eLife

Reviewed Preprint
v1 • May 5, 2026
Not revised

For correspondence:
dev.majumdar@uvm.edu

Competing interests: No competing interests declared

Funding: See page 29

Reviewing editor: Larissa D Cunha, University of Sao Paulo, Brazil

Programmed Delayed Splicing: A Mechanism for Timed Inflammatory Gene Expression

Jacob S Dearborn[1], Luke Frankiw[2], Damas W Limoge[3], Christian H Burns[2], Logan Vlach[2], Patricia Turpin[2], Tylar Kirch[1], Zachary D Miller[1], William Dowell[1], Sylvester Languon[1], Yvette Garcia-Flores[1], Robert C Cockrell[1], David Baltimore[1], Devdoot Majumdar[1,2]

[1]Dept of Surgical Research, Larner College of Medicine, University of Vermont, Burlington, United States • [2]Division of Biology, California Institute of Technology, Pasadena, United States • [3]Nanotronics, New York, United States

eLife Assessment

This study analyzes the temporal dynamics of gene expression following TNF stimulation in macrophages. The work brings **valuable** data and new methodological approaches to implicate the splicing rate of certain introns as a mechanism regulating mature mRNA expression. This will be of interest to audiences in RNA biology and innate immune response regulation. The experimental design is **solid** for the core findings, although in places the data limit the conclusions.

https://doi.org/10.7554/eLife.109726.1.sa3

Abstract

Inflammation involves timed gene expression, suggesting that the fine-tuned onset, amplitude, and termination of expression of hundreds of genes is of critical importance to organismal homeostasis. Recent study of post-transcriptional regulation of inflammatory gene expression led to the suggestion of a regulatory role for pre-mRNA splicing. Here, using a hybrid capture approach to purify incompletely spliced, chromatin-associated pre-mRNAs, we use deep sequencing to study pre-mRNA splicing of the NF-kB transcriptome. By freezing transcription and examining subsequent splicing of complete transcripts, we find many introns splice tens to hundreds of times slower than average. Investigating the basis of these delays, we focused on evolutionarily conserved introns with suboptimal splice donor sequences and found that strengthening these donor sites by as few as two nucleotides in minigene reporter assays markedly increased gene expression for several targets. This suggests that such sites can act as timing elements that both delay mRNA production and limit expression amplitude. To broaden this mechanistic view, we applied deep learning sequence-to-function models with feature attribution to identify additional regulatory sequences—both intronic and exonic—that may contribute to delayed splicing through mechanisms independent of donor site strength. This integrated approach revealed non-canonical motifs enriched in slow-splicing introns, pointing to a broader repertoire of cis-elements that can fine-tune transcript maturation during inflammation. Together, these findings support a model in which the temporal regulation of pre-mRNA splicing serves as a layer of control in inflammatory gene expression, and raise the possibility that similar timing mechanisms operate in other rapid-response transcriptional programs.

Introduction

Gene expression in response to an inflammatory stimulus begins rapidly and is tightly controlled by conventional means (transcription and protein turnover (1–3)) and by an expanding list of modalities that have gained in appreciation as being general regulatory strategies (RNA stabilization, RNA deadenylation, ribosomal regulation, microRNA regulation, as examples (4–7)).

We and others have recently investigated the role of RNA splicing kinetics – independently of alternative splicing – in gene expression (8–11). In macrophages, an inflammatory stimulus leads to upregulation of expression of pre-mRNAs from hundreds of genes, providing an experimentally favorable system to investigate whether differential kinetics of pre-mRNA splicing may control the timing of gene expression following an inducing stimulus.

Pre-mRNA conversion to mRNA has been implicated in regulation of gene expression in diverse systems. As part of the cellular response to various environmental stressors, mRNAs for ribosomal proteins were shown to be downregulated due to decreased splicing efficiency in yeast (12). Global changes in efficiency of pre-mRNA splicing have been shown to be a developmental prerequisite for *Drosophila* early embryonic development (13). The developing vertebrate embryo obeys a 'segmentation clock' determining body segment length whose very timing relies on delays attributable to control of the splicing rate of the *Hes7* transcriptional repressor (14).

In certain well-studied cases, as with the cytokine TNFα, regulatory mechanisms modulating RNA levels exert significant physiological effects (15–22). The insight that TNFα contains AU-rich elements in its 3' untranslated region that act as mRNA degradation signals (23), and subsequent observations that a mouse in which these AU-rich elements were removed resulted in a robust autoimmune phenotype (17), was an early indication of the importance of precisely tuned mRNA levels in the regulation of inflammation to avoid autoimmunity. Given the role of pre-mRNA splicing in biogenesis of mature mRNA, we and others (8, 24–27) suggested that regulation of splicing kinetics may influence the gene expression kinetics that define the inflammatory cascade. Consistent with this idea, Braunschweig et al. (28) demonstrated that intron retention is a widespread and regulated feature of mammalian transcriptomes, often associated with reduced mRNA abundance and subject to tissue-specific control.

Deep learning models have rapidly advanced our ability to interpret noncoding regions of the genome by learning complex regulatory logic directly from DNA sequence. Early models such as DeepSEA demonstrated that convolutional neural networks could predict the effects of noncoding variants on transcription factor binding and chromatin state, laying a foundation for sequence-to-function inference (29). Building on this, SpliceAI adapted deep convolutional architectures to model splicing dynamics directly from primary sequence, achieving high accuracy in predicting splice site usage and the effects of sequence variants on splicing (30). More recent models such as Enformer and Borzoi have expanded this paradigm further by predicting RNA-based outputs, including gene expression levels, transcript isoform usage, and transcription start site precision (31, 32). These models also support feature attribution though interpretable machine learning approaches such as saliency mapping and deepSHAP, allowing inference of sequence elements—such as transcription factor binding sites and splicing regulators—that most influence RNA output. Together, these approaches provide an unprecedented modeling context—spanning hundreds of kilobases—that enables the systematic dissection of sequence-level determinants of intron retention and the broader landscape of RNA processing.

To examine the timing of intron removal from 230 different transcripts induced by TNFα in macrophages, we have developed a method for highly enriching transcript populations for mRNAs of interest, which is followed by deep sequencing of the largely pre-mRNA populations we purify. The induction of transcripts and removal of introns can be quantified with precision, with lifetimes of introns determined by blocking transcription early after induction. Among genes whose mRNAs appear more slowly after induction, we identify ones containing introns with poor binding sites for splicing factor U1, usually finding one per transcript. We call these "bottleneck" introns. Among the most rapidly induced genes we find no such introns. We show that in these pre-mRNAs, the sequence of the U1 binding site is critical for the speed of intron removal by building mini-genes with "repaired" introns and showing that these splice at the canonical rate (identified as about 20 seconds after polymerase has passed that point). To complement these findings and explore regulatory determinants at larger scale, we apply a deep learning model trained on genomic sequence to predict splicing dynamics. Feature attribution methods identified sequence features associated with the splicing delays. We observe and model fine-tuning

improved detection of patterns associated with intron retention. We propose that bottleneck introns are important for determining either the rate of degradation of pre-mRNAs or the rate of appearance of mature mRNA or both.

Results

Hybrid Capture of Chromatin-Associated Transcripts

It has been established that pre-mRNA is highly enriched in the chromatin-associated, polyadenylated RNA fraction (33). To examine splicing events in pre-mRNAs, we performed time-course experiments with TNF-stimulated bone marrow-derived macrophages (BMDMs), isolating RNA after biochemical separation of chromatin-associated material (24). Using a hybrid-capture approach, we targeted sequencing toward transcripts of 230 genes previously identified as TNF-induced inflammatory mRNAs (34). Purification of cDNA corresponding to these inflammatory transcripts involved reverse transcription of chromatin-associated RNA using oligo(dT), and capture of desired cDNAs using biotinylated probes complementary to the last exon of each gene of interest. Oligo(dT) priming and the choice of the last exon as a capture target enabled us to sequence complete transcripts from the standpoint of the splicing machinery, because all introns will have been transcribed in such transcripts. The hybrid capture strategy, based on a published approach (35), involved: (1) microarray printing of 12,000 150-bp ssDNAs designed from tiled fragments of the last exon of each gene of interest; (2) conversion to a pool of biotinylated ssRNA probes using PCR followed by in vitro transcription with biotinylated ribonucleotides; (3) hybridization of ssRNA pools to cDNA from each biological experiment, and (4) and streptavidin-coated bead-mediated capture of the transcripts of interest (Fig. S1). This approach resulted in a ~30-fold enrichment of genes of interest, with 70% of the sequenced reads corresponding to the genes of interest; RNA submitted only to poly(A) selection contained only 2% of such reads (Fig. S2A). In this way we could analyze 1098 introns (after TPM filter >100) from genes induced by an inflammatory stimulus (Fig S2B).

The selected chromatin-associated transcripts collected across time points after induction are shown as read-density histograms, displaying sequencing read abundance along each gene to reveal exon–intron structure. In Fig. 1A, Nfκbia (encoding IκB-α) is shown over 1 hour after TNF induction, with $\log_{10}$-scaled densities normalized within each time point to highlight transcriptional kinetics. New transcription becomes evident by 6 minutes, especially on an unnormalized linear scale (Fig. S3).The corresponding log-scaled histograms permit visualization of the intronic signal as a function of time after induction. Furthermore, from Fig. 1A it is evident that at all time points following induction, the 5'-proximal introns have been totally removed from the sequenced transcripts, indicating that the selection against partial transcripts is quite complete. Whereas excision of the first intron is always observed, the middle three introns are seen at intermediate states of excision in all time points such that intron definition for these introns is readily observable from read density histograms. We attribute this first exon excision largely to co-transcriptional splicing, consistent with other genome-wide splicing studies (36). Strikingly, the final intron deviates significantly in its kinetic trajectory, as its read density does not obey a similar relative reduction. This might be due to a lag in terminal intron splicing (37) or a feature of splicing that accompanies transcript release from chromatin.

Quantifying Splice Completion Across the Transcriptome

To better quantify the observed dynamics, we adapted the Coefficient of Splicing (CoSI) (Figure 1B), which quantifies the extent of splicing as a ratio of spliced to total (spliced and unspliced) junction reads such that CoSI values of ~1 and ~0 imply near-complete splicing and virtually unspliced states, respectively (33). Though we observed a decrease in read density as a function of distance from the 3' of the gene (Fig. S3), presumably as a consequence of premature termination of the reverse transcriptase during copying of the pre-mRNA, the use of CoSI allows for an intron-specific splicing score regardless of read densities at neighboring introns. Using the CoSI metric, a time course plotting of the extent of splicing showed a dip in CoSI at ~6 minutes (Fig. 1C) corresponding to the aforementioned accumulation of new, unspliced transcripts. The splicing

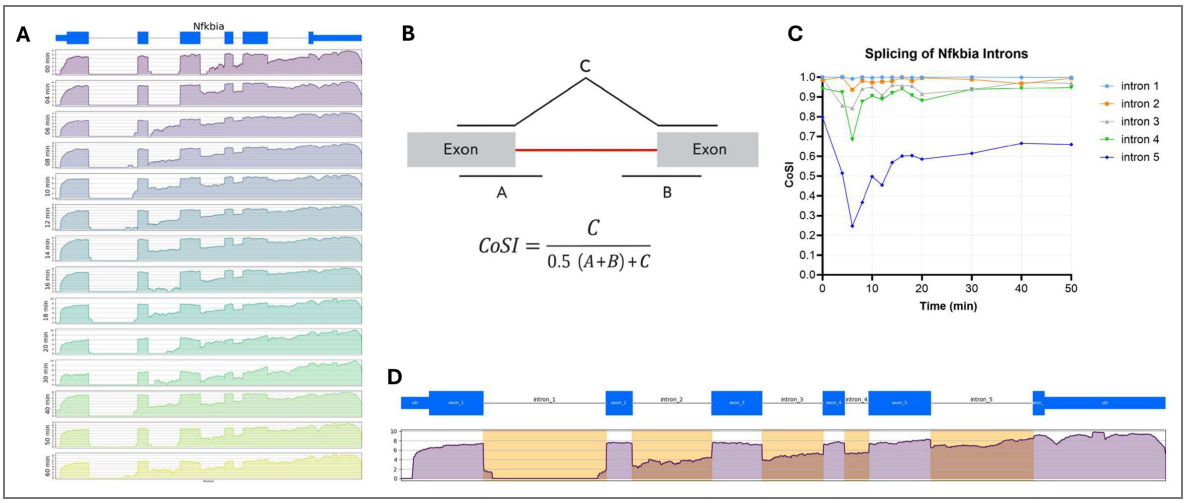


Fig. 1. Sequencing of complete, chromatin-associated pre-mRNA during inflammatory stimulus reveals differential splicing dynamics among introns of IKBα.

(A) Histogram of reads corresponding to the TNF-induced expression and splicing of IKBα pre-mRNA of BMDMs. RNA-Seq was performed on chromatin associated RNA, enriched for NFkB genes as a function of a TNF stimulation timecourse, time shown in minutes after stimulation. Reads are histogrammed in $\log_{10}$ scale and normalized to each time point's maximum value. (B) The Coefficient of Splicing (CoSI) metric quantifies extent of splicing as a function of time, expressed as a ratio of reads from each splice junction to total junctional reads. Dynamics of IKBα splicing as a function of each intron's CoSI is shown (C), where 1=spliced and 0=unspliced, with corresponding introns highlighted in sample timepoint. (D) Differential dynamics of splicing for each *Nfikbia* intron are further demonstrated in the coverage plot for the transcript.

dynamics of each *Nfkbia* intron can be inferred from the CoSI dynamics, and the notable difference in splicing between the 5' and 3' introns (Fig 1D) is demonstrated by the amplitude of the 'dip' at 6 minutes and the time required for each intron to return to Co-SI of 1.

A surprising heterogeneity in CoSI was observed among all inflammatory introns (Fig 2A), implying diversity in their propensity to be spliced (Fig. 2). When considering all 1,024 introns in the chromatin-associated TNF time course we find most introns very near to CoSI ~1 relatively soon after induction, indicating that most introns do not remain unspliced long after induction begins between 4-8minutes (Fig. 2A). The observation that many TNF-responsive transcripts initiate within 4–8 minutes is consistent with that of other related studies (24) Unexpectedly, although the median CoSI value remains high, we identified considerable heterogeneity among introns, some at and remaining near CoSI values <0.5, indicating relatively poor splicing, found very late into the time course.

As an example of a poorly spliced intron, chemokine *Cxcl10* intron 2 (Fig. 2A) is notable as it remains poorly spliced despite clear excision of neighboring introns, remaining quite unspliced even ~30 minutes after induction. It is possible that this intron undergoes splicing after its nascent chromatin-associated state, as is likely the case for the 3'-terminal intron of *Nfκbia*. It is also possible that this intron targets *Cxcl10* transcripts for degradation and the relatively fixed nature of intron 2's splicing status throughout the time course is a function of a constant rate of degradation. Introns with low CoSI at late time points post-TNF induction were considered putative 'bottleneck introns'–borrowing from the language that accompanied the discovery and characterization of slowly splicing U12-type introns (38). These introns were so slow to splice that they may intrinsically delay gene expression. Notably, the distribution of CoSI values of the entire dataset (Fig 2A) was very broad, and though most introns spliced immediately, many introns showed evidence of splicing bottlenecks, noticeable by their significant deviation from mean CoSI. At 10 and 60 minutes post-induction, 14% and 11% of introns, respectively, had CoSI values below one standard deviation of the mean (0.86±0.25 and 0.91±0.19, respectively). Shown as examples are *Cd40*, *Cxcl10*, *Daxx*, and *Irf7* (Fig. 2A-B), genes whose immunological and inflammatory importance is well-established in studies with knockout mice (39–41).

Measurements of Intron Splicing Half-Lives

To quantify splicing kinetics, we used Actinomycin D (actD) to freeze transcription and followed the loss of intron and accumulation of splice junctions. In these experiments, splicing was analyzed at many time points immediately following actD treatment on the same 230 transcripts of interest and selected using hybrid capture from the total pool of cellular RNA rather than chromatin-associated RNA. Fitting the accumulation of spliced transcripts (as measured by CoSI) with an exponential distribution, we were able to extrapolate intron excision half-lives (Fig. 3). Because total cellular RNA was used, observed rates were independent of chromatin localization. We find intron half-lives that range from 20-40s (56% of introns splicing in this timing window) to several minutes, reflecting the considerable heterogeneity that is observed from CoSI differences.

A Multifaceted Approach for Identification of Determinants of Slow Splicing

Building on our observation that certain introns exhibit markedly delayed splicing by both CoSI dynamics and direct half-life measurements, we sought to uncover the underlying sequence features responsible for these delays. To this end, we employed a multifaceted strategy: (1) computational analysis of 5' splice donor strength using MaxEntScan, (2) experimental testing of splicing efficiency with a minigene reporter assay, and (3) interpretable deep learning to identify additional cis-regulatory elements—beyond canonical splice sites—that may modulate splicing kinetics.

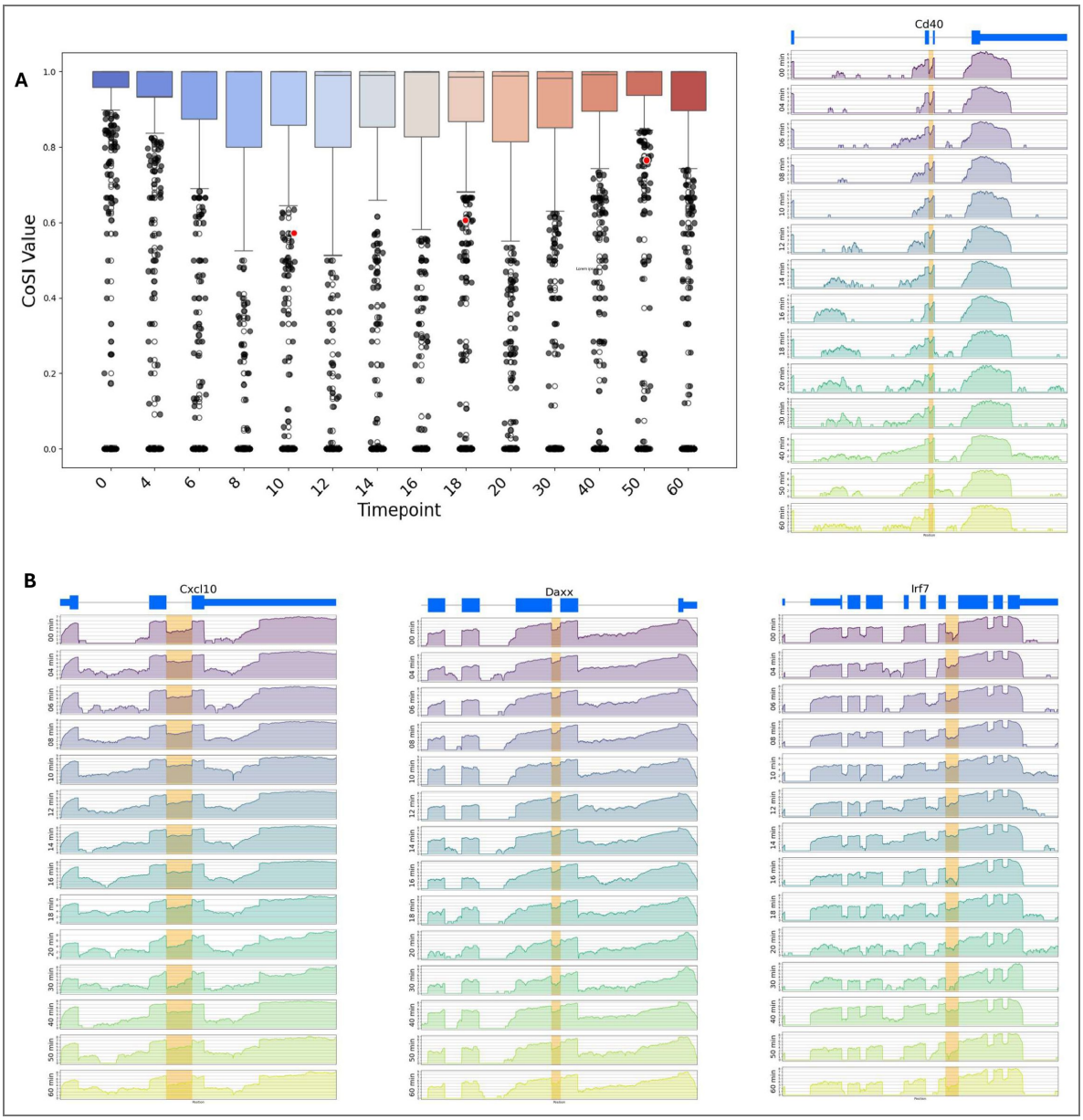


Fig. 2. Heterogeneity of splicing at each intron reveals splicing 'bottlenecks'.

