Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Neoantigens recognized by T and B cells are generated by aberrant mRNA splicing in lupus neutrophils

Tomas Mustelin

tomas2@uw.edu

UW https://orcid.org/0000-0001-5912-8840

Rayan Najjar

University of Washington https://orcid.org/0000-0002-9214-521X

Xiaoxing Wang

UW

Farheen Shaikh

UW

Ning Wang

UW

Ethan Le

UW

Jie An

UW https://orcid.org/0009-0000-7393-6406

Cecilia Mustelin

UW

Fatemeh Moadab

University of Washington

Vincent Gureghian

UW https://orcid.org/0000-0001-8878-1427

Sadie Van den Bogaerde

UW

Stephen Schoenberger

La Jolla Institute for Immunology

Biological Sciences - Article

Keywords: lupus, splicing, neoantigen, neutrophils, T cell activation, autoimmunity

Posted Date: September 26th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7668471/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Yes there is potential Competing Interest. TM reports consulting fees from Cugene, Evolved, Codify, and Rome Therapeutics. He serves on the Scientific Advisory Board of the Tri-Institutional Drug Discovery Institute (non-profit), and the University of Copenhagen LEO Skin Immunology Centre (non-profit).

Running title: neoantigens from dysregulated mRNA splicing in lupus

Neoantigens recognized by T and B cells are generated by aberrant mRNA splicing in lupus neutrophils

by

Rayan Najjar[1,*], Xiaoxing Wang[1], Farheen Shaikh[1], Ning Wang[1], Ethan Le[1], Jie An[1], Cecilia Mustelin[1], Fatemeh Moadab[1], Vincent Gurghian[1], Sadie J. Van den Bogaerde[1], Stephen P. Schoenberger[2] & Tomas Mustelin[1,*]

[1]Division of Rheumatology, Department of Medicine, University of Washington; Seattle, WA 99109, USA.

[2]La Jolla Institute for Immunology, 9420 Athena Circle, La Jolla CA 92037

*Corresponding authors. Emails: tomas2@uw.edu, najjar@uw.edu

Proteome diversity is generated through alternative mRNA splicing[1]. In tumors, mis-spliced transcripts that contain intronic sequences, skipped exons, or noncanonical donor/acceptor splice junctions resulting in frame-shifts are often translated into neoantigens that drive anti-tumor immunity[2,3]. Here we report that neutrophils from patients with systemic lupus erythematosus (SLE), unlike neutrophils from healthy donors or patients with rheumatoid arthritis, contain numerous aberrantly spliced gene transcripts, the translation of which was predicted to result in novel amino acid sequences absent in normal splice variants or any other proteins. 21 peptides derived from these novel amino acid sequences were detected by mass spectrometry on class I major histocompatibility complex (MHC) molecules in patient samples. Co-culture of patient T cells with overlapping sets of 15-mer peptides from the novel sequences resulted in the expansion of CD4 and CD8 T cells that were rapidly reactivated by them to secrete γ-interferon and IL-17. Serum from SLE patients also contained IgG autoantibodies that recognized epitopes within the novel amino acid sequences, particularly those patients predicted to harbor the corresponding cryptic splicing event. Our findings demonstrate that dysregulated mRNA splicing in SLE neutrophils generates neoantigens that have caused T and B cell activation, clonal expansion, IgG class-switch, and autoimmunity. (200 words)

Key words: lupus, splicing, neoantigen, neutrophils, T cell activation, autoimmunity

Introduction

Systemic lupus erythematosus (SLE) is a multifaceted and debilitating autoimmune disorder of unclear etiology[4-6]. SLE disproportionately affects young women, involves many organs, and is generally characterized by the presence of autoantibodies against nuclear proteins, nucleic acid-associated proteins and double-stranded DNA[7]. Although antigen-specific T cells responses[8] are instrumental, it is largely unknown what antigens T cells respond to[9]. Mounting evidence also suggests that innate immune cells play a central role in driving SLE's underlying mechanisms[10-14], including the elevated levels of type I interferons (IFN-I) found in most SLE patients[15]. Neutrophilic granulocytes (neutrophils) are also being increasingly implicated[11-13,16-18] as the sources of SLE autoantigens, both nucleic acids and proteins, released during inflammatory cell death mechanisms like NETosis[13] and ferroptosis[19]. Although neutropoiesis is accelerated in SLE patients[15], many patients are mildly neutropenic, indicating that neutrophils are consumed at a higher than normal rates in this disease. SLE-associated neutrophils also display widespread genomic hypomethylation, leading, in part, to a distinctive IFN-I-weighted transcriptional profile[20], which differs from that in other autoimmune diseases, but bears much resemblance to that in neutrophils from SARS-CoV2-infected individuals[21].

Alternative RNA transcript splicing begins during gene transcription and results in the nuclear export of mature mRNA variants that profoundly contribute to proteome complexity and molecular diversity[1]. Proteins translated from alternatively spliced mRNA variants from a single gene often display substantial differences in molecular functions, stability, and interactions with other proteins or nucleic acids. Alternative splicing mechanisms can also modulate mRNA disposition, including degradation through nonsense-mediated decay (NMD)[22]. Alternative splicing occurs more frequently in transcripts for autoantigens[20] compared to randomly selected proteins and escape from NMD has been shown to cause autoimmunity[23]. Neutrophils represent an abundant reservoir of intracellular proteins and nucleic acids that potentially function as autoantigenic sources. For example, autoantibodies targeting Ro52 (encoded by *TRIM21*) predominantly recognize a neutrophil-specific epitope generated through alternative mRNA splicing mechanisms[24].

To comprehensively analyze alternative mRNA splicing we built on Spladder, a tool for in-depth analysis of individual alternative splicing events, to create the RiboSplitter[25] pipeline, which additionally yields the translated protein products of splice variants. With this tool we identified a total of 65,730 high-confidence alternative splicing events in 5,697 protein-coding genes in neutrophils from SLE patients, of which 521 events in 356 genes were statistically significantly different from healthy donors[26]. Here, we focus on the SLE-restricted mRNA splicing events predicted to generate novel protein sequences, referred to as 'cryptic' splicing events.

RESULTS

Cryptic splicing events in SLE neutrophils

Our experimental approach is shown schematically in **Fig. 1a** and our computational pipeline in **Fig. 1b.** Thirtyfive SLE patients representing a typical academic medical center population (**Supplemental Table S1**) consented to blood donation. Healthy donors and patients with rheumatoid arthritis (RA) served as normal controls and disease controls, respectively. Immune cells and neutrophils were isolated, and the patients' sera stored frozen until use. RNA was isolated from neutrophils and used for stranded paired-end RNA sequencing after quality controls, to a sequencing depth of approximately 50 million reads with 101 bp read length. Using a lower limit of 5 reads to eliminate stochastic low-frequency errors in splicing (see Experimental Procedures), we detected a total of 271 cryptic splicing events in 130 genes in SLE neutrophils (n=35), all absent in neutrophils from the healthy donors (n=20) and RA patients (n=15). The most abundant events were abnormal exon skips, followed by multi-exon skips, atypical 5' or 3' splice sites, and (partial) intron retention (**Fig. 1c**). Several cryptic events were combinations of these mechanisms. A pathway analysis of the aberrantly spliced genes showed a strong bias for type I interferon-inducible genes (**Fig. 1d**), including the most frequently mis-spliced genes *IFI44L*, *MX1*, *OAS1*, *RSAD2*, *XAF1*, *IFI27, MX1,* and *IFIT1*. However, cryptic splicing events were also observed in genes not considered interferon-inducible, such as *ABCG1*, *PSMF1*, and *OGG1* (**Supplemental Fig. S1**)

showing that expression levels alone cannot explain the generation of these cryptic splicing events, although it could be a factor . As expected, read support for cryptic splicing events was lower than non-cryptic events (**Fig. 1e**), but read support for cryptic splicing events that generated novel amino acid sequences varied from 283 reads down to our cut-off threshold of 5 (median=10).

Cryptic splicing events were detected in 34 of the 35 SLE patients (**Fig. 1f**), but the total number of events, and the genes they occurred in, varied between patients from 45 events in 18 genes down to 1 event in 1 gene (**Fig. 1f**). These data also revealed that the precise pattern of cryptic splicing is relatively unique to each patient even if several genes are mis-spliced in up to 15 of the 35 patients. For comparison, we analyzed two publicly available RNA sequencing datasets of peripheral blood mononuclear cells (PBMC) from SLE patients (n=36 total) using our pipeline and did not find any of the cryptic splicing events, suggesting these patterns of cryptic splicing are unique to neutrophils in SLE. Similarly, the acute myeloid cell line AML14, which represents an immature myeloid phenotype, also lacked all cryptic splicing observed in SLE neutrophils.

The total number of 9-mer peptides optimal for class I MHC binding or 15mer for class II MHC the cryptic transcripts theoretically could generate also varied similarly between patients (**Fig. 1g**). To assess the likelihood that some of them would indeed bind and be presented to T cells, we determined each patient's major histocompatibility complex (MHC) class I and II haplotype using arcasHLA (**Supplemental Table S2**) and ranked all possible peptides derived from their individual set of novel amino acid sequences for binding using NetMHCpan-4.1 and NetMHCIIpan-4.0[27]. These prediction tools use neural networks and are trained on experimental binding affinity as well as mass spectrometry data. We only considered peptides that did not match any known human protein sequences in Ensemble and in secondary BLAST searches. Individual SLE patients varied in the number of potential neoantigen peptides with predicted high-affinity binding to their own haplotype of class I or II molecules (**Fig. 1h**): 86 peptides from 46 genes ranked in the top 2nd percentile, 65 of them in the top 0.5 percentile.

Validation of cryptic splice variants by RT-PCR and sequencing

To experimentally confirm cryptic mRNA splicing events, we designed polymerase chain reaction (PCR) primers flanking the splicing event designed to yield an amplicon of diagnostic size. In the mRNA sample from an SLE patient with the multi-exon skip in *USP18,* this variant was readily visualized by the appearance of an amplicon corresponding to the predicted 207 bp (**Fig. 2a**). Sanger sequencing of this band verified its identity as the joined exons 6 and 11. In addition, there were bands of ~300 bp and ~450 bp, representing the skip of 3 or 2 exons (rather than 4). The canonical splice form was also prominently present at the predicted 653 bp size. These data validate the presence of the cryptic splice variant in a patient our computational tools predicted to have it.

Similarly, the cryptic *ABCG1* exon skip was validated by the flanking PCR primers yielding a 203 bp band (**Fig. 2b**), the sequencing of which verified the joining of exon 8 to exon 10, as well as the canonical 352 bp band. The skip of exons 4 and 5 in the cryptic *IFI44L* variant was seen as a 101 bp amplicon (**Fig. 2c**) by flanking PCR primers and its sequence verified the joining of exon 3 to exon 6. The cryptic splicing of *OAS1* (**Fig. 2d**) and *MX1*(**Fig. 2e**), were validated by RT-PCR and Sanger sequencing of the relevant band in the agarose gel. All these data also illustrate that the cryptic splice variants are less abundant than the canonical or physiological splice variants, in good agreement with the computationally obtained read counts for the cryptic events compared to all the splicing events in the same genes (**Fig. 2f** and **g**).

