ACR Open Rheumatology
Vol. 7, No. 1, January 2025, e11770
DOI 10.1002/acr2.11770

American College of Rheumatology
Empowering Rheumatology Professionals

Altered Protein Structures and Neoepitopes in Lupus Neutrophils From Dysregulated Splicing of Messenger RNA

Rayan Najjar,[1] Xiaoxing Wang,[1] Jose Mario Bello Pineda,[1] Hugh Alessi,[1] Alison Bays,[1] Robert K. Bradley,[2] James N. Jarvis,[1] and Tomas Mustelin[1]

Objective. To test whether messenger RNA (mRNA) splicing is altered in neutrophils from patients with systemic lupus erythematosus (SLE) and can produce neoantigens.

Methods. RNA sequencing of neutrophils from patients with SLE (n = 15) and healthy donors (n = 12) were analyzed for mRNA splicing using the RiboSplitter pipeline, an event-focused tool based on SplAdder with subsequent translation and protein domain annotation. RNA sequencing from SARS-CoV2–infected individuals was used as an additional comparator.

Results. Neutrophils from patients with SLE contained 521 statistically significant altered mRNA splicing events compared with healthy donor neutrophils, many of them affecting important immunologic pathways, myeloid function, transcription factors, and proteins involved in mRNA splicing. A subset of splicing events were only present in SLE samples, and some of them occurred at unannotated splice acceptor or donor sites. Two patients were particularly rich in such events. Only a small number of dysregulated splicing events were more pronounced in patients with active disease or with high type I interferons but were not detected in SARS-CoV2–infected individuals, who also had high type I interferons. Besides causing a range of structural changes, 80 mRNA splice variants exclusive to SLE were predicted to translate into novel amino acid sequences. Peptides derived from these novel amino acid sequences were predicted to bind to the individual patients' class I and II major histocompatibility complex molecules with high affinity.

Conclusion. We conclude that aberrant mRNA splicing in SLE has the potential to affect both the function of granulocytes and to generate novel autoantigens.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a complex and heterogenous autoimmune disease of unknown etiology. In addition to the involvement of antigen-specific adaptive immunity, there is growing evidence that innate immune cells also are dysfunctional and contribute to the pathogenesis of SLE. Neutrophilic granulocytes display altered oxidative metabolism[1,2] and die by inflammatory programmed cell death pathways, including ferroptosis[3] and NETosis,[4–6] contributing to the presence of extracellular DNA, circulating immune complexes, and elevated expression of type I interferon-inducible genes.[4,6] SLE neutrophils display hypomethylation of their genomes,[7] and a neutrophil transcriptional signature in blood correlates with disease progression.[4] Neutrophils produce type I interferons[8] and express increased levels of interferon-inducible genes.[9] These findings underscore the contribution of neutrophils to SLE pathology, warranting further exploration. Research into the transcriptional biology of neutrophils has been limited, in part due to technical challenges working with neutrophils because of their short lifespan ex vivo and high levels of proteases and nucleases, as well as a previous assumption that terminally differentiated neutrophils are transcriptionally silent.

Alternative splicing of primary RNA transcripts to form mature messenger RNA (mRNA) for translation contributes greatly to proteome diversity. Different protein isoforms generated from a given gene can differ in stability, function, and interactions with other proteins or with nucleic acids. Alternative splicing can also affect mRNA fate, including turnover via nonsense-mediated decay.[10] In the context of autoimmune disease, alternative splicing has been reported to be more prevalent among autoantigens

Supported by the NIH (grant AR-007108 to Dr Najjar and grants AI188952, AR075134, AR074939, and AR081654 to T.M., AI-186337, and AI-183320 to Dr Mustelin).

[1]Rayan Najjar, MD, MPH, Xiaoxing Wang, PhD, Jose Mario Bello Pineda, MS, Hugh Alessi, BA, Alison Bays, MD, MPH, James N. Jarvis, MD, Tomas Mustelin, MD, PhD: University of Washington, Seattle; [2]Robert K. Bradley, PhD: University of Washington, Seattle, and Fred Hutchinson Cancer Research Center, Seattle, Washington.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr2.11770).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr2.11770.

Address correspondence via email to Tomas Mustelin, XX, at tomas2@uw.edu.

Submitted for publication January 20, 2024; accepted in revised form June 5, 2024.

compared with randomly selected proteins (100% vs 42%, respectively) with 80% of autoantigens having noncanonical transcripts (compared with <1% in controls).[11] This supports the notion that abnormal splicing could create proteins with novel sequences, for example, due to frameshifts or the translation of intronic sequences, potentially creating neoepitopes for which immunologic tolerance is weak or absent. This phenomenon is well recognized in immunogenic cancers.[12,13] Abnormal splicing can also result in proteins that cause cellular dysfunction and disease. As an example, a *FOXP3* transcript with exon 2 missing (ie, exon skip) causes regulatory T cell dysfunction leading to the severe immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome.[14] The exon 2–deficient isoform is also present in other autoimmune conditions, such as anti–neutrophil cytoplasmic antigen–associated vasculitis[15] and giant cell arteritis.[16] Other risk variants located in noncoding regions[17,18] found in genome-wide association studies of autoimmune diseases (eg, in SLE) may similarly affect alternative splicing in detrimental ways. Indeed, altered alternative splicing events have been reported in SLE,[19] but most of these studies were not comprehensive and did not include neutrophils.

PATIENTS AND METHODS

Patients with SLE. Freshly drawn blood was obtained from adult (n = 15) patients with SLE and healthy age-matched individuals (n = 12) recruited through the University of Washington, Division of Rheumatology Biorepository to participate in research studies at the University of Washington Medical Center, Seattle, WA. The study was approved by the institutional ethics board (STUDY00006196), and informed written consent was obtained from all participants according to the Declaration of Helsinki.

All data used in the analysis will be made available for research use. We are also committing to deposit our RNA-sequencing (RNA-seq) data set in a public database upon acceptance of our paper. The COVID-19 data we used are publicly available with accession (GSE261866). RiboSplitter is publicly available (github.com/R-Najjar/RiboSplitter).

Isolation of neutrophils and peripheral blood mononuclear cells. Polymorphonuclear (PMN) cells were isolated from freshly drawn venous blood (within one hour of venipuncture) by gradient centrifugation on PolymorphPrep according to the manufacturer's instructions. The PMN fraction was >95% $CD66b^+$ and >85% $CD15^+$. The cells were washed and suspended in Hanks' buffered salt solution at 10^7/mL.

RNA isolation and RNA-seq. Extraction of RNA was performed by using a hybrid protocol of Trizol Reagent (Invitrogen cat. no. 15596026) and RNeasy Micro Kit (Qiagen cat. no. 74004). Cells were lysed and homogenized with Trizol Reagent, followed by addition of chloroform to separate DNA and proteins from the RNA, to which β-mercaptoethanol was added for additional RNase inhibition. The RNA was precipitated with 70% ethanol, passed through an RNeasy MinElute spin column, and treated with DNase I on the column. The spin column was subsequently washed extensively with the kit-supplied wash buffers, followed by additional conditioning with 70% ethanol before eluting the RNA with RNase-free water. RNA was sent to the UW Northwest Genomics core for strand-specific paired-end RNA-seq at 50 million 101-bp reads per sample.

