

T and B cells recognize splicing-derived neutrophil neoantigens in lupus

Corresponding Author: Professor Tomas Mustelin

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)

This manuscript presents a substantial body of work and, for the first time, provides solid evidence for the presence of neoantigens and related autoantibodies in SLE. Overall, the findings are novel, compelling, and significantly advance the field. I have several comments and questions that I believe should be addressed:

1. The authors elegantly demonstrate that dysregulated mRNA splicing in SLE neutrophils generates neoantigens that contribute to immune tolerance breakdown. However, they suggest that these noncanonical RNA splicing events might arise later in life. While this is plausible, it is also possible that they result from aberrant expression or function of factors involved in RNA splicing. In this regard, autoantibodies against components of the spliceosome, which are frequently found in SLE patients, could alter spliceosome function and contribute to such abnormalities.
2. Products of alternatively spliced genes may disrupt or modulate distinct molecular pathways. Do the authors have any evidence indicating that any of the aberrantly spliced products identified in their study affect molecular pathways known to be linked to neutrophil abnormalities in SLE?
3. This is the first report to identify alternative splicing events in neutrophils from patients with SLE, although alternatively spliced mRNA species have been previously implicated in other immune cell types and associated with dysregulated cellular function.
4. The authors may also wish to discuss whether disease activity influences the extent or pattern of alternative splicing, and whether immune complexes—abundant in SLE sera—could promote alternative splicing through mechanisms known to affect splicing regulation. Were any patients examined during both active and inactive disease states?
5. Although the discussion is informative and enjoyable to read, it would benefit from some condensation and moderation of a few particularly strong statements.

Referee #2

(Remarks to the Author)

Najjar, Mustelin and colleagues present a report interrogating the idea that autoreactivity to peptides generated by aberrant splicing may importantly drive the autoimmune response in SLE. The authors show that neutrophils from SLE patients express non-canonical splice variants of multiple genes compared to neutrophils from controls, many of which are ISGs. Mass spec profiling of peptides from the relevant HLA molecules pulls down peptides corresponding to the aberrantly spliced transcripts. Serum samples from some SLE patients are shown to have antibodies that can bind peptides from aberrantly spliced transcripts, and peptide stimulation experiments suggest that T cells from 3 SLE patients may respond to these peptides as well.

The concept that autoimmunity in SLE derives in part from responses to peptides generated from aberrantly spliced products is interesting and provocative. It would provide another example of how aberrantly presented peptides may drive SLE autoimmunity, via a different mechanism than that described by Mori et al (PMID: 39276775).

While the concept is interesting, several key controls are needed to make the case compelling, and a more definitive demonstration of an antigen-specific response in SLE patients against the nominated peptides is critical to support the core

idea.

MAJOR:

1) The authors identify many splice variants that are found in SLE neutrophils but not neutrophils of controls, and not in annotated databases; many of these are ISGs. It seems that the analysis may be enriched for detecting any genes that are more highly expressed in SLE neutrophils compared to control neutrophils; presumably the canonically spliced versions of these same ISGs are also highly upregulated in SLE neutrophils compared to control neutrophils. To what extent is the detection of these non-canonical ISG variants a result of IFN response, rather than an inherent aspect of SLE autoimmunity? It seems plausible that an IFN response may involve broad shifts in RNA splicing, in addition to expected upregulation of ISGs.

ii) Are the 'cryptic' variants produced in response to type I IFN? Would neutrophils (or other cells) stimulated in vitro with type I IFN produce these splice variants?

ii) Would these splice variants been seen in other IFN-high diseases (e.g. Ro+ SSc, DM, SSc-ILD, or IFNopathies)?

iii) A semantic question: what do the authors consider necessary to consider these 'cryptic,' as opposed to 'not previously reported' splice variants?

iv) Are the ISG cryptic variants found only in IFN-high patients, and do the autoimmune T cell responses and autoantibodies segregate to the IFN-high patients?

2) If the majority of the 'cryptic' splicing is found in ISGs, does this imply that SLE patients would already need an IFN response to have been generated for this ISG-enriched neoantigen-specific adaptive immune response to occur? If so, is this a secondary effect, rather than causal? Alternatively, if aberrant splicing and generation of cryptic antigens predisposes to SLE, shouldn't it occur primarily in non-ISGs, which would be expressed at baseline before disease initiation? Are there examples of cryptic splice variants for which the cryptic variant is highly overrepresented in SLE patients, and the canonical variant is comparably expressed in controls and SLE patients (or lower in SLE patients)? Can the authors demonstrate autoimmunity to any of these neopeptides?

3) The antibody detection studies in Figure 4 are not compelling. Comparison to samples from only 8 controls is not sufficient to demonstrate specificity, esp when compared to 35 SLE samples. The HC are likely undersampled to detect positive signal in this group. Controls for the specificity of the protein detection are also needed. If the assay is run with a comparable set of canonical peptides from the same regions of these proteins (rather than the neopeptides), would the results be different?

4) If the concept is that neopeptide are presented on relevant MHC molecules, thus eliciting T cell activation, T cell help to B cells, and B cell class switch and production of IgGs against the antigen, then it seems that positive Ab signal should only come from patients who have the relevant MHC molecule to present the relevant, specific neopeptide. Is this the case?

5) The T cell evidence in Figure 5 is also not compelling. If the hypothesis is that a cryptic peptide binds MHC and activates an Ag-specific T cell, then the peptide stimulation should activate either CD4 clone(s) or CD8 clone(s), but not both. Expansion of both CD4s and CD8s suggests that non-Ag-specific responses are being detected. The CD69 staining is not convincing - going from 2% baseline to 4% response in both CD4s and CD8s does not strongly indicate an Ag-specific expanded clone response. Use of control peptides that represent native, and not cryptic, peptides as a comparator is also essential. For this set of experiments to be convincing, the authors should be able to expand out Ag-specific T clones, sequence and express the TCR, and show that the TCR reacts to the neo-peptide and not the native peptide. If the authors could show a couple examples of this would, this would be very strong evidence. It may also allow the authors to track phenotypes and quantities of neoag-specific T cells in SLE patient samples by matching the neoag-specific TCRs to those of T cells in PBMC samples.

6) Specificity for neutrophils. The authors analyze 2 PBMC bulk RNA-seq datasets and indicate that they do not find cryptic variants specific to SLE in those datasets. This point needs more detail and nuance. How robust are those analyses? Are the datasets of comparable quality? Are the ISG splice variants detected in neutrophils really not seen in monocytes or lymphocytes? Is it possible that the lack of signal in total PBMC datasets is because PBMCs are a mix of different cell types, which muddies or obscures signals, whereas a cell-type specific signal is more easily seen in the neutrophil dataset because it is a single cell type? Can one see aberrantly spliced genes in sorted immune cell subsets from SLE patients in the ImmunNeXUT dataset (PMID 35998627) or comparable data? Ideally, the authors would generate RNA-seq of monocytes from the same patients from which neutrophils were analyzed and demonstrate that the same enrichment does not occur.

MINOR:

1) 'All of these data also illustrate that the cryptic splice variants are less abundant than canonical or physiologic splice variants, in good agreement...' Figure 2f/g don't really give the message that cryptic splice reads are uniformly lower than canonical splice reads. The lower part of plots seems to have a decent portion of the genes with lower expression of the canonical reads?

2) A Venn Diagram or Table would be helpful to understand the overlap in the genes implicated from the different studies/experiments. I think the 21 peptides detected by mass spec were all also identified in the RNAseq analyses; how many of these were also confirmed by PCR, sequencing, and how many of these are tested in the Ab and T cell assays?

3) Are the immunopeptidome analyses using MHC pulldowns performed on neutrophils from SLE patients? This should be made explicit. Does this suggest that neutrophil presentation of the antigens is critical, rather than acquisition and presentation of phagocytosed neutrophil antigens by macrophages?

4) The extreme p-values in Fig 5g/h do not fit with a common-sense interpretation of the data presented in these plots. Are the X2 tests done at the level of counting single cells? This should be revisited w/ more typical statistics for this kind of analysis.

5) Figure 5b,c,d. The error bars are wide – what do they represent? It is hard to tell if these signals are real or noise.

6) There are several typos/minor grammatical quirks in the text/legends/figures. E.g. 'cryptic' is misspelled in Fig 1f. Could use another read-through for these things.

Referee #3

(Remarks to the Author)

The authors present follow-up work on their paper from earlier in the year in ACR open rheumatology. They use RNA-seq, RT-PCR and MS of neutrophils from 35 SLE patients to detect alternative splicing derived neoantigens. They detect IgG antibodies against and T-cells reactive to many of these neoantigens.

This is an interesting paper with some potentially compelling data. As currently presented it is much less convincing than I think it could be. The numbers (number of cryptic events, number of antigens confirmed by MS, number of synthetic antigens that antibodies respond to) are not given context: out of how many possibilities? How likely were these counts to occur by chance? Relatedly there are limited controls described for each experiment. More specifically:

- How strong is the agreement between RNA/MS-observed neoantigens and IgG antibody presence (by ELISA)? How likely is this by chance? How does this compare to random antigen sequences?

- How many cryptic events are seen in the PBMCs that are not seen in the neutrophils?

Given the complexity of the experiment design, and challenges such as the variable sensitivity of RNA-seq and MS, the authors would benefit from having a statistician (or at least someone with statistical training) work with them to think these analyses through. The one place a statistical analysis is mentioned (T-cell numbers in flow cytometry) it is poorly described, not specifying what is being compared or the hypotheses being tested. Without rigorous statistical analysis and controls it is hard to judge the biological significance of the findings.

The attempt to connect changes in RBP expression to cryptic splicing is also weaker than it could be. Are there binding sites (predicted or observed in public eCLIP) for the deregulated RBPs in or near the cryptic events? Are these cryptic events seen after KD of the RBP in the ENCODE K562 KD data?

Minor comments:

1-15 per million T cells: is this enough to cause autoimmune disease?

Why would the cryptic events show up in type I interferon-inducible genes?

The language could be made more accessible to non-immunologists in various ways. "neutropenic" is not a term I've seen before, sticking to CD8 or cytotoxic, and pointing out early that neutrophils are not PBMCs.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

The authors have addressed the reviewer's concern.

Referee #2

(Remarks to the Author)

Najjar and colleagues provide a revised version of the manuscript highlighting the presence of unique splice variants in neutrophils of SLE patients, some of which may elicit antigen-specific B cells and T cell responses in patients with SLE. The revisions add substantial and new data, including assessment of additional transcriptomic datasets (monocytes, SjD neutrophils, IFN-treated PMN) and new sequencing of TCRs. The new data are much appreciated, though in some key respects go only part way to what I would view as a compelling answer. Overall, the demonstration of Ag-specific B cell and T cell responses still seems to me somewhat weak, with responses of modest magnitude demonstrated from the SLE samples and absence of key controls.

1) Figure 5. Ag-specific T cell responses. We are shown responses of T cells to Ag-stimulation from 5 SLE patients and no control patients.

a) The signals of activated T cells across many Ags, as in Fig 5i, are small (as expected, I agree); however, there is no demonstration that these Ag-induced responses are unique to SLE patients and would not be similarly seen in cells from control patients.

b) A demonstration of lack of response to control peptides is now provided (Fig 5e), but it is from a different patient than was

used to demonstrate positive responses, with no positive signal to SLE-cryptic peptides shown for this patient, and reads out IL-17 rather than IFN γ ? This is not sufficient; would be required from multiple donors w/ Cryptic vs Canonical peptide comparison.

Given a lack of measurement in control donors (as was done for the Ab signals), and a lack of clear distinction between cryptic vs canonical signal in multiple patients, the case is not compelling that this is a SLE-relevant Ag-specific T cell response.

2) Fig 5j. It is good to see clonal expansion of specific TCRs in the stimulation cultures. But do these expanded TCRs respond to the peptide of interest? Expressing the TCR(s) and showing reactivity to cryptic peptide of interest, and not other relevant control peptides, is needed to get to the answer.