Detection of MHC-bound peptides derived from the neo-sequences

Next we used an immunopeptidomics[28] approach to determine if peptides derived from the computationally predicted novel amino acid sequences were indeed translated and proteolytically processed for presentation on MHC[29]. This is vital as transcripts with premature stop codons are frequently degraded by NMD[30-32] and because the proteolytic processing creates skewing in the presented peptide repertoire. We addressed this by immunoprecipitating class I MHC molecules using an anti-pan HLA antibody and analyzing the eluted peptides by targeted liquid chromatography-tandem mass spectrometry (LC-

MS/MS), using a custom reference database containing all computationally predicted neoantigen peptides. These experiments yielded spectra matching 21 of the predicted novel peptides and their calculated daughter ions for a high confidence of identification (**Fig. 3, TABLE I** and **Supplemental Fig. S2**). A parallel analysis using the non-redundant protein database routinely used for mass spectrometry yielded a long list of 8-11 amino acid peptides representing self-peptides presented on class I MHC on neutrophils (not shown). We refer to the novel peptides identified by LC-MS/MS in HLA immunoprecipitates as *bona fide* 'neoantigens'. By NetMHCpan-4.1 ranking, the majority of them fell into the highest-affinity group (**Supplemental Fig. S3**), with only one notable exception: the 10-residue peptide AGRRKEEEIL derived from a cryptic splicing event in *MX1* which was predicted to bind the patient's HLA haplotype with modest affinity, perhaps reflecting some post-translational modification not recognized in the LC-MS/MS, a mutation or modification of the HLA molecule in this patient, or a false negative prediction by NetMHCpan-4.1.

Mechanisms of neoantigen generation

Fig. 3 also illustrates in detail a set of cryptic splicing events that generated novel amino acid sequences from which peptides bound to HLA were detected. For example, a cryptic splicing event in *USP18* (**Fig. 3a**) involving a multi-exon skip to splice exon 6 to exon 11 resulted in a frame-shift of exon 11 and translation of 24 codons (purple box in **Fig. 3a**) until an in-frame stop codon (indicated by a red line). The translated USP18 protein has a novel C-terminus, from which an 11-amino acid peptide was identified with very high confidence (idotp 0.99) in anti-HLA immunoprecipitates. A similar event in the transcript from *MYO15B* (**Fig. 3b**) excised exons 41-43, resulting in a frame-shift in the translation of exon 44 (purple box in **Fig. 3b**), from which a 9-amino acid residue peptide was detected by LC-MS/MS in an anti-HLA immunoprecipitate from the patient predicted to have this cryptic splice variant.

A cryptic splicing event in *PHF11* in another SLE patient (**Fig. 3c**) generated a novel amino acid sequence by a different mechanism: an aberrant skip of exon 11, in which a stop codon normally terminates translation, resulted in a continued translation of 53 codons in

the normally untranslated exon 12. The junction of exon 10 and this novel amino acid sequence yields an HLA-bound peptide FQAGCHTQL (idotp 0.92) in which the first 4 residues are encoded by exon 10, but the following 5 represent neo-sequence unique to the cryptic splicing event with no match in a BLAST search. In the canonical isoform, exon 10 is translated in its intended reading-frame resulting in a completely different amino acid sequence (golden box).

A similar mechanism generated a novel amino acid product of an *OAS1* transcript (**Fig 3d**) where a cryptic splicing event representing both a skip of exon 6 (which occurs physiologically in transcript variant 10) and an altered 3' acceptor site for splicing located 3' of the stop codon in exon 7, resulting in the translation of 26 codons of exon 7 (purple box in **Fig. 3d**) until an in-frame stop codon (red line). Hence, the resulting OAS1 protein has a novel C-terminus, which can give rise to the FLLTGVDL peptide we detected by LC-MS/MS on HLA-A*02:01 in the neutrophils from the patient that harbors this cryptic splicing event with 15 read counts. We also note that OAS1 is normally spliced into alternative isoforms, *e.g.,* transcript variant 1, which includes exon 6, in which a stop codon terminates translation into a protein with a different C-terminus than variant 10 (golden boxes). However, only the cryptic isoform unique to SLE encodes the neoantigen sequence (purple box; no matches in a BLAST search). A total of 15 of the 35 SLE patients expressed cryptic splice variants encoding this particular neoantigen.

A different mechanism generated a 32-amino acid residue long novel sequence from the aberrant splicing of the *FCGRT* transcript, in which 96 nucleotides of an inverted 302-bp *AluJb* element located in the 10-kb intron (which contains 26 *Alu*s) between exons 4 and 5 was spliced in to generate the novel sequence (**Fig. 3e**). A peptide encoded by the junction of exon 4 and the Alu sequence (NLEWKETGF) was detected by LC-MS/MS in anti-HLA immunoprecipitates. Further illustrative examples of splicing events unique to SLE included a two-exon skip plus an aberrant 3' splice acceptor site in *MX1* (**Fig. 3f**), resulting in a frame-shift in the translation of exon 16, from which a peptide was identified by LC-MS/MS.

Supplemental Fig. S2 shows additional cryptic splicing events in *ABCG1*, *RSAD2*, *DAAM2*, *IFIT1*, *IFI44*, and *IFI44L*. The event in *RSAD2* includes part of an intron consisting of a long interspersed element-1 (LINE1)-derived HAL1 ('half-L1'[33]) retrotransposon, while the

cryptic event in *IFIT1* includes part of the intron containing an inverted AluSx element. Some of the shown events generated peptides identified in anti-HLA immunoprecipitates, while others did not.

It is worth noting that most of the cryptic splicing events that resulted in neoantigen peptides detected by immunopeptidomics reside near the 3' end of the open-reading frame and therefore do not introduce premature stop codons that would result in NMD. Another set of peptides were derived from events that spliced in a portion of an intron, *e.g.*, an inverted *Alu* element, but did not introduce a premature stop codon.

Autoantibodies against novel amino acid sequences created by cryptic splicing in SLE

Although we cannot categorically exclude the possibility that these noncanonical RNA splicing events with unannotated splice sites, occurred from infancy in our SLE patients, it is more likely that they arose later in life through dysregulation of the splicing machinery. This notion is supported by their complete absence in healthy donors or in patients with RA. Hence, the SLE patients would likely not have developed central tolerance against the cryptic splice variants. To test if class-switched IgG autoantibodies exist in our patient cohort against these neo-sequences, we synthesized a series of 15-residue peptides to cover, in a tiled fashion, the novel amino acid sequences generated by a set of cryptic splicing events observed in the highest numbers of SLE patients. ELISAs with these peptides revealed that many of the patients computationally predicted to express the selected neoantigens indeed had IgG reactive with peptides corresponding to them (**Fig. 4**). Using the 95th percentile of the background reactivity of healthy donors as a cut-off, 17 of the 35 SLE patients were positive for IgG recognizing the first IFI44L peptide, including 6 of the 10 patients observed to generate this neo-sequence at the time of blood draw for our RNA sequencing. These results support the notion that the IFI44L-neo-sequence was produced in these patients in a chronic or episodic manner for a sufficiently long time to generate a B cell response leading to circulating IgG antibodies against this sequence. Peptides from the novel amino acid sequences in *OAS1*, *MX1*, *IFI44*, and *RSAD2* were also recognized by IgG autoantibodies, with

the fifth peptide from the frame-shifted translation of exon 16 after the cryptic multi-exon skip with a 3' neojunction in exon 16 of *MX1* recognized by 19 of the 35 SLE patients.

Autoantibodies were also present in some patients in which the neoantigen-producing cryptic splicing event was not detected. These either represent patients who had the cryptic splice form earlier, but not at the time of venipuncture, or which were missed in our analysis because our RNA sequencing depth of 50 million reads is estimated to give us approximately 60% coverage of mRNA splicing events[34,35], indicating there might be more splicing-generated neoantigens than we observed.

Antigen-specific T cells recognizing neoantigens

The presence of class-switched IgG autoantibodies recognizing the neo-amino acid stretches suggests that neoantigen-specific T cells were also involved and able to provide B cell help for class switch and affinity maturation. To directly test this, we used a 14-day expansion and restimulation assay optimized for the detection of tumor antigen-specific T cells[36]. In this regimen, the antigen-specific T cells are activated and undergo several cell divisions and then return to a resting state before the restimulation on day 14 (**Fig. 5a**). Peripheral blood mononuclear cells (PBMCs) from patients included in the RNA sequencing cohort, as well as additional ones, were incubated for 14 days in the presence of peptide pools consisting of 3-7 distinct 15-mer peptides covering the novel amino acid sequence of a cryptic splice variant and then restimulated on day 14 with each individual peptide separately and assayed on day 15 for interferon-γ (IFNγ) and interleukin-17 (IL-17) production by ELISpot, as well as analysis of activation by flow cytometry. While T cells from one SLE patient (on immunosuppressive treatment: tacrolimus and prednisone) reacted weakly, T cells from three other tested patients (**Fig. 5b-e**), and 2 additional ones, showed moderate to strong reactivity against several peptides. While each patient's T cells displayed a unique profile, the third peptide from the MX1-derived neo-sequence was the strongest in several patients. IFNγ secretion was more frequent than IL-17, while IL-5 was rarely produced and not increased by any of the peptides (not shown). Both CD4+ and CD8+ T cells upregulated the activation marker CD69 (**Fig. 5f-h**), but many individual peptides

preferentially activated either subset. In contrast, T cells from healthy donors had little response to any of the neoantigen peptides (not shown). In the patient shown in **Fig. 5g** and **h**, peptides from the cryptic variants of *ABCG1*, *FCGRT, SERPING1, RSAD2*, *IFIT1, TBC1D15, MX1, IFI44L,* and *IFI44* stimulated CD8 T cells, while peptides from *SERPING1, IFIT1*, *IFI44L* and *IFI44* stimulated CD4 T cells. A statistical analysis based on T cell numbers in flow cytometry using χ^2 and adjusted for multiple comparisons, gave a highly significant p-values for many of these peptide responses. A complete list of all adjusted p-values for **Fig. 5h** is in **Supplemental Table S3**.

Since SLE neutrophils with cryptic splice variants have neoantigen peptides derived from them on their surface class I MHC, as evidenced by our immunopeptidomics experiments, we stained SLE neutrophils with a monoclonal antibody we recently developed recognizing the N-terminal knob exposed in 20–24 copies on the extracellular face of polymerized perforin pores generated by cytotoxic T cells. We have shown that this antibody detects poly-perforin pores on the surface of 15–45% of neutrophils freshly isolated from patients with rheumatoid arthritis as well as neutrophils treated with monomeric perforin. Indeed, 0–20% of SLE neutrophils contained dots (**Fig. 5i**) of the same size as rheumatoid arthritis neutrophils, while neutrophils from healthy donors did not (**Fig. 5j**), indicating that cytotoxic T cells have encountered antigen on their surface and attempted to kill them via perforin deposition. The percentage of SLE neutrophils with poly-perforin pores was lower than in rheumatoid arthritis (**Fig. 5k**).

