2008).

***Alu* editing index**

As most of the editing sites in the human genome (over 90%) are concentrated in *Alu* genomic regions, AEI gives a good approximation of the global editing level and ADAR1 editing efficiency in the sample (Bazak et al., 2014a).

Reads that were aligned to *Alu* elements were collected and mismatches between them and the reference genome (hg19) were identified. To get a reliable collection of mismatches, mismatches with a Phred score lower than 30 or those reported as SNPs in dbSNP (SNP build 135) were filtered out (Eisenberg et al., 2005). To eliminate putative sequence errors the results were further filtered using a probabilistic model assuming an a priori sequencing error rate of 0.001 (corresponding to Phred score of 30). This was achieved by applying Benjamini-Hochberg multiple testing correction for all *Alu* adenosines, with a false discovery rate (FDR) of 0.05. Over 90% of all *Alu* evaluated had A-to-G mismatch dominance (i.e. that the relative part of A-to-G mismatches is higher than all others combined), which further supports the cleanness of the signal. However, for the calculations of the editing levels we used all *Alu* elements (to avoid biases caused by choosing dominant A-to-G *Alu*s only) as described below .

A straight forward comparison of editing levels, based on the editing rates at all sites, is quite complicated and requires ultra-high coverage due to the fact that most adenosines in *Alu* elements are edited to at least some extent (though usually to a low degree, less than 1% of adenosines) (Bazak et al., 2014a, 2014b). Hence, any set of detected sites in a given sample may represent only a small random fraction of the editing sites. Thus, the number of reads supporting a specific mismatch is not a good measure of the effect in the whole sample.

To apply all the above to our data set, we used the AEI (*Alu* Editing Index). This index gives the average editing level across all adenosines in *Alu* elements weighted by their expression, which is the ratio of A-to-G mismatches (presumably due to deamination to inosines), over the total number of nucleotides aligned to an *Alu* repeat (both adenosines and presumed inosines). This index averages over millions of adenosines, which as previously described, (Bazak et al., 2014b) makes it rather robust to statistical noise.

**Recoding Editing Sites Analyses**

The examination of editing rates at non-synonymous, pre-known editing sites was limited to 60 sites, which were expressed by at least 10 samples in each group, so that n≥10 for each group of the Wilcoxon test.

Protein prediction was assessed using a list of edited sites filtered from known SNPs (dbSNP135 (Sherry et al., 2001) and highly polymorphic HLA genes, but with conserved editing sites (Pinto et al., 2014) included). The amino acid sequence (as found in the NCBI database) was queried and the "mutated" AA was planted in the right place of the sequence generating peptides of 21 AA long (where the edited AA is in the middle). The affinity of each peptide in the couple (edited and non-edited form) for the predicted HLA alleles of the samples were evaluated using NetMHCPan 4.0 (Jurtz et al., 2017) for K-mers (window) of between 8-11 AA long.

**Expression analysis**
The gene annotations for the expression analysis were obtained from the UCSC gene tables (Karolchik et al., 2003). The number of reads aligned to each gene was calculated by STAR.

**Number of Epitopes Prediction**
The number of epitopes per peptide chain was calculated by combining three algorithms: a proteasomal cleavage algorithm (Ginodi et al., 2008), a TAP binding algorithm (Peters et al., 2003), and the MLVO MHC binding algorithm (Vider-Shalit and Louzoun, 2011) as previously described in (Agranovich et al., 2013).

Since the used algorithms only apply to the 39 most common alleles in Caucasian population, the epitopes were computed for the 35 HLA alleles of the relevant samples. The results were weighted according to the allele frequency in the global human population. The efficiency of these algorithms and their quality has been systematically validated (previously) vs. epitope databases and was found to induce low false positive (FP) and false negative (FN) error rates (Maman et al., 2011a, 2011b). The number of epitopes was calculated for both the sequence and the average of three randomly scrambled versions of it, in order to assess whether the number of epitopes is affected by the sequence of the peptide or the AA composition (a biochemical effect).

**Supplemental References**

Agranovich, A., Maman, Y., and Louzoun, Y. (2013). Viral proteome size and CD8+ T cell epitope density are correlated: The effect of complexity on selection. Infect. Genet. Evol. *20*, 71–77.

Bazak, L., Levanon, E.Y., and Eisenberg, E. (2014a). Genome-wide analysis of Alu editability. Nucleic Acids Res. *42*, 6876–6884.

Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., Isaacs, F.J., Rechavi, G., Li, J.B., Eisenberg, E., et al. (2014b). A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. Genome Res. *24*, 365–376.

Boehm, U., Klamp, T., Groot, M., and Howard, J.C. (1997). CELLULAR RESPONSES TO INTERFERON-γ. Annu. Rev. Immunol. *15*, 749–795.