3) Figure 4. Ab detection. It is helpful to see additional control patients assessed for comparison to the SLE cohort. The SLE pts show modest and variable positivity for different neopeptides. However, there is no demonstration that SLE pts would not show the same low level positivity towards control (eg canonical) peptides compared to control donors, or that the binding is truly Ag-specific (can it be competed off?). This is critical given that SLE pts have many polyreactive autoabs and serum from SLE pts can often give low level signals that differ from serum from controls.

4) Addition of transcriptomic assessments of monocytes, SjD neutrophils, and IFN-treated neutrophils is very nice; however, there is minimal information provided about these datasets.

a) How long were neutrophils treated with IFN?

b) For SjD patients, this is a very valuable comparison. How many of these patients show an IFN signature (presumably several of the Ro+ patients?). Though underpowered, is there a hint that IFN-positive SjD patients may show a similar cryptic transcript accumulation? Are they all primary SjD? More information about the cohort would be helpful to underscore the value.

c) For the monocytes, how many samples are analyzed, and it is from the same patients as was used for the neutrophils? For each of the datasets above, it is indicated that there is virtually no overlap with the SLE neutrophil cryptic splice variants. However, for each of the datasets above, do they show their own unique collection of cryptic splice variants (i.e. do monocytes have their own cryptic splice variants? Do IFN-treated neutrophils have their own cryptic variants)?

5) The authors' answers to my previous questions highlight several times the point that IFN-low SLE patients in the cohort also have cryptic splice variants, supporting that this is an SLE-centric phenomenon and not primarily an IFN-centric phenomenon. I appreciate the authors' point here; it would be important to show these data in a manner that make this point clear.

MINOR:

The term 'cryptic' in this paper sounds overly mysterious to me. Based on the authors' rebuttal, it seems to me that 'SLE-associated' or 'SLE-specific' would be a more informative term.

The text in several of the figures is too small to read.

Referee #3

(Remarks to the Author)

I'm satisfied that the authors have addressed my concerns and recommend the work for publication.

Version 2:

Reviewer comments:

Referee #2

(Remarks to the Author)

The authors provide additional data to demonstrate that 1) a single healthy donor showed a lack of T cell responses to cryptic peptides (Supp Fig 9) and that 2) one TCR isolated from neoantigen-expanded T cells could be shown to respond to peptides from an IFI44L neoantigen.

From prior comments:

Prior 1a: "Figure 5. Ag-specific T cell responses. We are shown responses of T cells to Ag-stimulation from 5 SLE patients and no control patients.

a) The signals of activated T cells across many Ags, as in Fig 5i, are small (as expected, I agree); however, there is no demonstration that these Ag-induced responses are unique to SLE patients and would not be similarly seen in cells from control patients."

The authors now provide evidence from 1 control donor showing no responsiveness to the tested peptides. This is OK, but as the authors indicate, 'there is always patient-to-patient variability in human immunology experimentation.' Showing one single control donor does not make a convincing case that the varied antigen-specific responses detected in 3 SLE patients (Fig 5b-d) is robustly disease-specific. The authors may be correct, but it remains highly uncertain from the data shown.

Prior 1b: "A demonstration of lack of response to control peptides is now provided (Fig 5e), but it is from a different patient than was used to demonstrate positive responses, with no positive signal to SLE-cryptic peptides shown for this patient, and

reads out IL-17 rather than IFN γ ? This is not sufficient; would be required from multiple donors w/ Cryptic vs Canonical peptide comparison. Given a lack of measurement in control donors (as was done for the Ab signals), and a lack of clear distinction between cryptic vs canonical signal in multiple patients, the case is not compelling that this is a SLE-relevant Ag-specific T cell response."

The authors explain that patient responses vary, and there can be background IFN production detected, thus they show IL17 from Patient D rather than IFN. This is clearly not an adequate control. What do the responses of Patients A, B, C look like when stimulated with canonical peptides?

We are currently presented with :

- Positive but varied responses to a range of cryptic peptides in 3 SLE patients by IFN γ
- Negative response to canonical peptide in 1 SLE patient by IL17
- Negative response to cryptic peptides in 1 Control by IFN γ

The results text states: "In contrast, T cells from healthy donors had little response to any of the neoantigen peptides (not shown)."

The authors explain in their response that IFN γ readout can be noisy, that they have 'seen a lack of response to canonical peptides in additional patients' (not shown), and that sometimes IL-10 is detected in response to canonical peptides (not shown). In my view, this assessment of a small number of patients (n=1 in two key controls conditions, one with an alternate readout), is not sufficient to make a conclusion about the nature of autoantigen targets in patients with lupus. In total, I do not find this a compelling data package for the very impactful claim that the authors are making.

Prior 2: "Fig 5j. It is good to see clonal expansion of specific TCRs in the stimulation cultures. But do these expanded TCRs respond to the peptide of interest? Expressing the TCR(s) and showing reactivity to cryptic peptide of interest, and not other relevant control peptides, is needed to get to the answer."

I applaud the authors for the effort to get full a/b TCRs and test reactivity. The authors demonstrate a single TCR with reactivity to IFI44L-2 peptides. However, it is unclear how many TCRs were tested that did not show reactivity in these analyses – this would be typical to report and important to understand. Further, It's unclear why the authors have not gotten to a single peptide. It does not seem that they have demonstrated a specificity of the reactivity to cryptic, vs canonical, peptide, which would be expected (and critical) for the overall concept.

Prior 3: "Figure 4. Ab detection. It is helpful to see additional control patients assessed for comparison to the SLE cohort. The SLE pts show modest and variable positivity for different neopeptides. However, there is no demonstration that SLE pts would not show the same low level positivity towards control (eg canonical) peptides compared to control donors, or that the binding is truly Ag-specific (can it be competed off?). This is critical given that SLE pts have many polyreactive autoabs and serum from SLE pts can often give low level signals that differ from serum from controls."

There were two aspects to the point here: 1) non-specific signal in SLE serum samples, and 2) ability to distinguish between cryptic and canonical epitopes.

Aspect 1 is not addressed. Would these SLE samples show increased signal compared to control samples is assayed against a set of irrelevant control peptides? This is a critical missing control to assess the extent to which the signal is specific to the cryptic peptides.

Aspect 2 - the authors do not find it important to assess the specificity of the Ab response to cryptic epitopes versus canonical proteins. They anticipate that Ab-reactivity to the canonical proteins may be present by epitope spreading. This may be correct; in this case, we might conclude that SLE patients have antibodies against many IFN-induced proteins. The contribution of cryptic recognition is uncertain. Yet testing only binding to cryptic peptides implies a certain specificity to cryptic peptides that the authors recognize may not be the case. Given that the authors underscore autoimmunity to cryptic neoantigens as a core concept of this paper, it seems a highly relevant point that remains unclear.

Prior 5: "The authors' answers to my previous questions highlight several times the point that IFN-low SLE patients in the cohort also have cryptic splice variants, supporting that this is an SLE-centric phenomenon and not primarily an IFN-centric phenomenon. I appreciate the authors' point here; it would be important to show these data in a manner that make this point clear."

The authors response to this seems to be – 'go find it yourself.' The authors have not made a compelling case that this is phenomenon exists independently from a strong IFN response, and that the cryptic variants detected are not primarily a product of an ongoing IFN response (independent of SLE). The data may be in there somewhere, but it is not evident and not currently convincing.

Minor:

As indicated initially, I do not think a X2 test is appropriate in Fig 5h.

Version 4:

Reviewer comments:

Referee #2

(Remarks to the Author)

I am satisfied with the revisions. The additional controls are impactful. The identification of a defined TCR, with sequence provided, and the specific peptide that it recognizes, is impressive.

Two minor comments for the authors to consider:

1) The quantification of GFP+ cells in Figure 5L does not match the examples shown in Figure 5K (though the qualitative message is clear).

2) A chi-squared test assumes that the observations are independent. The observations are not independent when sampling many cells out of the same stimulated sample. This is why a chi-squared test is not appropriate in the context used.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to reviewers:

We are grateful to the three reviewers for their excellent and constructive questions and have done our best to answer them all comprehensively. This has resulted in revisions and additions to Figures 1, 4, and 5, as well as new supplemental figures S2, S3, S7 and S8. Several additions, edits, and clarifications have been added to the text. In detail:

Referee #1 (Remarks to the Author)

This manuscript presents a substantial body of work and, for the first time, provides solid evidence for the presence of neoantigens and related autoantibodies in SLE. Overall, the findings are novel, compelling, and significantly advance the field. I have several comments and questions that I believe should be addressed:

1. The authors elegantly demonstrate that dysregulated mRNA splicing in SLE neutrophils generates neoantigens that contribute to immune tolerance breakdown. However, they suggest that these noncanonical RNA splicing events might arise later in life. While this is plausible, it is also possible that they result from aberrant expression or function of factors involved in RNA splicing. In this regard, autoantibodies against components of the spliceosome, which are frequently found in SLE patients, could alter spliceosome function and contribute to such abnormalities.

We agree that aberrant expression of function of RNA splicing factors is a likely cause of cryptic splicing in SLE; in fact, our analysis of the ENCODE data base suggested by Reviewer #3 provides an indication that this is the case. As requested, we tested for associations between cryptic splicing events and antibodies against the spliceosome including RNP and Smith autoantibodies and found no association by Wilcoxon-Mann-Whitney test (p values >0.35). These data are in the new Supplemental figure S2c.

2. Products of alternatively spliced genes may disrupt or modulate distinct molecular pathways. Do the authors have any evidence indicating that any of the aberrantly spliced products identified in their study affect molecular pathways known to be linked to neutrophil abnormalities in SLE?

It is indeed possible that cryptic splicing events alter the function of the translated protein, but since the cryptic splice variants are generally considerably less abundant than the canonical splice forms, their impact on the biological process in which they participate, likely remains small. Besides the interferon-inducible genes present in most SLE patients, none of the other genes subjected to cryptic splicing are obvious components of SLE-related pathways, such as neutrophil NETosis, cell death pathways, or accelerated neutropoiesis.

Relevant to this question, we previously reported (ref. 26) that 521 transcripts in SLE neutrophils were subjected to statistically different annotated ('normal') alternative splicing events compared to healthy donor neutrophils; this set contains many examples of altered protein function (e.g., lacking a transmembrane region or an RNA-binding domain).

3. This is the first report to identify alternative splicing events in neutrophils from patients with SLE, although alternatively spliced mRNA species have been previously implicated in other immune cell types and associated with dysregulated cellular function.

We agree.

4. The authors may also wish to discuss whether disease activity influences the extent or pattern of alternative splicing, and whether immune complexes—abundant in SLE sera—could promote alternative splicing through mechanisms known to affect splicing regulation. Were any patients examined during both active and inactive disease states?

As suggested, we tested for correlations with disease activity (SLEDAI) and found no association between cryptic splicing and disease activity (See Supplemental figure S2c).

5. Although the discussion is informative and enjoyable to read, it would benefit from some condensation and moderation of a few particularly strong statements.

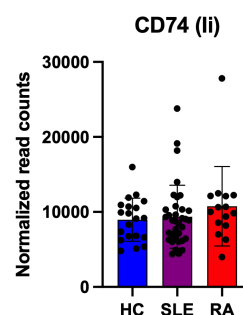
We have condensed the discussion and tempered our conclusions.

Referee #2 (Remarks to the Author):

Najjar, Mustelin and colleagues present a report interrogating the idea that autoreactivity to peptides generated by aberrant splicing may importantly drive the autoimmune response in SLE. The authors show that neutrophils from SLE patients express non-canonical splice variants of multiple genes compared to neutrophils from controls, many of which are ISGs. Mass spec profiling of peptides from the relevant HLA molecules pulls down peptides corresponding to the aberrantly spliced transcripts. Serum samples from some SLE patients are shown to have antibodies that can bind peptides from aberrantly spliced transcripts, and peptide stimulation experiments suggest that T cells from 3 SLE patients may respond to these peptides as well.