Alignment to reference genome and alternative splicing detection. After quality control (QC) of sequencing quality, FASTQ files were aligned to the reference genome (GRCh38.p13) using STAR (version 2.7.9a)[20] with two-pass mode, which improves detection of read mapping to novel splicing junctions. After postalignment QC, sorted and indexed BAM files were used to identify alternative splicing events by SplAdder (version 3.0.3),[21] which uses alignment and annotation data to create splicing graphs that are enriched with splicing events based on the provided RNA-seq data. Splicing events can then be quantified. SplAdder detects both annotated and novel splicing events and is capable of detecting six types of events: exon skip, alternative 3′/5′ splice site, intron retention, mutually exclusive exons, and multiple exon skip. SplAdder filters input alignment reads based on their quality and allows four levels of confidence levels; we ran the software with the highest confidence setting. Additionally, SplAdder uses event-specific criteria to confirm splicing events based on genomic coordinates and coverage; we only included confirmed events in our analysis.

Hypothesis testing and prediction of protein changes. We used our established methodology (RiboSplitter)[22] to identify statistically different alternative splicing, and to predict and visualize their protein changes. The frequencies of spliced reads supporting isoforms 1 and 2 will be referred to as "isoform counts." We consistently calculated percent spliced-in (PSI) values for isoform 2 so that a higher PSI always indicates a higher proportion of isoform 2, which makes it easier to interpret which group has more of which isoform. Events with more than three samples with missing PSI per group (indicating low coverage) were excluded. Additionally, we excluded splicing events with very low variability in overall PSI (SD ≤0.05) that are unlikely to be biologically or statistically significant.

We adjusted *P* values for multiple comparisons and used an alpha = 0.05 as cutoff for statistical significance. A subset of statistically significant events were considered more likely to be biologically meaningful when healthy neutrophils expressed a high percent of one isoform (median ≥95% of one isoform) and SLE neutrophils expressed more of the other isoform, when the difference in average PSI per group was high (≥0.2), or when SLE neutrophils had a higher PSI range than healthy neutrophils.

Subgroup hypothesis testing was performed for high-activity (Systemic Lupus Erythematosus Disease Activity Index [SLEDAI] score ≥6) versus low-activity (SLEDAI score <6) SLE and high versus low interferon signature subgroups.

Gene expression. We quantified gene expression using Rsubread featurecounts (version 2.12.3).[23] Next, we identified differentially expressed genes between SLE and healthy samples and calculated gene expression normalized counts using DESeq2 (version 1.36.0).[24] The average expression of 111 interferon-induced genes (ISGs)[25] was calculated per sample, and SLE samples were considered to have high interferon signatures when the average expression of ISGs was higher than the 95th percentile of ISG expression in healthy control samples.

Comparison with COVID-19 data. SLE is an autoimmune disease with aberrant activation of the immune system and overexpression of ISGs (the interferon signature) that is similar to what is seen in viral infections. To identify alternative splicing events that are more specific to SLE, rather than general immune activation, we applied our analysis pipeline for splicing event detection in COVID-19 RNA-seq data of whole blood samples, with the intent of subtracting shared events to arrive at more specific alternations in SLE. The COVID-19 data are publicly available from the Gene Expression Omnibus (GSE196822). We included moderate cases of COVID-19 (n = 10) and healthy controls (n = 9) because we are interested in the early viral infection period when interferons are rising. The data were paired-end stranded with ∼15 to 30 million reads per sample. We used two methodologies for comparing alternative splicing events of SLE and COVID-19. First, we ran our analysis pipeline on the COVID-19 data and identified statistically different alternative splicing, then compared these events with the SLE events looking for shared events. A second approach was to start with the SLE events that we identified and investigate whether they existed in the COVID-19 events, whether statistically significant or not.

Neoantigen prediction. Alternative splicing can create new peptides that are not expressed in healthy people, and some of these peptides can become neoantigens if they have high-affinity binding to the major histocompatibility complex (MHC). We searched for splicing events in which all healthy control neutrophils exclusively expressed a single isoform with at least some SLE neutrophils expressing the other isoform. We identified each patient with SLE's class I and class II MHC haplotypes from RNA-seq data using arcasHLA (version 0.5.0), which was shown to be highly accurate (100% accurate for class I MHC and >99.7% accurate for class II MHC).[26]

MHC affinity binding was predicted for peptides unique to SLE using NetMHCpan-4.1 and NetMHCIIpan-4.0,[27] which use neural networks trained on binding affinity as well as mass spectrometry data. We excluded any core peptides with exact matching to known proteins of the gene. Strong affinity binding was defined as percent rank of predicted affinity of 0.5% or lower (compared with a set of 100,000 random natural peptides).

Outlier analysis. SLE is a heterogenous disease, which means that individual patients may have different molecular biology than other patients and healthy people. To explore this aspect in alternative splicing, we conducted an outlier analysis in which we identified a single PSI that was ≥0.3 percentage points higher or lower than the next highest or lowest PSI value, respectively. Additionally, we required an interquartile range of ≤0.1 (therefore, all outliers are at least 3× IQR away from the rest of the data), and finally, we required that the count of reads supporting the outlier sample be ≥20. We then used our analysis pipeline to predict and visualize relative protein changes among the isoforms.

Cryptic splice sites discovery. We searched for splicing events that were present in SLE neutrophils and not in those from healthy controls. We used strict criteria in which not a single read supporting isoforms 1 or 2 was present in any healthy control sample and at least 10 reads supporting isoforms 1 or 2 were present in at least one SLE sample. Additionally, only events categorized as having novel introns in both isoforms 1 and 2 were included. To further test the specificity of these events to SLE, we looked for them in the COVID-19 data to exclude shared events to arrive at a more specific set of events for SLE. Protein changes were predicted using our bioinformatics pipeline, and gene expression for these events was quantified in SLE and COVID-19.

RESULTS

Alternative splicing events in SLE compared with healthy donor neutrophils. We undertook an unbiased systematic search for SLE-specific alternative splicing events in neutrophils freshly isolated from patients with SLE (n = 15) and healthy volunteers (n = 12). Seven of the patients had active disease, defined as having an SLEDAI score ≥6, whereas eight had inactive disease with an SLEDAI score <6. Using the analytical pipeline RiboSplitter,[22] a total of 65,730 high-confidence alternative splicing events in 5,697 protein-coding genes were identified in SLE and healthy control neutrophils (median 5, range 1–523 per gene). Although the majority of events were similar among samples, variability was greater among patients with SLE compared with healthy controls (13.9% vs 10.6% had an SD >0.1 in PSI, χ^2 P value = 2.2×10^{-16}).

Statistical testing with a beta binomial model and adjusting for multiple comparisons identified 521 high-confidence alternative mRNA splicing events (Figure 1A) in 356 genes that were significantly different between SLE and healthy control neutrophils. To focus on events that are more likely to be biologically relevant, we selected events predicted to cause larger differences in