Carmi, S., Borukhov, I., and Levanon, E.Y. (2011). Identification of widespread ultra-edited human RNAs. PLoS Genet. *7*, e1002317.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dohm, J.C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. *36*, e105.

Eisenberg, E., Adamsky, K., Cohen, L., Amariglio, N., Hirshberg, A., Rechavi, G., and Levanon, E.Y. (2005). Identification of RNA editing sites in the SNP database. Nucleic Acids Res. *33*, 4612–4617.

Ferrington, D.A., and Gregerson, D.S. (2012). Immunoproteasomes: structure, function, and antigen presentation. Prog. Mol. Biol. Transl. Sci. *109*, 75–112.

Ginodi, I., Vider-Shalit, T., Tsaban, L., and Louzoun, Y. (2008). Precise score for the prediction of peptides cleaved by the proteasome. Bioinformatics *24*, 477–483.

Gommans, W.M., Tatalias, N.E., Sie, C.P., Dupuis, D., Vendetti, N., Smith, L., Kaushal, R., and Maas, S. (2008). Screening of human SNP database identifies recoding sites of A-to-I RNA editing. RNA *14*, 2074–2085.

Groettrup, M., Kirk, C.J., and Basler, M. (2010). Proteasomes in immune cells: more than peptide producers? Nat. Rev. Immunol. *10*, 73–78.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J. Immunol. *199*, 3360–3368.

Karolchik, D., Baertsch, R., Diekhans, M., Furey, T.S., Hinrichs, A., Lu, Y.T., Roskin, K.M., Schwartz, M., Sugnet, C.W., Thomas, D.J., et al. (2003). The UCSC Genome Browser Database. Nucleic Acids Res. *31*, 51–54.

Lipsky, P.E. (2001). Systemic lupus erythematosus: an autoimmune disease of B cell hyperactivity. Nat. Immunol. *2*, 764–766.

Maman, Y., Blancher, A., Benichou, J., Yablonka, A., Efroni, S., and Louzoun, Y. (2011a). Immune-Induced Evolutionary Selection Focused on a Single Reading Frame in Overlapping Hepatitis B Virus Proteins. J. Virol. *85*, 4558–4566.

Maman, Y., Nir-Paz, R., and Louzoun, Y. (2011b). Bacteria modulate the CD8+ T cell epitope repertoire of host cytosol-exposed proteins to manipulate the host immune response. PLoS Comput. Biol. *7*, e1002220.

Paz-Yaacov, N., Bazak, L., Buchumenski, I., Porath, H.T., Danan-Gotthold, M., Knisbacher, B.A., Eisenberg, E., and Levanon, E.Y. (2015). Elevated RNA Editing Activity Is a Major Contributor to Transcriptomic Diversity in Tumors. Cell Rep. *13*, 267–276.

Peters, B., Bulik, S., Tampe, R., van Endert, P.M., and Holzhutter, H.-G. (2003). Identifying MHC Class I Epitopes by Predicting the TAP Transport Efficiency of Epitope Precursors. J. Immunol. *171*, 1741–1749.

Pinto, Y., Cohen, H.Y., and Levanon, E.Y. (2014). Mammalian conserved ADAR targets comprise only a small fragment of the human editosome. Genome Biol. *15*, R5.

Porath, H.T., Carmi, S., and Levanon, E.Y. (2014). A genome-wide map of hyper-edited RNA reveals numerous new sites. Nat. Commun. *5*, 4726.

Rai, R., Chauhan, S.K., Singh, V.V., Rai, M., and Rai, G. (2016). RNA-seq Analysis Reveals Unique Transcriptome Signatures in Systemic Lupus Erythematosus Patients with Distinct Autoantibody Specificities. PLoS One *11*, e0166312.

Reichenberger, E.J., Levine, M.A., Olsen, B.R., Papadaki, M.E., Lietman, S.A., Ohnishi, T., Matsuguchi, T., Shimokawa, H., Ohya, K., Amagasa, T., et al. (2012). The role of SH3BP2 in the pathophysiology of cherubism. Orphanet J. Rare Dis. *7*, S5.

Rusinova, I., Forster, S., Yu, S., Kannan, A., Masse, M., Cumming, H., Chapman, R., and Hertzog, P.J. (2013). INTERFEROME v2.0: an updated database of annotated interferon-regulated genes. Nucleic Acids Res. *41*, D1040–D1046.

Schroder, K., Hertzog, P.J., Ravasi, T., and Hume, D.A. (2004). Interferon-γ : an overview of signals, mechnisms and functions. J. Leukozyte Biol. *75*.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. *29*, 308–311.

Vider-Shalit, T., and Louzoun, Y. (2011). MHC-I prediction using a combination of T cell epitopes and MHC-I binding peptides. J. Immunol. Methods *374*, 43–46.