The Co-SI of each intron per timepoint is shown as a function of the entire inflammatory mRNA dataset as box-whisker plot (A). Each point represents an intron of one of 230 genes, revealing high rates of splicing (median CoSI indicated by bar near 1.0 for each timepoint) for most genes with significant outliers. As an example, Cxcl10 intron 2 (red arrowhead) is represented by the datapoint with arrowhead, and a histogram of reads is shown to demonstrate relative unspliced nature of this intron, which is not involved in alternative splicing. (B) Several similar introns that are relatively unspliced are found throughout the inflammatory transcriptome; shown are bottleneck introns within Cd40, Daxx, and Irf7 as examples in the context of their neighboring introns.

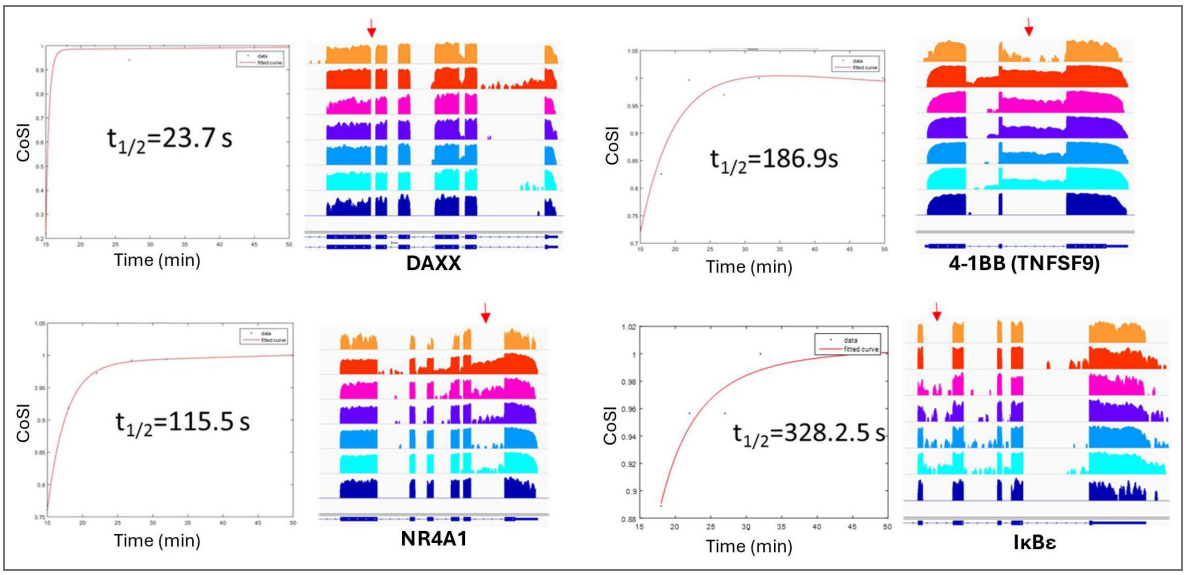


Fig. 3. Splicing kinetics of inflammatory introns are heterogeneous, ranging from seconds to minutes.

CoSI of introns representing various splicing rates are measured and fit to half-lives. Cells were treated with Actinomycin D-treated, from which hybrid capture of genes of interest and sequencing was performed on total (unfractionated) RNA. Shown are four representative samples of splicing kinetics.

Computational Analysis of Splice Donor Strength

A delay in splicing at certain sites could simply confer a delay in gene expression (by ~5 minutes in the slow case shown in Figure 3, IκBε), or, as is seen in yeast studies, it could result in both gene expression delay and gene expression diminution due to degradation of slowly splicing pre-mRNA (42). Prior studies placed IκBε in a delayed splicing category (8), suggesting a pronounced splicing delay relative to rapidly induced genes. To understand a potential mechanistic basis for these differences in splicing time, each intron within our dataset was assessed computationally for the concurrence of its 5' splice donor sequence to a consensus sequence (43). The 5' splice donor is a highly conserved sequence that directly base pairs with splicing factor U1 (44); deviation from consensus sequence confers a significantly reduced ability to engage the splicing machinery.

A maximum entropy model (45) was used to calculate an intron quality score measuring extent of deviation from consensus splice sequence (Fig. S4). Among the inflammatory transcripts studied, many examples of introns with poor 5' donor scores were identified such as *Irf7* and *Il12b*, where lower scores indicate significant deviation from consensus. We suggest that having non-consensus splice sites may be a regulatory mechanism affecting gene expression. We considered that splicing might show profound differences in the previously defined categories of induction (immediate/early/intermediate/late) characteristic of the inflammatory gene expression kinetics (24). We found that the 'immediate' genes showed consistently fast splicing (highest CoSI values) at all of their introns but the most 3', but that the other three groups shared similar CoSI distributions (Fig. S6). Using the bioinformatics "intron quality score" we also found that the introns of the later gene classes have significantly lower scoring 5' splice donor sequences (Fig. S4). Therefore, from experimental measurement of splicing and sequence-based prediction, genes expressed immediately following inflammatory stimulus are spliced faster, whereas all other inflammatory genes have a complex and heterogeneous distribution of splicing efficiency that does not stratify cleanly into the later kinetic categories (early/intermediate/late). Slowly splicing introns are found throughout these later kinetic categories in similar abundance, perhaps playing very gene-specific roles in diverse kinetic categories.

Experimental Validation with Minigene Fluorescent Reporter Assay

To test whether delays in splicing result in changes to gene expression, we identified a set of introns with the following criteria: (1) introns that splice poorly as defined by RNA-Seq, (2) introns that contain a low-scoring (non-consensus) 5' splice donor, and (3) introns whose weak 5' splice donor is evolutionarily conserved across many mammalian species. These introns were tested in the context of a splicing reporter expressed on a bidirectional promoter (46). For each intron of interest, the reporter construct consists of a single transcript containing: (1) the 5' neighboring exon from the gene of interest, the intron of interest, and the 3' neighboring exon; (2) a 2A 'self-cleaving' peptide; and (3) the GFP gene. In the opposite orientation but from the same promoter, a blue fluorescent protein (BFP) mRNA is made in equal amounts to the intron-GFP construct,GFP fluorescence of cells transfected with this reporter is a readout of splicing efficiency of the intron-GFP construct, when normalized to BFP fluorescence levels. Transfected into HEK293T cells and expressed for 24 hours, this bidirectional reporter enabled us to understand, at steady state, whether gene expression is affected by slow splicing introns. Measured using flow cytometry, the slope of the line corresponding to BFP:GFP ratios provides a relative metric of splicing efficiency, where slopes ~.9 and ~0.1 imply efficient and inefficient splicing, respectively.

To test the effect of a poor splice donor, we 'rescued' some poorly splicing introns in the context of the reporter by mutation to consensus splice donor 'GTAAG.' For instance, for *IL12* intron 3, the splice donor sequence of "GTAAT" that is conserved among many mammalian species, was altered to "GTAAG" (Fig. S5). Expression levels of the reporter construct with the wild-type *IL12* intron were found to be about half (57%) of the levels of the same construct with a single base pair alteration to make the stronger splice donor. *IRF7* intron 5 was tested against a 'fixed' intron as well as a wild-type intron from an actin gene, both resulting in two-fold improvements of gene expression.

In one case, expression of *TFEC* (transcription factor EC) was not altered by splice site repair, suggesting that other mechanisms beyond 5'-splice site deficiencies may be involved in mediating slow splicing. Generally, when the BFP:GFP slopes of wild-type and mutated introns were compared by taking a ratio of their slopes, a change of ~2 nucleotides dramatically altered the slope of the line (Fig. 4).

Deep Learning–Based Discovery of Non-Canonical Regulatory Motifs

To extend our understanding of regulatory features beyond canonical splice donor sequences, we applied a supervised regulatory sequence model (Borzoi) to investigate sequence determinants associated with slow splicing. Borzoi is a model originally trained on steady-state RNA profiles across diverse cell types and tissues, which may limit its ability to capture stimulus-specific regulatory dynamics at play during TNFα-induced intron retention. Fine-tuning of this model offers a way to adapt a large, pre-trained sequence model to a specialized context by continuing training on domain-specific datasets, leveraging the general regulatory knowledge already encoded in the model weights while learning new patterns relevant to the task at hand. Fine-tuning has been shown to enhance Borzoi's performance in other specialized contexts, including tissue-specific expression, transcription factor knockdowns, and cellular aging (47). We fine-tuned it using RNA-seq coverage data from BMDMs at 18, 20, and 30 minutes post-TNFα induction—timepoints that capture dynamic intron retention. Fine-tuning was performed over 20 epochs using the Adam optimizer (learning rate = 1e-6, MSE loss). To evaluate generalization, chromosomes 10 and 11 were held out for training and used for validation and testing, respectively. Model loss steadily decreased over the training period for both datasets (Fig. 5A), and Pearson correlation (reflecting the strength of their linear relationship) between predicted and observed expression improved from $r = 0.51$ ($p = 4.44 \times 10^{-16}$) for the pretrained model to $r = 0.61$ ($p = 4.84 \times 10^{-24}$) after fine-tuning (Fig. 5B).

To interpret the model's predictions and identify candidate regulatory elements, we applied a widely adopted backpropagation based technique, DeepLIFT (48), to compute nucleotide-resolution importance scores, capturing the contribution of each base to splicing of the retained intron. Attribution scores were computed for each intron along with sequence within the model input window (524kb). In these scores, larger positive values indicate nucleotides that the model predicts as contributing to intron retention (i.e., delaying splicing), whereas negative values indicate nucleotides predicted to be associated with splice completion. We analyzed these scores using modisco-lite, which clusters recurring high-importance patterns into position weight matrices (PWMs) representing putative regulatory motifs (49). To assess motif enrichment among delayed introns, we focused on the 50 slowest-splicing introns as ranked by area under the CoSI-time curve (Fig. S7C). Motif occurrences were identified using a conventional motif scanning approach, Finding Individual Motif Occurrences (FIMO) (50), and enrichment was calculated relative to either the remaining introns in our dataset or all introns in the transcriptome.

We compared the frequency of putative regulatory motifs in the 50 slowest-splicing introns to their frequency across all introns in the murine transcriptome, highlighting both the most enriched motifs overall and specific examples of their genomic locations within delayed introns. Among these, a GA-repeat–like motif showed enrichment in the slow-splicing set (FIMO, 80.0% vs 55.2%) (Fig. 6A). Representative instances of this motif were also recovered from the original source seqlets used to generate the PWM by modisco-lite, providing further support for its regulatory relevance (Fig. 6B-D). These findings illustrate how sequence-to-expression modeling, combined with attribution-based motif discovery, can uncover non-canonical cis-regulatory elements that may influence splicing efficiency during inflammatory gene activation.

Discussion

In this study we sought to understand splicing kinetics of the large number of genes that comprise the inflammatory response and to assess whether splicing itself might play a regulatory role in inflammation. We developed a targeted sequencing strategy, purifying transcripts containing each

Fig. 4. Bottleneck introns can be repaired, and account for significant alterations to gene expression.

(A) Intron-GFP splicing reporters for each wild-type intron (red) and modified intron (green) are shown as BFP:GFP ratio. (B) Ratio of WT:Fixed slopes is shown; whereas Tfec expression is not altered by improved 5′ sequence, Malt1 intron sequence is significantly impaired owing to its 5′ donor sequence, exhibiting a roughly 5-fold impairment in gene expression due to the 5′ splice donor.

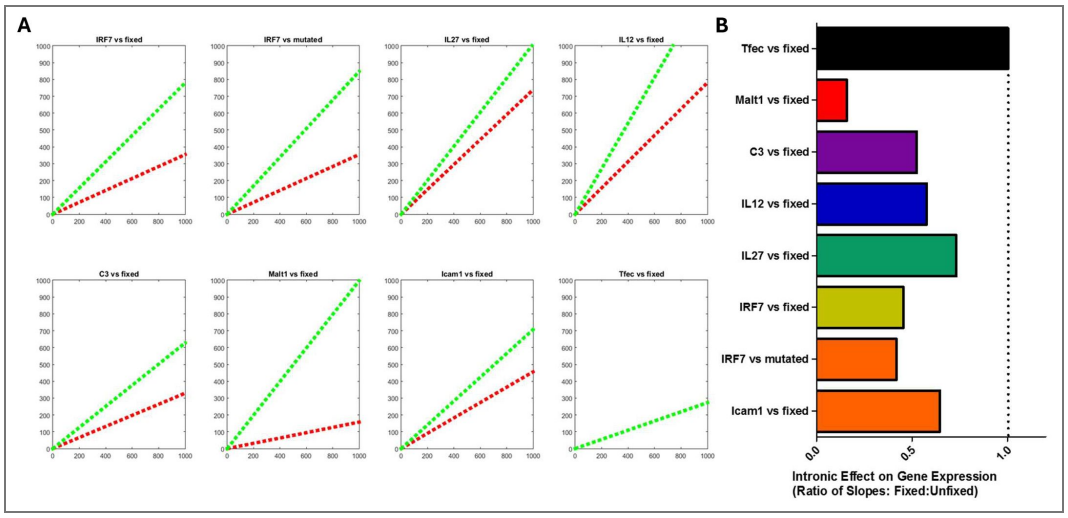
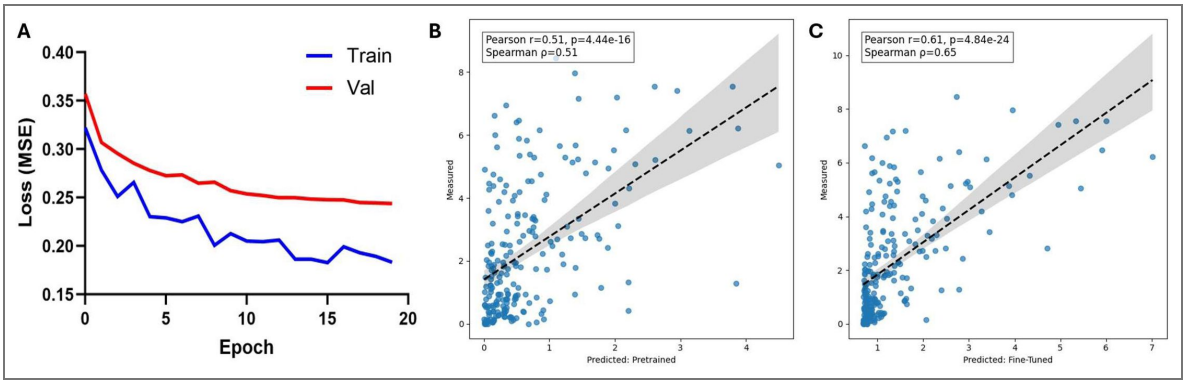


Fig. 5. Fine-tuning improves gene expression prediction in macrophages.

(A) Training and test loss curves over 20 epochs of model fine-tuning. (B-C) Predicted versus measured RNA expression values for 230 transcripts induced by TNFα in macrophages, shown for the pretrained model (B) and the fine-tuned model (C).



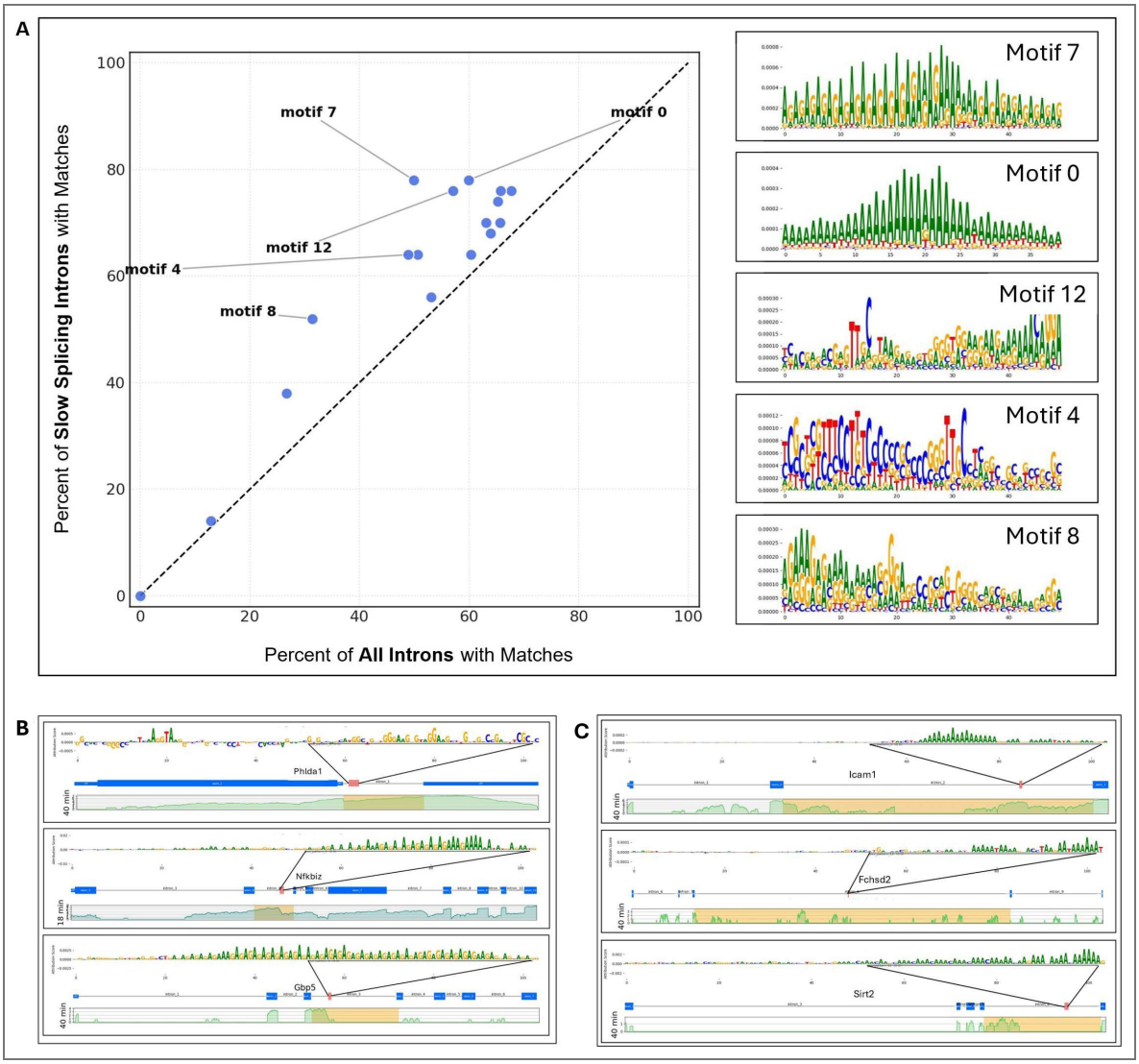


Fig. 6. Putative regulatory sequences are enriched in slow-splicing introns.