These experiments demonstrate that neoantigen-specific T cells exist in SLE patients at frequencies that can be estimated to reach 1-15 per million T cells, which is compatible with the frequencies of tumor antigen-specific T cells[36]. The presence of IgG autoantibodies against the neo-sequences also indicates that some of the antigen-specific CD4 T cells provide T cell help to B cells, *i.e.,* they were either follicular helper T cells (Tfh) or peripheral helper T cells (Tph). The former are known to be expanded in SLE.

Discussion

Our finding that neoantigens are generated from aberrant mRNA splicing in SLE neutrophils introduces a potential new paradigm for the pathogenesis of this disease. We propose that the appearance of neoantigens, rather than loss of tolerance, is the driving factor behind the escalating immune activation that characterizes this disease. In this scenario, the immune system acts as it is programmed to do: making every effort to eliminate the source of 'foreign' neoantigens. This scenario can better explain many features of the disease than the 'loss of tolerance' concept, such as the waxing and waning course with sudden exacerbations (precipitated by variable neoantigen production), the broad and heterogenous range of organ manifestations and symptoms (likely guided by cell type(s) expressing neoantigens), and the strong immunosuppression required for treating active disease. This paradigm is also supported by the striking similarity in neutrophil expression profiles in SLE and SARS-CoV2 infected individuals[21]. Many of the reported shifts in immune cell linages reported in SLE also occur during viral infections. While none of the 15-mer peptides used for IgG autoantibody testing matched anything beyond 4 amino acid resides encoded by the human genome, the MX1 peptides 4 and 5 showed instances of identity with bacteriophage sequences above the 6-residue lower limit of an IgG epitope. This may be entirely co-incidental but illustrates that aberrantly produced novel amino acid sequences may, by chance, resemble viral or bacterial proteins against which immunity may even pre-exist in SLE patients.

The importance of neoantigens is well-recognized in anti-tumor immunity, which is pivotal for cancer immunotherapy. Anti-tumor immunity is a form of autoimmunity, in which altered self-peptides are recognized as foreign antigen by CD4 and CD8 T cells ultimately resulting in tumor killing by $CD8^+$ cytotoxic T cells that recognize neoantigen peptides presented on class I MHC molecule on tumor cells. It is by now well recognized that these tumor-specific peptides can be generated from point-mutations in proteins[37], from aberrant post-translational modifications, from intergenic translation products, or generated by aberrant mRNA splicing[3,38,39], with the last category yielding the largest repertoire of neoantigens. Our findings in SLE are similar in that aberrant transcript splicing is widespread, *i.e.,* detected in 34 of 35 patients, and results in novel amino acid sequences with

the capacity to yield neoantigens presented on class I and II MHC molecules. With our sequencing depth of 50 million, the estimated coverage of mRNA splicing events is approximately 60%. Hence, it is likely that many aberrant slicing events remained undetected in our analysis. Our mRNA splicing data also represents a single time-point, while the presence of IgG autoantibodies against the novel amino acid sequences indicates that patients had been exposed to them over an extended time before our sample acquisition. Hence, it is likely that aberrant mRNA splicing is more widespread than our findings show. It may also vary over time, perhaps related to exacerbations interspersed with periods of low disease activity.

Neoantigens presented on class I MHC on the surface of SLE neutrophils would make them vulnerable to killing by cytotoxic $CD8^+$ T cells, which are activated[40] but decreased in number[41] in SLE. Indeed, we detect poly-perforin pores on the surface of SLE neutrophils, which are known to be consumed at higher rates in SLE[15]. Perforin-mediated killing is more inflammatory than neutrophil apoptosis in healthy individuals and may include other programmed cell death pathways[11,13,42]. In SLE, cytotoxic effector T cells have a $CD38^{hi}CD127^-$ phenotype similar to that of CD8 T cells in check-point inhibitor-induced arthritis and viral infections[43], a phenotype that can be induced *in vitro* by type I interferons[43].

The processing of nascent transcripts begins during transcription and involves a large number of splicing factors, RNA molecules, enzymes, and post-translational modifications. Hence, a multitude of different disturbances could underpin our finding that SLE neutrophils from all but one patient harbor cryptic splice variant transcripts with the capacity to yield neoantigens displayed in HLA molecules. Several core mRNA splicing factors and RNA-binding proteins were either decreased (*SF3B1*[44], *SF3B5*, *SF3A2*, *U2AF2*, *SRSF7*, and *DDX46*[45]) or increased (*SNRNP27*) in neutrophils from SLE patients compared to healthy and RA neutrophils (**Supplemental Figs. S4** and **S5**). SF3B1 is notable because of its relatively frequent mutations in myeloid malignancies, resulting in altered 3' splice junctions[46], which we also observe in cryptic splice variants of *OAS1*, *RSAD2*, and *MX1* (**Fig. 2**). The Ro60-associated poly-U binding protein encoded by *PUF60*[47] was also decreased, while the RNA N6-adenosine-methyltransferase *METTL14*[48] and the N6-methyladenosine binding protein

YTHDF1[49] were increased. Together with METTL3, METTL4 catalyzes N6-methylation of a multitude of transcripts to regulate their stability and splicing[50]. We also reported recently[26] that the alternative splicing of *RBM10* and *RBM48* were skewed towards isoforms lacking RNA-binding domains and hence likely less to be fully functional. RBM48 plays an important role in the minor spliceosome.

While aberrant mRNA splicing is common in hematological malignancies[51] and solid tumors[52], the exact molecular mechanisms underlying them remain unclear except in cases of specific splicing factor mutations, such as SF3B1 K700E[46,53,54] in myeloid leukemias. Interestingly, SF3B1 expression was somewhat reduced (log2 fold -0.24, p_{adj}=0.045) in SLE neutrophils, as were four other mRNA splicing factors (SF3A2, U2AF2, DDX41, and DDX46). However, the splicing of nascent gene transcripts is a complex event involving over 200 proteins and U RNA components, making it very challenging to pin-point the cause of aberrant splicing, which may not be the same in every SLE patient.

Our results fit with prior studies that explored a link between neoantigens formed by somatic mutations in caner and autoimmunity. A study of scleroderma patients reported humoral and cellular immune responses against antigens that arose through somatic mutation in cancer cells in the *POLR3A* gene that encodes RNA polymerase III. These patients had scleroderma and antibodies against RNA polymerase III, while patients lacking this antibody did not have the *POLR3A* mutation.[55] In another example in a different autoimmune disease, patients with autoimmune myositis with TIF1g antibodies and cancer had somatic mutations and loss of heterozygosity in TIF1 genes.[56] These studies show that autoimmunity can arise from immune responses against mutated forms of canonical proteins in cancer cells. Our study shows that neoantigens triggering cellular and humoral immune responses can be generated through aberrant splicing in SLE, without the presence of highly mutated cancer cells.

Neutropoiesis includes an extensive burst of proliferation in the bone marrow, during which deficiencies in epigenetic regulation could be amplified to significant levels in the mature neutrophils. Decreased DNA methylation and epigenetic gene silencing are hallmarks of SLE[57,58], leading to increased transcription of many loci and genes that remain repressed in healthy controls, such as endogenous retroviruses, L1 retrotransposons, *Alu*s,

and other transposable elements. In healthy individuals, intronic transposable elements in nascent transcripts are prevented from acting as splice donor or acceptor sites by specific RNA binding proteins encoded by *HNRNPC* (Heterogeneous Nuclear Ribonucleoprotein C), *ELAVL1*[59], *TIA*, and *PTBP1*[60]. SLE neutrophils express reduced levels of *PTBP1* (log2 fold -0.44, p_{adj}=1.75 x 10^{-5}) and *ELAVL1* (log2 fold -0.27, p_{adj}=0.010). This may explain the aberrant splicing in of an inverted Alu in *FCGRT* (**Fig. 3e**), an inverted Alu in *IFIT1* (**Supplemental Fig. S2d**), and a HAL1 in *RSAD2* (**Supplemental Fig. S2b**) and other cases of cryptic exonization. Tumor neoantigens have also been reported to arise from splice junctions between exons and intronic transposable elements[61]. However, the majority of the cryptic splicing events in SLE neutrophils were exon skips, and multi-exon skips, alone or in combination with altered 5' and 3' splice junctions. These likely arise by different mechanisms than intronic exonization of transposable elements.

Another explanation for our findings should also be considered, namely that the SLE-specific pathology lies in the quality control mechanism of nonsense-mediated decay. In this scenario, naturally occurring errors of the splicing machinery accumulate because they remain undetected during translation and therefore can give rise to the peptides that we detect in anti-HLA immunoprecipitates. This possibility is compatible with the somewhat reduced expression of one of the two pivotal NMD proteins, *UPF1* (log2 fold -0.31, p_{adj}=0.041). However, *UPF1* expression is similarly reduced in RA neutrophils, which lack all of the cryptic splicing events seen in SLE neutrophils.

In either case, our main conclusion remains unchanged: SLE neutrophils produce neoantigens for which immunological tolerance does not exist. Whether these neoantigens are truly important in SLE pathogenesis can only be ascertained in clinical trials with agents that block their production. In support of an important role, recent work with anti-CD19 chimeric antigen receptor (CAR) T cells has demonstrated that B cell depletion needs to be complete[62] for an apparent cure of SLE to be achieved[63]. The post-treatment re-emergence of B cells represent newly developed naïve B cells, which consequently are tolerant of self-epitopes, including neoantigens present in the patient. These trials emphasize the importance of B cells (rather than only dendritic cells) in neoantigen presentation to T cells

and subsequent immunity against neoantigens derived from altered mRNA splicing in neutrophils.

Experimental Procedures

SLE patients. We obtained freshly drawn blood from adult (n=35) patients with SLE, healthy individuals (n=20), and patients with RA (n=15) to serve as disease controls. Study participants were 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.

The 35 SLE patients included in our study represent the typical spectrum of SLE disease seen in an academic medical center with standard-of-care medication (**Supplemental Table S1**). 11 patients had active disease, defined as having an SLE disease activity index (SLEDAI ≥ 6), while 24_ had inactive disease with a SLEDAI < 6 at the time of venipuncture. Based on the 95th percentile of the expression of 111 interferon-inducible genes in healthy donors, 26 were considered interferon signature positive. All healthy donors and RA patients were interferon signature negative.

Isolation of neutrophils and peripheral blood mononuclear cells. Polymorphonuclear (PMN) cells were isolated from freshly drawn venous blood (within 1 h 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 Sequencing. Extraction of RNA was performed using Quick-RNA Whole Blood kit (Zymo Research, #R1201). Cells were lysed in DNA/RNA Shield and RNAs were isolated following manufacturer's instruction and eluted in RNase-free water. RNA was sent to the UW Northwest Genomics core for strand-specific paired-end RNA sequencing at 50 million 101-bp reads per sample.

Alignment to reference genome and alternative splicing detection. After QC of sequencing quality, FASTQ files were aligned to the reference genome (GRCh38) using STAR (v. 2.7.9a)[64] with two-pass mode which improves detection of reads mapping to novel splicing junctions. After post-alignment QC, sorted and indexed BAM files were used to

identify alternative splicing events by SplAdder (v. 3.0.3)[65] 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 6 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 4 levels of confidence; 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.