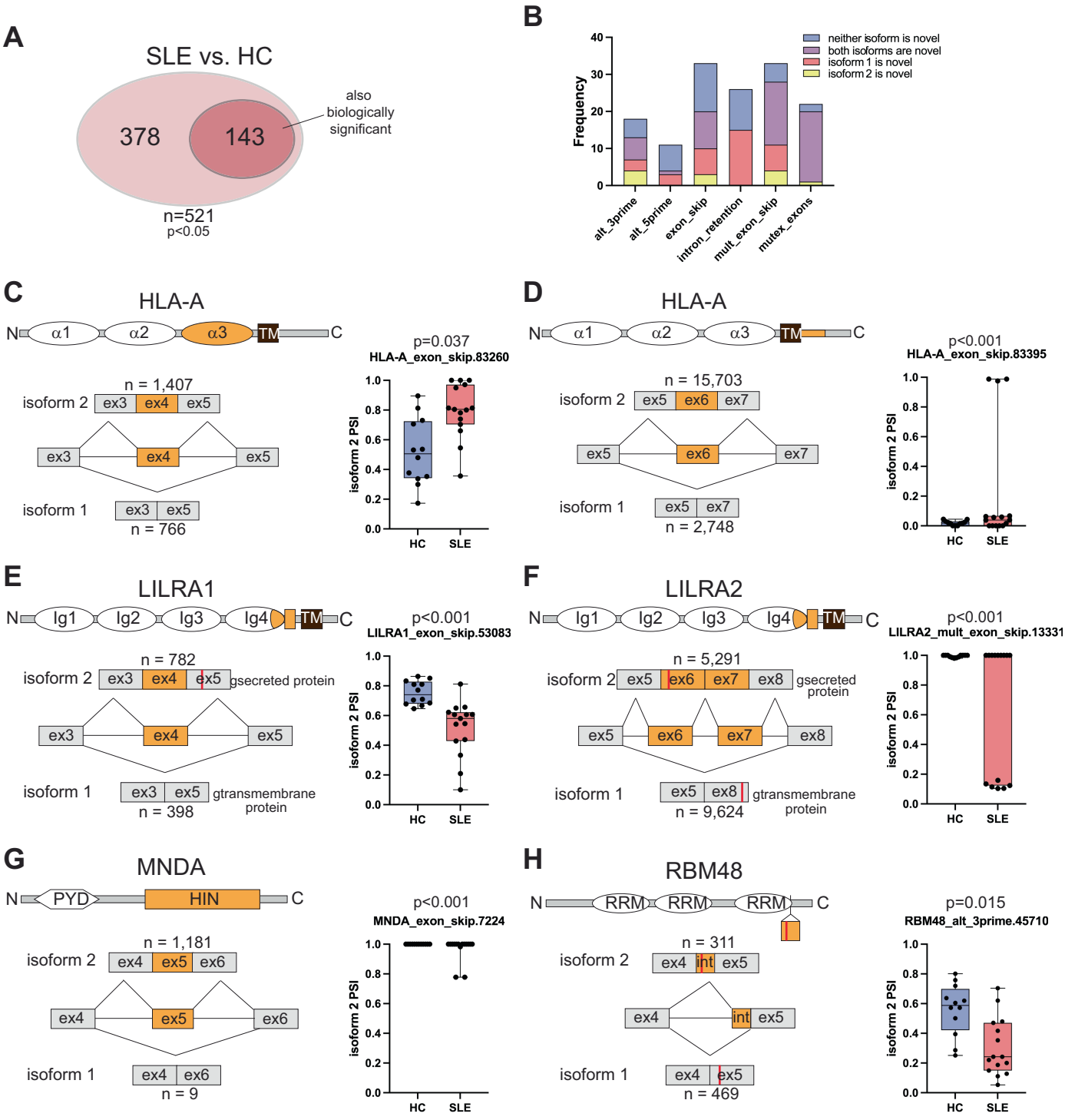

Figure 1. Representative altered alternative splicing events in SLE neutrophils. (A) Venn diagram of the number of statistically significant altered alternative splicing events and the smaller subset more likely to have biologic relevance is given. (B) Frequency of the six types of alternative splicing and their production of novel (unannotated) isoforms is shown. Examples are given of altered splice events in (C and D) *HLA-A*, (E) *LILRA1*, (F) *LILRA2*, (G) *MNDA*, and (H) *RBM48*. Statistical significance is shown for each event, and read counts (n) supporting isoform 1 and isoform 2 are indicated. Stop codons are shown as red lines. α, major histocompatibility α-domain; alt, alternative; ex, exon; HC, healthy control; HIN, HIN200 domain; IgC, immunoglobulin constant domain; int, intron; mult, multiple; mutex, mutually exclusive; PYD, pyrin domain; RRM, RNA-recognition motif; SLE, systemic lupus erythematosus; TM, transmembrane helix.

protein isoform distribution. This yielded 143 alternative splicing events in 105 genes (Figure 1A, Supplemental Table S1, and Supplemental Figure S1A–E). Exon skips and multiple exon skips were the most common events with 33 events (23.1%) each, followed by intron retention (n = 26, 18.2%), mutually exclusive exons (n = 22, 15.4%), and alternative 5′ or 3′ splice events (Figure 1B). Approximately a third of these events produced already annotated isoforms, but 53 events produced 2 novel (ie, not currently annotated) isoforms and 47 events 1 novel isoform.

Alterations in selected proteins. Collectively, these altered splicing events had a range of consequences for the translated proteins, from minimal and likely inconsequential to more dramatic truncations via the introduction of premature stop codons or the loss of domains. Many predicted translation products contained novel amino acid sequences encoded by retained introns or due to frameshifts by exon skips or altered 5′ or 3′ splice sites. Examples of splicing events altered in patients with SLE compared with healthy controls are illustrated in Figure 1 (all of them are shown in Supplemental Figure S1A–E with protein change annotations). They occurred in transcripts from genes encoding several immunologically important surface receptors, such as *HLA-A* (Figure 1C and D), *LILRA1* (Figure 1E), *LILRA2* (Figure 1F), *FCGR1A*, *FCGRT*, *CLEC7A*, and *TLR2*; proteins involved in receptor signaling, kinases, phosphatases, neutrophil activation, intracellular traffic, cytoskeleton and cell motility, myeloid-specific differentiation (eg, *MNDA*; Figure 1G); and in RNA splice factors such as those encoded by *RBM48* (Figure 1H) and *RBM10*. Read counts covering the exons and splice junctions are shown in the panels of Figure 1. In some cases, both isoforms were present in both healthy and SLE neutrophils, but in different ratios. In other cases, patients with SLE contained isoforms not seen in any healthy samples, eg, the exon 6 skip in *HLA-A* (Figure 1D) and the skip of two exons in *LILRA2* (Figure 1F), a receptor for cleaved immunoglobulin.[28] In *HLA-A*, exon 6 encodes a membrane-proximal intracellular segment, which was skipped in all healthy samples but retained in 3 SLE samples with high isoform counts but similar overall *HLA-A* expression (log2 fold change 0.3, $P = 0.312$). The multiple exon skip in *LILRA2* introduces a frameshift and an early stop codon, resulting in a secreted form of LILRA2. Healthy donor neutrophils contained only this isoform, while SLE neutrophils also produced the transmembrane receptor form. Expression of LILRA1 in neutrophils was confirmed by robust read counts (Supplemental Figure S2).

Outlier analysis. For an additional view into the molecular heterogeneity in SLE, we identified outliers—ie, events exclusive to individual patients—and found 56 such events (Figure 2A). All 56 are shown in Supplemental Figure S3A–C. A biologically relevant example is a multiple exon skip event in *B2M* encoding β2-microglobulin, in which a single patient with SLE produced a shorter isoform (isoform 1) predicted to result in a loss of the C-terminal end of the protein, followed by a novel sequence and a stop codon in the normally not translated exon 5 (the canonical open reading frame ends in exon 4) (Figure 2B). This altered structure of β2-microglobulin would compromise class I MHC surface expression, recognition of antigen by CD8 T cells, and the functions of CD1 and the neonatal Fc receptor. The presence of outlier splicing events was skewed among patients (Figure 2C), with 15 found in one patient with SLE with low disease activity and low ISG signature, including a multiple exon skip event in *IL1R2*, in which the first retained exon in isoform 2 introduces a stop codon, as well as multiple mutually exclusive events in the *FCGRT* gene (Supplemental Figure S3A). Another patient with active disease and high ISG signature had 10 outlier events (Figure 2C and Supplemental Figure S3B). These findings suggest that aberrant splicing can be unique to individual patients and appears to be enriched in a subset of patients with a more extensive mis-splicing phenotype.