(A) Scatterplot showing percent representation of position weight matrices (PWMs) scanned across slow-splicing introns versus all introns genome-wide using FIMO. Sequence logos of the top 5 enriched PWMs are shown to the right. (B–D) Attribution plots highlighting GA-rich sequences (source seqlets for enriched PWMs) and their locations mapped to gene schematics for GA-rich motif (B), and A-rich motif (C). Corresponding RNA-seq tracks are shown for each gene.

gene's terminal exon. This approach allowed us to sequence the 1,024 introns within inflammatory genes and permitted direct assessment of the structures of nearly-completed transcripts. We found considerable heterogeneity in splicing efficiency among these introns. In studying evolutionarily conserved weak 5' splice donors, we have isolated one cause of slow appearance of mRNA following a pulse of stimulus; many other slowly spliced introns without such sequences were also identified in this study and suggest other regulatory mechanisms may be responsible.

Crucially, the hybrid capture approach averts a common ambiguity in analyzing splicing kinetics of not being able to differentiate completed pre-mRNA from nascent transcripts during an induction pulse – this often leads to an overestimation of the unspliced status of early introns and complicates quantification of splicing kinetics. To the contrary, we rarely found genes containing unspliced first introns. This was true of chromatin-associated RNA and of whole cell RNA after inhibition of transcription with actD. These effects are consistent with the emerging model of co-transcriptional pre-mRNA splicing, where the splicing machinery has been suggested to lag 3-5kb behind the polymerase. Indeed, several recent global studies of RNA splicing bolster the claim that much pre-mRNA is spliced co-transcriptionally: 74% (yeast) or 75-84% (human) of introns are found to be at least 50% spliced by the time of transcription termination in several other studies (33, 37, 51–54). Surprisingly, this ~80% figure remains constant whether total RNA or chromatin-associated RNA is measured, implying that our choice to analyze chromatin-associated RNA does not significantly overrepresent splicing intermediates.

We found that most introns are spliced very efficiently, appearing and disappearing as a rapid dip of CoSI immediately following induction, returning to a CoSI of ~0.95 within minutes after induction. Notably, the distribution of CoSI values of the entire dataset (Figure 2) was very broad. Though most introns spliced immediately, there were several 'bottleneck introns'. In order to determine more specifically the rates of slowly splicing introns, studies employing actD to stall transcription and examine intron splicing half-lives corroborated the idea that there is tremendous intron-to-intron heterogeneity. Most delayed introns ultimately reached higher CoSI values over the time course, consistent with completion of splicing rather than stable, long-term retention (see Figs. 2–3). Whereas most introns spliced within 20-40s, some were delayed significantly (upwards of 5 minutes). Of note, however, is that our 20-40s rate of splicing is somewhat at odds with other figures in the literature of 8-10 minutes for intron excision (9) after a washout of the drug D-ribofuranosylbenzimidazole. There is some debate as to the perturbative role of actD in splicing, with one report observing that splicing intermediates in the context of the MS2 reporter system are prematurely liberated from chromatin upon actD treatment. Even in this case, the rapid actD-based rates, are likely underestimating even faster kinetics if one considers that the co-transcriptional splicing machinery targets chromatin mRNA faster than released mRNA (55). However, even in the absence of actD, stimulation revealed that most of IKBα's introns are spliced in less than two minutes when one takes into account the 'dip' in CoSI due to induction and the time to reach steady CoSI levels (Figure 2). While the terminal intron of IKBα appears to have a longer half-life, this unique feature of terminal introns is consistent with prior studies (37).

In testing gene expression differences in bottleneck introns among introns with poor splice sites that are also evolutionarily conserved, we found that steady state levels of reporter proteins were upregulated when the 5' splice donor sequence was mutated to the consensus sequence 'GTAAG' in all cases but one. Attenuated U1 binding provides a mechanistic insight for bottleneck introns that were chosen for their weak 5' splice donors. This implies that at the level of splicing, either due to delays in expression or perhaps degradation due to delayed expression, significant differences in gene expression arise from small differences in nucleotide sequence. These reporter assays measure steady-state expression influenced by intron sequence; while consistent with changes in splicing efficiency, other post-transcriptional mechanisms may be at work to preclude these introns from efficient splicing. We also find many slowly spliced introns not explained by weak 5' splice donors. In some cases, we find multiple bottlenecks introns per gene, as is the case of IRF7, where one bottleneck (intron 5) was attributable to an evolutionarily conserved weak 5' splice donor while another (intron 7) was observed experimentally but of unknown cause. This may be due to any of a number of potential mechanisms that may also serve in tuning the speed of

splicing: cis-regulatory protein recruitment, 3' splice acceptor sequence or other sequence elements, or alterations of RNA polymerase speed or chromatin marks or three-dimensional gene structure.

Central to our inquiry is the enigmatic nature of these bottlenecks remaining in physiologically critical genes, often evolutionarily conserved, and yet intrinsically mediating an inefficiency in gene expression. Importantly, the conservation of these weak donor sites suggests they confer a regulatory advantage—such as fine-tuning of expression timing or transcript abundance—rather than representing neutral or deleterious features merely tolerated by selection. We posit that the gene expression changes that are shown in bone marrow-derived macrophages offer a regulatory strategy to slow up and maybe restrict expression of genes in a manner dependent on the composition of mRNA processing factors in the cell ('the splicing landscape'), the cell type, or the stimulus type in question. Recent studies have demonstrated global changes in intron retention preferences in B cell lymphomagenesis and granulocyte differentiation (56). In a similar manner, we suggest that selection of splicing and kinetics of splicing might allow a previously unappreciated level of specificity to gene expression decisions in cells presented with an inflammatory stimulus (57–63).

Induction with TNF is a particularly favorable situation because many of the genes we examined were up-regulated in their transcription within 4-6 minutes of adding inducer (Supplemental Fig. 2B), allowing examination of large numbers of pre-mRNA transcripts. This, in concert with the hybrid capture approach that provides a large number of junctional sequencing reads, has permitted unique insight into the kinetics of splicing of mature transcripts and revealed surprising heterogeneity. We suggest that this methodology and analysis could have wider applicability for other gene induction situations.

Despite the promise of deep learning models, challenges remain in translating their predictions into biological insight. While performance on benchmark tasks continues to improve—often surpassing traditional motif-finding tools—generalization across conditions and cell types is still constrained by the scope of training data. Interpretability is another major hurdle, particularly for genomic sequences. Interpretable machine learning methods can highlight sequence elements that influence model predictions, but connecting these regions to underlying biological mechanisms often demands extensive experimental validation. In the case of RNA splicing, complexity of overlapping regulatory layers—such as RNA secondary structure, co-transcriptional dynamics, and nuclear export—can make it difficult to disentangle causal sequence features from correlated signals. Nonetheless, sequence-to-expression modeling provides a powerful framework for identifying sequence features correlated with delayed splicing and for guiding experimental discovery of novel regulators through fine-tuning and perturbation-informed learning.

Our deep learning-based analysis further supports the notion that splicing efficiency is shaped by a broad array of cis-regulatory features, many of which fall outside canonical splice site motifs. By leveraging, a sequence-to-expression transformer trained on multimodal regulatory data, we identified non-canonical sequence motifs enriched in the slowest-splicing introns—motifs that may act as silencers or delay elements involved in splicing. These results offer a complementary, agnostic perspective to our mutational analyses, and suggest that intronic bottlenecks can arise from diverse sequence architectures. The enrichment of these motifs in bottleneck introns—independent of donor site strength—points to additional layers of splicing control that may be particularly relevant in rapid-response transcriptional programs like inflammation. We propose that such regulatory elements may encode a form of temporal tuning, modulating transcript availability through fine control of intron excision. More broadly, our work demonstrates interpretable machine learning can uncover latent regulatory features that elude traditional sequence analysis, advancing efforts to decode the logic of splicing regulation in a cell-type or stimulus-specific context.

Experimental procedures

Cells

C56BL6/J mice were sacrificed via CO_2 euthanasia and sterilized with 70% ethanol. Femur and tibia bones harvested and stripped of muscle tissue. Bone marrow cells were resuspended in 20mL of fresh DMEM. 2.5e6 bone-marrow cells plated in a 15-cm dish in 20mL of BMDM Media (DMEM, 20% FBS, 30% L929 condition media, and 1% Pen/Strep) and grown at 5% CO_2 and 37°C. BMDM media completely replaced on day 3 as well as a supplemental addition of 5mL L929 condition media on day 5.

RNA fractionation

RNA was fractionated into cytoplasmic, nucleoplasmic, and chromatin-associated pools as previously described (24) with modifications. Confluent 15 cm dishes of mature BMDMs were scraped into 400 µL cold NP-40 lysis buffer (10 mM Tris-HCl pH 7.5, 0.08% NP-40, 150 mM NaCl) and layered onto a 1 mL sucrose cushion (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 24% w/v sucrose). Samples were centrifuged at 13,000 rpm for 10 min at 4 °C. The supernatant (cytoplasmic fraction) was mixed with 3 volumes of 100% ethanol and 2 volumes of buffer RLT (4 M GuSCN, 0.1 M β-mercaptoethanol, 0.5% N-lauroylsarcosine, 25 mM Na-citrate, pH 7.2) and stored at -80 °C.

The nuclear pellet was resuspended in 200 µL cold glycerol buffer (20 mM Tris-HCl pH 7.5, 75 mM NaCl, 0.5 mM EDTA pH 8.0, 50% glycerol, 0.85 mM DTT) and lysed with an equal volume of nuclear lysis buffer (20 mM HEPES pH 7.5, 7.5 mM $MgCl_2$, 0.2 mM EDTA pH 8.0, 1 M urea, 1% NP-40). After vortexing and centrifugation (14,000 rpm, 5 min, 4 °C), the supernatant (nucleoplasmic fraction) was processed as above.

The remaining chromatin pellet was hydrated in 1× PBS and dissolved in 500 µL TRIzol reagent at 50 °C with intermittent vortexing. After phase separation with 100 µL chloroform, the aqueous phase was recovered and processed as above. All RNA fractions were purified using the Qiagen RNeasy protocol and eluted in nuclease-free water. Typical yields were 300–500 ng/µL for cytoplasmic RNA, 100–250 ng/µL for nucleoplasmic RNA, and 300–500 ng/µL for chromatin-associated RNA. RNA was DNase-treated (TURBO DNase, Thermo Fisher) and stored at -80 °C.

Template-switch reverse transcription

One microgram of RNA was mixed with 1 µM oligo$(dT)_{30}$ (5'-AAGCAGTGGTATCAACGCAGAGTACT$_{30}$-3'), heated to 80 °C for 2.5 min, and snap-cooled on ice. A 10 µL reverse-transcription mix containing 10 µM template-switch oligo (5'-AAGCAGTGGTATCAACGCAGAGTACACArGrGrG-3'), 20 mM DTT, 2× First-Strand Buffer (Invitrogen), 1 mM dNTPs, 40 U Murine RNase Inhibitor (NEB), and 200 U SuperScript II (Invitrogen) was added. Reactions were incubated sequentially at 42 °C (30 min), 45 °C (30 min), 50 °C (10 min), then heat-inactivated at 80 °C (10 min).

RNA templates were degraded by adding NaOH (final 0.1 M) and EDTA (5 mM) and heating to 70 °C for 10 min, followed by neutralization with HCl. cDNA was purified using 2× Sera-Mag carboxylate-modified magnetic beads (GE Healthcare) with standard PEG/ethanol washes and eluted in nuclease-free water.

Hybrid capture probe design and synthesis

Biotinylated RNA probes were designed against the terminal exons of inflammatory genes (see Supplemental Table 2 ↗). For each gene, 100-bp overlapping oligonucleotides were synthesized (CustomArray Inc.) and pooled into nine expression-matched subgroups. Each subpool was PCR-amplified to append a T7 RNA polymerase promoter and transcribed in vitro using the AmpliScribe T7 Biotin Flash Kit (Epicentre). Purified RNA probe subpools were combined in weighted ratios (A–I) to normalize capture representation across expression levels as below:

Pool	Pool composition
A	0.01113
B	0.0211
C	0.03235
D	0.05514
E	0.10542
F	0.21908
G	0.4674
H	1
I	1

cDNA hybrid capture and elution

Biotinylated probes were hybridized to cDNA at 74 °C for 4.5 min, followed by addition of 2× hybridization buffer (1 M LiCl, 40 mM Tris-HCl pH 7.5, 20 mM EDTA pH 8.0, 4 M urea, 0.5% Triton X-100, 1% SDS, 0.2% Na-deoxycholate). Hybridization continued 30 min at 70 °C. Streptavidin BioMag beads (0.3 mg) were washed and incubated 20 min at 70 °C to capture cDNA–probe complexes. Beads were washed twice with 1× HYB, once each with Wash 4 and Wash 5 buffers, and eluted in 35 µL base elution buffer (125 mM NaOH, 10 mM EDTA pH 8.0, 10 mM Tris-HCl pH 7.5) at 74 °C for 5 min. The eluate was neutralized and purified with 1× Sera-Mag beads and eluted in 45µL and stored at -80°C.

Determining the efficiency of cDNA pulldown

Pulldown efficiency was evaluated by qPCR to quantify enrichment of target transcripts and depletion of background RNA. qPCR reactions (KAPA SYBR 2× Master Mix) compared pre- and post-pulldown cDNA at a 2:1 ratio. Primers:

- L32 (background control): F 5'-AAGCGAAACTGGCGGAAAC-3'; R 5'-TAACCGATGTTGGGCATCAG-3'
- NF-κBIA (spliced exon 5–6 junction): F 5'-ACGGAGTCAGAATTCACAGAGG-3'; R 5'-CACAAAGACA ACAGCCGAATC-3'

Cycle-threshold (Ct) values were used to calculate ΔCt between L32 and NF-κBIA. Successful pulldowns typically showed ΔL32 = –7 to –9 cycles and ΔNF-κBIA = –2 to –4, corresponding to ΔL32/ΔNF-κBIA > 2.0.

Post-pulldown cDNA amplification

Pulldown cDNA was amplified prior to tagmentation. PCR reactions contained Q5 High-Fidelity 2× Master Mix (NEB), 1 µM primer (5'-AAGCAGTGGTATCAACGCAGAGTACT-3'), and ~5% of the pulldown reaction. Cycling: 95 °C (2 min) → 20–25 cycles of 95 °C (30 s), 62.5 °C (30 s), 72 °C (150 s) → 72 °C (5 min). PCR products were purified (0.9× Sera-Mag) and eluted in 25 µL H_2O. Concentrations were determined using a Qubit HS dsDNA Assay.

Tagmentation of cDNA libraries with Tn5 transposase

Tn5 transposase was purified as in Picelli et al., Genome Res. (2014) and pre-loaded with hybridized adapter oligos (Tn5MErev 5'-[phos]CTGTCTCTTATACACATCT-3'; Tn5ME-A 5'-TCGTCGGC AGCGTCAGATGTGTATAAGAGACAG-3'; Tn5ME-B 5'- GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-

3'). Tagmentation reactions contained ~40 ng amplified cDNA, 0.2 µL Tn5, 5% PEG8000, 10 mM TAPS (pH 8.5), 5 mM $MgCl_2$, incubated 10 min at 55 °C. SDS (0.02%) was added and incubation continued 10 min to inactivate Tn5. Reactions were 1.4× Sera-Mag purified and eluted in 20 µL H_2O.

Library barcoding and sequencing

Tagmented libraries were PCR-amplified with paired barcode oligos using Q5 High-Fidelity 2× Master Mix (NEB). Cycling: 72 °C (3 min), 98 °C (2 min), then 25 cycles of 98 °C (10 s), 63 °C (30 s), 72 °C (30 s), and 72 °C (5 min). Libraries were double-purified (1.0×/1.4× Sera-Mag), quantified by Qubit HS dsDNA assay, pooled equimolarly, repurified, and analyzed with an Agilent Bioanalyzer 2100. Sequencing was performed on an Illumina HiSeq 2500 (50 bp, single-end mode).

RNA-seq alignment and analysis

Single-end 50 bp reads were aligned to the mm10 genome using STAR. Junctions were retained only if both exonic and intronic segments were ≥ 3 bp. Using pysam, splice and intron junctions were classified as types a, b, or c, and Completion of Splicing Index (CoSI) values were computed using custom Python scripts. All analyses were performed in Python and R. Sequencing data are being deposited to NCBI GEO, and all analysis code will be made available on GitHub.

Deep Learning Analysis of Intron Regulatory Features

Model architecture and initialization

We used Borzoi, a transformer-based sequence-to-function model trained to predict steady-state RNA abundance from genomic DNA sequence. The pretrained model captures general patterns of cis-regulatory architecture by integrating multi-omic training targets (CAGE, ATAC-seq, and ChIP-seq tracks). However, because Borzoi is optimized for baseline expression rather than stimulus-responsive contexts, we adapted it to our macrophage TNFα time course through targeted fine-tuning.

Fine-tuning on TNFα-stimulated macrophage data

We generated genome-wide RNA-seq coverage tracks from bone marrow–derived macrophages (BMDMs) collected at 18, 20, and 30 minutes post-TNFα induction—timepoints representative of dynamic intron retention. These coverage values served as supervised training targets to refine the model's prediction of RNA abundance under inflammatory conditions. Fine-tuning was performed for 20 epochs using the Adam optimizer (learning rate = 1×10^{-6}, mean-squared-error loss). To evaluate generalization, chromosomes 10 and 11 were withheld for validation and testing, respectively. Model loss steadily decreased over training for both sets, and the correlation between predicted and observed expression improved from Pearson r = 0.51 ($p = 4.44 \times 10^{-16}$) for the pretrained model to r = 0.61 ($p = 4.84 \times 10^{-24}$) after fine-tuning (Fig. 5A–B).

Feature attribution and motif discovery

To interpret the fine-tuned model's predictions, we applied DeepLIFT to compute nucleotide-resolution importance scores, quantifying each base's contribution to the predicted RNA signal within a 524 kb genomic window centered on each intron of interest. Positive importance scores indicate sequence positions contributing to higher predicted intronic RNA signal (i.e., delayed splicing), whereas negative scores correspond to features predictive of splice completion.

Importance profiles were analyzed using modisco-lite, which clusters recurrent high-importance patterns across examples into position weight matrices (PWMs) representing putative regulatory motifs. Enrichment of discovered motifs was assessed using FIMO, scanning the 50 (or 150) slowest-splicing introns—ranked by area under the CoSI–time curve—against all annotated murine introns (GENCODE M23). For targeted motif discovery, attributions input to modisco-lite were masked under different configurations to restrict pattern identification to either intronic or exonic regions adjacent to slow-splicing introns.

Motif enrichment and visualization

Relative motif frequencies were plotted as a scatter of percent representation in slow- versus all-intron backgrounds. Sequence logos of the top five enriched motifs were visualized alongside genomic examples where GA-rich and A-rich motifs coincided with regions of high attribution in DeepLIFT profiles (Fig. 6A–D).