The concept that autoimmunity in SLE derives in part from responses to peptides generated from aberrantly spliced products is interesting and provocative. It would provide another example of how aberrantly presented peptides may drive SLE autoimmunity, via a different mechanism than that described by Mori et al (PMID: 39276775). **Yes, Mori et al reported that the absence of the invariant chain of class II MHC results in presentation of neoantigens. However, we do not observe decreased expression of the invariant chain (adjacent figure) in our 35 patients.**

While the concept is interesting, several key controls are needed to make the case compelling, and a more definitive demonstration of an antigen-specific response in SLE patients against the nominated peptides is critical to support the core idea.



MAJOR:

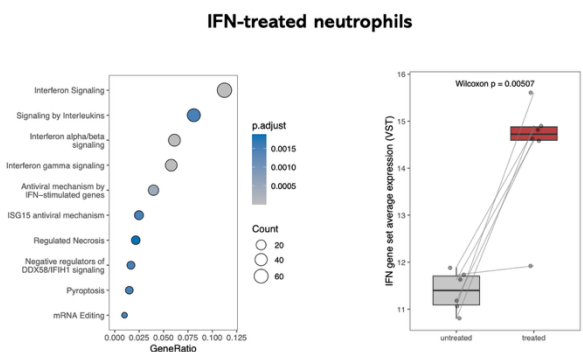
1) The authors identify many splice variants that are found in SLE neutrophils but not neutrophils of controls, and not in annotated databases; many of these are ISGs. It seems that the analysis may be enriched for detecting any genes that are more highly expressed in SLE neutrophils compared to control neutrophils; presumably the canonically spliced versions of these same ISGs are also highly upregulated in SLE neutrophils compared to control neutrophils. To what extent is the detection of these non-canonical ISG variants a result of IFN response, rather than an inherent aspect of SLE autoimmunity?

This is an important question. We performed new analyses to explore this and now show that both gene expression levels and splicing patterns contributed to the generation of cryptic splicing that we observed. The new [Supplemental figure S3a-d](#) shows gene expression levels for controls, RA, and SLE samples without and SLE with the cryptic events, showing a similar gene expression for most affected transcripts and higher expression for others, mostly ISGs, in the patients with cryptic events. In [Supplemental figure S3f](#), we show the percent spliced-in for cryptic events, demonstrating large shifts in splicing pattern in many cryptic splicing events and no change in PSI in the other end of the spectrum. In other words, in highly expressed ISGs an unchanged PSI can result in detectable (>5reads) cryptic variants. In the majority of genes, the expression level is unchanged while the cryptic variant is generated by increased PSI.

It seems plausible that an IFN response may involve broad shifts in RNA splicing, in addition to expected upregulation of ISGs.

i) Are the ‘cryptic’ variants produced in response to type I IFN? Would neutrophils (or other cells) stimulated in vitro with type I IFN produce these splice variants?

To address this directly, we generated a new RNA-seq data of healthy neutrophils treated with interferon (adjacent figure) and searched for the cryptic splicing events that we observed in SLE. We did not find any of them ([new Fig. 1j](#)). This shows that the SLE-specific cryptic events are not simply induced by interferon. Although there is a positive correlation between ISG expression and cryptic splicing events ([new Supplemental Fig. S2b](#)), this is likely driven by the cryptic splicing in the ISGs themselves. SLE patients that lack an interferon signature, still have cryptic splicing of some ISGs, as well as in many non-ISGs.



This result is also consistent with our previous study based on a series of analyses in SLE versus COVID19 neutrophils or whole blood ([Najjar et al Lupus Sci. Med. 2024 Jan 31;11\(1\):e001059. doi: 10.1136/lupus-2023-001059](#)), in which none of the splicing events unique to SLE were present in COVID19 patients although the expression profile of genes were nearly identical.

ii) Would these splice variants been seen in other IFN-high diseases (e.g. Ro+ SSc, DM, SSc-ILD, or IFNopathies)?

We analyzed neutrophil data from patients with Sjögren’s disease (n=10, including 7 with positive Ro autoantibodies) and found no overlap with the SLE cryptic events ([see new Fig. 1j](#)).

iii) A semantic question: what do the authors consider necessary to consider these ‘cryptic,’ as opposed to ‘not previously reported’ splice variants?

We use the term ‘cryptic’ to mean both ‘not previously reported’ and ‘not present in healthy donors or rheumatoid arthritis patients’. In other words, unique to SLE.

iv) Are the ISG cryptic variants found only in IFN-high patients, and do the autoimmune T cell responses and autoantibodies segregate to the IFN-high patients?

No, we observed cryptic events in the IFN-low lupus patients as well and T cell responses do not seem to segregate by ISG levels.

2) If the majority of the 'cryptic' splicing is found in ISGs, does this imply that SLE patients would already need an IFN response to have been generated for this ISG-enriched neoantigen-specific adaptive immune response to occur?

If so, is this a secondary effect, rather than causal?

Our data do not support this possibility: there are patients without an IFN signature who have cryptic splicing of IFI44L and MX1. In addition, our new data of IFN-treated neutrophils did not show any SLE-cryptic splicing, indicating that IFN are not a likely cause for the generation of this cryptic splicing. This suggests that the cryptic splicing might well precede the interferons in lupus pathogenesis.

Alternatively, if aberrant splicing and generation of cryptic antigens predisposes to SLE, shouldn't it occur primarily in non-ISGs, which would be expressed at baseline before disease initiation?

The presence of cryptic splicing events in non-ISG genes that are expressed similarly to healthy donors, cryptic splicing in ISGs in IFN-low patients, as well as the IFN-treated neutrophil data mentioned above are compatible with neoantigens existing before high IFNs. However, this does not rule out the possibility that high ISG expression may increase aberrant splicing of ISGs adding to the neoantigen burden as the disease progresses.

Are there examples of cryptic splice variants for which the cryptic variant is highly overrepresented in SLE patients, and the canonical variant is comparably expressed in controls and SLE patients (or lower in SLE patients)? Can the authors demonstrate autoimmunity to any of these neopeptides?

Yes, we do observe this with non-ISGs, which are the majority of the cryptic splicing events. Fig. 5i shows T cell responses to peptides that are ISGs as well as non-ISGs.

3) The antibody detection studies in Figure 4 are not compelling. Comparison to samples from only 8 controls is not sufficient to demonstrate specificity, esp when compared to 35 SLE samples. The HC are likely undersampled to detect positive signal in this group. Controls for the specificity of the protein detection are also needed. If the assay is run with a comparable set of canonical peptides from the same regions of these proteins (rather than the neopeptides), would the results be different?

This is an excellent point and have now repeated all the ELISA assays with a larger cohort of healthy controls (n=40) in the revised Fig. 4. We have also now included a set of canonical IFI44L peptides in Fig. 5e.

4) If the concept is that neopeptide are presented on relevant MHC molecules, thus eliciting T cell activation, T cell help to B cells, and B cell class switch and production of IgGs against the antigen, then it seems that positive Ab signal should only come from patients who have the relevant MHC molecule to present the relevant, specific neopeptide. Is this the case?

The IgG results do show a skewing towards the patients computationally predicted to generate the relevant neoantigen (indicated with blue dots in Fig. 4). However, the RNAseq data represent a single point in time and cover approximately 60% of splicing events. An antibody response requires much longer time. Hence it is not surprising that there are also patients with IgGs who did not contain the cryptic splice variant (at the time of blood draw) and there are patients who had it but did not have IgGs against the neoantigen. Knowing each patient's MHC haplotype does not help here because the novel amino acid stretches in some cryptic variants are long and may, theoretically, be processed into numerous different peptides fitting onto many different class I and II alleles. In addition, IgGs typically bind different epitopes than the TCR recognizes, which is why we designed the peptides in Fig. 4 to cover most or all of the novel amino acid sequence of the cryptic variant.

5) The T cell evidence in Figure 5 is also not compelling. If the hypothesis is that a cryptic peptide binds MHC and activates an Ag-specific T cell, then the peptide stimulation should activate either CD4 clone(s) or CD8 clone(s), but not both. Expansion of both CD4s and CD8s suggests that non-Ag-specific responses are being detected.

This assumption is not entirely accurate: We used 15-amino acid peptides because they can bind to both class I and II MHC alleles; a 15mer can theoretically bind to class I MHC in 7-8 different positions. Nevertheless, many neoantigen peptides activate CD4 or CD8 T cells, but some activate both.

Our new sequencing of TCRs in T cells expanded by neoantigen shows that more than a single unique TCR was increased in frequency, although one and two dominated the response to the two tested neoantigens, respectively. TCR sequencing also makes the T cell evidence more molecularly defined.

The CD69 staining is not convincing - going from 2% baseline to 4% response in both CD4s and CD8s does not strongly indicate an Ag-specific expanded clone response. Use of control peptides that represent native, and not cryptic, peptides as a comparator is also essential.

We understand this concern but believe that a larger increase would be surprising. A 2% increase within a pool of 50,000 T cells analyzed by flow cytometry is 1,000 T cells, a substantial number for Ag-specific T cells. We now show the lack of a response to three canonical and three random peptides in the new Fig. 5e. We have also used OX40 and 4-1BB as activation markers with very similar results to CD69 (not shown).

For this set of experiments to be convincing, the authors should be able to expand out Ag-specific T clones, **sequence and express the TCR**, and show that the TCR reacts to the neo-peptide and not the native peptide. If the authors could show a couple examples of this would, this would be very strong evidence. It may also allow the authors to track phenotypes and quantities of neoag-specific T cells in SLE patient samples by matching the neoag-specific TCRs to those of T cells in PBMC samples.

As requested, we have now sequenced TCRBs expanded by two neoantigen peptides and added these results to the paper as the new Fig. 5j.

6) Specificity for neutrophils. The authors analyze 2 PBMC bulk RNA-seq datasets and indicate that they do not find cryptic variants specific to SLE in those datasets. This point needs more detail and

nuance. How robust are those analyses? Are the datasets of comparable quality? Are the ISG splice variants detected in neutrophils really not seen in monocytes or lymphocytes? Is it possible that the lack of signal in total PBMC datasets is because PBMCs are a mix of different cell types, which muddies or obscures signals, whereas a cell-type specific signal is more easily seen in the neutrophil dataset because it is a single cell type? Can one see aberrantly spliced genes in sorted immune cell subsets from SLE patients in the [ImmunNeXUT dataset \(PMID 35998627\)](#) or comparable data? Ideally, the authors would generate RNA-seq of monocytes from the same patients from which neutrophils were analyzed and demonstrate that the same enrichment does not occur.

We have added new panels l, j, and k to Fig. 1. We also show sequencing depths of the datasets we used in the new Supplemental Fig. S2a. The PBMC data had higher sequencing depths than the SLE neutrophil data. Despite that, only two of the SLE cryptic events were seen in the PBMC:GSE122459 data, none in the PBMC:GSE211700 data. This strengthens our claim of the specificity of the cryptic events.

As suggested, we downloaded the ImmuNexUT data and attempted to analyze them. Unfortunately, during our QC, we discovered that the ImmuNexUT data were unstranded, and therefore unsuitable for our analysis pipeline built on Spladder, which requires stranded data.

As requested, we generated new data of SLE monocytes (new Fig. 1i) and ran our analysis pipeline. We found a single shared event with the SLE neutrophil cryptic events, suggesting specificity of these events for neutrophils.

MINOR:

1) 'All of these data also illustrate that the cryptic splice variants are less abundant than canonical or physiologic splice variants, in good agreement...' Figure 2f/g don't really give the message that cryptic splice reads are uniformly lower than canonical splice reads. The lower part of plots seems to have a decent portion of the genes with lower expression of the canonical reads?

Yes! These numbers are reads spanning splice junctions (not all reads used in expression analysis).

2) A Venn Diagram or Table would be helpful to understand the overlap in the genes implicated from the different studies/experiments. I think the 21 peptides detected by mass spec were all also identified in the RNAseq analyses; how many of these were also confirmed by PCR, sequencing, and how many of these are tested in the Ab and T cell assays?