Cryptic splice events. We identified unannotated splice events exclusive to SLE, which we refer to as "cryptic," and found 30 such splicing events in SLE neutrophils, but none in those from healthy donors. One cryptic event was also present in the COVID-19 data (see below). The remaining 29 cryptic splice events (Figure 2D and Supplemental Figure S4) were present in different frequencies among the patients, one patient having 11 of them (Figure 2E). Interestingly, nearly all of these events occurred in ISGs (eg, *IFI44*, *IFI44L*, *OAS1*, *OAS3*, *RSAD2*, *MX1*, and *LY6E*) and in *C5AR1*, which encodes for the receptor for complement C5a and is an important chemotactic and priming receptor for neutrophils. Importantly, the transcripts in which these cryptic splice events occurred in SLE neutrophils were also well expressed in healthy donor samples, albeit at lower levels and devoid of the cryptic splicing events. These findings further underscore the dysregulation of primary transcript processing in SLE. Together with the outlier splice events, the cryptic splice events add more individual patient variability to the alterations in multiple proteins and the production of neoepitopes in this disease. The presence of cryptic alternative splicing events correlated with average ISG expression, whereas outlier events correlated with disease duration (Supplemental Figure S5).

Altered splicing events in patients with active disease or with high type I interferons. To investigate whether dysregulated mRNA splicing is associated with active disease (akin to complement C3 and C4 consumption) or a feature of SLE disease regardless of its activity state (as are many autoantibodies), we segregated the patients with SLE into those with active (n = 7) or inactive (n = 8) disease. We also extracted a set of ISGs from each SLE sample and designated those patients with average expression above the 95th percentile of the same genes in healthy samples as ISG_hi (n = 11) and those within the

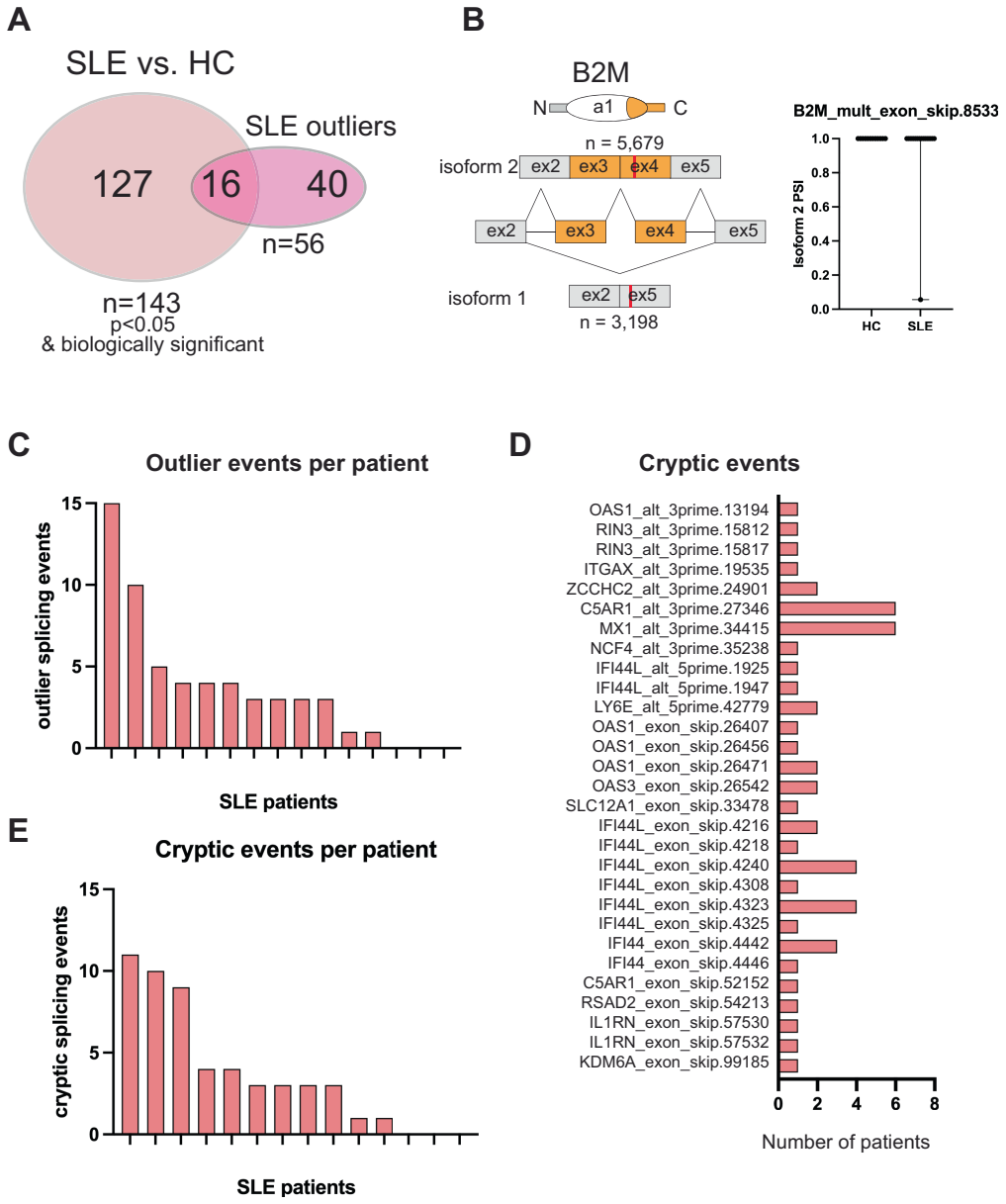

Figure 2. Outliers and cryptic splicing events in SLE neutrophils. (A) Venn diagram of the number of SLE-specific outlier splicing events and overlap with the statistically significant and biologically relevant splicing events in the total SLE group is given. (B) Example of a unique outlier splicing event in *B2M* present in a single patient with SLE is shown. Read counts (n) supporting the existence of the two isoforms are indicated. The first stop codon is indicated by a red line. (C) Distribution of outlier splicing events among the patients with SLE is shown. (D) The 29 SLE-specific cryptic splicing events and how many patients had each of them are shown. (E) Distribution of cryptic splicing events among the patients with SLE is shown. alt, alternative; ex, exon; HC, healthy control; mult, multiple; SLE, systemic lupus erythematosus.

healthy control range as ISG_lo (n = 4). As illustrated in Figure 3, some dysregulated splicing events were more evident in patients with active disease, as well as in patients with high type I interferon gene signature. However, few of the differences were statistically significant, and many splicing events were equally altered compared with healthy donors regardless of disease activity or type I interferons (Figure 3). We conclude that most altered splice events were associated with SLE disease irrespective of disease activity or type I interferons.