Discovery and translation of cryptic alternative splicing. We detected cryptic alternative splicing in SLE, defined as splicing events with read support in SLE neutrophils and either absence or rare read support in healthy and RA disease controls. We used our established methodology (RiboSplitter[25]) to translate splicing events into peptide sequences to allow downstream analysis and validation. Splicing events with valid unambiguous translations were used for downstream analyses.

Because an alternative splicing event includes two isoforms, we identified cryptic splicing events (both isoforms are cryptic), as well as cryptic splicing isoforms (one of the event's isoforms is cryptic). We used strict filtering criteria as follows: first, we required cryptic events and isoforms to have ≥10 read support in SLE and no more than two read support in every healthy or RA sample; or ≥5 reads in SLE and zero reads in every healthy or RA sample. Second, we excluded annotated splicing events and isoforms based on the Hg38 reference. Third, cryptic peptides were excluded if a match existed among known annotated proteins of the gene. These filtering criteria add specificity in detecting events that are cryptic and unique to SLE neutrophils.

We consistently calculated PSI values for isoform 2, so that a higher PSI always indicates higher proportion of isoform 2, which makes it easier to interpret which group has more of which isoform.

MHC affinity prediction. We identified each SLE patient's class I and class II MHC haplotypes from RNA-seq data using arcasHLA (v. 0.5.0), which was shown to be highly accurate (100% accurate for class I MHC and >99.7% accurate for class II MHC).[66]

MHC affinity binding was predicted for cryptic SLE peptides using NetMHCpan-4.1 (mers length=8-11) and NetMHCIIpan-4.0 [27] (mers length=12-18) that use neural networks trained on binding affinity as well as mass spectrometry data. Strong affinity binding was defined as percent rank of predicted affinity of 2% or lower (compared to a set of 100,000 random natural peptides) with predicted binding affinity <500.

HLA immunoprecipitation. Peptides that bind to class I MHC and class II MHC were immunoprecipitated using a method described previously (Ref: Comparison of the MHC I Immunopeptidome repertoire of B-cell lymphoblasts using two isolation methods. Proteomics. 2018 Jun; 18 (12):e1700251. Doi: 10.1002/pmic.201700251). Briefly, HLA-A B C (W6/32), HLA-DR (L243), or isotype antibody was conjugated to protein A magnetic beads (Millipore-Sigma) at 1mg/ml ratio. Cell lysate from 10 million cells was mixed with 100 ul of anti-HLA beads (or isotype control beads). After washing, bound peptides were eluted with 0.2% TFA followed by a C-18 column clean-up and further analyzed by LC-MS.

LC-MS and analysis. Mass spectrometry was performed at the University of Washington's Proteomics Resource (UWPR). Peptide separation was performed on a Thermo EASY-nLC system using a 120-min gradient at 300 nL/min. Mobile phase B was increased from 6% to 45% over 89 min, followed by a ramp to 100% B and column re-equilibration. Samples were loaded at 2.5 µL/min. Mass spectrometric analysis was carried out on an Orbitrap Exploris 480 operated in positive ion mode. The full MS scans were acquired at 120,000 resolution (at m/z 200) across an m/z range of 375–1575, with a normalized AGC target of 200%. Targeted parallel reaction monitoring (PRM) MS/MS scans were acquired at 30,000 resolution using HCD fragmentation with a normalized collision energy of 30%. The isolation window was set to 1.6 m/z. Scheduled acquisition windows of 0–120 min was used for targeted precursor ions listed in the inclusion list. Individual peptide was further analyzed in Skyline.

Mass Spectra Visualization and Processing. The Skyline Chromatogram Data was imported into and visualized with Skyline (64-bit) 24.1.0.199. Data were filtered to include only precursors with an isotope dot product (idotp) ≥ 0.85 and mass error < 10 ppm. Then the match of peak shape and retention time of the spectra from the daughter ions to those from precursors were visually inspected and confirmed.

ELISA of neoantigen detection. Selected neoantigen peptides synthesized by Shanghai RoyoBiotechnology (China) were coated onto medium-binding EIA plates at 1 ng/ul in 0.1M bicarbonate coating buffer, pH 9.5. Plates were then blocked in 2% BSA in PBS. Sera from SLE patients and healthy donors were added to the blocked wells followed by anti-human IgG HRP. TMB substrates were added to the wells, and the reaction was stopped by 2N sulfuric acid solution. OD450nm and OD630nm were measured by BioTek Synergy 2 plate reader.

Gene expression. We quantified gene expression using Rsubread featurecounts[67] (v. 2.12.3). We detected differentially expressed genes and calculated gene expression normalized counts using DESeq2 (v. 1.36.0)[68]. The average expression of 111 interferon-induced genes (ISGs)[69] was calculated per sample and SLE samples were considered having high interferon signatures when the average expression of ISGs was higher than the 95% percentile of ISG expression in HC.

Comparison to PBMC splicing. To test if cryptic splicing events that we discovered in SLE neutrophils were also seen in PBMC, we used two publicly available RNA-seq data of PBMC from SLE patients. One study (GSE122459)[70] included 20 SLE patients and 6 healthy controls, and another study (GSE211700)[71] included 20 SLE patients and 10 healthy controls. We ran the same analysis pipeline as described above to align the data and detect alternative splicing events. Then we searched for neutrophil SLE cryptic splicing in the PBMC datasets.

T-Cell Activation Assay. PBMCs (2×10^6 per well) from SLE patients were cultured in 24-well plates in RPMI 1640 medium (Gibco; Cat. No. 11-875-093) supplemented with 5% autologous patient serum that was syringe filtered (0.22 μm), heat-inactivated at 56 °C for 20 min, and re-filtered (0.22 μm). Cells were stimulated with 5 μg/ml peptide pools

containing three to six neopeptides and supplemented with IL-2 (10 U/ml; R&D, Cat. No. BT-002-050) on days 4 and 9. Positive controls consisted of stimulation with phytohemagglutinin-M (10 μg/ml; Sigma, Cat. No. NC2034236) on day 1, followed by restimulation with CD3/CD28 Dynabeads (Gibco; Cat. No. 11131D) on day 14. On day 14, PBMCs were restimulated with individual peptides from each pool, and ELISPOT plates were coated with capture antibody for overnight incubation. Following 18–24 h coculture, triplicate aliquots of 3×10^5 cells were resuspended in CTL medium supplemented with 1% L-glutamine (Gibco; Cat. No. 2503008), plated into coated ELISPOT wells, and incubated for 4 h at 37 °C. Spots were developed using the CTL-Human IFN-γ/IL-17 Double-Color ELISPOT kit according to the manufacturer's instructions and quantified with a CTL S6 Universal M2 reader. The remaining cells were collected for flow cytometry analysis.

Flow Cytometry Following 14-day stimulation, cells were transferred to a 96 well round bottom plate and .washed with 1x phosphate buffered saline (PBS). To prevent non-specific binding, cells were resuspended in 1x PBS with 1% bovine serum albumin (BSA) and incubated at room temperature for 20mins. Cells were then stained with a cocktail of fluorochrome-labeled mAbs, including: CD3 conjugated to Alex Fluor 488 (biolegend, catalog #300320) at 1:50, CD4 BV510 (biolegend, catalog # 317444) at 1:50, CD8 PE-CY7 (biolegend, catalog # 344712) at 1:50 and CD69 APC (biolegend, catalog # 985206) at 1:200 in 1xPBS with 1% BSA for 30mins at 4 °C in the dark. Flow cytometry was run on CytoFlex Flow Cytometer (Beckman Coulter) and analysis was done with FlowJo (Becton, Dickinson and Company).

Immunofluorescence microscopy. Neutrophils from healthy donors, RA, and SLE patients were seeded onto glass-bottom 96-well plates and incubated at 37 °C for 30 minutes to adhere to the plate. Cells were then fixed with 4% paraformaldehyde for 15 minutes, followed by three PBS washes and blocked with 10% normal goat serum overnight at 4 °C. Neutrophils were incubated with perforin 4G10 antibody (1:20 dilution) overnight, washed three times with PBS, and then stained with goat anti-mouse IgG AF647-conjugated secondary antibody (ThermoFisher Scientific, #A-21236; 1:1000 dilution) and DAPI (ThermoFisher Scientific, #62248) for 45 minutes. After three additional PBS washes, cells

were resuspended in PBS. Images were captured using a Keyence BZ-X810 fluorescence microscope and analyzed with Keyence Analyzer software. Control conditions included cells stained with primary antibody only and secondary antibody only in healthy, RA ,and SLE neutrophils.

DATA ACCESS

All data used in the analysis will be made available for research use. We are also committing to deposit our RNAseq data set in a public data base upon acceptance of our paper. RiboSplitter is publicly available (github.com/R-Najjar/RiboSplitter)

COMPETING INTEREST STATEMENT

TM reports consulting fees from Cugene, Evolved, Codify, and Rome Therapeutics. He serves on the Scientific Advisory Board of the Tri-Institutional Drug Discovery Institute (non-profit), and the University of Copenhagen LEO Skin Immunology Centre (non-profit).

ACKNOWLEDGMENTS

We thank the patients who donated blood for this research. This work was supported by NIH grants AR075134, AR077266, AR074939, and AR081654 to T.M. and AR007108 to R.N.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version. Drs. Mustelin and Najjar had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. SJVB consented patients for research blood samples.

XW, FS, EL, and CM purified neutrophils and PBMC. RN created the analysis pipeline and carried out all of the computational analyses with NW and VG. XW and CM performed the ELISAs. XW isolated RNA and carried out molecular biology and proteomics experiments. JA provided in-depth mass spectrometry analysis. FS and EL performed the T cell experiments with expert advice by SPS, who also participated in data analysis and interpretation. TM supervised the study and wrote the first version of the manuscript together with RN.