All our analyses start with the RNAseq data, which gives the splicing events. These are then translated, identified to correct reading-frame and any stop codons, and subjected to BLAST searches to determine the beginning and end of novel amino acid sequences. These novel amino acid sequences were then used as a data base for mass spectrometry in the immunopeptidomics experiments, as well as for the design of 15mer peptides, which overlap in a tiled fashion and cover all of the novel sequence. We focused the search for IgGs and neoantigen-specific T cells on novel amino acid sequences shared by as many patients as possible, as indicated in Fig. 4. For example, the IFI44L neosequence is 22 amino acid residues and is present in 10 of our 35 patients; this sequence is covered in two 15aa peptides with an 8aa overlap.

3) Are the immunopeptidome analyses using MHC pulldowns performed on neutrophils from SLE patients? This should be made explicit. Does this suggest that neutrophil presentation of the

antigens is critical, rather than acquisition and presentation of phagocytosed neutrophil antigens by macrophages?

Yes, it was performed on SLE neutrophils (now explicitly stated). Neutrophils express class I but are not recognized as good antigen-presenting cells for CD4 T cells. Our data do not illuminate how neoantigens are presented or transferred to myeloid dendritic cells or B cells in the patients. Our T cell assays contain few, if any, neutrophils, but we have observed that depleting the B cells will reduce the response to some, but not all neoantigen peptides.

4) The extreme p-values in Fig 5g/h do not fit with a common-sense interpretation of the data presented in these plots. Are the X2 tests done at the level of counting single cells? This should be revisited w/ more typical statistics for this kind of analysis.

For each neoantigen, we constructed two-by-two tables with frequencies of cells with and without activation in flow cytometry and tested the difference in proportions using the Chi-square test. Then we adjusted the p-values for multiple comparisons. The comparisons are with the controls for CD4 and CD8 cells, and not CD4 vs CD8 cells. We have adjusted the figure for clarity.

Importantly, the inference is strictly within-donor (stimulated vs. no-stimulation comparisons). We are not using these p-values to make population-level claims, and although the percentage change is modest, it reflects a substantial number of activated T cells and is biologically interpretable

5) Figure 5b,c,d. The error bars are wide – what do they represent? It is hard to tell if these signals are real or noise.

Error bars represent the standard deviation in triplicate determinations. They are real and represent numbers of responding T cells, now indicated on top of the bars.

6) There are several typos/minor grammatical quirks in the text/legends/figures. E.g. 'cryptic' is misspelled in Fig 1f. Could use another read-through for these things.

Done.

Referee #3 (Remarks to the Author):

The authors present follow-up work on their paper from earlier in the year in ACR open rheumatology. They use RNA-seq, RT-PCR and MS of neutrophils from 35 SLE patients to detect alternative splicing derived neoantigens. They detect IgG antibodies against and T-cells reactive to many of these neoantigens.

This is an interesting paper with some potentially compelling data. As currently presented it is much less convincing than I think it could be. The numbers (number of cryptic events, number of antigens confirmed by MS, number of synthetic antigens that antibodies respond to) are not given context: out of how many possibilities?

How likely were these counts to occur by chance?

This is a difficult question to answer mathematically, but we are convinced that these events are extraordinarily unlikely to occur by chance. Given that neutrophils express close to 10,000 genes with primary RNA transcripts having an average length of 7-15kb, which translated in all 6 reading-frames would collectively yield polypeptides of 70-150 million amino acid residues. This could yield 69,999,985-149,999,985 million 15mer peptides.

Our methodology of filtering against normal control is not new and has been used in cryptic splicing analysis, for example in this paper (Britson et al, Science translational medicine 2022 DOI: 10.1126/scitranslmed.abi9196). We additionally filtered against multiple datasets of disease controls and found only minimal overlap, bolstering our hypothesis that this cryptic splicing is indeed unique to lupus neutrophils, otherwise, we would have seen it in the other datasets as well.

We detect 162,184 splicing events in HCs and 162,579 events in SLE neutrophils, which is an average of 16 per gene transcript. Each splicing event could theoretically occur at $7-15 \times 10^3$ sites, which with 16 events in 10,000 genes, gives a total of 1.12×10^9 random splicing events. Based on the assumption that such random background errors do occur, we used 5 read counts as a cut-off to ensure that each cryptic splice event we recognize had occurred repeatedly. Cryptic events detailed in our paper (Figs. 2 and 3) have read counts of 12 – 48, which is well within the range of canonical splicing events (Fig. 2f and g).

By our definition of 'cryptic' none are present in healthy controls but turning the definition around to ask if there are unannotated splicing events in healthy controls that do not occur in SLE, the answer is 2 events out of 162,184 total splicing events, compared to 271 cryptic splicing events in lupus (of 162,579 events), $p < 0.001$ Fisher's test. We conclude that the rate of cryptic splicing is significantly increased in SLE neutrophils. This is further visualized in the new Fig. S3a-d which compared expression levels of the relevant genes in HC, RA, SLE without the cryptic event and SLE with the event. The new Fig. S3e shows the standard deviation of the percent sliced-in (PSI) as a measure of how much the ratio has changed from HC to SLE.

To further minimize the risk that any of the 271 cryptic events we describe in the paper are stochastic errors of the splicing machinery, we validated a set of them by PCR and Sanger sequencing. Our immunopeptidomics was also designed to determine if the computationally predicted events were real and resulted in sufficient quantities of the translation product to be processed for binding to class I MHC in quantities that can be detected on mass spectrometry.

For all of these reasons, we find it extremely unlikely that the cryptic events we observe were created by chance.

Relatedly there are limited controls described for each experiment.

We have added a set of control peptides to the T cell assays in Fig. 5: 3 peptides from the canonical amino acid sequence of IFI44L selected for high affinity binding to HLA-DR3 and 3 peptides generated by random amino acid residues (out of 20^{15} possible peptides), selected for low scores in pBLAST searches. These 6 peptides did not provoke any response.

More specifically:

- How strong is the agreement between RNA/MS-observed neoantigens and IgG antibody presence (by ELISA)? How likely is this by chance? How does this compare to random antigen sequences?

Peptide identification by mass spectrometry usually relies on the non-redundant human protein data base, which cannot be used for novel sequences. Instead, we used a data base consisting of all computationally derived novel amino acid sequences. Hence, the agreement between RNA and MS is complete.

The group of patients with IgG against neoantigens was enriched for patients who had the relevant neoantigen by RNAseq – indicated with blue dots in Fig. 4. However, some patients had IgGs, but not > 5 read counts of the relevant cryptic event at the time of blood draw and some who did have the cryptic event did not have IgGs. This is not surprising: the development of IgGs against novel antigens requires time and it is well established that the autoantibodies that characterize SLE develop of several years before symptoms develop and a diagnosis can be made (Arbuckle et al, NEJM 2003;349:1526-33). Hence, the IgG antibodies reflect neoantigens that have been present in the past but not necessarily at the moment of blood donation.

- How many cryptic events are seen in the PBMCs that are not seen in the neutrophils?

We ran our analysis pipeline on two PBMC datasets and found zero cryptic events.

Given the complexity of the experiment design, and challenges such as the variable sensitivity of RNA-seq and MS, the authors would benefit from having a statistician (or at least someone with statistical training) work with them to think these analyses through. The one place a statistical analysis is mentioned (T-cell numbers in flow cytometry) it is poorly described, not specifying what is being compared or the hypotheses being tested. Without rigorous statistical analysis and controls it is hard to judge the biological significance of the findings.

The first author is a statistician. We have improved the description of T cell number in the flow cytometry experiments.

The attempt to connect changes in RBP expression to cryptic splicing is also weaker than it could be. Are there binding sites (**predicted or observed in public eCLIP**) for the deregulated RBPs in or near the cryptic events? Are these cryptic events seen after KD of the RBP in the ENCODE K562 KD data?

This was an excellent idea! We have now analyzed our own data against both the eCLIP and shRNA-seq data from ENCODE. The outcomes are presented in a new Supplemental material section and in new Supplemental figures S7 and S8. First, we searched for RBPs binding to coordinates associated with the cryptic splicing events in SLE using a shuffle-based null model and calculated an enrichment score that is normalized for peak counts (to control for RBPs that have relatively widespread binding across the genome). This resulted in a list of 43 RBPs with binding peaks that are statistically significantly correlated with cryptic splicing (heatmap in Fig. S7a and b). Among them were DDX24, SF3B4, and TROVE2 (Ro60), a known target of autoantibodies in SLE. Out of the four patients with cryptic splicing events correlated with Ro60, three had SSA autoantibodies (= Ro52, Ro60, or both). DDX24 was expressed at relatively low levels in the patient with cryptic splicing of TTC14 compared to all SLE samples (Fig. S7c) and SF4B4 was very low in the patient with cryptic splicing of MON2 (Fig. S7c). Fig. S7d illustrates were these two latter proteins bound in the transcript in the eCLIP data.

Next, we used data of shRNA knockdown of RBPs followed by RNA-seq of the K562 cell line. We ran our analysis pipeline and found the cryptic splicing present in SLE neutrophils in 18 RBPs knockdowns (Supplemental Fig. S8a), suggesting that reduced expression (or

dysfunction) of these RBPs causes some of the cryptic splicing events. Among this list, we highlight SF3B4 and DDX24 that also had eCLIP binding peak overlaps (Fig. S7d).

It should also be pointed out that the ENCODE data does not cover all RBPs and that the eCLIP and shRNAi data sets have limited overlap (Fig. S8c) and cover only a portion of all RBPs. For this reason, these data do not provide a conclusive molecular mechanisms for cryptic splicing in SLE but gives us a priority list of candidate RBPs to be explored (beyond the scope of the present paper).

Minor comments:

1-15 per million T cells: is this enough to cause autoimmune disease?

This is a great question! Although a very small number of distinct TCRs can be sufficient to cause a devastating immune response, exemplified by the hapten effect (a small molecule drug conjugated covalently to a protein), we find it more likely that neoantigens work in concert in SLE. The higher the number of distinct T cell clones that are engaged, the stronger the overall immune response will be. It is plausible that cryptic splicing persists for extended periods of time, either in a chronic or episodic fashion, giving the T cell immune response repeated boosts, eventually providing B cell help resulting in anti-neoantigen antibodies, the formation of immune complexes, more inflammation, and worse disease.

Neoantigen-specific T cells in an anti-tumor immune response have frequencies of a very similar magnitude. After release from inhibitory constraints, e.g. by checkpoint inhibitors, they are sufficient for tumor control, sometimes complete eradication of the tumor.

Why would the cryptic events show up in type I interferon-inducible genes?

Our data do not provide a clear answer to this. We speculate that the very high expression of a given ISGs in a patient with cryptic splicing of that ISG (Supplemental figure S3) may exceed the available pool of splice factors, resulting in imperfect splicing. This may not be an early event in SLE but may add to the pool of neoantigens over time or perhaps following a viral infection.

The language could be made more accessible to non-immunologists in various ways. "neutropenic" is not a term I've seen before, sticking to CD8 or cytotoxic, and pointing out early that neutrophils are not PBMCs.

We have followed this recommendation (neutropenia is a common medical term for below normal neutrophil numbers).

Referee #1 (Remarks to the Author):

The authors have addressed the reviewer's concern.

Thank you!

Referee #2 (Remarks to the Author):

Najjar and colleagues provide a revised version of the manuscript highlighting the presence of unique splice variants in neutrophils of SLE patients, some of which may elicit antigen-specific B cells and T cell responses in patients with SLE. The revisions add substantial and new data, including assessment of additional transcriptomic datasets (monocytes, SjD neutrophils, IFN-treated PMN) and new sequencing of TCRs. The new data are much appreciated, though in some key respects go only part way to what I would view as a compelling answer. Overall, the demonstration of Ag-specific B cell and T cell responses still seems to me somewhat weak, with responses of modest magnitude demonstrated from the SLE samples and absence of key controls.

1) Figure 5. Ag-specific T cell responses. We are shown responses of T cells to Ag-stimulation from 5 SLE patients and no control patients.

a) The signals of activated T cells across many Ags, as in Fig 5i, are small (as expected, I agree); however, there is no demonstration that these Ag-induced responses are unique to SLE patients and would not be similarly seen in cells from control patients.