Comparison with a different form of immune activation with high type I interferons. To further test whether the altered splicing we observed in SLE may be related to neutrophil activation and type I interferons, we used our pipeline to analyze a publicly available RNA-seq data set of whole blood from patients infected with the SARS-CoV2 virus (GSE196822). Neutrophil-specific genes were high in this data set, indicating the expected large proportion of this lineage in the samples. There were 372 statistically significant altered splicing

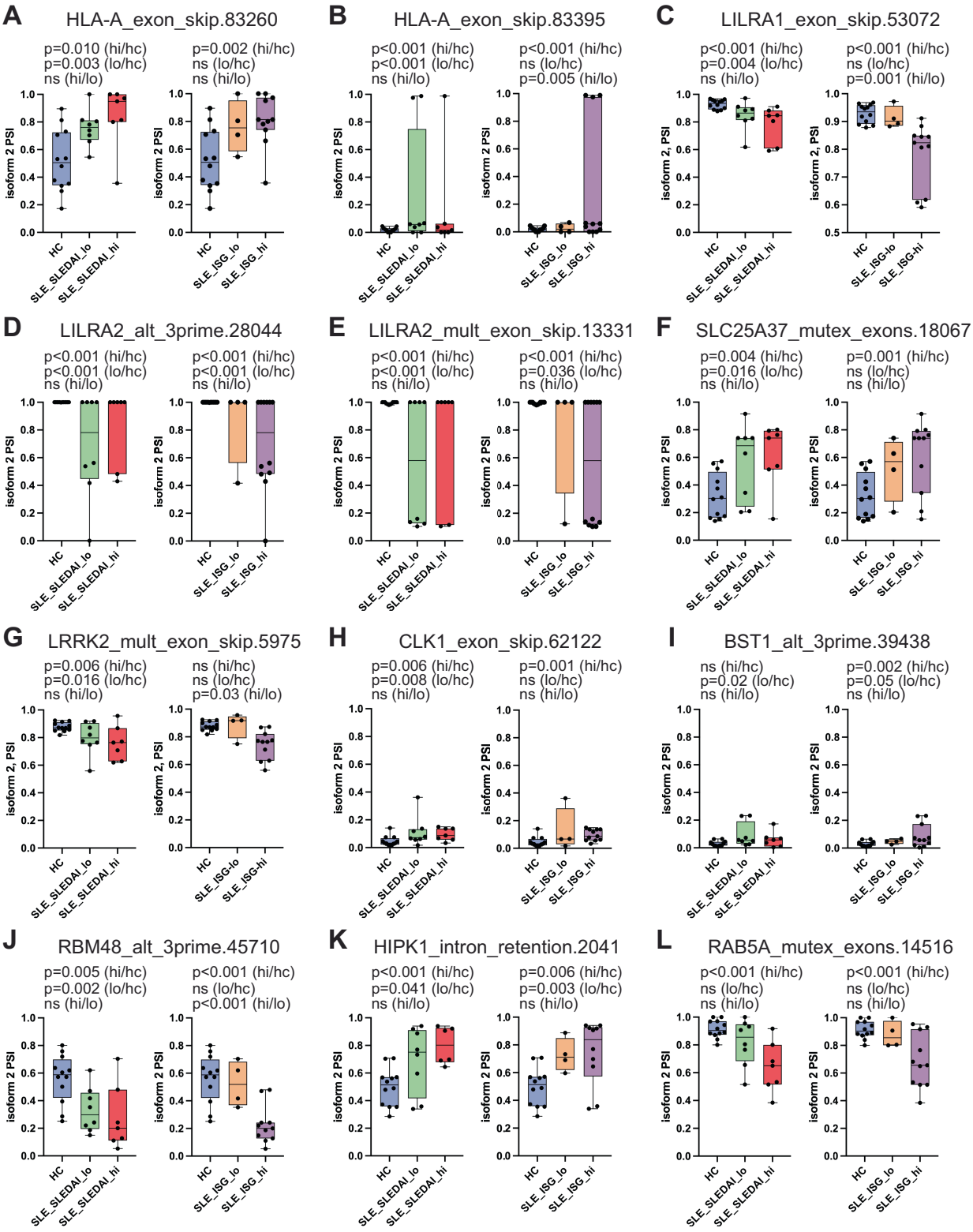

Figure 3. Representative altered alternative splicing events in neutrophils from healthy donors and patients with SLE segregated by disease activity or interferon signature. (A–L) Box and whiskers plots of the indicated event expressed as PSI. In each pair of graphs, the donors with SLE are segregated by disease activity (left panels) or interferon signature (right panels). *P* values for the comparisons among the three groups in each graph are indicated. alt, alternative; HC, healthy controls; ISG, interferon-induced gene; mult, multiple; mutex, mutually exclusive; ns, not significant; PSI, percent spliced-in; SLE, systemic lupus erythematosus; SLE_ISG_hi, patients with elevated ISG expression; SLE_ISG_lo, patients with ISGs similar to healthy controls; SLE_SLEDAI_hi, active disease; SLE_SLEDAI_lo, inactive disease.

events in patients with COVID-19 compared with healthy controls (Figure 4A). Given the presence of other immune cells in whole blood, only a subset of these events can be ascribed to neutrophils. Nevertheless, comparing these events with all the statistically significant SLE events (n = 521), only four events were shared between the two diseases (but not all in the same direction) (Figure 4A–F). Additionally, we searched within the events that we identified as likely more biologically relevant in SLE (n = 143), and 25 of them were detected in at least 1 patient with COVID-19, but only 3 of them were statistically significant