1 Park, E., Pan, Z., Zhang, Z., Lin, L. & Xing, Y. The Expanding Landscape of Alternative Splicing Variation in Human Populations. *Am J Hum Genet* **102**, 11-26 (2018). https://doi.org/10.1016/j.ajhg.2017.11.002

2 Lu, S. X. *et al*. Pharmacologic modulation of RNA splicing enhances anti-tumor immunity. *Cell* **184**, 4032-4047 e4031 (2021). https://doi.org/10.1016/j.cell.2021.05.038

3 Li, G. *et al*. Splicing neoantigen discovery with SNAF reveals shared targets for cancer immunotherapy. *Sci Transl Med* **16**, eade2886 (2024). https://doi.org/10.1126/scitranslmed.ade2886

4 Cooper, G. S. *et al*. Hormonal, environmental, and infectious risk factors for developing systemic lupus erythematosus. *Arthritis Rheum* **41**, 1714-1724 (1998). https://doi.org/10.1002/1529-0131(199810)41:10<1714::AID-ART3>3.0.CO;2-U

5 Shipman, L. Systemic lupus erythematosus: New GWAS loci and insights into ancestry. *Nat Rev Rheumatol* **12**, 499 (2016). https://doi.org/10.1038/nrrheum.2016.128

6 Ghodke-Puranik, Y., Olferiev, M. & Crow, M. K. Systemic lupus erythematosus genetics: insights into pathogenesis and implications for therapy. *Nat Rev Rheumatol* **20**, 635-648 (2024). https://doi.org/10.1038/s41584-024-01152-2

7 Barnett, E. V. Specificity of anti--double-stranded DNA antibodies for the diagnosis of systemic lupus erythematosus. *Arthritis Rheum* **24**, 1452-1453 (1981). https://doi.org/10.1002/art.1780241125

8 Dorner, T. & Lipsky, P. E. The essential roles of memory B cells in the pathogenesis of systemic lupus erythematosus. *Nat Rev Rheumatol* **20**, 770-782 (2024). https://doi.org/10.1038/s41584-024-01179-5

9 Robinson, S. & Thomas, R. Potential for Antigen-Specific Tolerizing Immunotherapy in Systematic Lupus Erythematosus. *Front Immunol* **12**, 654701 (2021). https://doi.org/10.3389/fimmu.2021.654701

10 Rodriguez, M. A. & Blasini, A. M. Just Autoimmunity? The Role of the Innate Immune Response in Lupus. *J Clin Rheumatol* **31**, 71-77 (2025). https://doi.org/10.1097/RHU.0000000000002209

11 Knight, J. S. & Kaplan, M. J. Lupus neutrophils: 'NET' gain in understanding lupus pathogenesis. *Curr Opin Rheumatol* **24**, 441-450 (2012). https://doi.org/10.1097/BOR.0b013e3283546703

12 Lorenz, G. & Anders, H. J. Neutrophils, Dendritic Cells, Toll-Like Receptors, and Interferon-alpha in Lupus Nephritis. *Semin Nephrol* **35**, 410-426 (2015). https://doi.org/10.1016/j.semnephrol.2015.08.003

13 Smith, C. K. & Kaplan, M. J. The role of neutrophils in the pathogenesis of systemic lupus erythematosus. *Curr Opin Rheumatol* **27**, 448-453 (2015). https://doi.org/10.1097/BOR.0000000000000197

14 Tay, S. H., Celhar, T. & Fairhurst, A. M. Low-Density Neutrophils in Systemic Lupus Erythematosus. *Arthritis Rheumatol* **72**, 1587-1595 (2020). https://doi.org/10.1002/art.41395

15 Bennett, L. *et al*. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med* **197**, 711-723 (2003).

16 Yamamoto, T. Role of neutrophils in cutaneous lupus erythematosus. *J Dermatol* **51**, 180-184 (2024). https://doi.org/10.1111/1346-8138.17036

17 Ma, S., Jiang, W., Zhang, X. & Liu, W. Insights into the pathogenic role of neutrophils in systemic lupus erythematosus. *Curr Opin Rheumatol* **35**, 82-88 (2023). https://doi.org/10.1097/BOR.0000000000000912

18 Fresneda Alarcon, M., McLaren, Z. & Wright, H. L. Neutrophils in the Pathogenesis of Rheumatoid Arthritis and Systemic Lupus Erythematosus: Same Foe Different M.O. *Front Immunol* **12**, 649693 (2021). https://doi.org/10.3389/fimmu.2021.649693

19 Li, P. *et al*. Glutathione peroxidase 4-regulated neutrophil ferroptosis induces systemic autoimmunity. *Nat Immunol* **22**, 1107-1117 (2021). https://doi.org/10.1038/s41590-021-00993-3

20 Ng, B. *et al*. Increased noncanonical splicing of autoantigen transcripts provides the structural basis for expression of untolerized epitopes. *J Allergy Clin Immunol* **114**, 1463-1470 (2004). https://doi.org/10.1016/j.jaci.2004.09.006

21 Najjar, R. *et al*. Large overlap in neutrophil transcriptome between lupus and COVID-19 with limited lupus-specific gene expression. *Lupus Sci Med* **11** (2024). https://doi.org/10.1136/lupus-2023-001059

22 Fair, B. *et al*. Global impact of unproductive splicing on human gene expression. *Nat Genet* **56**, 1851-1861 (2024). https://doi.org/10.1038/s41588-024-01872-x

23 Bachmann, M. P. *et al*. Autoimmunity as a result of escape from RNA surveillance. *J Immunol* **177**, 1698-1707 (2006). https://doi.org/10.4049/jimmunol.177.3.1698

24 Gomez-Banuelos, E. *et al*. Alternative exon usage in TRIM21 determines the antigenicity of Ro52/TRIM21 in systemic lupus erythematosus. *JCI Insight* **7** (2022). https://doi.org/10.1172/jci.insight.163795

25 Najjar, R. & Mustelin, T. Prediction of alternative pre-mRNA splicing outcomes. *Sci Rep* **13**, 20000 (2023). https://doi.org/10.1038/s41598-023-47348-6

26 Najjar, R. *et al*. Altered Protein Structures and Neoepitopes in Lupus Neutrophils From Dysregulated Splicing of Messenger RNA. *ACR Open Rheumatol* **7**, e11770 (2025). https://doi.org/10.1002/acr2.11770

27 Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. 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* **48**, W449-W454 (2020). https://doi.org/10.1093/nar/gkaa379

28 Purcell, A. W., Ramarathinam, S. H. & Ternette, N. Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics. *Nat Protoc* **14**, 1687-1707 (2019). https://doi.org/10.1038/s41596-019-0133-y

29 Pastor, F., Kolonias, D., Giangrande, P. H. & Gilboa, E. Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. *Nature* **465**, 227-230 (2010). https://doi.org/10.1038/nature08999

30 Tan, K., Stupack, D. G. & Wilkinson, M. F. Nonsense-mediated RNA decay: an emerging modulator of malignancy. *Nat Rev Cancer* **22**, 437-451 (2022). https://doi.org/10.1038/s41568-022-00481-2

31 Lykke-Andersen, S. & Jensen, T. H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* **16**, 665-677 (2015). https://doi.org/10.1038/nrm4063

32 Brogna, S. & Wen, J. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol* **16**, 107-113 (2009). https://doi.org/10.1038/nsmb.1550

33 Bao, W. & Jurka, J. Origin and evolution of LINE-1 derived "half-L1" retrotransposons (HAL1). *Gene* **465**, 9-16 (2010). https://doi.org/10.1016/j.gene.2010.06.005

34 Blencowe, B. J., Ahmad, S. & Lee, L. J. Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes. *Genes Dev* **23**, 1379-1386 (2009). https://doi.org/10.1101/gad.1788009

35 Tsoy, O. *et al*. RNA sequencing depth guidelines for the study of alternative splicing *bioRxiv* (2024). https://doi.org/https://doi.org/10.1101/2024.10.09.617406

36 Miller, A. M. *et al*. A functional identification platform reveals frequent, spontaneous neoantigen-specific T cell responses in patients with cancer. *Sci Transl Med* **16**, eabj9905 (2024). https://doi.org/10.1126/scitranslmed.abj9905

37 Hsiue, E. H. *et al*. Targeting a neoantigen derived from a common TP53 mutation. *Science* **371** (2021). https://doi.org/10.1126/science.abc8697

38 Lang, F. *et al*. Prediction of tumor-specific splicing from somatic mutations as a source of neoantigen candidates. *Bioinform Adv* **4**, vbae080 (2024). https://doi.org/10.1093/bioadv/vbae080

39 Ji, M. *et al*. Long-range alternative splicing contributes to neoantigen specificity in glioblastoma. *Brief Bioinform* **25** (2024). https://doi.org/10.1093/bib/bbae503

40 Fujita, Y. *et al*. Association of peripheral CD8(+) T cell activation with disease activity and treatment resistance in systemic lupus erythematosus. *RMD Open* **11** (2025). https://doi.org/10.1136/rmdopen-2024-005122

41 Lao, J., Huang, R., Wu, R. & Yuan, Y. Changes of Peripheral T Cells in Systemic Lupus Erythematosus Patients. *Immun Inflamm Dis* **13**, e70156 (2025). https://doi.org/10.1002/iid3.70156

42 Garcia-Romo, G. S. *et al*. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. *Sci Transl Med* **3**, 73ra20 (2011). https://doi.org/10.1126/scitranslmed.3001201

43 Wang, R. *et al*. Clonally expanded CD38(hi) cytotoxic CD8 T cells define the T cell infiltrate in checkpoint inhibitor-associated arthritis. *Sci Immunol* **8**, eadd1591 (2023). https://doi.org/10.1126/sciimmunol.add1591

44 Bak-Gordon, P. & Manley, J. L. SF3B1: from core splicing factor to oncogenic driver. *RNA* **31**, 314-332 (2025). https://doi.org/10.1261/rna.080368.124

45 Yang, F. *et al*. Mechanisms of the RNA helicases DDX42 and DDX46 in human U2 snRNP assembly. *Nat Commun* **14**, 897 (2023). https://doi.org/10.1038/s41467-023-36489-x

46 Darman, R. B. *et al*. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Rep* **13**, 1033-1045 (2015). https://doi.org/10.1016/j.celrep.2015.09.053

47 Hsiao, H. T. *et al*. Unraveling the mechanism of recognition of the 3' splice site of the adenovirus major late promoter intron by the alternative splicing factor PUF60. *PLoS One* **15**, e0242725 (2020). https://doi.org/10.1371/journal.pone.0242725

48 Zhang, P. *et al*. The m6A methyltransferase METTL14 promotes oncogenic Kras induced juvenile myelomonocytic leukemia through dysregulating autophagy. *Cell Death Differ* (2025). https://doi.org/10.1038/s41418-025-01561-0

49 Wang, C. *et al*. YTHDF1 promotes pancreatic cancer cell progression by enhancing SF3B2 translation though m6A modification. *J Biochem* **177**, 425-435 (2025). https://doi.org/10.1093/jb/mvaf018

50 Ramasamy, D. *et al*. Transcriptome-wide profiling identifies colon cancer-associated m6A transcripts and potential RNA methyl modifiers. *Mol Biol Rep* **51**, 299 (2024). https://doi.org/10.1007/s11033-024-09217-x

51 Yoshimi, A. *et al*. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. *Nature* **574**, 273-277 (2019). https://doi.org/10.1038/s41586-019-1618-0

52 Kahles, A. *et al*. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* **34**, 211-224 e216 (2018). https://doi.org/10.1016/j.ccell.2018.07.001

53 Inoue, D. *et al*. Spliceosomal disruption of the non-canonical BAF complex in cancer. *Nature* **574**, 432-436 (2019). https://doi.org/10.1038/s41586-019-1646-9

54 Madan, V. *et al*. Distinct and convergent consequences of splice factor mutations in myelodysplastic syndromes. *Am J Hematol* **95**, 133-143 (2020). https://doi.org/10.1002/ajh.25673

55 Joseph, C. G. *et al*. Association of the autoimmune disease scleroderma with an immunologic response to cancer. *Science* **343**, 152-157 (2014). https://doi.org/10.1126/science.1246886

56 Pinal-Fernandez, I. *et al*. Tumour TIF1 mutations and loss of heterozygosity related to cancer-associated myositis. *Rheumatology (Oxford)* **57**, 388-396 (2018). https://doi.org/10.1093/rheumatology/kex413