We have remedied this concern by including an experiment with a healthy donor, which shows a complete lack of response to SLE-specific neoantigen peptides. New Supplemental Fig. S9.

b) A demonstration of lack of response to control peptides is now provided (Fig 5e), but it is from a different patient than was used to demonstrate positive responses, with no positive signal to SLE-cryptic peptides shown for this patient, and reads out IL-17 rather than IFN γ ? This is not sufficient; would be required from multiple donors w/ Cryptic vs Canonical peptide comparison.

Given a lack of measurement in control donors (as was done for the Ab signals), and a lack of clear distinction between cryptic vs canonical signal in multiple patients, the case is not compelling that this is a SLE-relevant Ag-specific T cell response.

There is always patient-to-patient variability in human immunology experimentation; IFN γ sometimes gives more non-T cell background, while IL-17 is more specific for T cells, although more lineage restricted. We have now seen a lack of response to canonical peptides in additional patients. Interestingly, there are occasional IL-10 producing T cells, likely Treg, that respond to canonical peptides, as expected.

2) Fig 5j. It is good to see clonal expansion of specific TCRs in the stimulation cultures. But do these expanded TCRs respond to the peptide of interest? Expressing the TCR(s) and

showing reactivity to cryptic peptide of interest, and not other relevant control peptides, is needed to get to the answer.

Yes, we have now utilized TIRTL technology to obtain TCR α/β pair sequences from neoantigen-expanded T cells, rather than only TCR β . This allowed us to transfect T cells containing a NFAT-GFP reporter gene with several specific expanded TCR pairs. One of them responded vigorously to the IFI44L-2 neoantigen. This is now shown in the **new Fig. 5k. Interestingly, a pool of 10mer peptides based on the same IFI44L-2 peptide sequence, did not activate this TCR, indicating that it recognizes the peptide on class II MHC, *i.e.* it came from a CD4 T cell.**

3) Figure 4. Ab detection. It is helpful to see additional control patients assessed for comparison to the SLE cohort. The SLE pts show modest and variable positivity for different neopeptides. However, there is no demonstration that SLE pts would not show the same low level positivity towards control (eg canonical) peptides compared to control donors, or that the binding is truly Ag-specific (can it be competed off?). This is critical given that SLE pts have many polyreactive autoabs and serum from SLE pts can often give low level signals that differ from serum from controls.

It would not be surprising if a SLE patients with a cryptic splice variant of a protein and antibodies against the novel sequence, through epitope-spreading also make some antibodies against other parts of the same protein. Hence, it would not illuminate the immunogenicity of the novel sequence if a patient also can recognize some canonical epitopes in the same protein. We believe that polyreactivity would look like reactivity against all peptides, which is far from the selective reactivity that we see.

4) Addition of transcriptomic assessments of monocytes, SjD neutrophils, and IFN-treated neutrophils is very nice; however, there is minimal information provided about these datasets.

a) How long were neutrophils treated with IFN?

4 hours.

b) For SjD patients, this is a very valuable comparison. How many of these patients show an IFN signature (presumable several of the Ro+ patients?). Though underpowered, is there a hint that IFN-positive SjD patients may show a similar cryptic transcript accumulation? Are they all primary SjD? More information about the cohort would be helpful to underscore the value.

Sjogren's patients had primary disease. 7 patients were Ro+, and 6 of them had high IFN signature. There were 12 cryptic events. The majority of them in a single patient, with one of the events also seen in another patient. These numbers are much lower than in SLE, even relative to sample sizes. Given the many similarities between the

two diseases, one has to wonder if the patient with the 12 events may have had SLE as well.

c) For the monocytes, how many samples are analyzed, and it is from the sample patients as was used for the neutrophils?

There were monocyte samples from 10 SLE patients. 4 of them also had neutrophil data.

For each of the datasets above, it is indicated that there is virtually no overlap with the SLE neutrophil cryptic splice variants. However, for each of the datasets above, do they show their own unique collection of cryptic splice variants (i.e. do monocytes have their own cryptic splice variants? Do IFN-treated neutrophils have their own cryptic variants)?

The IFN-treated neutrophils (n=6 samples) showed 6 unique events, not seen in SLE. In the first revision, we reported that the lupus PBMC had no cryptic splicing.

5) The authors' answers to my previous questions highlight several times the point that IFN-low SLE patients in the cohort also have cryptic splice variants, supporting that this is an SLE-centric phenomenon and not primarily an IFN-centric phenomenon. I appreciate the authors' point here; it would be important to show these data in a manner that make this point clear.

All our data are in the supplemental files and/or will be send to qualified researchers requesting them.

MINOR:

The term 'cryptic' in this paper sounds overly mysterious to me. Based on the authors' rebuttal, it seems to me that 'SLE-associated' or 'SLE-specific' would be a more informative term.

'Cryptic' is a standard term in mRNA splicing and also denotes that the splice variant has not been reported before in any cell type or condition. We used multiple filters to ensure that the cryptic variant had not been reported before in any cell type, gene annotation, or condition. Nevertheless, we will gladly replace 'cryptic' with 'SLE-specific' if the editor asks.

Referee #3 (Remarks to the Author):

I'm satisfied that the authors have addressed my concerns and recommend the work for publication.

Thank you!

Data availability:

Publicly available SLE PBMC data were downloaded from GEO (GSE211700) and (GSE122459). In addition, our RNA-seq data were deposited in GEO (GSE327018).

Upon acceptance, we will make all our data publicly available. At this time, the data can be accessed using the token: wjgtmcaodbqvluv

Referee #1 (Remarks to the Author):

The authors have addressed the reviewer's concern.

Thank you!

Referee #2 (Remarks to the Author):

Najjar and colleagues provide a revised version of the manuscript highlighting the presence of unique splice variants in neutrophils of SLE patients, some of which may elicit antigen-specific B cells and T cell responses in patients with SLE. The revisions add substantial and new data, including assessment of additional transcriptomic datasets (monocytes, SjD neutrophils, IFN-treated PMN) and new sequencing of TCRs. The new data are much appreciated, though in some key respects go only part way to what I would view as a compelling answer. Overall, the demonstration of Ag-specific B cell and T cell responses still seems to me somewhat weak, with responses of modest magnitude demonstrated from the SLE samples and absence of key controls.

1) Figure 5. Ag-specific T cell responses. We are shown responses of T cells to Ag-stimulation from 5 SLE patients and no control patients.

a) The signals of activated T cells across many Ags, as in Fig 5i, are small (as expected, I agree); however, there is no demonstration that these Ag-induced responses are unique to SLE patients and would not be similarly seen in cells from control patients.

We have remedied this concern by including an experiment with a healthy donor, which shows a complete lack of response to SLE-specific neoantigen peptides. New Supplemental Fig. S9.

b) A demonstration of lack of response to control peptides is now provided (Fig 5e), but it is from a different patient than was used to demonstrate positive responses, with no positive signal to SLE-cryptic peptides shown for this patient, and reads out IL-17 rather than IFN γ ? This is not sufficient; would be required from multiple donors w/ Cryptic vs Canonical peptide comparison.

Given a lack of measurement in control donors (as was done for the Ab signals), and a lack of clear distinction between cryptic vs canonical signal in multiple patients, the case is not compelling that this is a SLE-relevant Ag-specific T cell response.

There is always patient-to-patient variability in human immunology experimentation; IFN γ sometimes gives more non-T cell background, while IL-17 is more specific for T cells, although more lineage restricted. We have now seen a lack of response to canonical peptides in additional patients. Interestingly, there are occasional IL-10 producing T cells, likely Treg, that respond to canonical peptides, as expected.

2) Fig 5j. It is good to see clonal expansion of specific TCRs in the stimulation cultures. But do these expanded TCRs respond to the peptide of interest? Expressing the TCR(s) and

showing reactivity to cryptic peptide of interest, and not other relevant control peptides, is needed to get to the answer.

Yes, we have now utilized TIRTL technology to obtain TCR α/β pair sequences from neoantigen-expanded T cells, rather than only TCR β . This allowed us to transfect T cells containing a NFAT-GFP reporter gene with several specific expanded TCR pairs. One of them responded vigorously to the IFI44L-2 neoantigen. This is now shown in the **new Fig. 5k. Interestingly, a pool of 10mer peptides based on the same IFI44L-2 peptide sequence, did not activate this TCR, indicating that it recognizes the peptide on class II MHC, *i.e.* it came from a CD4 T cell.**

3) Figure 4. Ab detection. It is helpful to see additional control patients assessed for comparison to the SLE cohort. The SLE pts show modest and variable positivity for different neopeptides. However, there is no demonstration that SLE pts would not show the same low level positivity towards control (eg canonical) peptides compared to control donors, or that the binding is truly Ag-specific (can it be competed off?). This is critical given that SLE pts have many polyreactive autoabs and serum from SLE pts can often give low level signals that differ from serum from controls.

It would not be surprising if a SLE patients with a cryptic splice variant of a protein and antibodies against the novel sequence, through epitope-spreading also make some antibodies against other parts of the same protein. Hence, it would not illuminate the immunogenicity of the novel sequence if a patient also can recognize some canonical epitopes in the same protein. We believe that polyreactivity would look like reactivity against all peptides, which is far from the selective reactivity that we see.

4) Addition of transcriptomic assessments of monocytes, SjD neutrophils, and IFN-treated neutrophils is very nice; however, there is minimal information provided about these datasets.

a) How long were neutrophils treated with IFN?

4 hours.

b) For SjD patients, this is a very valuable comparison. How many of these patients show an IFN signature (presumable several of the Ro+ patients?). Though underpowered, is there a hint that IFN-positive SjD patients may show a similar cryptic transcript accumulation? Are they all primary SjD? More information about the cohort would be helpful to underscore the value.

Sjogren's patients had primary disease. 7 patients were Ro+, and 6 of them had high IFN signature. There were 12 cryptic events. The majority of them in a single patient, with one of the events also seen in another patient. These numbers are much lower than in SLE, even relative to sample sizes. Given the many similarities between the

two diseases, one has to wonder if the patient with the 12 events may have had SLE as well.

c) For the monocytes, how many samples are analyzed, and it is from the sample patients as was used for the neutrophils?

There were monocyte samples from 10 SLE patients. 4 of them also had neutrophil data.

For each of the datasets above, it is indicated that there is virtually no overlap with the SLE neutrophil cryptic splice variants. However, for each of the datasets above, do they show their own unique collection of cryptic splice variants (i.e. do monocytes have their own cryptic splice variants? Do IFN-treated neutrophils have their own cryptic variants)?

The IFN-treated neutrophils (n=6 samples) showed 6 unique events, not seen in SLE. In the first revision, we reported that the lupus PBMC had no cryptic splicing.

5) The authors' answers to my previous questions highlight several times the point that IFN-low SLE patients in the cohort also have cryptic splice variants, supporting that this is an SLE-centric phenomenon and not primarily an IFN-centric phenomenon. I appreciate the authors' point here; it would be important to show these data in a manner that make this point clear.

All our data are in the supplemental files and/or will be send to qualified researchers requesting them.

MINOR:

The term 'cryptic' in this paper sounds overly mysterious to me. Based on the authors' rebuttal, it seems to me that 'SLE-associated' or 'SLE-specific' would be a more informative term.

'Cryptic' is a standard term in mRNA splicing and also denotes that the splice variant has not been reported before in any cell type or condition. We used multiple filters to ensure that the cryptic variant had not been reported before in any cell type, gene annotation, or condition. Nevertheless, we will gladly replace 'cryptic' with 'SLE-specific' if the editor asks.

Referee #3 (Remarks to the Author):

I'm satisfied that the authors have addressed my concerns and recommend the work for publication.

Thank you!