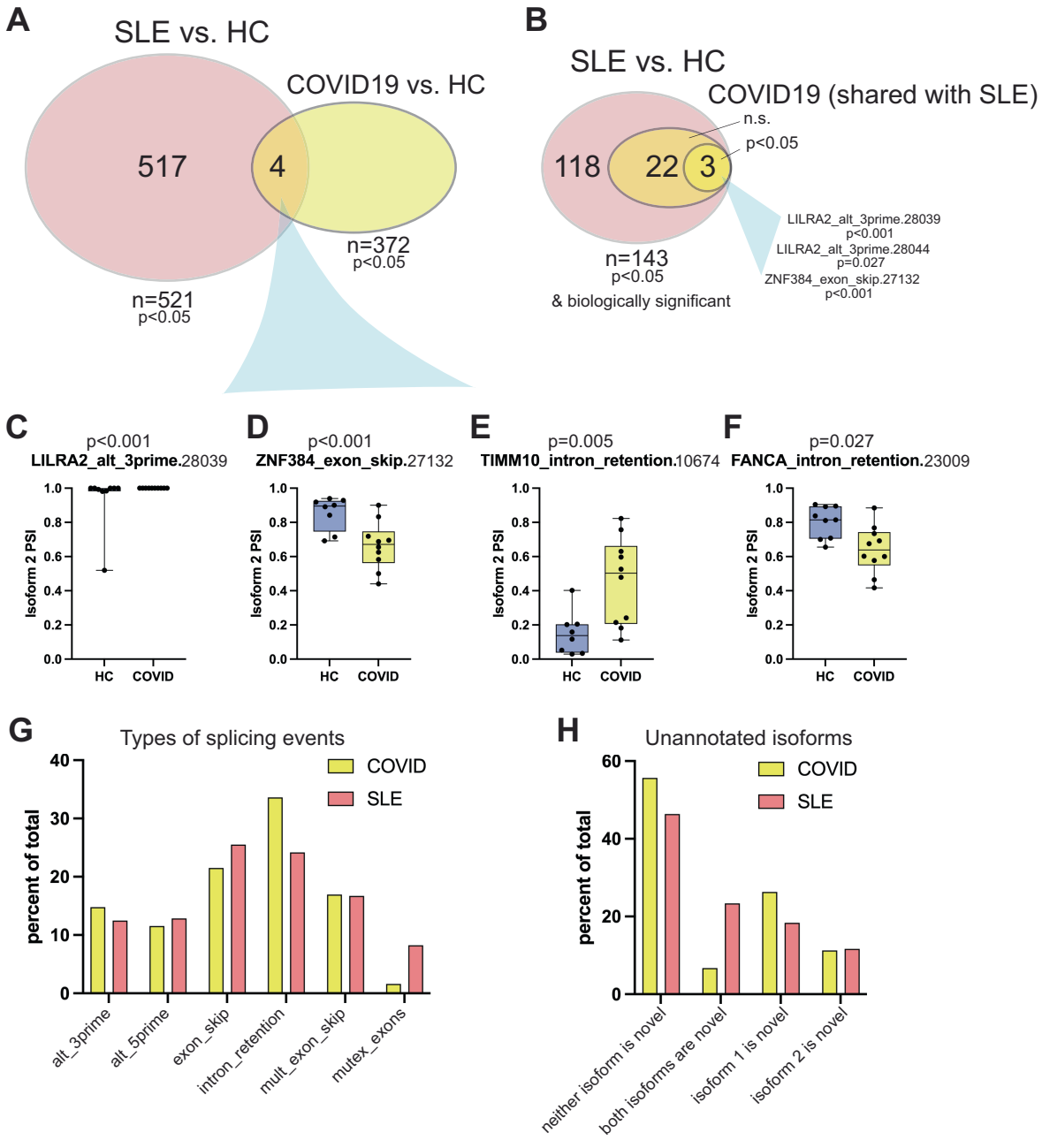

Figure 4. Alternative splicing events altered in neutrophils from patients with SLE compared with patients with COVID-19. (A) Venn diagram of the number of statistically significant alternative splicing events in SLE neutrophils and COVID-19 whole blood compared with respective healthy controls is given. (B) Venn diagram of the splicing events in COVID-19 neutrophils within the subset of 143 events in SLE that were statistically significant and likely biologically meaningful is given. Note that only 3 of the 25 events present in COVID-19 reached statistical significance. (C–F) Individual distribution in PSI of the four statistically significantly altered splice variants in the genes shared between SLE and COVID-19 is shown: (C) *LILRA2*, (D) *ZNF384*, (E) *TIMM10*, and (F) FANCA. (G) Categories of splice events in COVID19 (yellow bars) and SLE (red bars) are shown. (H) Percent of novel (unannotated) isoforms produced by splice events in COD19 (yellow bars) versus SLE (red bars) is given. alt, alternative; HC, healthy control; mult, multiple; mutex, mutually exclusive; n.s., not significant; PSI, percent spliced-in; SLE, systemic lupus erythematosus.

(Figure 4B). Compared with SLE, COVID-19 also had a different distribution of splicing event types (χ^2 $P = 5.89 \times 10^{-5}$), with fewer mutually exclusive events (1.6% vs 8.3%) and fewer intron retentions (24.2% vs 33.6%) (Figure 4G). Many fewer events produced novel (ie, unannotated) isoforms in COVID-19 compared with SLE (both isoforms novel, 6.7% vs 23.4%, χ^2 $P = 6.096 \times 10^{-11}$, and any isoform novel, 44.4% vs 53.6%, χ^2 $P = 0.008$) (Figure 4H). We conclude that the alternative mRNA splicing pattern in SLE is disease-specific and not simply due to neutrophil activation or elevated type I interferons.

Computational prediction of neoepitope presentation on MHC. To better understand whether the dysregulated mRNA splicing in SLE has the potential to produce neoantigens—ie, peptides not present in healthy individuals and with high affinity for MHC—we searched for splicing events in which SLE neutrophils, at least in some patients, expressed an isoform not seen in any of the healthy donors (Figure 5A). We identified 80 such cases with an unambiguously identified reading frame (Supplemental Table S2). In parallel, we used arcasHLA to determine each patient's class I and II MHC haplotypes.

Next, all amino acid sequences present only in SLE were subjected to NetMHCpan-4.1 and NetMHCIIpan-4.0[27] for ranking of peptides with high-affinity binding to the individual MHC alleles present in each patient. These prediction tools use neural networks and are trained on binding affinity as well as mass spectrometry data. Peptide sequences with exact matches to known proteins were excluded.

All 15 patients with SLE had at least one mis-spliced event; 8 of them had >10 such events (Figure 5B), and 1 patient had 36 neoepitope-producing splice events in 12 genes. The most frequently mis-spliced genes (Figure 5C) include 20 ISGs among the 156 genes with cryptic events or isoforms in lupus (12.8%), enriched to 10 of 33 (30.3%) when only including neoantigen generating events. Among the outliers, only 1 of 47 genes was an ISG. Cryptic events (including neoantigen generating events) were correlated with higher ISG expression, but not with activity or disease duration (Supplemental Figure S5), whereas outliers were correlated with disease duration. *RBM48* expression was lowest in the two patients with the highest numbers of outliers (Supplemental Figure S5). Figure 5C shows the highest-ranking peptides and the MHC haplotype they are predicted to bind to with high affinity (percentile binding <0.5). These data predict that abnormal transcript splicing in SLE neutrophils can create neoepitope peptides that the patient's T cells would see as a "foreign" antigen. However, it is likely that some of these mis-spliced transcripts will be degraded by the nonsense-mediated decay machinery[29–31] and that only a subset of them will be translated into immunogenic polypeptides.[32]

DISCUSSION

Our findings that SLE neutrophils contain a large number of mRNA splicing events that differ from those in healthy control samples, as well as those in patients with SARS-CoV2, suggest that dysregulated mRNA splicing is a feature of SLE disease. Subgroup comparisons revealed that relatively few splicing events were associated with active disease or elevated type I interferons, leaving it less likely that altered splicing is a consequence of SLE activity or its treatment. The minimal overlap with splice events in COVID-19 also argues against a role of normal neutrophil activation or elevated type I interferons in driving them. This SLE-specific dysregulated splicing could well contribute to the changes in neutrophils observed in SLE, or it could be the consequence of undefined upstream events that independently cause neutrophil abnormalities. It is possible that individual nucleotide polymorphisms within splicing factor recognition sites or in splice donor or acceptor sites create additional variability within the SLE group. All these possibilities are not mutually exclusive.

The alterations observed in the splicing factors encoded by *RBM10* and *RBM48* introduces the possibility that dysfunction of these proteins could be responsible for the mis-splicing of other genes. The retention of the intron between exons 12 and 13 in RBM10 creates a novel sequence and a stop codon, which will terminate the protein before its octamer repeat, Zinc finger, and G-patch domain, which are important for its function.[33] Perturbation of RBM10 leads to diverse changes in alternative splicing, most prominently alteration in cassette exon inclusion. Enhanced UV cross-linking immunoprecipitation experiments in human cells demonstrate that RBM10 binding to exon-adjacent intronic regions regulates cassette exon splicing.[33,34] Hence, this mis-spliced form of RBM10 may cause some of the exon or multiexon skips we observe in SLE. RBM48 plays a role in splicing of U12-type minor introns,[35] which are present in genes regulating myeloid differentiation. This possibly relates to the altered splicing we observe in *MNDA* and *NCF4*.

The creation of proteins with altered structure and function, eg, those with lost domains such as a transmembrane helix, could well contribute to cellular dysfunction, particularly if numerous such events co-exist. In fact, many of the affected proteins interact physically or functionally (eg, as revealed by STRING [STRING Consortium 2024] analysis); hence, even subtle alterations in several proteins in a shared pathway could amplify their impact on the behavior of neutrophils. For example, signaling from even subtly altered surface lectins, immunoglobulin and pattern-recognition receptors, and Fc receptors, which jointly govern the response of neutrophils to extracellular stimuli mediated by alternatively spliced kinases and phosphatases, could profoundly affect neutrophil maturation, adhesion, migration, intracellular transport, and transcription. It should also be noted that mis-spliced transcripts, particularly those with premature stop codons, may be rapidly degraded by the non–sense-mediated decay machinery.[29–31] As a consequence, such splice forms create a loss of the protein to a degree that matches the portion of the altered mRNA of the total mRNA for the gene in question.