57 Zhou, X., Zhou, S. & Li, Y. An updated review on abnormal epigenetic modifications in the pathogenesis of systemic lupus erythematosus. *Front Immunol* **15**, 1501783 (2024). https://doi.org/10.3389/fimmu.2024.1501783

58 Araki, Y. & Mimura, T. Epigenetic Dysregulation in the Pathogenesis of Systemic Lupus Erythematosus. *Int J Mol Sci* **25** (2024). https://doi.org/10.3390/ijms25021019

59 Rothamel, K. *et al.* ELAVL1 primarily couples mRNA stability with the 3' UTRs of interferon-stimulated genes. *Cell Rep* **35**, 109178 (2021). https://doi.org/10.1016/j.celrep.2021.109178

60 Sibley, C. R., Blazquez, L. & Ule, J. Lessons from non-canonical splicing. *Nat Rev Genet* **17**, 407-421 (2016). https://doi.org/10.1038/nrg.2016.46

61 Burbage, M. *et al.* Epigenetically controlled tumor antigens derived from splice junctions between exons and transposable elements. *Sci Immunol* **8**, eabm6360 (2023). https://doi.org/10.1126/sciimmunol.abm6360

62 Tur, C. *et al.* CD19-CAR T-cell therapy induces deep tissue depletion of B cells. *Ann Rheum Dis* **84**, 106-114 (2025). https://doi.org/10.1136/ard-2024-226142

63 Taubmann, J. *et al.* CD19 Chimeric Antigen Receptor T Cell Treatment: Unraveling the Role of B Cells in Systemic Lupus Erythematosus. *Arthritis Rheumatol* **76**, 497-504 (2024). https://doi.org/10.1002/art.42784

64 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013). https://doi.org/10.1093/bioinformatics/bts635

65 Kahles, A., Ong, C. S., Zhong, Y. & Ratsch, G. SplAdder: identification, quantification and testing of alternative splicing events from RNA-Seq data. *Bioinformatics* **32**, 1840-1847 (2016). https://doi.org/10.1093/bioinformatics/btw076

66 Orenbuch, R. *et al.* arcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics* **36**, 33-40 (2020). https://doi.org/10.1093/bioinformatics/btz474

67 Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**, e47 (2019). https://doi.org/10.1093/nar/gkz114

68 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014). https://doi.org/10.1186/s13059-014-0550-8

69 Chiche, L. *et al.* Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. *Arthritis Rheumatol* **66**, 1583-1595 (2014). https://doi.org/10.1002/art.38628

70 Tokuyama, M. *et al.* ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc Natl Acad Sci U S A* **115**, 12565-12572 (2018). https://doi.org/10.1073/pnas.1814589115

71 Zou, X. *et al.* Uncovering lupus nephritis-specific genes and the potential of TNFRSF17-targeted immunotherapy: a high-throughput sequencing study. *Front Immunol* **15**, 1303611 (2024). https://doi.org/10.3389/fimmu.2024.1303611

Table I. Class I MHC bound peptides derived from neo-amino acid sequences created by cryptic mRNA splicing detected by mass spectrometry

peptide	sample_id	Idotp score	event_id	gene_id	iso_reads	HLA_allele	HLA_percentile
GLHEDGVLMEM	SLE1466	0.99	mult_exon_skip.22553	*USP18*	12	HLA-B*37:01	4.7
LTPRNRYSIC	SLE1483	0.99	mult_exon_skip.1194	*IFI44L*	22	HLA-B*08:01	1.6
SLLSVPPTT	SLE1483	0.97	mult_exon_skip.23428	*OGG1*	20	HLA-A*02:01	0.93
SIRTNRELA	SLE1483	0.97	mult_exon_skip.22400	*TRAPPC10*	16	HLA-B*08:01	3.2
KVFGYSPLF	SLE1419	0.97	exon_skip.39005	*OAS3*	22	HLA-B*13:01	0.06
GLASVATAPF	SLE1419	0.97	exon_skip.60589	*XAF1*	12	HLA-B*13:01	0.51
NLEWKETGF	SLE1422	0.97	exon_skip.74029	*FCGRT*	41	HLA-C*08:02	4.4
LTLADILYKAG	SLE1422	0.94	exon_skip.107132	*ANXA3*	283	HLA-A*02:01	4.7
AGRRKEEEIL	SLE1422	0.93	mult_exon_skip.22337	*MX1*	28	HLA-B*14:02	26
LPQHHEGL	SLE1490	0.93	exon_skip.93637	*ABCG1*	12	HLA-B*07:02	0.88
TEIPFCYFVP	SLE1466	0.93	exon_skip.76299	*RSAD2*	11	HLA-B*37:01	1.6
PADPTGHLL	SLE1466	0.92	mult_exon_skip.15426	*MYO15B*	12	HLA-A*01:01	2.1
FQAGCHTQL	SLE1419	0.92	exon_skip.41668	*PHF11*	14	HLA-C*03:04	0.13
VFGYSPLF	SLE1419	0.91	exon_skip.39005	*OAS3*	22	HLA-C*14:02	0.72
NLAPLDVCFLL	SLE1419	0.91	exon_skip.38909	*OAS1*	77	HLA-A*02:01	0.69
SHLCLRLFPA	SLE1466	0.91	alt_3prime.3817	*ADAMTSL4*	10	HLA-B*08:01	2
FLLTGDVL	SLE1490	0.91	exon_skip.38909	*OAS1*	15	HLA-A*02:01	2
LLDLGSSSLMM	SLE1483	0.9	mult_exon_skip.23428	*OGG1*	20	HLA-A*01:01	0.97
KEAVKPSF	SLE1466	0.9	exon_skip.76299	*RSAD2*	11	HLA-B*37:01	1.4
TPRNRYSI	SLE1483	0.89	mult_exon_skip.1194	*IFI44L*	22	HLA-B*08:01	0.09
LAPLDVCFLL	SLE1419	0.89	exon_skip.38909	*OAS1*	77	HLA-B*51:01	0.26
SIRTNREL	SLE1483	0.89	mult_exon_skip.22400	*TRAPPC10*	16	HLA-B*08:01	1.3
NLAPLDVCFL	SLE1419	0.88	exon_skip.38909	*OAS1*	77	HLA-A*02:01	0.47
LLDLGSSSL	SLE1483	0.88	mult_exon_skip.23428	*OGG1*	20	HLA-A*02:01	0.61

Fig. 1

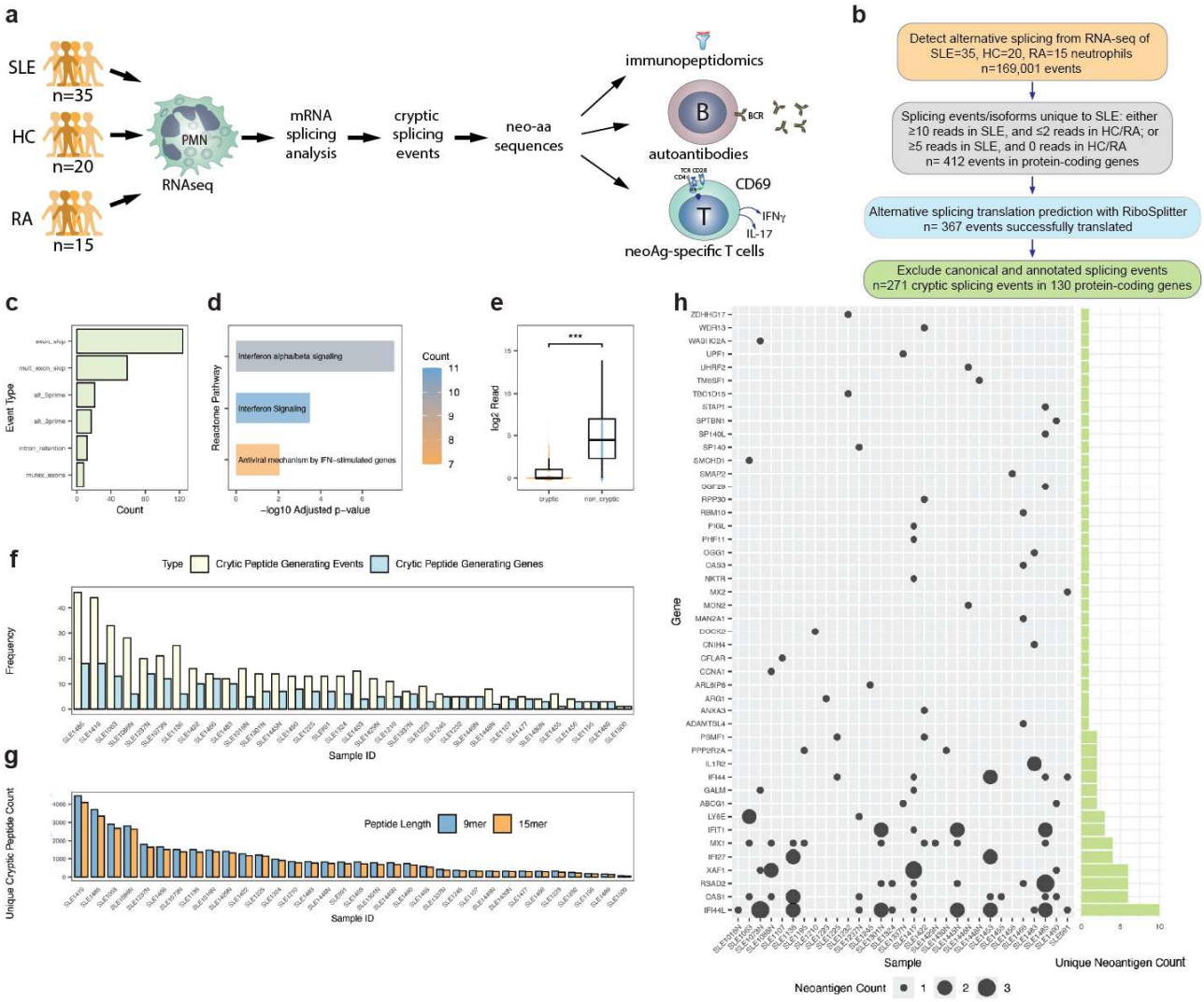


Fig. 1 | Cryptic RNA splicing events in SLE neutrophils. a, Schematic overview of our experimental approach. **b**, Our computational pipeline. **c**, Categories of the cryptic splicing events. **d,** Pathway analysis of the aberrantly spliced genes. **e**, Average read counts for cryptic splicing events that generate novel amino acid sequences compared to all splicing events in the same gene transcripts. **f**, Number of cryptic splicing events and genes they affected in the individual SLE patients. **g**, Total number of 9mer of 15mer peptides the cryptic transcripts theoretically could generate. **h**, in-silico predicted neoantigens in genes subjected to cryptic splicing in each patient with the number of cryptic events in each. Right side, the number of predicted neoantigen peptides with high-affinity binding to their own haplotype of class I or II molecules.