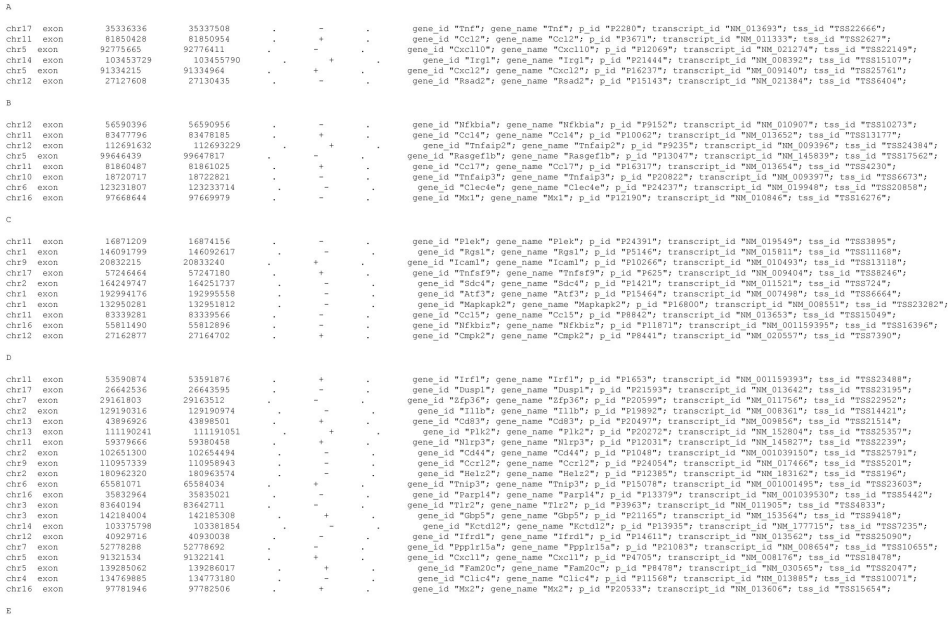
Data availability

Data will be uploaded to GEO after government shutdown

Supplemental materials

Oligo	Sequence	Use
d14-952	5'-AAGCAGTGGTATCAACGCAGAGTACACArGrGrG	Template-Switch Oligo
d14-954	5'-AAGCAGTGGTATCAACGCAGAGTACT(30)	RT dT Oligo
d14-955	5'-AAGCAGTGGTATCAACGCAGAGTAC	Post Pulldown Amplification Oligo
L32-For	5'-AAGCGAAACTGGCGGAAAC	qPCR Primer
L32-Rev	5'-TAACCGATGTTGGGCATCAG	qPCR Primer
NF-KBIA Ex5-6-For	5'-ACGGAGTCAGAATTCACAGAGG	qPCR Primer
NF-KBIA Ex5-6-Rev	5'-CACAAAGACAACAGCCGAATC	qPCR Primer
Tn5MErev	5′-[phos]CTGTCTCTTATACACATCT	Tn5 Adaptor Oligo
Tn5ME-A	5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG	Tn5 Adaptor Oligo

Table 1. **Oligos used**



A

```
chr17  exon    35336336    35337508    .  -  .  gene_id "Tnf"; gene_name "Tnf"; p_id "P2280"; transcript_id "NM_013693"; tss_id "TSS22666";
chr11  exon    81850428    81850954    .  +  .  gene_id "Ccl2"; gene_name "Ccl2"; p_id "P3671"; transcript_id "NM_011333"; tss_id "TSS2627";
chr5   exon    92775665    92776411    .  -  .  gene_id "Cxcl10"; gene_name "Cxcl10"; p_id "P12069"; transcript_id "NM_021274"; tss_id "TSS22149";
chr14  exon    103453729   103455790   .  +  .  gene_id "Irg1"; gene_name "Irg1"; p_id "P21444"; transcript_id "NM_008392"; tss_id "TSS15107";
chr5   exon    91334215    91334964    .  +  .  gene_id "Cxcl2"; gene_name "Cxcl2"; p_id "P16237"; transcript_id "NM_009140"; tss_id "TSS25761";
chr12  exon    27127608    27130435    .  -  .  gene_id "Rsad2"; gene_name "Rsad2"; p_id "P15143"; transcript_id "NM_021384"; tss_id "TSS6404";
```

B

```
chr12  exon    56590396    56590956    .  -  .  gene_id "Nfkbia"; gene_name "Nfkbia"; p_id "P9152"; transcript_id "NM_010907"; tss_id "TSS10273";
chr11  exon    83477796    83478185    .  +  .  gene_id "Ccl4"; gene_name "Ccl4"; p_id "P10062"; transcript_id "NM_013652"; tss_id "TSS13177";
chr12  exon    112691632   112693229   .  +  .  gene_id "Tnfaip2"; gene_name "Tnfaip2"; p_id "P9235"; transcript_id "NM_009396"; tss_id "TSS24384";
chr5   exon    99646439    99647817    .  -  .  gene_id "Rasgef1b"; gene_name "Rasgef1b"; p_id "P13047"; transcript_id "NM_145839"; tss_id "TSS17562";
chr11  exon    81860487    81861025    .  +  .  gene_id "Ccl7"; gene_name "Ccl7"; p_id "P16317"; transcript_id "NM_013654"; tss_id "TSS4230";
chr10  exon    18720717    18722821    .  -  .  gene_id "Tnfaip3"; gene_name "Tnfaip3"; p_id "P20822"; transcript_id "NM_009397"; tss_id "TSS6673";
chr6   exon    123231807   123233714   .  -  .  gene_id "Clec4e"; gene_name "Clec4e"; p_id "P24237"; transcript_id "NM_019948"; tss_id "TSS20858";
chr16  exon    97668644    97669979    .  -  .  gene_id "Mx1"; gene_name "Mx1"; p_id "P12190"; transcript_id "NM_010846"; tss_id "TSS16276";
```

C

```
chr11  exon    16871209    16874156    .  -  .  gene_id "Plek"; gene_name "Plek"; p_id "P24391"; transcript_id "NM_019549"; tss_id "TSS3895";
chr1   exon    146091799   146092617   .  -  .  gene_id "Rgs1"; gene_name "Rgs1"; p_id "P5146"; transcript_id "NM_015811"; tss_id "TSS11168";
chr9   exon    20832215    20833240    .  +  .  gene_id "Icam1"; gene_name "Icam1"; p_id "P10266"; transcript_id "NM_010493"; tss_id "TSS13118";
chr17  exon    57246464    57247180    .  +  .  gene_id "Tnfsf9"; gene_name "Tnfsf9"; p_id "P625"; transcript_id "NM_009404"; tss_id "TSS8246";
chr2   exon    164249747   164251737   .  -  .  gene_id "Sdc4"; gene_name "Sdc4"; p_id "P1421"; transcript_id "NM_011521"; tss_id "TSS724";
chr1   exon    192994176   192995558   .  -  .  gene_id "Atf3"; gene_name "Atf3"; p_id "P15464"; transcript_id "NM_007498"; tss_id "TSS6664";
chr1   exon    132950281   132951812   .  -  .  gene_id "Mapkapk2"; gene_name "Mapkapk2"; p_id "P16800"; transcript_id "NM_008551"; tss_id "TSS23282";
chr11  exon    83339281    83339566    .  -  .  gene_id "Ccl5"; gene_name "Ccl5"; p_id "P8842"; transcript_id "NM_013653"; tss_id "TSS15049";
chr16  exon    55811490    55812896    .  -  .  gene_id "Nfkbiz"; gene_name "Nfkbiz"; p_id "P11871"; transcript_id "NM_001159395"; tss_id "TSS16396";
chr12  exon    27162877    27164702    .  +  .  gene_id "Cmpk2"; gene_name "Cmpk2"; p_id "P8441"; transcript_id "NM_020557"; tss_id "TSS7390";
```

D

```
chr11  exon    53590874    53591876    .  +  .  gene_id "Irf1"; gene_name "Irf1"; p_id "P1653"; transcript_id "NM_001159393"; tss_id "TSS23488";
chr17  exon    26642536    26643595    .  -  .  gene_id "Duspl"; gene_name "Duspl"; p_id "P21593"; transcript_id "NM_013642"; tss_id "TSS23195";
chr7   exon    29161803    29163512    .  -  .  gene_id "Zfp36"; gene_name "Zfp36"; p_id "P20599"; transcript_id "NM_011756"; tss_id "TSS22952";
chr2   exon    129190316   129190974   .  -  .  gene_id "Il1b"; gene_name "Il1b"; p_id "P19892"; transcript_id "NM_008361"; tss_id "TSS14421";
chr13  exon    43896926    43898501    .  +  .  gene_id "Cd83"; gene_name "Cd83"; p_id "P20497"; transcript_id "NM_009856"; tss_id "TSS21514";
chr13  exon    111190241   111191051   .  +  .  gene_id "Plk2"; gene_name "Plk2"; p_id "P20272"; transcript_id "NM_152804"; tss_id "TSS25357";
chr11  exon    59379666    59380458    .  +  .  gene_id "Nlrp3"; gene_name "Nlrp3"; p_id "P12031"; transcript_id "NM_145827"; tss_id "TSS2239";
chr2   exon    102651300   102654494   .  -  .  gene_id "Cd44"; gene_name "Cd44"; p_id "P1048"; transcript_id "NM_001039150"; tss_id "TSS25791";
chr9   exon    110957339   110958943   .  -  .  gene_id "Ccrl2"; gene_name "Ccrl2"; p_id "P24054"; transcript_id "NM_017466"; tss_id "TSS5201";
chr2   exon    180962320   180963574   .  -  .  gene_id "Helz2"; gene_name "Helz2"; p_id "P12385"; transcript_id "NM_183162"; tss_id "TSS196";
chr6   exon    65581071    65584034    .  +  .  gene_id "Tnip3"; gene_name "Tnip3"; p_id "P15078"; transcript_id "NM_001001495"; tss_id "TSS23603";
chr16  exon    35832964    35835021    .  -  .  gene_id "Parp14"; gene_name "Parp14"; p_id "P13379"; transcript_id "NM_001039530"; tss_id "TSS5442";
chr3   exon    83640194    83642711    .  -  .  gene_id "Tlr2"; gene_name "Tlr2"; p_id "P3963"; transcript_id "NM_011905"; tss_id "TSS4833";
chr3   exon    142184004   142185308   .  +  .  gene_id "Gbp5"; gene_name "Gbp5"; p_id "P21165"; transcript_id "NM_153564"; tss_id "TSS9418";
chr14  exon    103375798   103381854   .  -  .  gene_id "Kctd12"; gene_name "Kctd12"; p_id "P13935"; transcript_id "NM_177715"; tss_id "TSS7235";
chr12  exon    40929716    40930038    .  -  .  gene_id "Ifrd1"; gene_name "Ifrd1"; p_id "P14611"; transcript_id "NM_013562"; tss_id "TSS25090";
chr7   exon    52778288    52778692    .  -  .  gene_id "Ppp1r15a"; gene_name "Ppp1r15a"; p_id "P21083"; transcript_id "NM_008654"; tss_id "TSS10655";
chr5   exon    91321534    91322141    .  +  .  gene_id "Cxcl1"; gene_name "Cxcl1"; p_id "P4705"; transcript_id "NM_008176"; tss_id "TSS18478";
chr5   exon    139285062   139286017   .  +  .  gene_id "Fam20c"; gene_name "Fam20c"; p_id "P8478"; transcript_id "NM_030565"; tss_id "TSS2047";
chr4   exon    134769885   134773180   .  -  .  gene_id "Clic4"; gene_name "Clic4"; p_id "P11568"; transcript_id "NM_013885"; tss_id "TSS10071";
chr16  exon    97781946    97782506    .  +  .  gene_id "Mx2"; gene_name "Mx2"; p_id "P20533"; transcript_id "NM_013606"; tss_id "TSS15654";
```