Data availability:

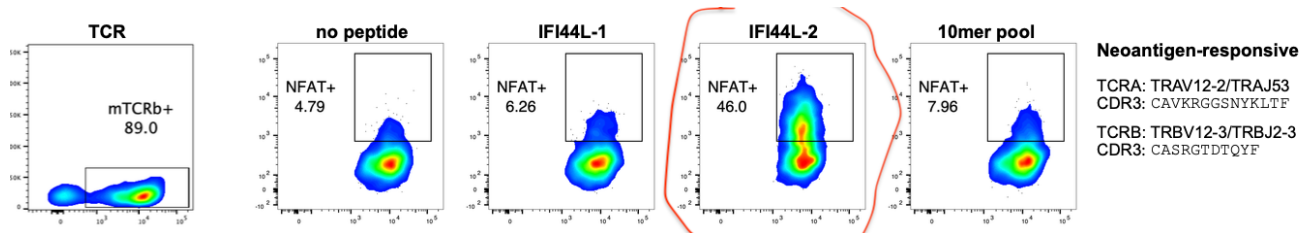
Publicly available SLE PBMC data were downloaded from GEO (GSE211700) and (GSE122459). In addition, our RNA-seq data were deposited in GEO (GSE327018).

Upon acceptance, we will make all our data publicly available. At this time, the data can be accessed using the token: wjgtmcaodbqvluv

Point-by-point response to Referee #2 (partly prospective):

The authors provide additional data to demonstrate that 1) a single healthy donor showed a lack of T cell responses to cryptic peptides (Supp Fig 9) and that 2) one TCR isolated from neoantigen-expanded T cells could be shown to respond to peptides from an IFI44L neoantigen.

Response: We have performed the TCR sequencing and recognition analysis as requested by reviewer 2. The TCR recognized one specific peptide (IFI44L-2; DLTPRNRYSICGLAY) and did not react with another neoantigen peptide (VNPQFKPDLTPRNRY).



In addition, we showed that this peptide must have been presented by class II MHC, since 10mer versions of it were not recognized. This means that the patient T cell with this TCR was a CD4 T cell. We will point this out more clearly in the revised manuscript.

From prior comments:

Prior 1a: “Figure 5. Ag-specific T cell responses. We are shown responses of T cells to Ag-stimulation from 5 SLE patients and no control patients.

a) The signals of activated T cells across many Ags, as in Fig 5i, are small (as expected, I agree); however, there is no demonstration that these Ag-induced responses are unique to SLE patients and would not be similarly seen in cells from control patients.”

The authors now provide evidence from 1 control donor showing no responsiveness to the tested peptides. This is OK, but as the authors indicate, ‘there is always patient-to-patient variability in human immunology experimentation.’ Showing one single control donor does not make a convincing case that the varied antigen-specific responses detected in 3 SLE patients (Fig 5b-d) is robustly disease-specific. The authors may be correct, but it remains highly uncertain from the data shown.

Response: We already have an additional independent SLE patient to include and plan to include additional patients. Assuming that ‘control patients’ means patients with another disease, we will include rheumatoid arthritis as a control disease.

Prior 1b: “A demonstration of lack of response to control peptides is now provided (Fig 5e), but it is from a different patient than was used to demonstrate positive responses, with no positive signal to SLE-cryptic peptides shown for this patient, and reads out IL-17 rather than IFN γ ? This is not sufficient; would be required from multiple donors w/ Cryptic vs Canonical peptide comparison. Given a lack of measurement in control donors (as was done for the Ab signals), and a lack of clear distinction between cryptic vs canonical signal in multiple patients, the case is not compelling that this is a SLE-relevant Ag-specific T cell response.”

The authors explain that patient responses vary, and there can be background IFN production detected, thus they show IL17 from Patient D rather than IFN. This is clearly not an adequate control. What do the responses of Patients A, B, C look like when stimulated with canonical peptides?

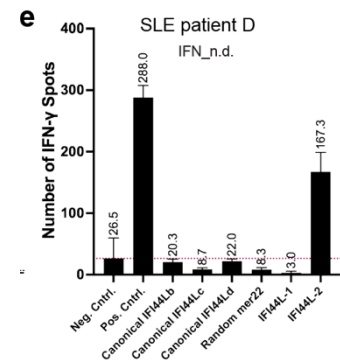
We are currently presented with :

- Positive but varied responses to a range of cryptic peptides in 3 SLE patients by IFN γ
- Negative response to canonical peptide in 1 SLE patient by IL17
- Negative response to cryptic peptides in 1 Control by IFN γ

Response: We will replace the IL-17 data with the IFN γ response in another SLE patient stimulated with both canonical (no response) and neoantigen peptides (response). We already have these data in hand.

The results text states: “In contrast, T cells from healthy donors had little response to any of the neoantigen peptides (not shown).”

The authors explain in their response that IFN γ readout can be noisy, that they have ‘seen a lack of response to canonical peptides in additional patients’ (not shown), and that sometimes IL-10 is detected in response to canonical peptides (not shown). In my view, this assessment of a small number of patients (n=1 in two key controls conditions, one with an alternate readout), is not sufficient to make a conclusion about the nature of autoantigen targets in patients with lupus. In total, I do not find this a compelling data package for the very impactful claim that the authors are making.



Response: We will add two additional healthy controls (although these neoantigens do NOT exist in healthy donors, which therefore cannot have an immune response against them).

Prior 2: “Fig 5j. It is good to see clonal expansion of specific TCRs in the stimulation cultures. But do these expanded TCRs respond to the peptide of interest? Expressing the TCR(s) and showing reactivity to cryptic peptide of interest, and not other relevant control peptides, is needed to get to the answer.”

I applaud the authors for the effort to get full a/b TCRs and test reactivity. The authors demonstrate a single TCR with reactivity to IFI44L-2 peptides. However, it is unclear how many TCRs were tested that did not show reactivity in these analyses – this would be typical to report and important to understand. Further, It’s unclear why the authors have not gotten to a single peptide. It does not seem that they have demonstrated a specificity of the reactivity to cryptic, vs canonical, peptide, which would be expected (and critical) for the overall concept.

Response: This TCR recognized one single peptide (IFI44L-2) and did not react with another neoantigen peptide (see first point above). 500,000 T cells were evaluated for antigen-specific expansion, 6,328 paired TCR clones isolated, 18 of them (0.0036%) prioritized for testing, and found 1 reactive with the IFI44L-2 peptide. We are now proceeding to test more TCRs and to show that the first TCR does not react with several additional neoantigen peptides.

Prior 3: “Figure 4. Ab detection. It is helpful to see additional control patients assessed for. comparison to the SLE cohort. The SLE pts show modest and variable positivity for

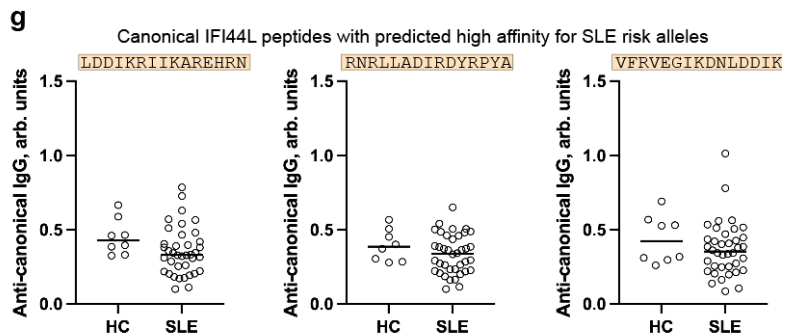
different neopeptides. However, there is no demonstration that SLE pts would not show the same low level positivity towards control (eg canonical) peptides compared to control donors, or that the binding is truly Ag-specific (can it be competed off?). This is critical given that SLE pts have many polyreactive autoabs and serum from SLE pts can often give low level signals that differ from serum from controls.”

There were two aspects to the point here: 1) non-specific signal in SLE serum samples, and 2) ability to distinguish between cryptic and canonical epitopes.

Aspect 1 is not addressed. Would these SLE samples show increased signal compared to control samples is assayed against a set of irrelevant control peptides? This is a critical missing control to assess the extent to which the signal is specific to the cryptic peptides.

Aspect 2 - the authors do not find it important to assess the specificity of the Ab response to cryptic epitopes versus canonical proteins. They anticipate that Ab-reactivity to the canonical proteins may be present by epitope spreading. This may be correct; in this case, we might conclude that SLE patients have antibodies against may IFN-induced proteins. The contribution of cryptic recognition is uncertain. Yet testing only binding to cryptic peptides implies a certain specificity to cryptic peptides that the authors recognize may not be the case. Given that the authors underscore autoimmunity to cryptic neoantigens as a core concept of this paper, it seems a highly relevant point that remains unclear.

We have done these controls and will add these 3 panels to Fig. 4 showing that 3 canonical peptides from IFI44L are NOT recognized by IgGs in SLE patients or healthy donors.



Prior 5: “The authors’ answers to my previous questions highlight several times the point that IFN-low SLE patients in the cohort also have cryptic splice variants, supporting that this is an SLE-centric phenomenon and not primarily an IFN-centric phenomenon. I appreciate the authors’ point here; it would be important to show these data in a manner that make this point clear.”

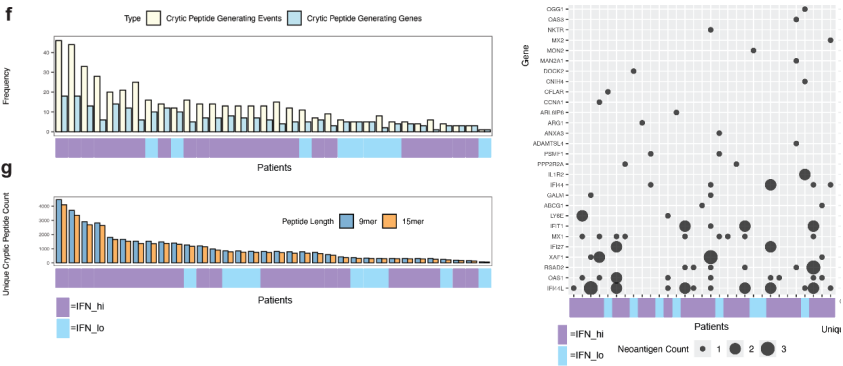
The authors response to this seems to be – ‘go find it yourself.’ The authors have not made a compelling case that this is phenomenon exists independently from a strong IFN response, and that the cryptic variants detected are not primarily a product of an ongoing IFN response (independent of SLE). The data may be in there somewhere, but it is not evident and not currently convincing.

Response: We are responding to this point by adding the IFN status to every patient in Fig. 1 and Fig. 5 and adding a Supplemental Table that summarizes the IFN status of our SLE patients, their cryptic splicing of ISGs, and peptides from them on class I MHC by mass spec. This will better illustrate the relationship (but not causality) between IFN and cryptic mRNA splicing.

Patient	SLEDAI	avg_ISG	ISG status	cryptic ISG	neoAg on MHC-I by mass spec (ISG)
1	2	14781	ISG_high	yes	
2	2	4206	ISG_low	yes	MX1, ANXA3, FCGRT
3	5	17262	ISG_high	yes	PHF, XAF1, OAS1, OAS3
4	2	14035	ISG_high	yes	
5	17	8869	ISG_high	no	
6	12	13405	ISG_high	yes	
7	2	5852	ISG_high	no	
8	8	8880	ISG_high	no	
9	4	16439	ISG_high	yes	
10	12	17871	ISG_high	yes	
11	0	17871	ISG_high	yes	
12	2	26772	ISG_high	yes	
13	19	2388	ISG_low	no	
14	9	17098	ISG_high	yes	
15	4	2987	ISG_low	no	
16	36	6958	ISG_high	yes	
17	4	21636	ISG_high	yes	
18	2	12305	ISG_high	yes	
19	6	12981	ISG_high	no	
20	2	15516	ISG_high	yes	
21	2	14207	ISG_high	yes	
22	4	10714	ISG_high	yes	
23	4	2632	ISG_low	no	
24	2	10939	ISG_high	yes	USP18, RSAD2, MYO15B, ADAMTSL4
25	5	24823	ISG_high	yes	
26	16	4318	ISG_low	yes	
27	0	5679	ISG_high	no	
28	14	4105	ISG_low	no	
29	12	4317	ISG_low	yes	OGG1, IFI44L, TRAPPC10
30	4	15119	ISG_high	yes	
31	4	30574	ISG_high	yes	
32	4	6936	ISG_high	no	
33	4	12726	ISG_high	yes	OAS1, ABCG1
34	1	3744	ISG_low	no	
35	4	3686	ISG_low	no	

HC 95%ile =4828

active



Minor:

As indicated initially, I do not think a X2 test is appropriate in Fig 5h.