Fig. 2

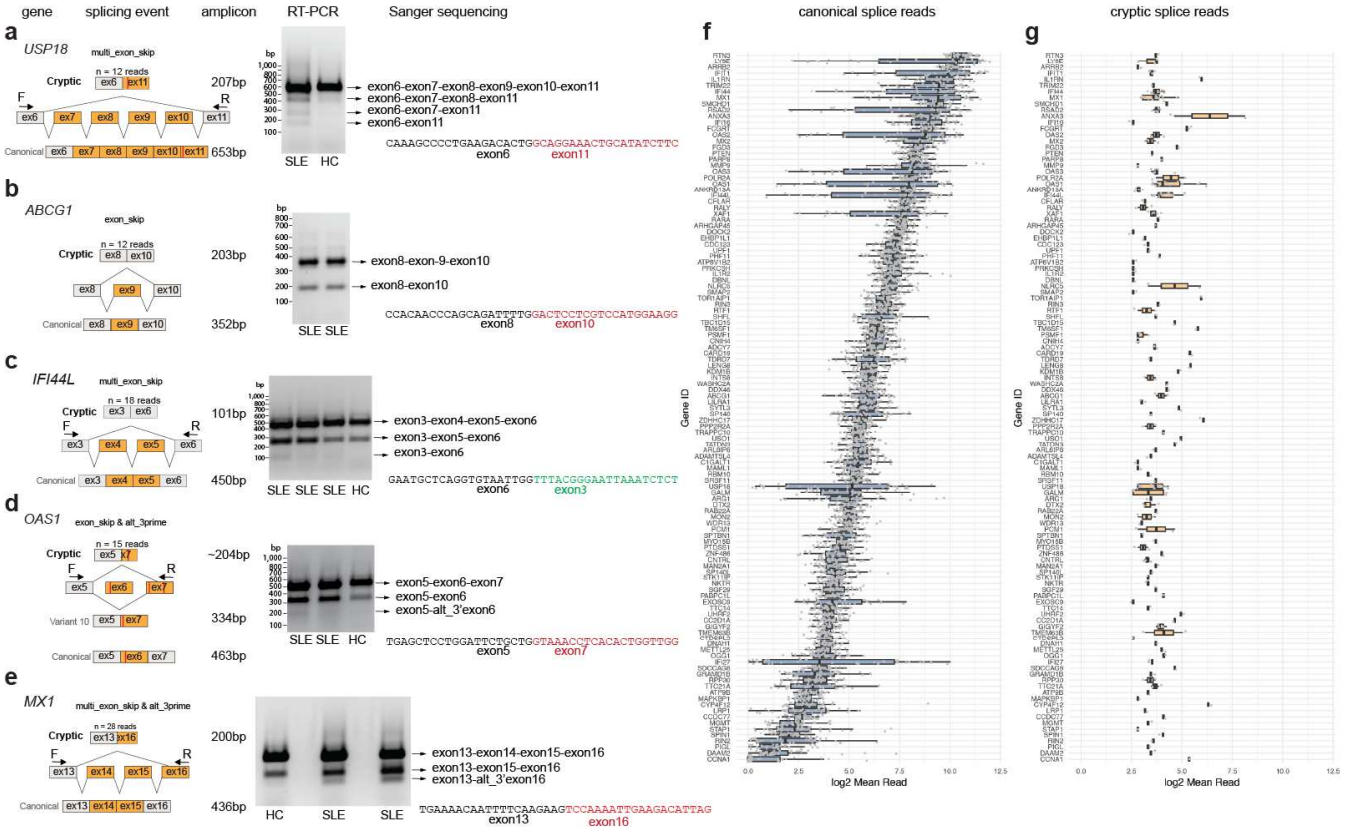


Fig. 2 | Validation of a set of cryptic splicing by RT-PCR and sequencing. The location of forward (F) and reverse (R) PCR primers to amplify diagnostic-sized fragments of 5 cryptic mRNA splicing events (left side) in: **a**, *USP18*, **b**, *ABCG1*, **c**, *IFI44L,* **d**, *OAS1,* and **e**, *MX1.* The PCR reaction products (middle) with the spliced exons in each band indicated and the sequence of the band corresponding to the cryptic variant. **f**, average reads of canonical splice junctions (left panel) compared to **g**, the read counts if the cryptic splice junctions in the same genes.

Fig. 3

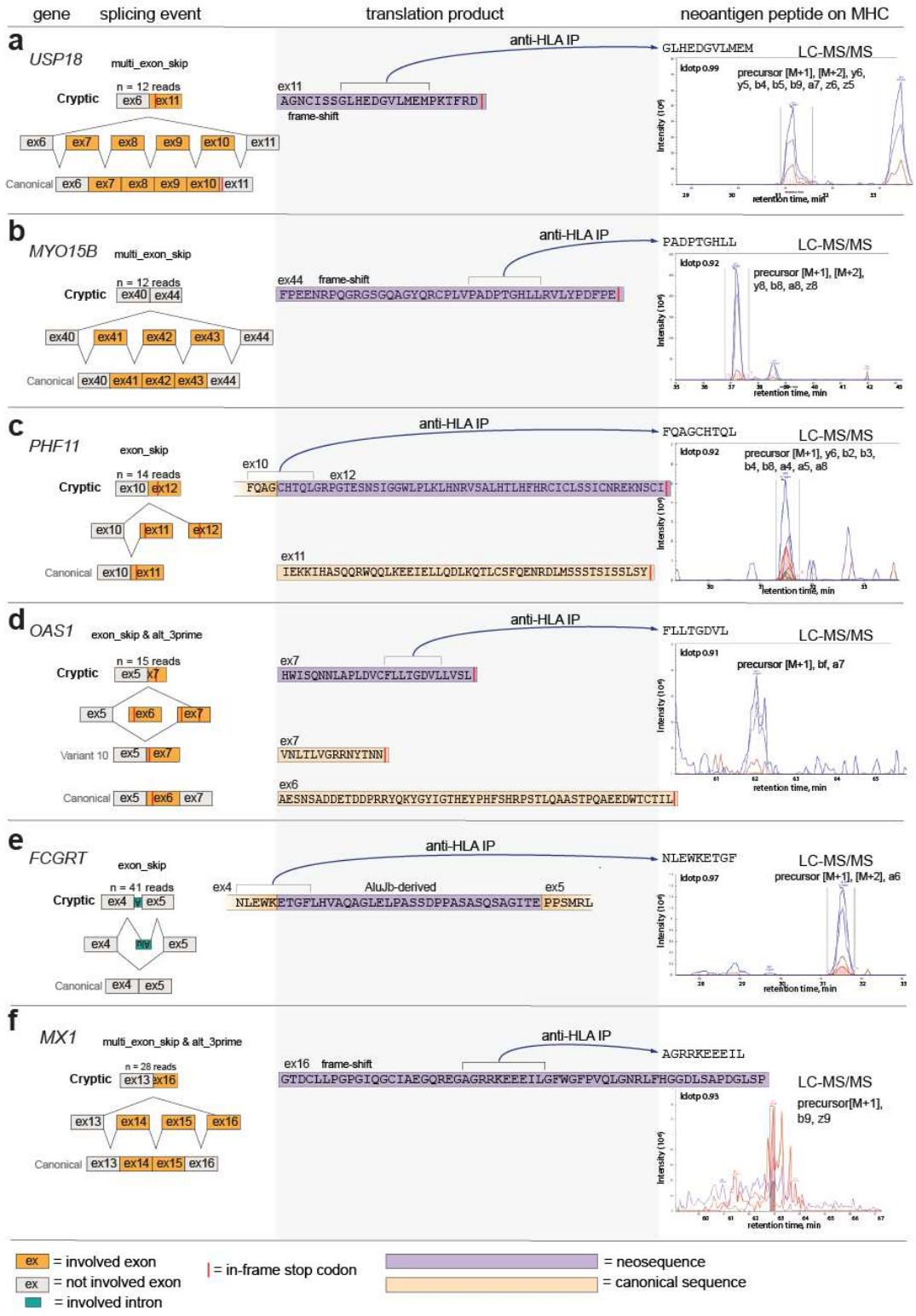

gene
splicing event
translation product
neoantigen peptide on MHC
a
USP18
multi_exon_skip
Cryptic
n = 12 reads
ex6 ex11
ex6 ex7 ex8 ex9 ex10 ex11
Canonical ex6 ex7 ex8 ex9 ex10 ex11
anti-HLA IP
ex11
AGNCISSGLHEDGVLMEMPKTFRD
frame-shift
GLHEDGVLMEM
LC-MS/MS
Idotp 0.99
precursor [M+1], [M+2], y6, y5, b4, b5, b9, a7, z6, z5
Intensity (10⁶)
retention time, min
b
MYO15B
multi_exon_skip
Cryptic
n = 12 reads
ex40 ex44
ex40 ex41 ex42 ex43 ex44
Canonical ex40 ex41 ex42 ex43 ex44
anti-HLA IP
ex44 frame-shift
FPEENRPQGRGSGQAGYQRCPLVPADPTGHLLRVLYPDFPE
PADPTGHLL
LC-MS/MS
Idotp 0.92
precursor [M+1], [M+2], y8, b8, a8, z8
Intensity (10⁶)
retention time, min
c
PHF11
exon_skip
Cryptic
n = 14 reads
ex10 ex12
ex10 ex11 ex12
Canonical ex10 ex11
anti-HLA IP
ex10
ex12
FQAGCHTQLGRPGTESNSIGGWLPLKLHNRVSALHTLHFHRCICLSSICNREKNSCI
ex11
IEKKIHASQQRWQQLKEEIELLQDLKQTLCSFQENRDLMSSSTSISSLSY
FQAGCHTQL
LC-MS/MS
Idotp 0.92
precursor [M+1], y6, b2, b3, b4, b8, a4, a5, a8
Intensity (10⁶)
retention time, min
d
OAS1
exon_skip & alt_3prime
Cryptic
n = 15 reads
ex5 ex7
ex5 ex6 ex7
Variant 10 ex5 ex7
Canonical ex5 ex6 ex7
anti-HLA IP
ex7
HWISQNNLAPLDVCFLLTGDVLLVSL
ex7
VNLTLVGRRNYTNN
ex6
AESNSADDETDDPRRYQKYGYIGTHEYPHFSHRPSTLQAASTPQAEEDWTCTIL
FLLTGDVL
LC-MS/MS
Idotp 0.91
precursor [M+1], bf, a7
Intensity (10⁶)
retention time, min
e
FCGRT
exon_skip
Cryptic
n = 41 reads
ex4 ex5
ex4 iny ex5
Canonical ex4 ex5
anti-HLA IP
ex4
AluJb-derived
ex5
NLEWKETGFLHVAQAGLELPASSDPPASASQSAGITE PPSMRL
NLEWKETGF
LC-MS/MS
Idotp 0.97
precursor [M+1], [M+2], a6
Intensity (10⁶)
retention time, min
f
MX1
multi_exon_skip & alt_3prime
Cryptic
n = 28 reads
ex13 ex16
ex13 ex14 ex15 ex16
Canonical ex13 ex14 ex15 ex16
anti-HLA IP
ex16 frame-shift
GTDCLLPGPGIQGCIAEGQREGAGRRKEEEILGFWGFPVQLGNRLFHGGDLSAPDGLSP
AGRRKEEEIL
LC-MS/MS
Idotp 0.93
precursor[M+1], b9, z9
Intensity (10⁶)
retention time, min
ex = involved exon
ex = not involved exon
= involved intron
= in-frame stop codon
= neosequence
= canonical sequence

Fig. 3 | Six examples of cryptic splicing events that produced neoantigen peptides detected on HLA. Schematic representations of the cryptic events (left side) in: **a**, *USP18*, **b**, *MYO15B*. **c**, *PHF11.* **d**, *OAS1.* **e**, *FCGRT* and, **f**, *MX1.* The translation products (middle) with the predicted neo-sequences indicated as purple boxes, normal sequences by golden boxes, and stop codons by a vertical red line. The origin within the neo-sequences of the peptides detected in anti-HLA immunoprecipitates are indicated and their mass spectra (right side) with precursor and daughter ions and idotp values are indicated.