E

Table 2. **Exons of interest for hybrid capture**

```
chr7  exon  128207581  128209031  .  -  .  gene_id "Igsf6"; gene_name "Igsf6"; p_id "P11261"; transcript_id "NM_030691"; tss_id "TSS16286";
chr2  exon  62433850   62436145  .  -  .  gene_id "Ifih1"; gene_name "Ifih1"; p_id "P4870"; transcript_id "NM_001164477"; tss_id "TSS19275";
chr19 exon  6298074    6300096   .  +  .  gene_id "Ehd1"; gene_name "Ehd1"; p_id "P19125"; transcript_id "NM_0101119"; tss_id "TSS24002";
chr19 exon  34721961   34724499  .  +  .  gene_id "Ifit1"; gene_name "Ifit1"; p_id "P22750"; transcript_id "NM_008331"; tss_id "TSS13500";
chr11 exon  117827401  117829632 .  -  .  gene_id "Socs3"; gene_name "Socs3"; p_id "P3923"; transcript_id "NM_007707"; tss_id "TSS25657";
chr11 exon  21047960   21050330  .  +  .  gene_id "Pel11"; gene_name "Pel11"; p_id "P23395"; transcript_id "NM_023324"; tss_id "TSS17987";
chr8  exon  73153293   73154073  .  -  .  gene_id "Gdf15"; gene_name "Gdf15"; p_id "P5185"; transcript_id "NM_0118119"; tss_id "TSS17076";
chr8  exon  87500808   87502647  .  -  .  gene_id "Junb"; gene_name "Junb"; p_id "P14294"; transcript_id "NM_008416"; tss_id "TSS14734";
chr10 exon  77540638   77542270  .  +  .  gene_id "Icosl"; gene_name "Icosl"; p_id "P12741"; transcript_id "NM_015790"; tss_id "TSS25520";
chr2  exon  127983967  127988283 .  +  .  gene_id "Bcl2l11"; gene_name "Bcl2l11"; p_id "P21609"; transcript_id "NM_009754"; tss_id "TSS12620";
chr19 exon  29459880   29462584  .  +  .  gene_id "Cd274"; gene_name "Cd274"; p_id "P24521"; transcript_id "NM_0218937"; tss_id "TSS25796";
chr2  exon  164896222  164897154 .  +  .  gene_id "Cd40"; gene_name "Cd40"; p_id "P289"; transcript_id "NM_011611"; tss_id "TSS22026";
chr3  exon  51053619   51055576  .  +  .  gene_id "Ccrn41"; gene_name "Ccrn41"; p_id "P42"; transcript_id "NM_009834"; tss_id "TSS17093";
chr7  exon  133732809  133733063 .  -  .  gene_id "Il27"; gene_name "Il27"; p_id "P17495"; transcript_id "NM_145636"; tss_id "TSS15214";
chr7  exon  148449082  148449279 .  -  .  gene_id "Irf7"; gene_name "Irf7"; p_id "P23208"; transcript_id "NM_016850"; tss_id "TSS20549";
chr5  exon  28403331   28405202  .  +  .  gene_id "Insig1"; gene_name "Insig1"; p_id "P6141"; transcript_id "NM_153526"; tss_id "TSS14384";
chr18 exon  65635642   65638446  .  +  .  gene_id "Mallt1"; gene_name "Mallt1"; p_id "P17757"; transcript_id "NM_172833"; tss_id "TSS23592";
chr1  exon  135971442  135974280 .  -  .  gene_id "Btg2"; gene_name "Btg2"; p_id "P9815"; transcript_id "NM_007570"; tss_id "TSS3364";
chr6  exon  129217343  129218387 .  -  .  gene_id "Cd69"; gene_name "Cd69"; p_id "P9137"; transcript_id "NM_001033122"; tss_id "TSS439";
chr17 exon  13200705   13200934  .  +  .  gene_id "Sod2"; gene_name "Sod2"; p_id "P6002"; transcript_id "NM_013671"; tss_id "TSS18083";
chr7  exon  31213153   31213765  .  +  .  gene_id "Mfkbid"; gene_name "Mfkbid"; p_id "P22378"; transcript_id "NM_172142"; tss_id "TSS20875";
chr4  exon  128760972  128761770 .  +  .  gene_id "Rnf19b"; gene_name "Rnf19b"; p_id "P5648"; transcript_id "NM_029219"; tss_id "TSS5924";
chr3  exon  142300347  142300972 .  +  .  gene_id "Gbp2"; gene_name "Gbp2"; p_id "P18816"; transcript_id "NM_010260"; tss_id "TSS1441";
chr17 exon  35958983   35959856  .  +  .  gene_id "Ier3"; gene_name "Ier3"; p_id "P16019"; transcript_id "NM_133662"; tss_id "TSS23149";
chr3  exon  31017377   31021845  .  +  .  gene_id "Skil"; gene_name "Skil"; p_id "P16978"; transcript_id "NM_0010390900"; tss_id "TSS20990";
chr9  exon  51976136   51976216  .  -  .  gene_id "Zc3h12c"; gene_name "Zc3h12c"; p_id "P19129"; transcript_id "NM_0011629211"; tss_id "TSS15697";
chr18 exon  35022119   35024610  .  +  .  gene_id "Egr1"; gene_name "Egr1"; p_id "P3177"; transcript_id "NM_007913"; tss_id "TSS5270";
chr6  exon  16783381   16784330  .  -  .  gene_id "Tfec"; gene_name "Tfec"; p_id "P3938"; transcript_id "NM_031198"; tss_id "TSS12610";
chr2  exon  68946499   68949347  .  +  .  gene_id "Cers6"; gene_name "Cers6"; p_id "P23342"; transcript_id "NM_172856"; tss_id "TSS24901";
chr17 exon  45698951   45700117  .  +  .  gene_id "Mfkbia"; gene_name "Mfkbia"; p_id "P16819"; transcript_id "NM_008690"; tss_id "TSS2517";
chr11 exon  81916612   81916901  .  +  .  gene_id "Ccl12"; gene_name "Ccl12"; p_id "P21246"; transcript_id "NM_011331"; tss_id "TSS4090";
chr17 exon  34052288   34052535  .  +  .  gene_id "Daxx"; gene_name "Daxx"; p_id "P7384"; transcript_id "NM_007829"; tss_id "TSS7741";
chr14 exon  21661496   21662610  .  +  .  gene_id "Plau"; gene_name "Plau"; p_id "P5851"; transcript_id "NM_008873"; tss_id "TSS14484";
chr10 exon  61957176   61957854  .  -  .  gene_id "Srgn"; gene_name "Srgn"; p_id "P15966"; transcript_id "NM_011157"; tss_id "TSS17012";
chr11 exon  44226726   44227519  .  +  .  gene_id "Il12b"; gene_name "Il12b"; p_id "P19893"; transcript_id "NM_008352"; tss_id "TSS584";

F

chr11 exon  97850894   97851653  .  -  .  gene_id "Arl5c"; gene_name "Arl5c"; p_id "P14239"; transcript_id "NM_207231"; tss_id "TSS9964";
chr2  exon  83643397   83647073  .  +  .  gene_id "Itgav"; gene_name "Itgav"; p_id "P22324"; transcript_id "NM_008402"; tss_id "TSS21510";
chr1  exon  151952503  151955142 .  +  .  gene_id "Ptgs2"; gene_name "Ptgs2"; p_id "P8580"; transcript_id "NM_011198"; tss_id "TSS16611";
chr4  exon  102277579  102279867 .  +  .  gene_id "Pde4b"; gene_name "Pde4b"; p_id "P3538"; transcript_id "NM_0011779817"; tss_id "TSS8995";
chr9  exon  5336535    5336791   .  +  .  gene_id "Casp4"; gene_name "Casp4"; p_id "P4979"; transcript_id "NM_007609"; tss_id "TSS8769";
chr3  exon  107543966  107546063 .  -  .  gene_id "Csf1"; gene_name "Csf1"; p_id "P4042"; transcript_id "NM_007778"; tss_id "TSS22705";
chr19 exon  34647557   34651024  .  +  .  gene_id "Ifit2"; gene_name "Ifit2"; p_id "P8949"; transcript_id "NM_008332"; tss_id "TSS2685";
chr19 exon  32197217   32199014  .  -  .  gene_id "Sgms1"; gene_name "Sgms1"; p_id "P20224"; transcript_id "NM_0011685626"; tss_id "TSS7412";
chr3  exon  142235438  142236176 .  +  .  gene_id "Gbp3"; gene_name "Gbp3"; p_id "P13206"; transcript_id "NM_018734"; tss_id "TSS15170";
chr11 exon  48908891   48910489  .  +  .  gene_id "Ifi47"; gene_name "Ifi47"; p_id "P1015"; transcript_id "NM_008330"; tss_id "TSS10281";
chr17 exon  57343396   57343577  .  -  .  gene_id "C3"; gene_name "C3"; p_id "P2319"; transcript_id "NM_009778"; tss_id "TSS1016";
chr2  exon  24204935   24207011  .  +  .  gene_id "Illrn"; gene_name "Illrn"; p_id "P526"; transcript_id "NM_0010397017"; tss_id "TSS6957";
chr8  exon  35881452   35882948  .  +  .  gene_id "Dusp4"; gene_name "Dusp4"; p_id "P7460"; transcript_id "NM_176933"; tss_id "TSS1452";
chr9  exon  7282018    7283333   .  +  .  gene_id "Mmp13"; gene_name "Mmp13"; p_id "P13295"; transcript_id "NM_008607"; tss_id "TSS15376";
chr2  exon  127163317  127164113 .  +  .  gene_id "Dusp2"; gene_name "Dusp2"; p_id "P9444"; transcript_id "NM_0100990"; tss_id "TSS2906";
chr13 exon  75908128   75909806  .  +  .  gene_id "Ell2"; gene_name "Ell2"; p_id "P9427"; transcript_id "NM_138953"; tss_id "TSS19027";
chr12 exon  86979723   86980828  .  +  .  gene_id "Jdp2"; gene_name "Jdp2"; p_id "P1564"; transcript_id "NM_001205052"; tss_id "TSS2002";
chr3  exon  122241637  122245201 .  -  .  gene_id "Fnbp11"; gene_name "Fnbp11"; p_id "P23393"; transcript_id "NM_0011146655"; tss_id "TSS1974";
chr4  exon  16050522   16051038  .  -  .  gene_id "Ripk2"; gene_name "Ripk2"; p_id "P9162"; transcript_id "NM_138952"; tss_id "TSS16489";
chr4  exon  129191961  129193225 .  +  .  gene_id "Marcksl1"; gene_name "Marcksl1"; p_id "P12626"; transcript_id "NM_0100807"; tss_id "TSS2577";
chr10 exon  39753341   39757210  .  -  .  gene_id "Slc16a10"; gene_name "Slc16a10"; p_id "P791"; transcript_id "NM_001114332"; tss_id "TSS14295";
chr2  exon  34798778   34799611  .  -  .  gene_id "Traf1"; gene_name "Traf1"; p_id "P23707"; transcript_id "NM_009421"; tss_id "TSS6544";
chr3  exon  10204343   10204570  .  -  .  gene_id "Fabp4"; gene_name "Fabp4"; p_id "P2902"; transcript_id "NM_024406"; tss_id "TSS6338";
chr7  exon  109706436  109708177 .  -  .  gene_id "Trim21"; gene_name "Trim21"; p_id "P23556"; transcript_id "NM_001082552"; tss_id "TSS11938";
chr19 exon  29386231   29387570  .  +  .  gene_id "Jak2"; gene_name "Jak2"; p_id "P5298"; transcript_id "NM_001048177"; tss_id "TSS16849";
chr4  exon  150240425  150242989 .  +  .  gene_id "Errfi1"; gene_name "Errfi1"; p_id "P9372"; transcript_id "NM_133753"; tss_id "TSS4716";
chr9  exon  119880284  119882329 .  -  .  gene_id "Csrnp1"; gene_name "Csrnp1"; p_id "P1038"; transcript_id "NM_153287"; tss_id "TSS10786";
chr5  exon  136623332  136626822 .  -  .  gene_id "Orai2"; gene_name "Orai2"; p_id "P8692"; transcript_id "NM_1787511"; tss_id "TSS11015";
chr15 exon  36169689   36171820  .  -  .  gene_id "Rnf19a"; gene_name "Rnf19a"; p_id "P15766"; transcript_id "NM_013923"; tss_id "TSS3635";
chr12 exon  86816767   86818219  .  +  .  gene_id "Fos"; gene_name "Fos"; p_id "P6877"; transcript_id "NM_010234"; tss_id "TSS15173";
chr5  exon  115386927  115387920 .  +  .  gene_id "Gas1l"; gene_name "Gas1l"; p_id "P22147"; transcript_id "NM_145209"; tss_id "TSS4493";
chr9  exon  20834826   20834975  .  +  .  gene_id "Icam4"; gene_name "Icam4"; p_id "P16107"; transcript_id "NM_023892"; tss_id "TSS5040";
chrX  exon  106393115  106395784 .  +  .  gene_id "Gm6377"; gene_name "Gm6377"; p_id "P22506"; transcript_id "NM_001037917"; tss_id "TSS20699";
chr8  exon  11453977   11454526  .  -  .  gene_id "Rab20"; gene_name "Rab20"; p_id "P9178"; transcript_id "NM_011227"; tss_id "TSS9231";
chr4  exon  155573533  155574175 .  -  .  gene_id "Isg15"; gene_name "Isg15"; p_id "P60899"; transcript_id "NM_015783"; tss_id "TSS19779";
chr7  exon  25260258   25260892  .  +  .  gene_id "Plaur"; gene_name "Plaur"; p_id "P9755"; transcript_id "NM_011113"; tss_id "TSS24418";
chr1  exon  64082247   64082338  .  -  .  gene_id "Klf7"; gene_name "Klf7"; p_id "P9777"; transcript_id "NM_033563"; tss_id "TSS7500";
chr1  exon  175677425  175677751 .  -  .  gene_id "Ifi204"; gene_name "Ifi204"; p_id "P15649"; transcript_id "NM_008329"; tss_id "TSS17138";
chr4  exon  135700795  135701307 .  +  .  gene_id "Id3"; gene_name "Id3"; p_id "P21755"; transcript_id "NM_008321"; tss_id "TSS23418";
chr19 exon  11122102   11122583  .  -  .  gene_id "AW112010"; gene_name "AW112010"; p_id "P21084"; transcript_id "NM_001177351"; tss_id "TSS12924";

G

chr16 exon  95940530   95942656  .  +  .  gene_id "Ets2"; gene_name "2015-03-02Ets2"; p_id "P20783"; transcript_id "NM_011809"; tss_id "TSS12733";
chr11 exon  78773156   78773626  .  +  .  gene_id "Nos2"; gene_name "Nos2"; p_id "P23999"; transcript_id "NM_010927"; tss_id "TSS5656";
chr8  exon  87185230   87186751  .  -  .  gene_id "Ier2"; gene_name "Ier2"; p_id "P3914"; transcript_id "NM_010499"; tss_id "TSS17043";
chr6  exon  13785074   13787073  .  -  .  gene_id "Gpr85"; gene_name "Gpr85"; p_id "P9006"; transcript_id "NM_145066"; tss_id "TSS13111";
chr2  exon  157942141  157943687 .  -  .  gene_id "Tgm2"; gene_name "Tgm2"; p_id "P20583"; transcript_id "NM_009373"; tss_id "TSS2703";
chr11 exon  58019546   58021094  .  +  .  gene_id "Igtp"; gene_name "Igtp"; p_id "P10081"; transcript_id "NM_018738"; tss_id "TSS3154";
chr4  exon  53562861   53563036  .  +  .  gene_id "Slc44a1"; gene_name "Slc44a1"; p_id "P9505"; transcript_id "NM_0011596633"; tss_id "TSS19315";
chr13 exon  114382387  114385070 .  +  .  gene_id "Snx18"; gene_name "Snx18"; p_id "P18333"; transcript_id "NM_130796"; tss_id "TSS9534";
chr10 exon  12659787   12661787  .  -  .  gene_id "Stx11"; gene_name "Stx11"; p_id "P1435"; transcript_id "NM_001163590"; tss_id "TSS20471";
chr8  exon  91211459   91212373  .  +  .  gene_id "Nod2"; gene_name "Nod2"; p_id "P13132"; transcript_id "NM_145857"; tss_id "TSS7189";
chr7  exon  53967368   53967634  .  -  .  gene_id "Saa3"; gene_name "Saa3"; p_id "P5870"; transcript_id "NM_011315"; tss_id "TSS9385";
chr6  exon  57612694   57615130  .  +  .  gene_id "Herc6"; gene_name "Herc6"; p_id "P15778"; transcript_id "NM_025992"; tss_id "TSS14171";
chr6  exon  121220490  121220935 .  +  .  gene_id "Usp18"; gene_name "Usp18"; p_id "P19317"; transcript_id "NM_011909"; tss_id "TSS3918";
chr10 exon  67002493   67004936  .  +  .  gene_id "Egr2"; gene_name "Egr2"; p_id "P18543"; transcript_id "NM_010118"; tss_id "TSS12595";
chr5  exon  30345921   30346508  .  +  .  gene_id "Il6"; gene_name "Il6"; p_id "P9970"; transcript_id "NM_031168"; tss_id "TSS6882";
chr7  exon  108430964  108432919 .  +  .  gene_id "Fcho2"; gene_name "Fcho2"; p_id "P13022"; transcript_id "NM_0011440110"; tss_id "TSS1726";
chr6  exon  146598633  146599122 .  +  .  gene_id "Med21"; gene_name "Med21"; p_id "P199"; transcript_id "NM_025315"; tss_id "TSS1643";
chr11 exon  58001241   58003700  .  +  .  gene_id "Gm12250"; gene_name "Gm12250"; p_id "P6220"; transcript_id "NM_0011351115"; tss_id "TSS1143";
chr1  exon  175576122  175578059 .  +  .  gene_id "Pyhin1"; gene_name "Pyhin1"; p_id "P19336"; transcript_id "NM_175026"; tss_id "TSS14108";
chr10 exon  33845873   33846372  .  -  .  gene_id "Fam26f"; gene_name "Fam26f"; p_id "P10436"; transcript_id "NM_175449"; tss_id "TSS7315";
chr7  exon  91257866   91259311  .  -  .  gene_id "Fam108c"; gene_name "Fam108c"; p_id "P22064"; transcript_id "NM_133722"; tss_id "TSS1034";
chr5  exon  20881072   20884204  .  +  .  gene_id "Fgl2"; gene_name "Fgl2"; p_id "P19695"; transcript_id "NM_008013"; tss_id "TSS20933";
chr17 exon  8125748    8127762   .  +  .  gene_id "Tagap"; gene_name "Tagap"; p_id "P8999"; transcript_id "NM_145968_1"; tss_id "TSS4240";
chr17_random exon 39677 41691 . - . gene_id "Tagap"; gene_name "Tagap"; p_id "P23244"; transcript_id "NM_145968"; tss_id "TSS6373";
chr7  exon  29551263   29551543  .  -  .  gene_id "Nfkbib"; gene_name "Nfkbib"; p_id "P10826"; transcript_id "NM_010906"; tss_id "TSS22699";
chr2  exon  129125346  129126635 .  -  .  gene_id "Il1a"; gene_name "Il1a"; p_id "P9941"; transcript_id "NM_010554"; tss_id "TSS14538";
chr19 exon  34661551   34663472  .  +  .  gene_id "Ifit3"; gene_name "Ifit3"; p_id "P93803"; transcript_id "NM_010501"; tss_id "TSS22430";
chr2  exon  122286156  122286866 .  +  .  gene_id "Slc28a2"; gene_name "Slc28a2"; p_id "P8445"; transcript_id "NM_172980"; tss_id "TSS16659";
chrX  exon  166412975  166414722 .  -  .  gene_id "G530011006Rik"; gene_name "G530011006Rik"; transcript_id "NR_029457"; tss_id "TSS23870";
chr15 exon  79187851   79189506  .  +  .  gene_id "Maff"; gene_name "Maff"; p_id "P21898"; transcript_id "NM_010755"; tss_id "TSS7572";
chr7  exon  52094953   52096179  .  +  .  gene_id "Il4i1"; gene_name "Il4i1"; p_id "P14475"; transcript_id "NM_010215"; tss_id "TSS21824";
chr18 exon  60549446   60552283  .  +  .  gene_id "Ilgp1"; gene_name "Ilgp1"; p_id "P24220"; transcript_id "NM_0011462775"; tss_id "TSS22447";
chr3  exon  142209284  142213047 .  +  .  gene_id "Gbp7"; gene_name "Gbp7"; p_id "P10586"; transcript_id "NM_001083312"; tss_id "TSS23222";
chr4  exon  88167929   88168698  .  -  .  gene_id "Ifnb1"; gene_name "Ifnb1"; p_id "P21845"; transcript_id "NM_010510"; tss_id "TSS2442";
chr3  exon  50168858   50176288  .  -  .  gene_id "Slc7a11"; gene_name "Slc7a11"; p_id "P4123"; transcript_id "NM_011990"; tss_id "TSS24352";
chr16 exon  75859039   75859884  .  -  .  gene_id "Samsn1"; gene_name "Samsn1"; p_id "P14457"; transcript_id "NM_023380"; tss_id "TSS20090";
chr7  exon  81420304   81420974  .  -  .  gene_id "Slco3a1"; gene_name "Slco3a1"; p_id "P121"; transcript_id "NM_001038643"; tss_id "TSS8114";
chr11 exon  69212020   69213351  .  -  .  gene_id "Kdm6b"; gene_name "Kdm6b"; p_id "P9055"; transcript_id "NM_001017426"; tss_id "TSS18373";
chr11 exon  13868718   13871184  .  +  .  gene_id "Hivep2"; gene_name "Hivep2"; p_id "P22703"; transcript_id "NM_010477"; tss_id "TSS6381";
chr9  exon  92166367   92167399  .  +  .  gene_id "Plscr1"; gene_name "Plscr1"; p_id "P17054"; transcript_id "NM_011636"; tss_id "TSS11480";
chr6  exon  54873936   54875003  .  -  .  gene_id "Nod1"; gene_name "Nod1"; p_id "P17761"; transcript_id "NM_172729"; tss_id "TSS13431";
chr5  exon  137537376  137539081 .  -  .  gene_id "Serpine1"; gene_name "Serpine1"; p_id "P16074"; transcript_id "NM_008871"; tss_id "TSS17025";
chr17 exon  18029697   18030916  .  +  .  gene_id "Ppz2"; gene_name "Ppz2"; p_id "P2200"; transcript_id "NM_008039"; tss_id "TSS20689";
chr10 exon  44156981   44160048  .  -  .  gene_id "Prdm1"; gene_name "Prdm1"; p_id "P5182"; transcript_id "NM_007548"; tss_id "TSS16926";
chr8  exon  28249245   28253866  .  -  .  gene_id "Rabl1fip1"; gene_name "Rabl1fip1"; p_id "P220"; transcript_id "NM_001080813"; tss_id "TSS19574";
chr9  exon  44494391   44495781  .  -  .  gene_id "Phldb1"; gene_name "Phldb1"; p_id "P8634"; transcript_id "NM_153537"; tss_id "TSS9456";
chr15 exon  103138666  103140087 .  -  .  gene_id "Qpr84"; gene_name "Qpr84"; p_id "P18218"; transcript_id "NM_030720"; tss_id "TSS17218";
chr5  exon  124313579  124315525 .  -  .  gene_id "Niacr1"; gene_name "Niacr1"; p_id "P4625"; transcript_id "NM_030701"; tss_id "TSS707";
chr8  exon  86734841   86736704  .  -  .  gene_id "Zswim4"; gene_name "Zswim4"; p_id "P24135"; transcript_id "NM_172503"; tss_id "TSS887";
chr5  exon  92750357   92752932  .  -  .  gene_id "Cxcl9"; gene_name "Cxcl9"; p_id "P19130"; transcript_id "NM_008599"; tss_id "TSS2167";
chr8  exon  69000245   69000808  .  -  .  gene_id "Tma16"; gene_name "Tma16"; p_id "P22272"; transcript_id "NM_025465"; tss_id "TSS4039";
chr15 exon  61820901   61821916  .  +  .  gene_id "Myc"; gene_name "Myc"; p_id "P18425"; transcript_id "NM_010849"; tss_id "TSS15915";
chr4  exon  63385637   63391077  .  -  .  gene_id "Tnfsf15"; gene_name "Tnfsf15"; p_id "P12252"; transcript_id "NM_177371"; tss_id "TSS3724";
chr19 exon  53615812   53615912  .  +  .  gene_id "Dusp5"; gene_name "Dusp5"; p_id "P6646"; transcript_id "NM_0010953390"; tss_id "TSS19363";
chr2  exon  173032114  173032727 .  -  .  gene_id "Zbp1"; gene_name "Zbp1"; p_id "P4281"; transcript_id "NM_021394"; tss_id "TSS13610";
chr4  exon  15924268   15926147  .  -  .  gene_id "Osgin2"; gene_name "Osgin2"; p_id "P24566"; transcript_id "NM_145950"; tss_id "TSS5672";
chr4  exon  11640623   11642140  .  +  .  gene_id "Gem"; gene_name "Gem"; p_id "P845"; transcript_id "NM_010276"; tss_id "TSS16095";
chr5  exon  92788571   92789832  .  -  .  gene_id "Cxcl10"; gene_name "Cxcl10"; p_id "P17002"; transcript_id "NM_019494"; tss_id "TSS8759";
chr15 exon  59485791   59488654  .  +  .  gene_id "Trib1"; gene_name "Trib1"; p_id "P21350"; transcript_id "NM_1445497"; tss_id "TSS14254";
chr11 exon  48947790   48948889  .  +  .  gene_id "Olfr56"; gene_name "Olfr56"; p_id "P9227"; transcript_id "NM_010999"; tss_id "TSS4614";
chr19 exon  34548274   34549625  .  -  .  gene_id "Ch25h"; gene_name "Ch25h"; p_id "P882"; transcript_id "NM_009890"; tss_id "TSS17189";
chr19 exon  6171306    6172475   .  +  .  gene_id "Batf3"; gene_name "Batf3"; p_id "P13629"; transcript_id "NM_028967"; tss_id "TSS15611";
chr2  exon  11654777   11655612  .  +  .  gene_id "Il15ra"; gene_name "Il15ra"; p_id "P10569"; transcript_id "NM_133836"; tss_id "TSS13507";
chr17 exon  12411851   12412506  .  +  .  gene_id "Agpat4"; gene_name "Agpat4"; p_id "P20641"; transcript_id "NM_026644"; tss_id "TSS3138";
chr18 exon  60405029   60407474  .  +  .  gene_id "Gm4951"; gene_name "Gm4951"; p_id "P3089"; transcript_id "NM_001033767"; tss_id "TSS3296";
chr1  exon  175942129  175942419 .  -  .  gene_id "Ifi205"; gene_name "Ifi205"; p_id "P4164"; transcript_id "NM_172648"; tss_id "TSS16317";
chr16 exon  92392198   92393873  .  -  .  gene_id "Rcan1"; gene_name "Rcan1"; p_id "P16767"; transcript_id "NM_019466"; tss_id "TSS831";
```

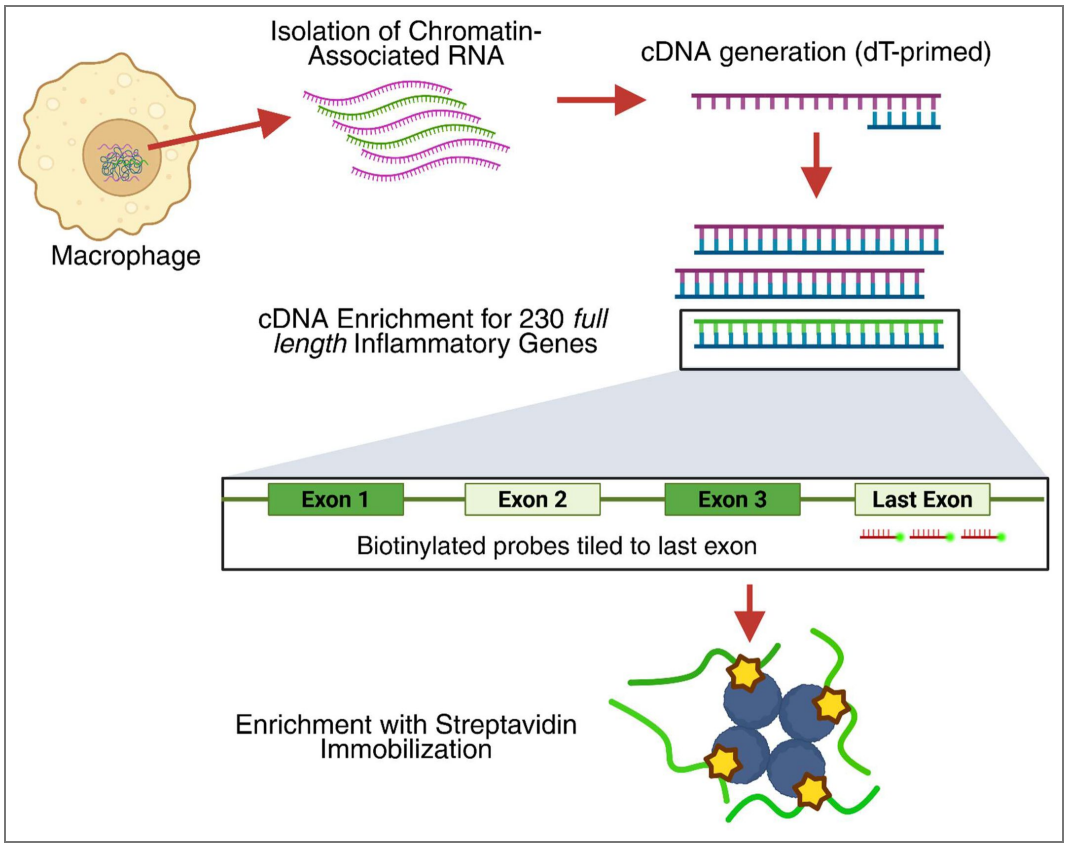
Table 2. (continued)