Furthermore, alternative splicing can introduce frameshifts or exonization of previously intronic segments that, if translated,

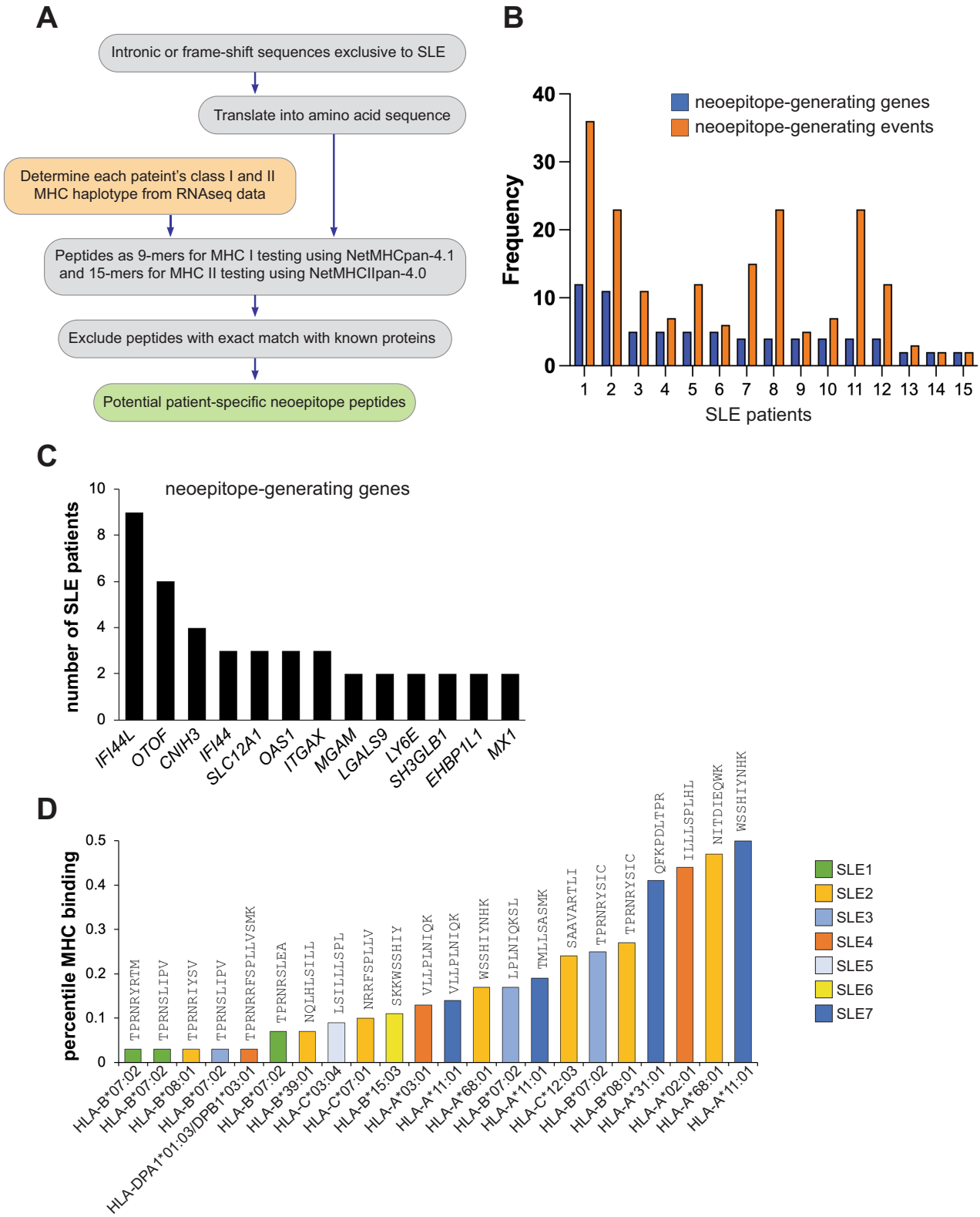

Figure 5. Analysis of potential neoantigens generated by dysregulated splicing in SLE. (A) Schematic representation of our analysis pipeline is given. (B) Frequency of neoepitope-generating genes and events in the 15 patients with SLE is shown. (C) Ranking of the neoepitope-generating genes by number of patients they are mis-spliced in is given. (D) Ranking of neoepitope peptides by binding affinity percentile rank below the 0.5 cutoff for high-affinity binding to HLA using NetMHCpan-4.1 (class I) and NetMHCIIpan-4.0 (class II) is given. Peptide sequences are shown above the bars with the relevant HLA alleles below the *x* axis, and the color codes indicate the individual patients. MHC, major histocompatibility complex; RNA-seq, RNA sequencing; SLE, systemic lupus erythematosus.

result in novel amino acid sequences for which immunological tolerance may be incomplete or absent. It is well documented in cancer that dysregulated mRNA splicing creates numerous tumor-specific neoantigens recognized in anti-tumor immune responses.[36,37] We propose that this phenomenon may also occur in SLE. Modeling all SLE-specific noncanonical novel peptide sequences for their ability to bind the corresponding patients' MHC alleles revealed a set of peptides with predicted high-affinity binding. Future work will need to verify which of these peptides are generated and presented on MHC and whether T cells specific for these neoantigens exist in patients.

Because SLE is a clinically remarkably heterogenous disease, it is expected that it is molecularly heterogenous as well.[38–40] The present findings add support for this notion: some of the SLE-specific splicing events display a clear dichotomy. For example, in exon_skip.83395 in *HLA-A*, most patients with SLE express isoform 1, whereas three patients with SLE show isoform 2 at essentially 100%. Similarly, three alternative 3′ splicing events in *LILRA2* express isoform 2 in healthy controls, but either all isoform 1 or around 50% isoform 2 in patients with SLE. However, other events are more variable and with more overlap between patients and controls. Interestingly, segregating the patients with SLE by type I interferon signature revealed a more nuanced pattern in which a minority of splicing events were concentrated in the patients with high interferon. This could mean either that interferons can influence the splicing of transcripts or, conversely, that the high type I interferons are a consequence of altered splicing of one or several regulatory proteins. The absence of these events in patients with COVID-19 does not support the former. Nevertheless, either scenario is compatible with the notion that SLE with high ISGs represents a molecularly distinct form of SLE. However, this subgroup comparison is limited by the small number of patients in the low ISG group. It is also interesting that nearly all cryptic splice events were in transcripts from ISGs.

Although altered alternative splicing may well occur in many different cell types, our study focused on neutrophils for several reasons. There are mounting indications that this versatile cell type plays a more active role in SLE than previously assumed. Neutropoiesis is accelerated in SLE,[41] suggesting that neutrophils are also consumed at an elevated rate in SLE. This implies that dying neutrophils represent a rich source of intracellular proteins and nucleic acids[42] that can serve as autoantigens. Indeed, a recent article[43] showed that autoantibodies against Ro52 (encoded by *TRIM21*) predominantly target an epitope produced in neutrophils through alternative mRNA splicing. Future studies will be needed to determine whether dysregulated mRNA splicing also occurs in other immune or nonimmune cells. We cannot exclude the possibility that what appear to be SLE-specific splicing events in neutrophils are, in fact, physiologic in some other cell lineage, although this seems particularly unlikely for the cryptic splicing events. We also cannot assume that all noncanonical neoepitopes found here truly are unique to SLE or that they have not been presented to T or B cells during their repertoire selection in the thymus or bone marrow, respectively. However, even a single one could appear as a persistent foreign antigen to the adaptive immune system and drive an escalating autoimmune response.