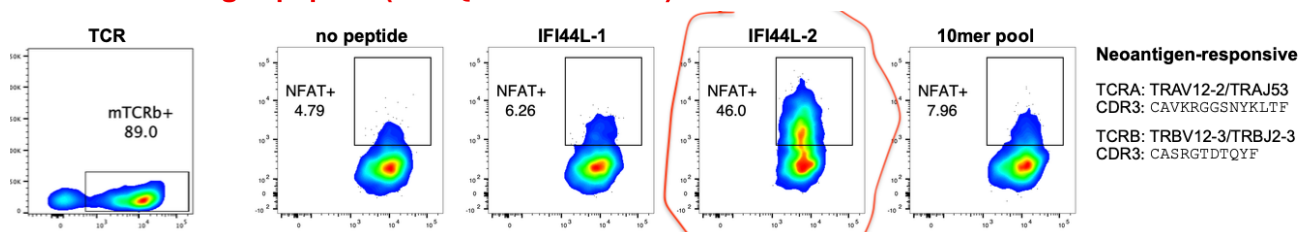
The first author of our paper is a statistician and feels that c^2 is the most appropriate statistical analysis method when comparing cell numbers.

We are looking forward to the your and the editorial team’s view of these responses. Which points are critical versus not necessary? Are there additional data we should generate?

Point-by-point response to Referee #2:

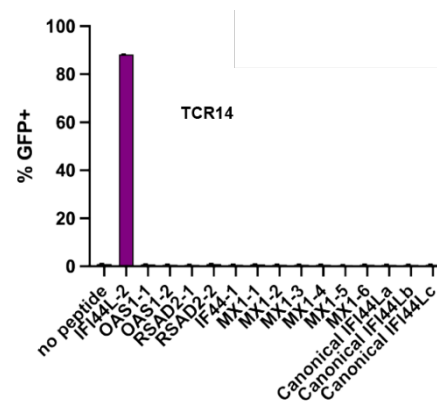
The authors provide additional data to demonstrate that 1) a single healthy donor showed a lack of T cell responses to cryptic peptides (Supp Fig 9) and that 2) one TCR isolated from neoantigen-expanded T cells could be shown to respond to peptides from an IFI44L neoantigen.

Response: We have performed the TCR sequencing and recognition analysis as requested by reviewer 2. The TCR recognized one specific peptide (IFI44L-2; DLTPNRYSICGLAY) and did not react with another neoantigen peptide (VNPQFKPDLTNRNY).



Now we have also tested the recognition by this TCR of 11 additional neoantigens and 3 canonical IFI44L peptides and find that it is exquisitely selective for the IFI44L-2 peptide. This is now panel I in Fig. 5.

In addition, we found that the IFI44L-2 peptide must have been presented by class II MHC, since 10mer versions of it were not recognized. This means that the patient T cell with this TCR was a CD4 T cell. We are pointing this out more clearly in the revised manuscript.



From prior comments:

Prior 1a: “Figure 5. Ag-specific T cell responses. We are shown responses of T cells to Ag-stimulation from 5 SLE patients and no control patients.

a) The signals of activated T cells across many Ags, as in Fig 5i, are small (as expected, I agree); however, there is no demonstration that these Ag-induced responses are unique to SLE patients and would not be similarly seen in cells from control patients.”

The authors now provide evidence from 1 control donor showing no responsiveness to the tested peptides. This is OK, but as the authors indicate, ‘there is always patient-to-patient variability in human immunology experimentation.’ Showing one single control donor does not make a convincing case that the varied antigen-specific responses detected in 3 SLE patients (Fig 5b-d) is robustly disease-specific. The authors may be correct, but it remains highly uncertain from the data shown.

Response: We have increased the shown data to 9 SLE patients (Fig. 5 + Suppl. Fig. S9 and S11) and added 3 additional healthy donors (Suppl. Fig. S10), as well as 2 disease control patients with rheumatoid arthritis (Suppl. Fig S10). These show more reactivity with neoantigen peptides in SLE patients and no reactivity to any of the neoantigens or control peptides in healthy and RA subjects.

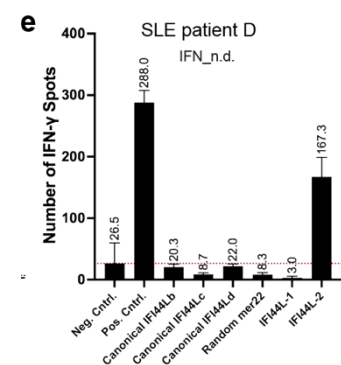
Prior 1b: “A demonstration of lack of response to control peptides is now provided (Fig 5e), but it is from a different patient than was used to demonstrate positive responses, with no positive signal to SLE-cryptic peptides shown for this patient, and reads out IL-17 rather than IFN γ ? This is not sufficient; would be required from multiple donors w/ Cryptic vs Canonical peptide comparison. Given a lack of measurement in control donors (as was done for the Ab signals), and a lack of clear distinction between cryptic vs canonical signal in multiple patients, the case is not compelling that this is a SLE-relevant Ag-specific T cell response.”

The authors explain that patient responses vary, and there can be background IFN production detected, thus they show IL17 from Patient D rather than IFN. This is clearly not an adequate control. What do the responses of Patients A, B, C look like when stimulated with canonical peptides?

We are currently presented with :

- Positive but varied responses to a range of cryptic peptides in 3 SLE patients by IFN γ
- Negative response to canonical peptide in 1 SLE patient by IL17
- Negative response to cryptic peptides in 1 Control by IFN γ

Response: We have replaced the IL-17 data with the IFN γ response in another SLE patient stimulated with both canonical (no response) and neoantigen peptides (response). We also added several patients and control subjects, as noted above.



The results text states: “In contrast, T cells from healthy donors had little response to any of the neoantigen peptides (not shown).”

The authors explain in their response that IFN γ readout can be noisy, that they have ‘seen a lack of response to canonical peptides in additional patients’ (not shown), and that sometimes IL-10 is detected in response to canonical peptides (not shown). In my view, this assessment of a small number of patients (n=1 in two key controls conditions, one with an alternate readout), is not sufficient to make a conclusion about the nature of autoantigen targets in patients with lupus. In total, I do not find this a compelling data package for the very impactful claim that the authors are making.

Response: We have added 3 additional healthy controls (HC2-4) with no response to SLE neoantigens as well as 2 patients with rheumatoid arthritis, also with no response. We have also added SLE patients and data to Supplemental Fig. S9 and S11 to a total of 8 patients all with response to cryptic splicing-derived neoantigens.

Prior 2: “Fig 5j. It is good to see clonal expansion of specific TCRs in the stimulation cultures. But do these expanded TCRs respond to the peptide of interest? Expressing the TCR(s) and showing reactivity to cryptic peptide of interest, and not other relevant control peptides, is needed to get to the answer.”

I applaud the authors for the effort to get full a/b TCRs and test reactivity. The authors demonstrate a single TCR with reactivity to IFI44L-2 peptides. However, it is unclear how many TCRs were tested that did not show reactivity in these analyses – this would be typical to report and important to understand. Further, It’s unclear why the authors have not gotten to a single peptide. It does not seem that they have demonstrated

a specificity of the reactivity to cryptic, vs canonical, peptide, which would be expected (and critical) for the overall concept.

Response: This TCR recognized one single peptide (IFI44L-2) and did not react with another neoantigen peptide (see first point above). We proceeded to show that this TCR does not react with 11 additional neoantigen peptides or 3 canonical peptides (new panel I in Fig. 5). 500,000 T cells were evaluated for antigen-specific expansion, 6,328 paired TCR clones isolated, 18 of them (0.0036%) prioritized for testing, and 1 found to be reactive with the IFI44L-2 peptide.

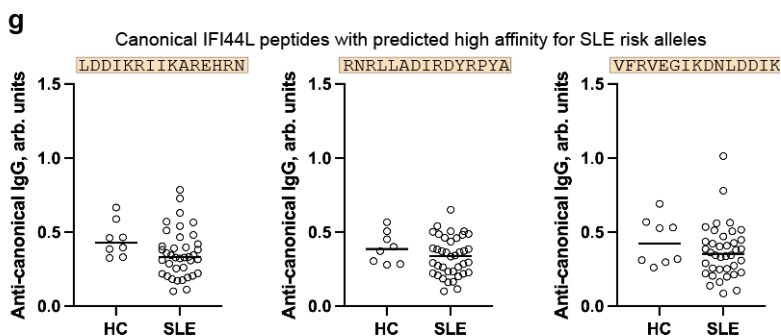
Prior 3: “Figure 4. Ab detection. It is helpful to see additional control patients assessed for. comparison to the SLE pts show modest and variable positivity for different neopeptides. However, there is no demonstration that SLE pts would not show the same low level positivity towards control (eg canonical) peptides compared to control donors, or that the binding is truly Ag-specific (can it be competed off?). This is critical given that SLE pts have many polyreactive autoabs and serum from SLE pts can often give low level signals that differ from serum from controls.”

There were two aspects to the point here: 1) non-specific signal in SLE serum samples, and 2) ability to distinguish between cryptic and canonical epitopes.

Aspect 1 is not addressed. Would these SLE samples show increased signal compared to control samples is assayed against a set of irrelevant control peptides? This is a critical missing control to assess the extent to which the signal is specific to the cryptic peptides.

Aspect 2 - the authors do not find it important to assess the specificity of the Ab response to cryptic epitopes versus canonical proteins. They anticipate that Ab-reactivity to the canonical proteins may be present by epitope spreading. This may be correct; in this case, we might conclude that SLE patients have antibodies against may IFN-induced proteins. The contribution of cryptic recognition is uncertain. Yet testing only binding to cryptic peptides implies a certain specificity to cryptic peptides that the authors recognize may not be the case. Given that the authors underscore autoimmunity to cryptic neoantigens as a core concept of this paper, it seems a highly relevant point that remains unclear.

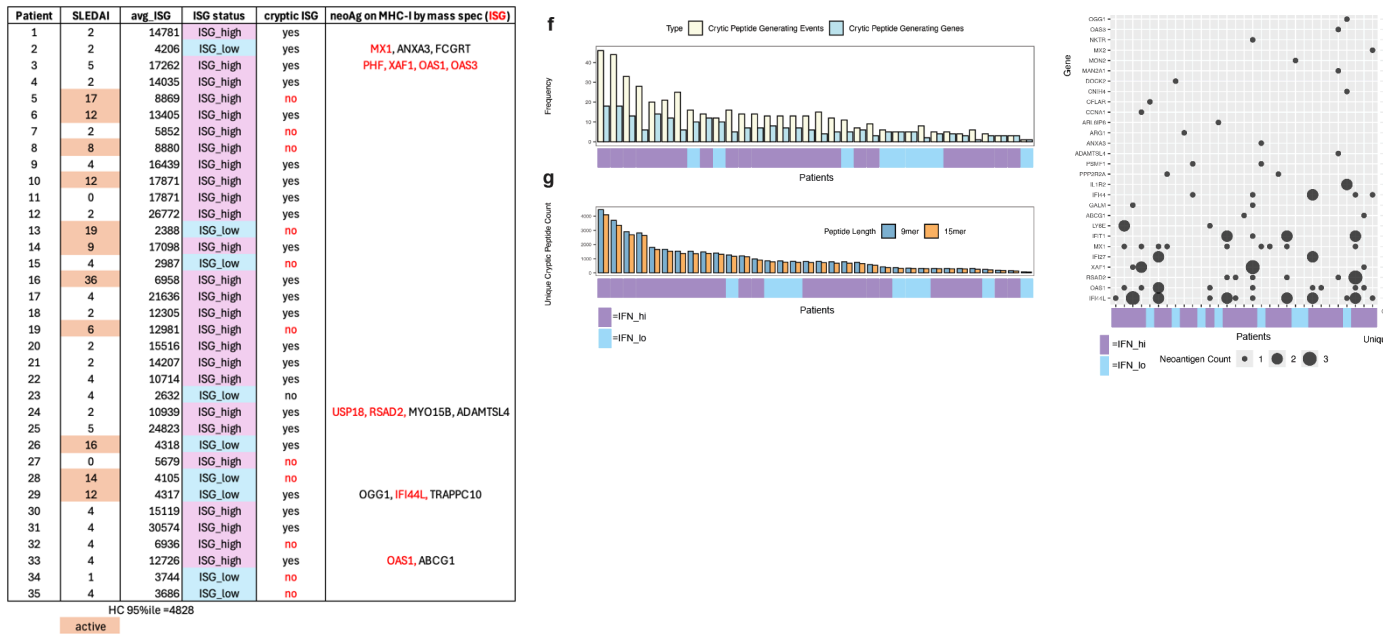
We have performed these control experiments and have added these 3 panels to Fig. 4 showing that 3 canonical peptides from IFI44L are NOT recognized by IgGs in SLE patients or healthy donors.