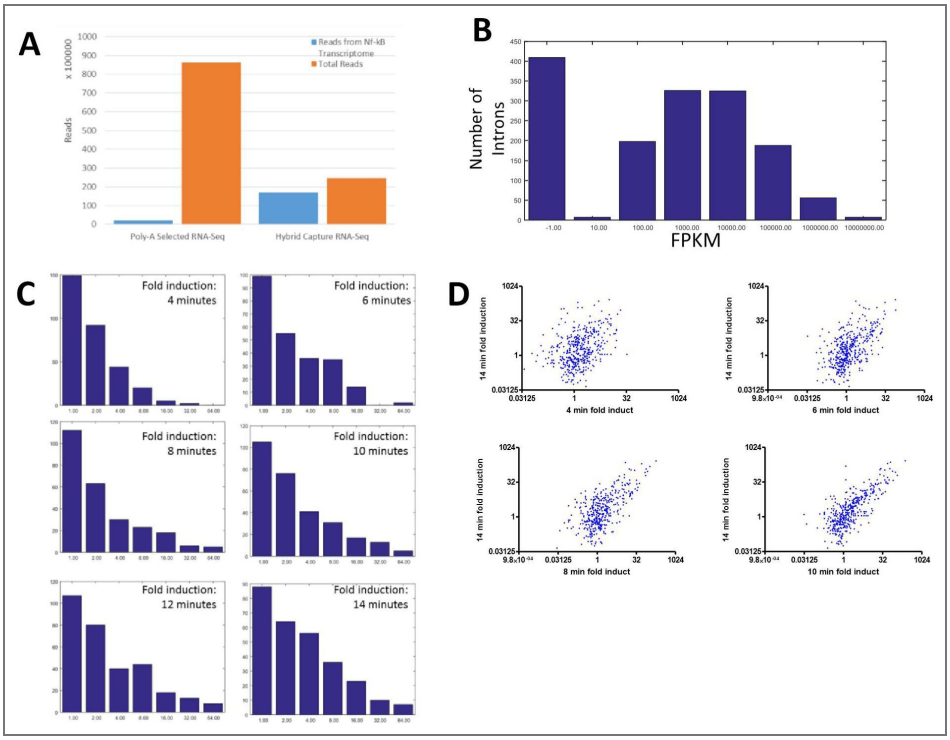
Table 2. (continued)

```
chr13  exon      89794915      89796963      .      -      .      gene_id "Vcan"; gene_name "Vcan"; p_id "P15421"; transcript_id "NM_001081249"; tss_id "TSS20677";
chr4   exon     155539399     155540762      .      -      .      gene_id "Agrn"; gene_name "Agrn"; p_id "P16425"; transcript_id "NM_021604"; tss_id "TSS15328";

H

chr17  exon      34993319      34993527      .      -      .      gene_id "Cfb"; gene_name "Cfb"; p_id "P13056"; transcript_id "NM_001142706"; tss_id "TSS23526";
chr17  exon      48004082      48005752      .      -      .      gene_id "Foxp4"; gene_name "Foxp4"; p_id "P16134"; transcript_id "NM_001110824"; tss_id "TSS19822";
chr9   exon     107202699     107204292      .      +      .      gene_id "Ciah"; gene_name "Ciah"; p_id "P83211"; transcript_id "NM_009899"; tss_id "TSS23598";
chr3   exon      27234144      27238587      .      +      .      gene_id "Tnfsf10"; gene_name "Tnfsf10"; p_id "P6001"; transcript_id "NM_009425"; tss_id "TSS7370";
chr14  exon       8782819       8784101      .      +      .      gene_id "Flnb"; gene_name "Flnb"; p_id "P6881"; transcript_id "NM_134080"; tss_id "TSS13225";
chr14  exon      59966244      59967888      .      -      .      gene_id "Phf11d"; gene_name "Phf11d"; p_id "P19480"; transcript_id "NM_199015"; tss_id "TSS13050";
chr3   exon     151393887     151395517      .      -      .      gene_id "Ifi44"; gene_name "Ifi44"; p_id "P47107"; transcript_id "NM_133871"; tss_id "TSS11790";
chr5   exon     105507413     105509283      .      -      .      gene_id "Gbp9"; gene_name "Gbp9"; p_id "P23637"; transcript_id "NM_172777"; tss_id "TSS22932";
chr13  exon      42402146      42403358      .      +      .      gene_id "Edn1"; gene_name "Edn1"; p_id "P15001"; transcript_id "NM_010104"; tss_id "TSS16855";
chr12  exon      72000371      72003221      .      +      .      gene_id "Frmd6"; gene_name "Frmd6"; p_id "P2929"; transcript_id "NM_028127"; tss_id "TSS4836";
chr7   exon      86064932      86065282      .      +      .      gene_id "Isg20"; gene_name "Isg20"; p_id "P23651"; transcript_id "NM_020583"; tss_id "TSS17804";
chr2   exon     122466437     122466812      .      +      .      gene_id "AA467197"; gene_name "AA467197"; p_id "P11166"; transcript_id "NM_001004174"; tss_id "TSS5413";
chr7   exon      83017584      83019016      .      +      .      gene_id "Klhl25"; gene_name "Klhl25"; p_id "P22799"; transcript_id "NM_001122780"; tss_id "TSS12992";
chr11  exon      86582351      86583763      .      -      .      gene_id "Dhx40"; gene_name "Dhx40"; p_id "P23840"; transcript_id "NM_026191"; tss_id "TSS24769";
chr18  exon      66622924      66625212      .      +      .      gene_id "Pmaip1"; gene_name "Pmaip1"; p_id "P7859"; transcript_id "NM_021451"; tss_id "TSS22625";
chr15  exon      57670146      57671387      .      +      .      gene_id "Zhx2"; gene_name "Zhx2"; p_id "P21154"; transcript_id "NM_199449"; tss_id "TSS3741";
chr19  exon      34667378      34669513      .      -      .      gene_id "Gm14446"; gene_name "Gm14446"; p_id "P16440"; transcript_id "NM_001101605"; tss_id "TSS23646";
chr8   exon      10986964      10987014      .      -      .      gene_id "Irs2"; gene_name "Irs2"; p_id "P3108"; transcript_id "NM_001081212"; tss_id "TSS25904";
chr8   exon     108497420     108497751      .      -      .      gene_id "Dpep3"; gene_name "Dpep3"; p_id "P1255"; transcript_id "NM_027960"; tss_id "TSS20475";
chr10  exon     110945116     110945705      .      +      .      gene_id "Phldal"; gene_name "Phldal"; p_id "P19275"; transcript_id "NM_009344"; tss_id "TSS19408";
chr2   exon      52990874      52993236      .      +      .      gene_id "Fmnl2"; gene_name "Fmnl2"; p_id "P4156"; transcript_id "NM_172409"; tss_id "TSS2704";
chr11  exon      82934523      82936161      .      +      .      gene_id "Slfn1"; gene_name "Slfn1"; p_id "P22608"; transcript_id "NM_011407"; tss_id "TSS25201";
chr19  exon      11500686      11501736      .      +      .      gene_id "Ms4a4c"; gene_name "Ms4a4c"; p_id "P22794"; transcript_id "NM_029499"; tss_id "TSS21383";
chr5   exon     105544786     105547518      .      -      .      gene_id "Gbp4"; gene_name "Gbp4"; p_id "P24799"; transcript_id "NM_008620"; tss_id "TSS16713";
chr3   exon      19892509      19892998      .      +      .      gene_id "Cp"; gene_name "Cp"; p_id "P6496"; transcript_id "NM_007752"; tss_id "TSS19796";
chr11  exon      83002165      83003718      .      +      .      gene_id "Slfn4"; gene_name "Slfn4"; p_id "P16897"; transcript_id "NM_011410"; tss_id "TSS23253";
chr14  exon      59999670      59999984      .      -      .      gene_id "Phf11c"; gene_name "Phf11c"; p_id "P2484"; transcript_id "NM_001164289"; tss_id "TSS18997";
chr16  exon      11072325      11075078      .      +      .      gene_id "Snn"; gene_name "Snn"; p_id "P22666"; transcript_id "NM_009223"; tss_id "TSS15465";
chr3   exon      32482450      32484148      .      -      .      gene_id "Gnb4"; gene_name "Gnb4"; p_id "P5107"; transcript_id "NM_013531"; tss_id "TSS18140";
chr6   exon     127075727     127080505      .      -      .      gene_id "Ccnd2"; gene_name "Ccnd2"; p_id "P4356"; transcript_id "NM_009829"; tss_id "TSS24028";
chr14  exon      70478703      70479964      .      +      .      gene_id "Egr3"; gene_name "Egr3"; p_id "P10547"; transcript_id "NM_018781"; tss_id "TSS18772";

I

chr3   exon     115729931     115731934      .      +      .      gene_id "Ext12"; gene_name "Ext12"; p_id "P23186"; transcript_id "NM_001163515"; tss_id "TSS13032";
chr15  exon     101104427     101105225      .      +      .      gene_id "Nr4a1"; gene_name "Nr4a1"; p_id "P11564"; transcript_id "NM_010444"; tss_id "TSS1779";
chr11  exon      83325238      83325629      .      -      .      gene_id "Gm11435"; gene_name "Gm11435"; p_id "P3561"; transcript_id "NM_001045543"; tss_id "TSS21732";
chr14  exon      59939801      59940120      .      -      .      gene_id "Phf11b"; gene_name "Phf11b"; p_id "P21005"; transcript_id "NM_001164327"; tss_id "TSS6272";
chr12  exon      17557413      17558308      .      +      .      gene_id "Odc1"; gene_name "Odc1"; p_id "P13979"; transcript_id "NM_013614"; tss_id "TSS312";
chr7   exon     149265392     149268935      .      -      .      gene_id "Dusp8"; gene_name "Dusp8"; p_id "P5729"; transcript_id "NM_008748"; tss_id "TSS6971";
chr10  exon      76720518      76720731      .      -      .      gene_id "Gm10941"; gene_name "Gm10941"; transcript_id "NR_026944"; tss_id "TSS17723";
chr16  exon      85794073      85796345      .      -      .      gene_id "Adamts1"; gene_name "Adamts1"; p_id "P4336"; transcript_id "NM_009621"; tss_id "TSS24958";
chr12  exon      40902973      40903815      .      -      .      gene_id "Gm889"; gene_name "Gm889"; p_id "P23562"; transcript_id "NM_001033437"; tss_id "TSS7723";
```

Supplemental Figure 1. Hybrid capture strategy for isolating chromatin-associated inflammatory transcripts.

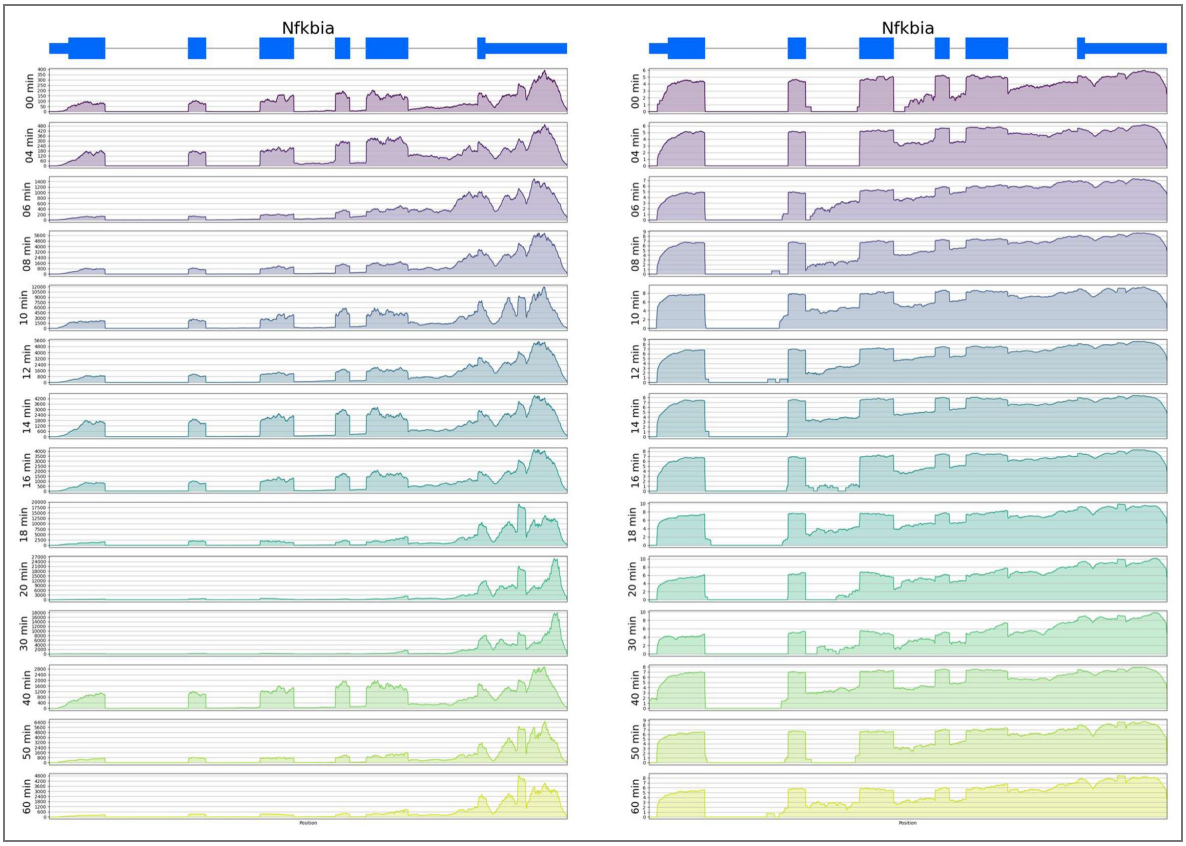
RNA was purified from chromatin-associated bone marrow–derived macrophages (BMDMs), and cDNA was generated using oligo(dT) priming to enrich for polyadenylated transcripts. Biotinylated RNA oligonucleotides complementary to the terminal exons of inflammatory genes were hybridized to the cDNAs, allowing for selective enrichment of these transcripts via streptavidin bead capture.





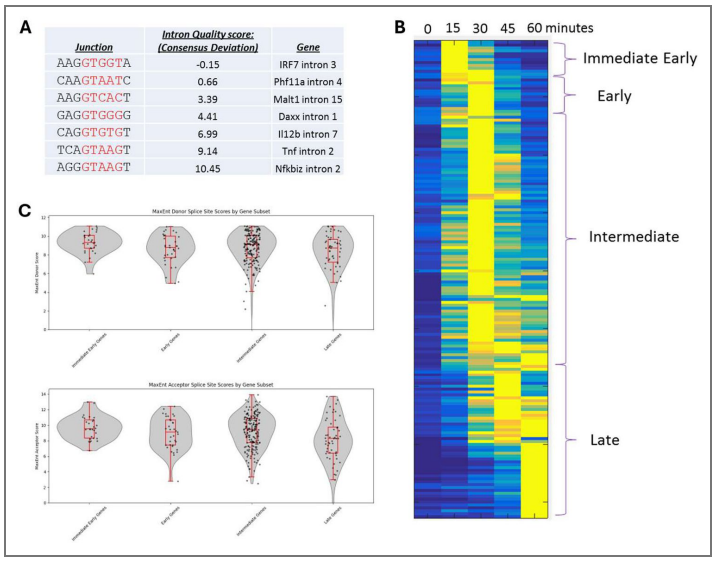
Supplemental Figure 2. Hybrid capture enriches NF-κB–responsive transcripts and yields robust intron coverage.

(A) Fraction of sequencing reads corresponding to NF-κB–responsive genes (blue) in hybrid-captured versus poly(A)-selected RNA (orange), showing substantial enrichment after hybrid capture. (B) Histogram of read counts per intron at the 6-min TNF induction time point. Reads were detected for 1,024 introns out of 1,508 targeted introns, with undetected introns largely corresponding to transcripts induced at later time points (>60 min). (C) Distribution of intron fold induction across time points reveals that many NF-κB target genes begin to show induction as early as 4 min post-TNF stimulation, with both the number of induced introns and magnitude of induction increasing markedly by 14 min. (D) Scatterplot comparing induction at 4 min versus 14 min for individual introns demonstrates that most early-induced genes are further upregulated at later time points.



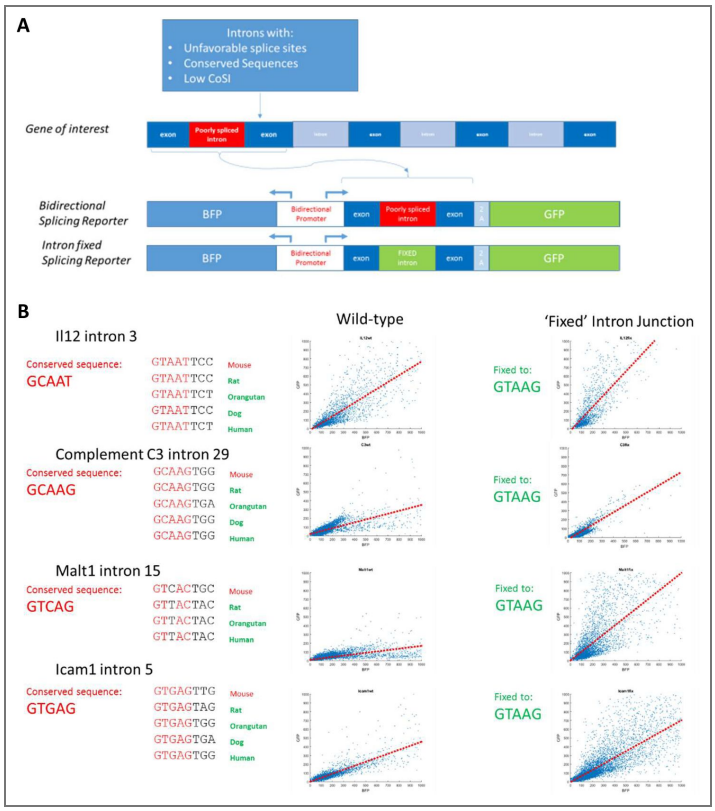
Supplemental Figure 3. Gene track visualization of Nfkbia induction dynamics.

Genome browser tracks show Nfkbia (IκBα) transcript induction over time in chromatin-associated, hybrid-captured RNA. Signal intensity is displayed on a linear scale (left) normalized to the maximum height at 20 min, and on a log scale (right) normalized to the maximum at each time point.