ACKNOWLEDGMENTS

We thank the patients who donated blood for this research and Megan Tran for the acquisition of these patient samples.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Mustelin confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Elloumi N, Ben Mansour R, Marzouk S, et al. Differential reactive oxygen species production of neutrophils and their oxidative damage in patients with active and inactive systemic lupus erythematosus. Immunol Lett 2017;184:1–6.
2. Perazzio SF, Salomão R, Silva NP, et al. Increased neutrophil oxidative burst metabolism in systemic lupus erythematosus. Lupus 2012; 21(14):1543–1551.
3. Li P, Jiang M, Li K, et al. Glutathione peroxidase 4-regulated neutrophil ferroptosis induces systemic autoimmunity. Nat Immunol 2021; 22(9):1107–1117.
4. Garcia-Romo GS, Caielli S, Vega B, et al. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. Sci Transl Med 2011;3(73):73ra20.
5. Lood C, Blanco LP, Purmalek MM, et al. Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease. Nat Med 2016;22(2):146–153.
6. Lande R, Ganguly D, Facchinetti V, et al. Neutrophils activate plasmacytoid dendritic cells by releasing self-DNA-peptide complexes in systemic lupus erythematosus. Sci Transl Med 2011;3(73):73ra19.
7. Coit P, Yalavarthi S, Ognenovski M, et al. Epigenome profiling reveals significant DNA demethylation of interferon signature genes in lupus neutrophils. J Autoimmun 2015;58:59–66.
8. Denny MF, Yalavarthi S, Zhao W, et al. A distinct subset of proinflammatory neutrophils isolated from patients with systemic lupus erythematosus induces vascular damage and synthesizes type I IFNs. J Immunol 2010;184(6):3284–3297.
9. Mistry P, Nakabo S, O'Neil L, et al. Transcriptomic, epigenetic, and functional analyses implicate neutrophil diversity in the pathogenesis of systemic lupus erythematosus. Proc Natl Acad Sci U S A 2019; 116(50):25222–25228.
10. Rogalska ME, Vivori C, Valcarcel J. Regulation of pre-mRNA splicing: roles in physiology and disease, and therapeutic prospects. Nat Rev Genet 2023;24(4):251–269.
11. Ng B, Yang F, Huston DP, et al. Increased noncanonical splicing of autoantigen transcripts provides the structural basis for expression

of untolerized epitopes. J Allergy Clin Immunol 2004;114(6):1463–1470.

12. Kahles A, Lehmann KV, Toussaint NC, et al; Cancer Genome Atlas Research Network. Comprehensive analysis of alternative splicing across tumors from 8,705 patients. Cancer Cell 2018;34(2):211–224.e6.
13. Jayasinghe RG, Cao S, Gao Q, et al; Cancer Genome Atlas Research Network. Systematic analysis of splice-site-creating mutations in cancer. Cell Rep 2018;23(1):270–281.e3.
14. Du J, Wang Q, Yang S, et al. FOXP3 exon 2 controls T_{reg} stability and autoimmunity. Sci Immunol 2022;7(72):eabo5407.
15. Free ME, Bunch DO, McGregor JA, et al. Patients with antineutrophil cytoplasmic antibody-associated vasculitis have defective Treg cell function exacerbated by the presence of a suppression-resistant effector cell population. Arthritis Rheum 2013;65(7):1922–1933.
16. Miyabe C, Miyabe Y, Strle K, et al. An expanded population of pathogenic regulatory T cells in giant cell arteritis is abrogated by IL-6 blockade therapy. Ann Rheum Dis 2017;76(5):898–905.
17. Vaughn SE, Kottyan LC, Munroe ME, et al. Genetic susceptibility to lupus: the biological basis of genetic risk found in B cell signaling pathways. J Leukoc Biol 2012;92(3):577–591.
18. Elghzaly AA, Sun C, Looger LL, et al. Genome-wide association study for systemic lupus erythematosus in an egyptian population. Front Genet 2022;13:948505.
19. Ren P, Lu L, Cai S, et al. Alternative splicing: a new cause and potential therapeutic target in autoimmune disease. Front Immunol 2021; 12:713540.
20. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29(1):15–21.
21. Kahles A, Ong CS, Zhong Y, et al. SplAdder: identification, quantification and testing of alternative splicing events from RNA-seq data. Bioinformatics 2016;32(12):1840–1847.
22. Najjar R, Mustelin T. Prediction of alternative pre-mRNA splicing outcomes. Sci Rep 2023;13(1):20000.
23. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res 2019;47(8):e47.
24. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15(12):550.
25. Chiche L, Jourde-Chiche N, Whalen E, et al. Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. Arthritis Rheumatol 2014;66(6):1583–1595.
26. Orenbuch R, Filip I, Comito D, et al. arcasHLA: high-resolution HLA typing from RNAseq. Bioinformatics 2020;36(1):33–40.
27. Reynisson B, Alvarez B, Paul S, et al. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res 2020;48(W1):W449–W454.
28. Hirayasu K, Saito F, Suenaga T, et al. Microbially cleaved immunoglobulins are sensed by the innate immune receptor LILRA2. Nat Microbiol 2016;1(6):16054.
29. Tan K, Stupack DG, Wilkinson MF. Nonsense-mediated RNA decay: an emerging modulator of malignancy. Nat Rev Cancer 2022;22(8): 437–451.
30. Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol 2015;16(11):665–677.
31. Brogna S, Wen J. Nonsense-mediated mRNA decay (NMD) mechanisms. Nat Struct Mol Biol 2009;16(2):107–113.
32. Pastor F, Kolonias D, Giangrande PH, et al. Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. Nature 2010;465(7295):227–230.
33. Inoue A. RBM10: Structure, functions, and associated diseases. Gene 2021;783:145463.
34. Wang E, Pineda JMB, Kim WJ, et al. Modulation of RNA splicing enhances response to BCL2 inhibition in leukemia. Cancer Cell 2023;41(1):164–180.e8.
35. Siebert AE, Corll J, Paige Gronevelt J, et al. Genetic analysis of human RNA binding motif protein 48 (RBM48) reveals an essential role in U12-type intron splicing. Genetics 2022;222(2):iyac129.
36. Matsushima S, Ajiro M, Iida K, et al. Chemical induction of splice-neoantigens attenuates tumor growth in a preclinical model of colorectal cancer. Sci Transl Med 2022;14(673):eabn6056.
37. Lu SX, De Neef E, Thomas JD, et al. Pharmacologic modulation of RNA splicing enhances anti-tumor immunity. Cell 2021;184(15): 4032–4047.e31.
38. Nehar-Belaid D, Hong S, Marches R, et al. Mapping systemic lupus erythematosus heterogeneity at the single-cell level. Nat Immunol 2020;21(9):1094–1106.
39. Bashant KR, Aponte AM, Randazzo D, et al. Proteomic, biomechanical and functional analyses define neutrophil heterogeneity in systemic lupus erythematosus. Ann Rheum Dis 2021;80(2):209–218.
40. Zheng M, Hu Z, Mei X, et al. Single-cell sequencing shows cellular heterogeneity of cutaneous lesions in lupus erythematosus. Nat Commun 2022;13(1):7489.
41. Bennett L, Palucka AK, Arce E, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. J Exp Med 2003; 197(6):711–723.
42. Ukadike KC, Ni K, Wang X, et al. IgG and IgA autoantibodies against L1 ORF1p expressed in granulocytes correlate with granulocyte consumption and disease activity in pediatric systemic lupus erythematosus. Arthritis Res Ther 2021;23(1):153.
43. Gomez-Bañuelos E, Wahadat MJ, Li J, et al. Alternative exon usage in TRIM21 determines the antigenicity of Ro52/TRIM21 in systemic lupus erythematosus. JCI Insight 2022;7(19):e163795.