Prior 5: “The authors’ answers to my previous questions highlight several times the point that IFN-low SLE patients in the cohort also have cryptic splice variants, supporting that this is an SLE-centric phenomenon and not primarily an IFN-centric phenomenon. I appreciate the authors’ point here; it would be important to show these data in a manner that make this point clear.”

The authors response to this seems to be – ‘go find it yourself.’ The authors have not made a compelling case that this is phenomenon exists independently from a strong IFN response, and that the cryptic variants detected are not primarily a product of an ongoing IFN response (independent of SLE). The data may be in there somewhere, but it is not evident and not currently convincing.

Response: We have responded to this point by adding the IFN status to every patient in Fig. 1 and Fig. 5 and adding a Supplemental Table that summarizes the IFN status of our SLE patients, their cryptic splicing of ISGs, and peptides from them on class I MHC by mass spec. This will better illustrate the relationship (but not causality) between IFN and cryptic mRNA splicing.



Minor:

As indicated initially, I do not think a X2 test is appropriate in Fig 5h.

We tested cells from antigen vs positive control which were either activated or not activated. This difference can be tested with a Chi-square test, an approach that has been used before¹.

We hope that you will now find our responses and much additional data in response to reviewer 2 sufficient for accepting our paper for publication.

Yours sincerely,

Tomas

1. Feng B, Jorgensen JL, Jones D, et al. Flow cytometric detection of peripheral blood involvement by mycosis fungoides and Sezary syndrome using T-cell receptor Vbeta chain antibodies and its application in blood staging. *Mod Pathol*. Feb 2010;23(2):284-95. doi:10.1038/modpathol.2009.175

Point-by-point response to Referee #2:

I am satisfied with the revisions. The additional controls are impactful. The identification of a defined TCR, with sequence provided, and the specific peptide that it recognizes, is impressive.

Thank you!

Two minor comments for the authors to consider:

1) The quantification of GFP+ cells in Figure 5L does not match the examples shown in Figure 5K (though the qualitative message is clear).

We agree, there is a difference (46% in Fig. 5k and 88% in panel l). We also agree that the qualitative message is clear.

2) A chi-squared test assumes that the observations are independent. The observations are not independent when sampling many cells out of the same stimulated sample. This is why a chi-squared test is not appropriate in the context used.

We contend that they are independent because each subculture was stimulated with a different neoantigen peptide i.e. they are not derived from a positive control stimulation condition.

In response to points raised by editorial assistants:

1. Please reduce the length of the title to 75 characters (with spaces) or less, so that it fits on two lines in the final layout.

We have shortened the title to: “T and B cells recognize neoantigens from aberrant mRNA splicing in lupus neutrophils”, which is 72 characters

2. The number of references should generally not exceed 60. Currently you have 73 references

We have removed references to have 60 for the main article

3. Please create a separate reference list for any methods references, making sure that the numbering continues from the main text references.

Methods references are now numbers 61-75

4. Please remove the main figures from the article file and re-supply them individually in an acceptable format such as EPS, AI, PS, PDF, PPT, PSD or XLS (for graphs) with editable vector files.

Submitted as .ai

5. Please note that the legends for the main figures should not exceed 300 words. If it is not possible to reduce the length accordingly, please ensure that the final legends are as close as possible to 300 words

We have shortened the legend for Fig. 5 to 292 words. The other figure legends are also <300.

6. Please reduce subheadings to 40 characters (with spaces) or less.

neoantigens from cryptic splicing in SLE

7. Please provide a supplementary information guide as a separate word document.

8. There are potential third party rights issues in the figures. Please check the sources of all illustrations and clarify whether permissions are needed to adapt or reproduce them. Please make sure to include the relevant details in third party rights table when you resubmit (more information below). If Biorender or a similar software has been used, please also ensure to provide relevant licenses. In particular please check figure/s: Figure 1a,

There are no third party rights issues. All schematic parts of figures (e.g. Fig. 1a) were created *de novo* using Adobe Illustrator

9. Please make sure to provide a third party rights table (more information below) when you resubmit. If Biorender or similar software has been used, please also ensure to provide relevant licenses.

N/A. Biorender was not used.

10. Flagging that there are display items in the SI - please check if any of these materials should be extended data and notify us if they require further reproducibility checks.

Checked.

11. Please remove Extended data from article file.

Done.

12. Please ensure that the text size in all figures is at least 5 pt Arial.

It is now (we enlarged some)

13. The funding information has been merged with the acknowledgment section; please provide it as a separate section.

done

	EDITORIAL REQUESTS:	AUTHOR RESPONSE:
1.	Data presentation: Please ensure that data presented in a plot, chart or other visual representation format shows data distribution clearly (e.g. dot plots, box-and-whisker plots). When using bar charts, please overlay the corresponding data points (as dot plots) whenever possible and always for $n \leq 10$. (Please see the following editorial for the rationale behind this request and an example https://www.nature.com/articles/s41551-017-0079). Panels requiring revision: Please note that data presentation has to be revised to comply with our policy in figure(s) 5b-e	We have revised Fig. 5b-e to show individual data points
2.	Statistics: Wherever statistics have been derived (e.g. error bars, box plots, statistical significance) the legend needs to provide and define the n number (i.e. the sample size used to derive statistics) as a precise value (not a range), using the wording “n=X biologically independent samples/animals/cells/independent experiments/n= X cells examined over Y independent experiments” etc. as applicable. Legends requiring revision: Please note that this information is missing in the legend(s) of figure(s) 1e,2f-g,4g,5b-e	The legends to Figs. 1e, 2f-g, 4g, and 5b-e have been revised
3.	Please note that statistics such as error bars significance and p values cannot be derived from $n < 3$ and must be removed from all such cases. We strongly discourage deriving statistics from technical replicates, unless there is a clear scientific justification for why providing this information is important. Conflating technical and biological variability, e.g., by pooling technically replicates samples across independent experiments is strongly discouraged. (For examples of expected description of statistics in figure legends, please see the following https://www.nature.com/articles/s41467-019-11636-5 or https://www.nature.com/articles/s41467-019-11510-4). All error bars need to be defined in the legends (e.g. SD, SEM) together with a measure of centre (e.g. mean, median). For example, the legends should state something along the lines of “Data are presented as mean values $\pm$ SEM” as appropriate. All box plots need to be defined in the legends in terms of minima, maxima, centre, bounds of box and whiskers and percentile. Legends requiring revision:5o Please note that the error bars/error bands need to be defined in the legend(s) of figure(s) 4a-g,5b-e,5o Please note that the box plots need to be defined in terms of minima, maxima, centre, bounds of box and whiskers and percentile in the legend(s) of figure(s) 1e,2f-g	The legends have been revised to state that they are mean $\pm$ SD
4.	The figure legends must indicate the statistical test used. Where appropriate, please indicate in the figure legends whether the statistical tests were one-sided or two-sided and whether adjustments were made for multiple comparisons.	

	For null hypothesis testing, please indicate the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P values noted. Please provide the test results (e.g. P values) as exact values whenever possible and with confidence intervals noted.	
	Legends requiring revision: Please indicate the statistical test used for data analysis and where appropriate, please specify whether it was one-sided or two-sided and whether adjustments were made for multiple comparisons, in the legend(s) of figure(s) 5h-i Please note that the information on whether the statistical test used was one-sided or two-sided, where appropriate, is missing in the legend(s) of figure(s) 5o Please note that the exact p value should be provided, when possible, in the legend(s) of figure(s) 5h-i,5o Please indicate what */ **/ ***/ **** represents; if this represents p value(s), please indicate the statistical test used and where appropriate, specify whether it was one-sided or two-sided and whether adjustments were made for multiple comparisons and the exact p value in the legend(s) of figure(s) 1e	We have added this. We also added a ‘Statistical testing’ section to the Methods. All were two-sided. Yes, these indicate *p<0.05, **p,0.01, ***p<0.005, ****p,0.001 (adding this to the legend will make it >300 words)
5.	Reproducibility: Please state in the legends how many times each experiment was repeated independently with similar results. This is needed for all experiments, but is particularly important wherever results from representative experiments (such as micrographs) are shown. If space in the legends is limiting, this information can be included in a section titled “Statistics and Reproducibility” in the methods section.	
6.	Legends requiring revision: revise this information in the legend(s) of figure(s) 2a-e	Revised. Gels were run 3 times and with additional patient samples (hence 3-5 replicates), but more important for reproducibility was obtaining the expected DNA sequence by Sanger sequencing.
	Data availability: This journal strongly supports public availability of data and custom code associated with the paper in a persistent repository where they can be freely and enduringly accessed or as a supplementary data file when no appropriate repository is available. If data and code can only be shared on request, please explain why in your data Availability Statement, and also in the correspondence with your editor. For more information, please	

	refer to https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data	
	Please ensure that datasets deposited in public repositories are now publicly accessible, and that accession codes or DOI are provided in the "Data Availability" section. As long as these datasets are not public, we cannot proceed with the acceptance of your paper. For data that have been obtained from publicly available sources, please provide a URL and the specific data product name in the data availability statement. Data with a DOI should be further cited in the methods reference section.	done
7.	Gels and Blots: Quantitative comparisons between samples on different gels/blots are discouraged; if this is unavoidable, the figure legend must state that the samples derive from the same experiment and that gels/blots were processed in parallel. Vertically sliced images that juxtapose lanes that were non-adjacent in the gel must have a clear separation or a black line delineating the boundary between the gels. Loading controls (e.g. GAPDH, actin) must be run on the same blot. Sample processing controls run on different gels must be identified as such in the figure legends, and distinctly from loading controls. All blots and gels must be accompanied by the locations of molecular weight/size markers. Blots should be cropped such that at least one marker position is present. Please also supply uncropped and unprocessed scans of the most important blots in the Source Data file or as a supplementary figure in the Supplementary Information. This should be cited once in the Methods section. For an example of presentation of full scan blots, see the Source Data file of https://www.nature.com/articles/s41467-020-16984-1#Sec35 and for more information, please refer to https://www.nature.com/nature-research/editorial-policies/image-integrity These are now Supplementary Fig. S12	
8.	Micrographs: Please ensure that all micrographs include a scale bar and this scale bar is defined on the panels or in the figure legends.	
	Panels requiring revision: Please note that scale bar is missing for figure(s) 5	A white 5 µm scale bar has been added
9.	Flow cytometry data: Please provide a Supplementary Figure to graphically account for all FACS sequential gating/sorting strategies, or provide gating/sorting strategies in-figure. If the former, please be sure to indicate, in the Supplementary Figure legend, which gating panel(s) correspond to which FACS data panel(s) in the manuscript figures. (For an example, please see https://www.nature.com/articles/ncomms15067#supplementary-information). This is now Supplementary Fig. S13	