Fig. 4

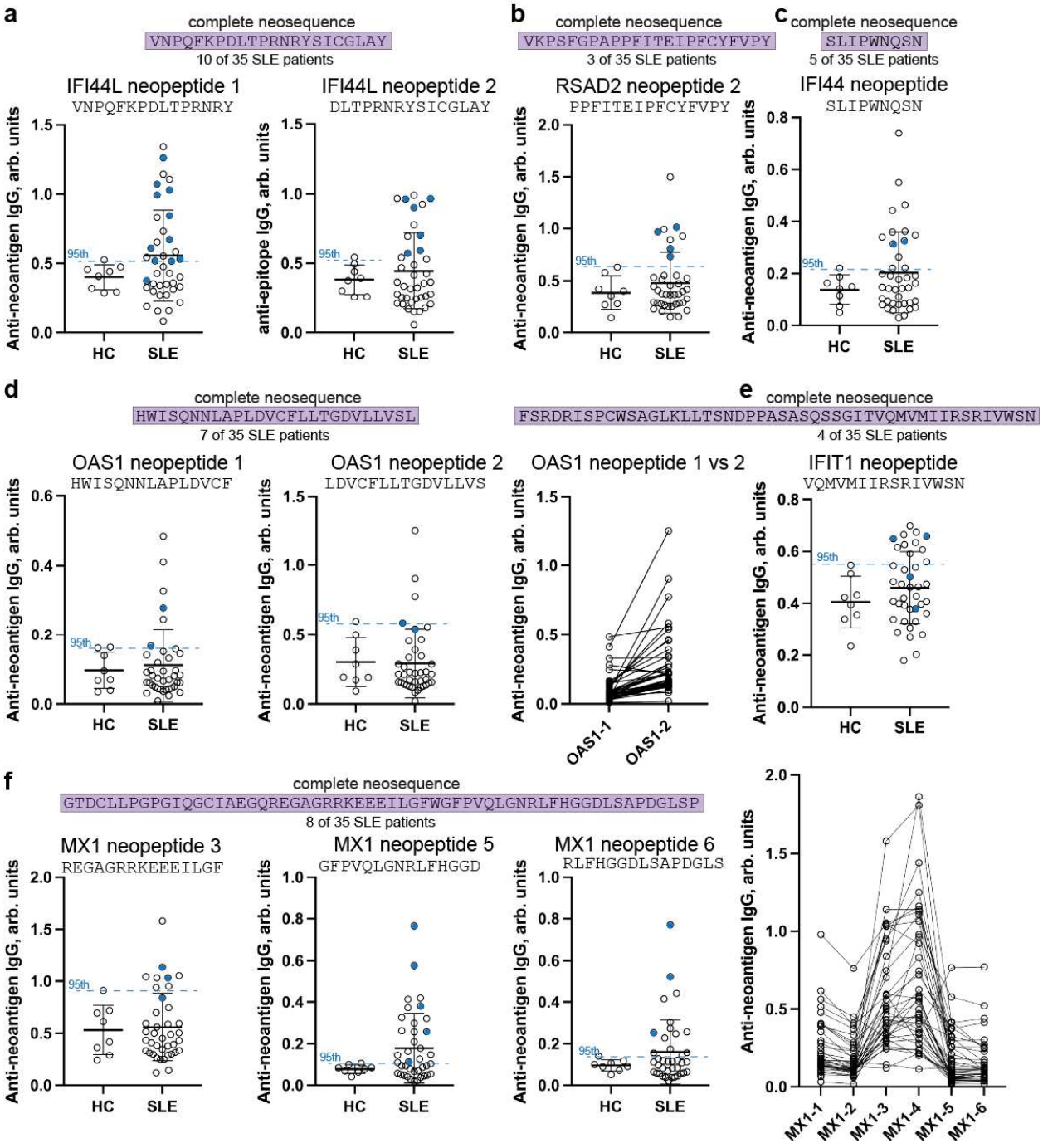


Fig. 4 | Autoantibodies against novel amino acid sequences created by cryptic splicing in SLE. a, ELISA with all 35 SLE patients' serum and the 2 indicated overlapping 15-amino acid residue peptides derived from the total novel amino acid sequence (purple box) generated from the transcript of *IFI44L*. Blue dots indicate the patients that were computationally predicted to generate this neo-sequence at the time of venipuncture. **b**, ELISA for the second peptide (indicated) covering most of the neo-sequence (purple box) generated by cryptic splicing of *RSAD2*. **c**, Same for the single 10-residue peptide covering the cryptic splice event in *IFI44*. **d**,

Same for the two peptides covering the neo-sequence (purple box) from *OAS1*. **e**, One peptide from the neo-sequence (purple box) generated by a cryptic splicing event in *IFIT1*. This peptide was chosen because it is predicted to bind the SLE-associated HLA-DRB1*15:01 with high affinity. **f**, Same for 3 of the 6 peptides covering the second half of the neo-sequence (purple box) generated from MX1. The fourth panel connects the reactivity towards the 6 peptides in each individual patient.

Fig. 5

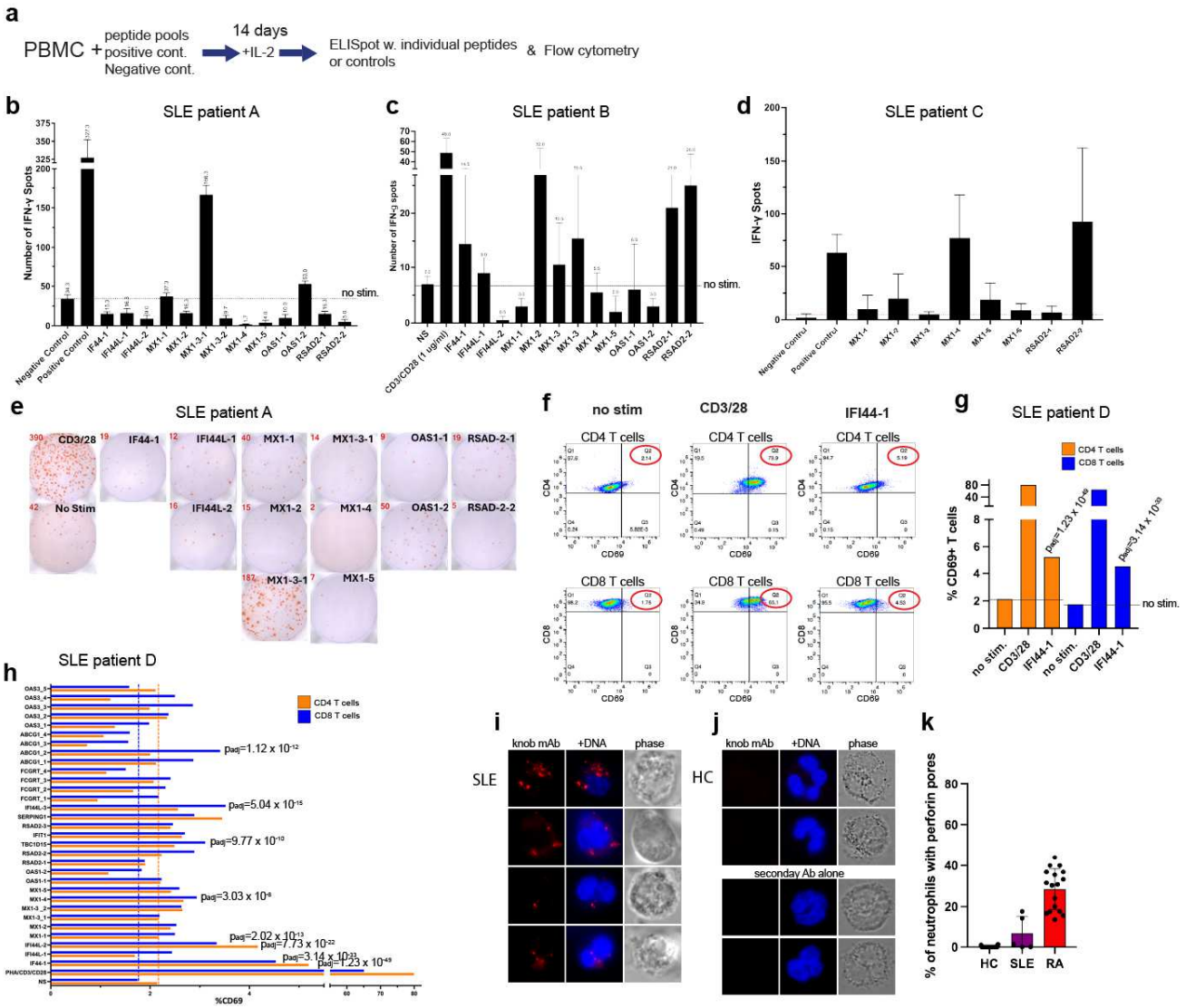


Fig. 5 | Antigen-specific T cells that are expanded and restimulated by neoantigens. a, experimental scheme. **b,** ELISpot assay of T cells restimulated overnight with the indicated individual peptides in patient A. **c**, ELISpot assay of T cells from a second SLE patient restimulated overnight with the indicated individual peptides. **d**, ELISpot assay of T cells from a third SLE patient restimulated overnight with the indicated individual peptides. Note that the two patients have different profiles of reactivity to this set of peptides. **d,** image of the ELISPOT assay with patient A. **f,** flow cytometry for expression of the CD69 activation marker in CD4 and CD8 T cells in the negative control, the positive control, and T cells stimulated with the first IFI44 peptide. **g,** summary of the flow cytometry in panel f; note that the IFI44 peptide increased the portion of activated CD69+ T cells 2-3 fold. **h,** CD69 expression in CD4 (orange) and CD8 (blue) T cells in a different SLE patients in response to a range of neoantigen peptides. Note that some peptides predominantly simulate CD4 T cells, other CD8 T cells. **i,** the presence of poly-perforin pores (red) on neutrophils from SLE patients. **j,** absence of poly-perforin pores on healthy donor neutrophils (upper two rows). A negative staining control with secondary antibody alone (two lower rows). **k,** quantitation of the poly-perforin pore-positive neutrophils from several healthy controls, SLE (purple bar), or RA patients (red bar) from counting neutrophils with one or more poly-perforin pores divided by the total number of neutrophils in 10 or more fields of view at 100X magnification. Statistical significance in panels g and h

was calculated using χ^2 and adjusted for multiple comparisons. A complete table of p-values for panels g and h can be found in **Supplemental Table S3**.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementalmaterial.pdf