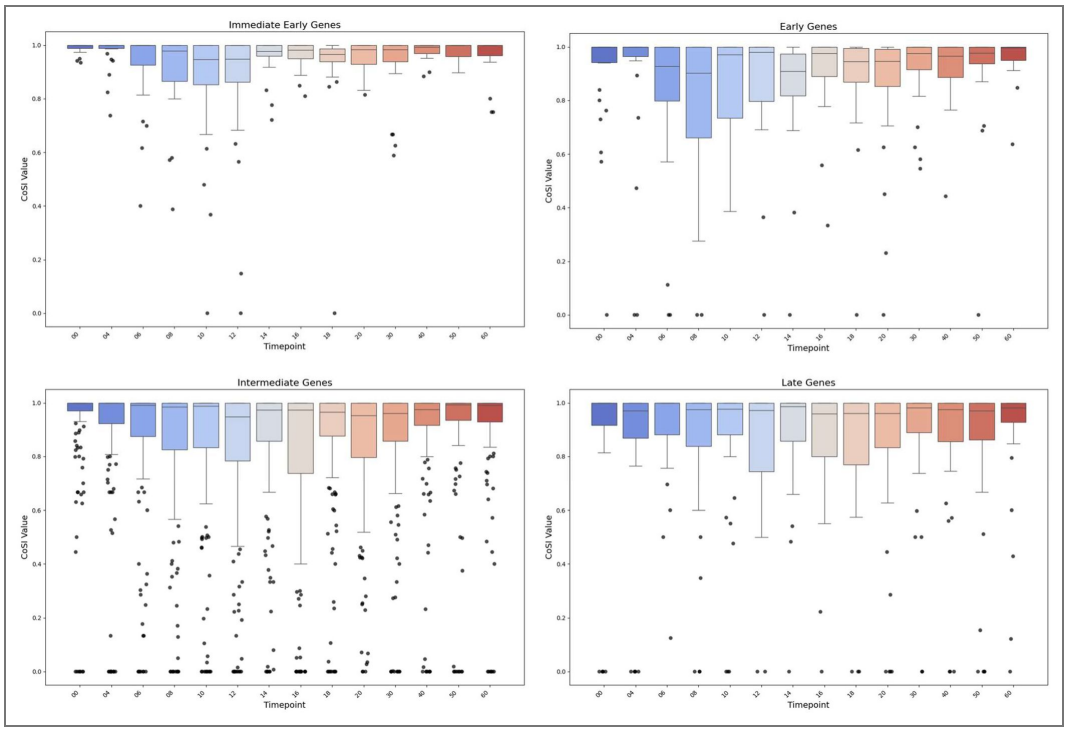
Supplemental Figure 4. Splice site strength and expression kinetics among NF-κB–induced genes.

(A) Representative introns from selected genes were scored for 5' splice donor strength based on similarity to the canonical 'GTAAG' motif; introns with weaker matches (e.g., Irf7 intron 3) received lower scores. (B) NF-κB–induced genes display well-characterized variability in expression kinetics (RNA-seq data from Reference 25). Heatmap shows temporal expression profiles of NF-κB target genes following lipid A stimulation, categorized into immediate-early, early, and later expression groups. (C) All introns in the NF-κB transcriptome were scored using MaxEntScan for both 5' splice donor (top) and 3' splice acceptor (bottom) sequences, stratified by their gene expression group.



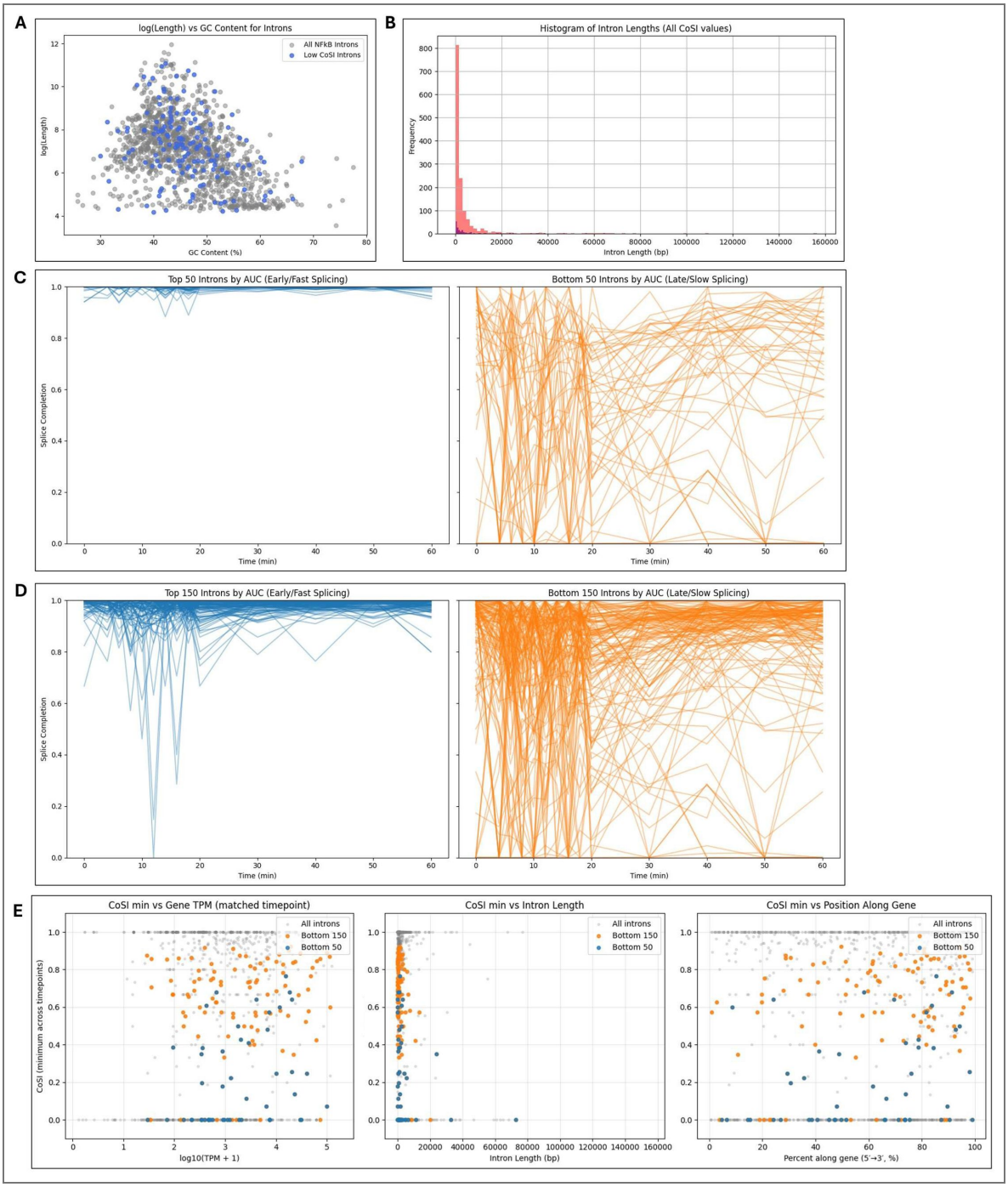
Supplemental Figure 5. Reporter assay design for assessing intron splicing efficiency.

(A) Schematic of the bidirectional reporter assay. Individual introns were cloned into a bidirectional promoter context together with their flanking exons, positioned upstream of a self-cleaving 2A peptide and GFP reporter. In the opposite transcriptional direction, a BFP reporter served as a transcriptional control. (B) Evolutionarily conserved weak 5' splice donors were "repaired" to the canonical GTAAG sequence within this reporter construct, and splicing efficiency was quantified by flow cytometry (FACS).



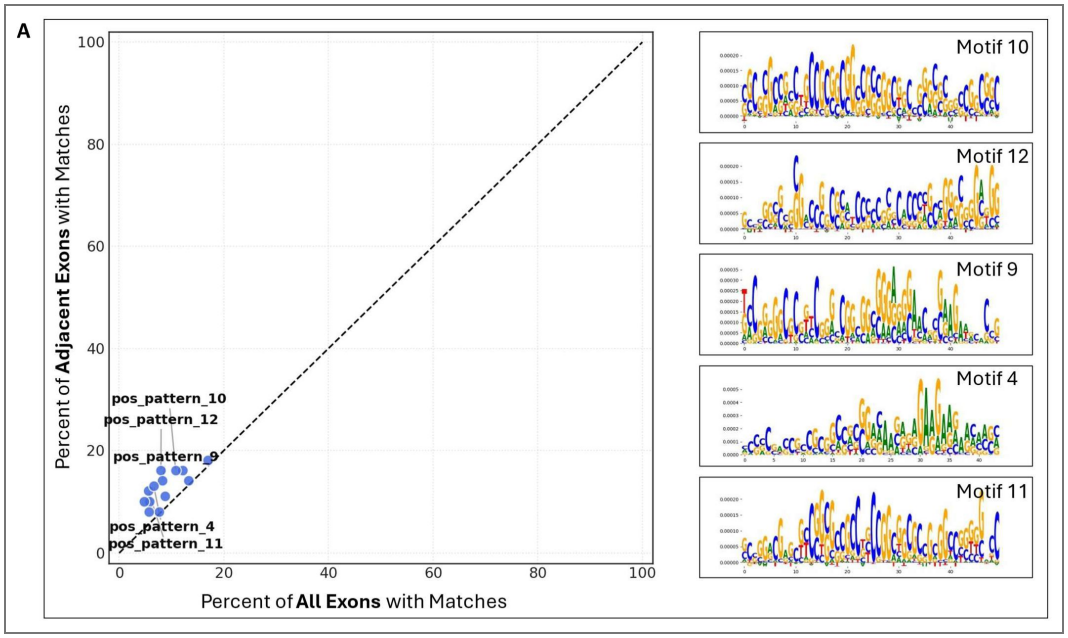
Supplemental Figure 6. Splicing completion across kinetic gene expression groups.

Box-and-whisker plots depict time-course CoSI values for introns from inflammatory gene cohorts. Each point represents the CoSI value of an individual intron at a given time point. Plots are grouped by expression category: Immediate Early (top left), Early (top right), Intermediate (bottom left), and Late (bottom right).



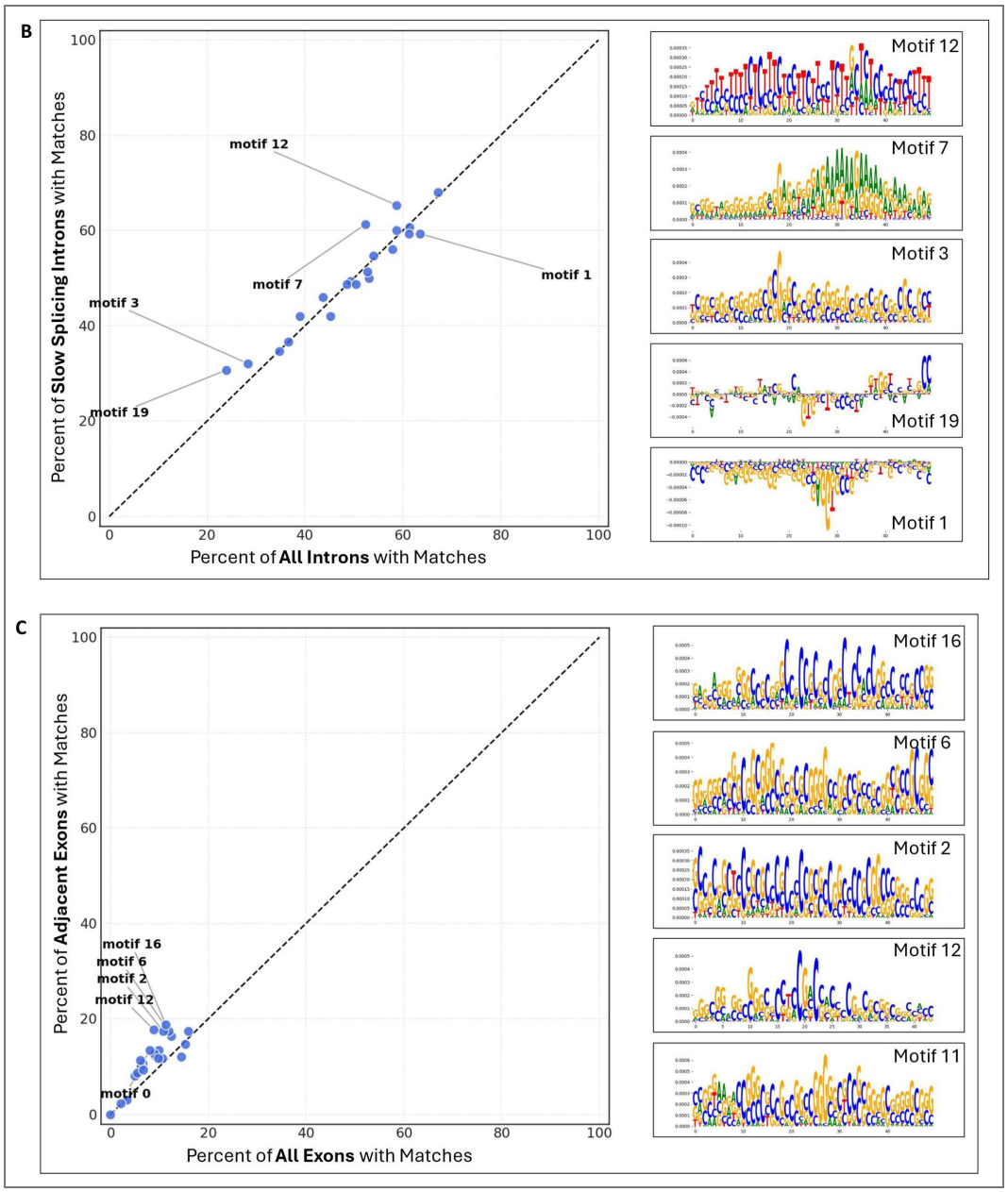
Supplemental Figure 7. Relationships between intron length, GC content, transcript position, and splicing kinetics.

(A) Scatterplot of intron length (y-axis) versus GC content (%) (x-axis) for the 150 slowest-splicing introns. Pearson r = −0.44 ($p = 2.76 \times 10^{-45}$) and Spearman r = −0.44 ($p = 1.34 \times 10^{-46}$). (B) Histogram of intron length distributions (100-nt bins) for all 1,098 introns (red) and the 150 slowest-splicing introns (blue). Length distributions are largely similar between cohorts. (C) Splice completion (CoSI) trajectories across the time course for the 50 fastest (left, blue) and 50 slowest (right, orange) introns, ranked by area under the CoSI curve (AUC). (D) Same analysis as in (C) but extended to the top 150 fastest and bottom 150 slowest introns. (E) Minimum CoSI values for each intron plotted against transcript abundance (TPM, left), intron length (middle), and position along the transcript (5' → 3' %, right) for the bottom 150 and bottom 50 introns. Correlations: TPM — Bottom 150: Pearson r = 0.204 ($p = 1.24 \times 10^{-2}$), Spearman r = 0.203 ($p = 1.27 \times 10^{-2}$); Bottom 50: Pearson r = 0.276 ($p = 5.25 \times 10^{-2}$), Spearman r = 0.316 ($p = 2.53 \times 10^{-2}$). Length — Bottom 150: Pearson r = −0.242 ($p = 2.88 \times 10^{-3}$), Spearman r = −0.266 ($p = 9.92 \times 10^{-4}$); Bottom 50: Pearson r = −0.160 (p = 0.268), Spearman r = −0.171 (p = 0.235). Transcript position — Bottom 150: Pearson r = 0.260 ($p = 1.33 \times 10^{-3}$), Spearman r = 0.231 ($p = 4.44 \times 10^{-3}$); Bottom 50: Pearson r = 0.259 ($p = 6.90 \times 10^{-2}$), Spearman r = 0.318 ($p = 2.43 \times 10^{-2}$).



Supplemental Figure 8. Enrichment of sequence motifs near slow-splicing introns.

Scatterplots show the percent representation of position weight matrices (PWMs) scanned across slow-splicing introns compared to all introns genome-wide using FIMO. Sequence logos of the top five enriched PWMs are displayed to the right of each scatterplot. (A) Scatterplot and enriched motifs identified in exons adjacent to the 50 slowest-splicing introns. (B) Scatterplot and enriched motifs identified within the 150 slowest-splicing introns. (C) Scatterplot and enriched motifs identified in exons adjacent to the 150 slowest-splicing introns.



Supplemental Figure 8. (continued)

Acknowledgements

The authors would like to thank Alex Shishkin and Mitchell Guttman (Dept. of Biology, Caltech) for assistance with hybrid capture strategy design; and Ann-Jay Tong Stephen Smale, Doug Black and Amy-Pandya Jones (Dept. of Biology, University of California, Los Angeles) for insights and advice; and Sergei Manakov, Evelyn Stuwe, Dubravka Pezic, Igor Antoshechkin, Sagar Damle, and Alok Joglekar (Dept. of Biology, California Institute of Technology) for experimental and computational assistance. This work was funded from a grant from NIH and from an endowment provided by the Raymond and Beverly Sackler Foundation. Research reported in this publication was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health under award number P20GM125498

Additional information

Author contributions

J.S. Dearborn, Conceptualization, Investigation, Data curation, Software, Formal analysis, Visualization, Methodology, Writing – review & editing. L. Frankiw, Conceptualization, Investigation, Resources, Methodology, Validation, Writing – review & editing. D.W. Limoge, Data curation, Formal analysis, Visualization, Writing – review & editing. C.H. Burns, L. Vlach, P. Turpin, T. Kirch, Z.D. Miller, W. Dowell, S. Languon, Y. Garcia-Flores, Investigation, Data curation, Validation, Writing – review & editing. R.C. Cockrell, Resources, Supervision, Project administration, Writing – review & editing. D. Baltimore, Conceptualization, Supervision, Funding acquisition, Project administration, Writing – review & editing. D. Majumdar, Conceptualization, Supervision, Funding acquisition, Project administration, Methodology, Writing – review & editing.

Funding

Funder	Grant reference number	Author
Raymond and Beverly Sackler Foundation		Dev Majumdar
HHS | NIH | National Institute of General Medical Sciences (NIGMS)	P20GM125498	Dev Majumdar

Author ORCID iDs

Jacob S Dearborn: https://orcid.org/0000-0001-8894-1574
David Baltimore: https://orcid.org/0000-0001-8723-8190

References

1. **Chen J.,** Chen Z. J (2013) Regulation of NF-κB by ubiquitination. *Curr. Opin. Immunol* **25**:4-12 https://doi.org/10.1016/j.coi.2012.12.005 | PubMed
2. **Gautier E. L.,** Shay T., Miller J., Greter M., Jakubzick C., Ivanov S., Helft J., Chow A., Elpek K. G., Gordonov S., *et al.* (2012) Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nat. Immunol* **13**:1118-1128 https://doi.org/10.1038/ni.2419 | PubMed
3. **Smale S. T.,** Tarakhovsky A., Natoli G (2014) Chromatin contributions to the regulation of innate immunity. *Annu. Rev. Immunol* **32**:489-511 https://doi.org/10.1146/annurev-immunol-031210-101303 | PubMed
4. **Hao S.,** Baltimore D (2009) The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. *Nat. Immunol* **10**:281-288 https://doi.org/10.1038/ni.1699 | PubMed

5. **Leppek K.,** Schott J., Reitter S., Poetz F., Hammond M. C., Stoecklin G (2013) Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. *Cell* **153**:869-881 https://doi.org/10.1016/j.cell.2013.04.016 | PubMed

6. **O'Connell R. M.,** Rao D. S., Baltimore D (2012) microRNA regulation of inflammatory responses. *Annu. Rev. Immunol* **30**:295-312 https://doi.org/10.1146/annurev-immunol-020711-075013 | PubMed

7. **Wan F.,** Anderson D. E., Barnitz R. A., Snow A., Bidere N., Zheng L., Hegde V., Lam L. T., Staudt L. M., Levens D., *et al.* (2007) Ribosomal protein S3: a KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. *Cell* **131**:927-939 https://doi.org/10.1016/j.cell.2007.10.009 | PubMed

8. **Hao S.,** Baltimore D (2013) RNA splicing regulates the temporal order of TNF-induced gene expression. *Proc. Natl. Acad. Sci. U. S. A* **110**:11934-11939 https://doi.org/10.1073/pnas.1309990110 | PubMed

9. **Pandya-Jones A.,** Bhatt D. M., Lin C.-H., Tong A.-J., Smale S. T., Black D. L (2013) Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. *RNA* **19**:811-827 https://doi.org/10.1261/rna.039081.113 | PubMed

10. **Rabani M.,** Levin J. Z., Fan L., Adiconis X., Raychowdhury R., Garber M., Gnirke A., Nusbaum C., Hacohen N., Friedman N., *et al.* (2011) Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat. Biotechnol* **29**:436-442 https://doi.org/10.1038/nbt.1861 | PubMed

11. **Rabani M.,** Raychowdhury R., Jovanovic M., Rooney M., Stumpo D. J., Pauli A., Hacohen N., Schier A. F., Blackshear P. J., Friedman N., *et al.* (2014) High-resolution sequencing and modeling identifies distinct dynamic RNA regulatory strategies. *Cell* **159**:1698-1710 https://doi.org/10.1016/j.cell.2014.11.015 | PubMed

12. **Bergkessel M.,** Whitworth G. B., Guthrie C (2011) Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. *RNA* **17**:1461-1478 https://doi.org/10.1261/rna.2754011 | PubMed

13. **Guilgur L. G.,** Prudêncio P., Sobral D., Liszekova D., Rosa A., Martinho R. G (2014) Requirement for highly efficient pre-mRNA splicing during Drosophila early embryonic development. *eLife* **3**:e02181 https://doi.org/10.7554/eLife.02181 | PubMed

14. **Takashima Y.,** Ohtsuka T., González A., Miyachi H., Kageyama R (2011) Intronic delay is essential for oscillatory expression in the segmentation clock. *Proc. Natl. Acad. Sci. U. S. A* **108**:3300-3305 https://doi.org/10.1073/pnas.1014418108 | PubMed

15. **Eissa N. T.,** Strauss A. J., Haggerty C. M., Choo E. K., Chu S. C., Moss J (1996) Alternative splicing of human inducible nitric-oxide synthase mRNA. tissue-specific regulation and induction by cytokines. *J. Biol. Chem* **271**:27184-27187 https://doi.org/10.1074/jbc.271.43.27184 | PubMed

16. **Hargreaves D. C.,** Horng T., Medzhitov R (2009) Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* **138**:129-145 https://doi.org/10.1016/j.cell.2009.05.047 | PubMed

17. **Kontoyiannis D.,** Pasparakis M., Pizarro T. T., Cominelli F., Kollias G (1999) Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies. *Immunity* **10**:387-398 https://doi.org/10.1016/s1074-7613(00)80038-2 | PubMed

18. **Mahtani K. R.,** Brook M., Dean J. L., Sully G., Saklatvala J., Clark A. R (2001) Mitogen-activated protein kinase p38 controls the expression and posttranslational modification of tristetraprolin, a regulator of tumor necrosis factor alpha mRNA stability. *Mol. Cell. Biol* **21**:6461-6469 https://doi.org/10.1128/mcb.21.9.6461-6469.2001 | PubMed

19. **Mino T.,** Murakawa Y., Fukao A., Vandenbon A., Wessels H.-H., Ori D., Uehata T., Tartey S., Akira S., Suzuki Y., *et al.* (2015) Regnase-1 and Roquin Regulate a Common Element in Inflammatory mRNAs by Spatiotemporally Distinct Mechanisms. *Cell* **161**:1058-1073 https://doi.org/10.1016/j.cell.2015.04.029 | PubMed

20. **Rao N.,** Nguyen S., Ngo K., Fung-Leung W.-P (2005) A novel splice variant of interleukin-1 receptor (IL-1R)-associated kinase 1 plays a negative regulatory role in Toll/IL-1R-induced inflammatory signaling. *Mol. Cell. Biol* **25**:6521-6532 https://doi.org/10.1128/mcb.25.15.6521-6532.2005 | PubMed

21. **Ruggiero T.,** Trabucchi M., De Santa F., Zupo S., Harfe B. D., McManus M. T., Rosenfeld M. G., Briata P., Gherzi R. (2009) LPS induces KH-type splicing regulatory protein-dependent processing of microRNA-155 precursors in macrophages. *The FASEB Journal* **23**:2898-2908 https://doi.org/10.1096/fj.09-131342 | PubMed

22. **Stoecklin G.,** Lu M., Rattenbacher B., Moroni C (2003) A constitutive decay element promotes tumor necrosis factor alpha mRNA degradation via an AU-rich element-independent pathway. *Mol. Cell. Biol* **23**:3506-3515 https://doi.org/10.1128/mcb.23.10.3506-3515.2003 | PubMed

23. **Han J.,** Brown T., Beutler B (1990) Endotoxin-responsive sequences control cachectin/tumor necrosis factor biosynthesis at the translational level. *J. Exp. Med* **171**:465-475 https://doi.org/10.1084/jem.171.2.465 | PubMed

24. **Bhatt D. M.,** Pandya-Jones A., Tong A.-J., Barozzi I., Lissner M. M., Natoli G., Black D. L., Smale S. T (2012) Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150**:279-290 https://doi.org/10.1016/j.cell.2012.05.043 | PubMed

25. **Cho V.,** Mei Y., Sanny A., Chan S., Enders A., Bertram E. M., Tan A., Goodnow C. C., Andrews T. D (2014) The RNA-binding protein hnRNPLL induces a T cell alternative splicing program delineated by differential intron retention in polyadenylated RNA. *Genome Biol* **15**:R26 https://doi.org/10.1186/gb-2014-15-1-r26 | PubMed

26. **Davis-Turak J. C.,** Allison K., Shokhirev M. N., Ponomarenko P., Tsimring L. S., Glass C. K., Johnson T. L., Hoffmann A (2015) Considering the kinetics of mRNA synthesis in the analysis of the genome and epigenome reveals determinants of co-transcriptional splicing. *Nucleic Acids Res* **43**:699-707 https://doi.org/10.1093/nar/gku1338 | PubMed

27. **Grabherr M. G.,** Haas B. J., Yassour M., Levin J. Z., Thompson D. A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol* **29**:644-652 https://doi.org/10.1038/nbt.1883 | PubMed

28. **Braunschweig U.,** Barbosa-Morais N. L., Pan Q., Nachman E. N., Alipanahi B., Gonatopoulos-Pournatzis T., Frey B., Irimia M., Blencowe B. J (2014) Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res* **24**:1774-1786 https://doi.org/10.1101/gr.177790.114 | PubMed

29. **Zhou J.,** Troyanskaya O. G (2015) Predicting effects of noncoding variants with deep learning–based sequence model. *Nat. Methods* **12**:931-934 https://doi.org/10.1038/nmeth.3547 | PubMed

30. **Jaganathan K.,** Kyriazopoulou Panagiotopoulou S., McRae J. F., Darbandi S. F., Knowles D., Li Y. I., Kosmicki J. A., Arbelaez J., Cui W., Schwartz G. B., *et al.* (2019) Predicting Splicing from Primary Sequence with Deep Learning. *Cell* **176**:535-548. https://doi.org/10.1016/j.cell.2018.12.015 | PubMed

31. **Avsec Ž.,** Agarwal V., Visentin D., Ledsam J. R., Grabska-Barwinska A., Taylor K. R., Assael Y., Jumper J., Kohli P., Kelley D. R. (2021) Effective gene expression prediction from sequence by integrating long-range interactions. *Nat. Methods* **18**:1196-1203 https://doi.org/10.1038/s41592-021-01252-x | PubMed

32. **Linder J.,** Srivastava D., Yuan H., Agarwal V., Kelley D. R (2025) Predicting RNA-seq coverage from DNA sequence as a unifying model of gene regulation. *Nat. Genet* **57**:949-961 https://doi.org/10.1038/s41588-024-02053-6 | PubMed

33. **Tilgner H.,** Knowles D. G., Johnson R., Davis C. A., Chakrabortty S., Djebali S., Curado J., Snyder M., Gingeras T. R., Guigó R (2012) Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res* **22**:1616-1625 https://doi.org/10.1101/gr.134445.111 | PubMed

34. **Ramirez-Carrozzi V. R.,** Braas D., Bhatt D. M., Cheng C. S., Hong C., Doty K. R., Black J. C., Hoffmann A., Carey M., Smale S. T (2009) A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* **138**:114-128 https://doi.org/10.1016/j.cell.2009.04.020 | PubMed

35. **Engreitz J. M.,** Pandya-Jones A., McDonel P., Shishkin A., Sirokman K., Surka C., Kadri S., Xing J., Goren A., Lander E. S., *et al.* (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* **341**:1237973 https://doi.org/10.1126/science.1237973 | PubMed

36. **Bentley D. L** (2014) Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet* **15**:163-175 https://doi.org/10.1038/nrg3662 | PubMed

37. **Carrillo Oesterreich F.,** Preibisch S., Neugebauer K. M. (2010) Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol. Cell* **40**:571-581 https://doi.org/10.1016/j.molcel.2010.11.004 | PubMed

38. **Patel A. A.,** McCarthy M., Steitz J. A (2002) The splicing of U12-type introns can be a rate-limiting step in gene expression. *EMBO J* **21**:3804-3815 https://doi.org/10.1093/emboj/cdf297 | PubMed

39. **Honda K.,** Yanai H., Negishi H., Asagiri M., Sato M., Mizutani T., Shimada N., Ohba Y., Takaoka A., Yoshida N., *et al.* (2005) IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature* **434**:772-777 https://doi.org/10.1038/nature03464 | PubMed

40. **Lei X. F.,** Ohkawara Y., Stämpfli M. R., Mastruzzo C., Marr R. A., Snider D., Xing Z., Jordana M (1998) Disruption of antigen-induced inflammatory responses in CD40 ligand knockout mice. *J. Clin. Invest* **101**:1342-1353 https://doi.org/10.1172/jci1662 | PubMed

41. **Michaelson J. S.,** Bader D., Kuo F., Kozak C., Leder P (1999) Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development. *Genes Dev* **13**:1918-1923 https://doi.org/10.1101/gad.13.15.1918 | PubMed

42. **Koodathingal P.,** Novak T., Piccirilli J. A., Staley J. P (2010) The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5′ splice site cleavage during pre-mRNA splicing. *Mol. Cell* **39**:385-395 https://doi.org/10.1016/j.molcel.2010.07.014 | PubMed

43. **Pessa H. K. J.,** Ruokolainen A., Frilander M. J (2006) The abundance of the spliceosomal snRNPs is not limiting the splicing of U12-type introns. *RNA* **12**:1883-1892 https://doi.org/10.1261/rna.213906 | PubMed

44. **Freund M.,** Hicks M. J., Konermann C., Otte M., Hertel K. J., Schaal H (2005) Extended base pair complementarity between U1 snRNA and the 5′ splice site does not inhibit splicing in higher eukaryotes, but rather increases 5′ splice site recognition. *Nucleic Acids Res* **33**:5112-5119 https://doi.org/10.1093/nar/gki824 | PubMed

45. **Yeo G.,** Burge C. B (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol. J. Comput. Mol. Cell Biol* **11**:377-394 https://doi.org/10.1089/1066527041410418 | PubMed

46. **Mukherji S.,** Ebert M. S., Zheng G. X. Y., Tsang J. S., Sharp P. A., van Oudenaarden A. (2011) MicroRNAs can generate thresholds in target gene expression. *Nat. Genet* **43**:854-859 https://doi.org/10.1038/ng.905 | PubMed

47. **Yuan H.,** Linder J., Kelley D. R. (2025) Parameter-Efficient Fine-Tuning of a Supervised Regulatory Sequence Model. *bioRxiv* 2025.05.26.656171 https://doi.org/10.1101/2025.05.26.656171

48. **Shrikumar A.,** Greenside P., Kundaje A. (2019) Learning Important Features Through Propagating Activation Differences. *arXiv* https://doi.org/10.48550/arXiv.1704.02685

49. **Shrikumar A.,** Tian K., Avsec Ž., Shcherbina A., Banerjee A., Sharmin M., Nair S., Kundaje A. (2020) Technical Note on Transcription Factor Motif Discovery from Importance Scores (TF-MoDISco) version 0.5.6.5. *arXiv* https://doi.org/10.48550/arXiv.1811.00416

50. **Grant C. E.,** Bailey T. L., Noble W. S (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**:1017-1018 https://doi.org/10.1093/bioinformatics/btr064 | PubMed

51. **Ameur A.,** Zaghlool A., Halvardson J., Wetterbom A., Gyllensten U., Cavelier L., Feuk L (2011) Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat. Struct. Mol. Biol* **18**:1435-1440 https://doi.org/10.1038/nsmb.2143 | PubMed

52. **Brugiolo M.,** Herzel L., Neugebauer K. M (2013) Counting on co-transcriptional splicing. *F1000prime Rep* **5**:9 https://doi.org/10.12703/p5-9 | PubMed

53. **Girard C.,** Will C. L., Peng J., Makarov E. M., Kastner B., Lemm I., Urlaub H., Hartmuth K., Lührmann R (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat. Commun* **3**:994 https://doi.org/10.1038/ncomms1998 | PubMed

54. **Khodor Y. L.,** Rodriguez J., Abruzzi K. C., Tang C.-H. A., Marr M. T., Rosbash M (2011) Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. *Genes Dev* **25**:2502-2512 https://doi.org/10.1101/gad.178962.111 | PubMed

55. **Martin R. M.,** Rino J., Carvalho C., Kirchhausen T., Carmo-Fonseca M (2013) Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity. *Cell Rep* **4**:1144-1155 https://doi.org/10.1016/j.celrep.2013.08.013 | PubMed

56. **Koh C. M.,** Bezzi M., Low D. H. P., Ang W. X., Teo S. X., Gay F. P. H., Al-Haddawi M., Tan S. Y., Osato M., Sabò A., *et al.* (2015) MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. *Nature* **523**:96-100 https://doi.org/10.1038/nature14351 | PubMed

57. **Brinster R. L.,** Allen J. M., Behringer R. R., Gelinas R. E., Palmiter R. D (1988) Introns increase transcriptional efficiency in transgenic mice. *Proc. Natl. Acad. Sci. U. S. A* **85**:836-840 https://doi.org/10.1073/pnas.85.3.836 | PubMed

58. **Damgaard C. K.,** Kahns S., Lykke-Andersen S., Nielsen A. L., Jensen T. H., Kjems J (2008) A 5′ splice site enhances the recruitment of basal transcription initiation factors in vivo. *Mol. Cell* **29**:271-278 https://doi.org/10.1016/j.molcel.2007.11.035 | PubMed

59. **Furger A.,** O'Sullivan J. M., Binnie A., Lee B. A., Proudfoot N. J (2002) Promoter proximal splice sites enhance transcription. *Genes Dev* **16**:2792-2799 https://doi.org/10.1101/gad.983602 | PubMed

60. **Juneau K.,** Miranda M., Hillenmeyer M. E., Nislow C., Davis R. W (2006) Introns regulate RNA and protein abundance in yeast. *Genetics* **174**:511-518 https://doi.org/10.1534/genetics.106.058560 | PubMed

61. **Kornblihtt A. R.,** de la Mata M., Fededa J. P., Munoz M. J., Nogues G. (2004) Multiple links between transcription and splicing. *RNA* **10**:1489-1498 https://doi.org/10.1261/rna.7100104 | PubMed

62. **Parenteau J.,** Durand M., Véronneau S., Lacombe A.-A., Morin G., Guérin V., Cecez B., Gervais-Bird J., Koh C.-S., Brunelle D., *et al.* (2008) Deletion of many yeast introns reveals a minority of genes that require splicing for function. *Mol. Biol. Cell* **19**:1932-1941 https://doi.org/10.1091/mbc.e07-12-1254 | PubMed

63. **Shabalina S. A.,** Ogurtsov A. Y., Spiridonov A. N., Novichkov P. S., Spiridonov N. A., Koonin E. V (2010) Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Mol. Biol. Evol* **27**:1745-1749 https://doi.org/10.1093/molbev/msq086 | PubMed

Peer reviews

Reviewer #1 (Public review):

Summary:

In this work, the authors revisit a well-defined experimental system for studying temporal gene expression mechanisms in TNF-alpha-stimulated macrophages, bringing new tools to the process. Using a hybrid-capture approach, they are able to obtain deeper RNA sequencing of target genes, which allows them to identify potential differences in splicing kinetics of individual introns. Further implementing transcriptional blocks to measure intron half-lives, and predictive machine learning models to identify potential contributing cis-acting RNA elements, they define a group of 'bottleneck' introns whose delayed splicing is a rate-limiting step in mRNA maturation.

Strengths:

(1) The hybrid-capture approach enables deeper RNA sequencing of target transcripts.

(2) The neural network application to identify motifs outside of splice sites could be related to intron removal kinetics.

(3) The paper uses splicing reporters with modulation of 5' splice sites to test the effect on reporter gene expression in the context of 'bottleneck' introns.

Weaknesses:

(1) While evidence is provided that these introns are distinct from previously published splicing kinetics studies, 'bottleneck' introns are not adequately placed in context for assessment of how they are similar or different.

(2) Splicing reporters are a good approach, but the complexities of post-transcriptional gene expression regulation are not adequately addressed

(3) Deep learning models are a potentially powerful tool for identifying novel regulatory sequences; however, their use here is underdeveloped.

https://doi.org/10.7554/eLife.109726.1.sa2

Reviewer #2 (Public review):

Summary:

The authors analyzed the temporal dynamics of gene expression patterns within the inflammatory response transcriptome following TNF stimulation, and proposed that the splicing rate of certain introns is a key mechanism of regulating mature mRNA expression rate.

Strengths:

The measurement strategy is generally well-designed to understand the core question of splicing rate and gene expression. The following computation analysis, as well as the mutation or repair studies, further supported the claims. The writing and presentation of the results are also generally clear and easy to follow. I think this manuscript will be of interest to a wide audience.

Weaknesses:

I do have some questions regarding some of the results and conclusions, and I think either more analysis or more explanation and discussion can make the claims more solid. Please see below for details:

(1) On the hybrid capture method and the RNA coverage results: The strategy of enriching for the last exon before sequencing does have significance in linking pre-mRNA and mature mRNA. If I understand correctly, this enriches for pre-mRNA molecules that are about to finish the full-length elongation of RNA polymerase. However, is this strategy biased towards measuring the splicing rate variation on introns closer to the 3-prime end? For example, if a gene takes 5 minutes for the RNA polymerase to elongate through the full length of the gene, for intron #1 that's very close to the 5' end, you can't tell if it takes 20s to be spliced out or 4 minutes, as both will show as fully spliced out in the sequencing library. In other words, for introns near the 5' end, a consistent "CoSI=1" pattern in the data doesn't necessarily suggest a true consistent fast splicing of that intron. Do you observe any general pattern of the measured "slowliness" in relation to the 5'-3' location of the introns? If so, should the 5' introns be specially considered or even excluded from certain analyses that use all introns?

(2) Following on my last point, it may benefit the readers if the author can provide a more detailed comparison of possible sequencing library construction choices. For example, is it

feasible to also enrich for other exons for the sequencing library, etc?

(3) Figure 1C: Are there biological replicates, and should there be error bars and statistics on the plot? Similarly, in places like Figure 2, Supplemental Figure 4C, Supplemental Figure 6, etc., is there any statistical analysis that can be done to show if the claimed differences are statistically significant?

(4) The logic behind measuring the half-lives of introns seems a little unclear to me. From the time-dependent RNA coverage plots in Figure 2, it seems that, if we assume a constant transcription elongation rate, then the splicing rate of a specific intron can vary across time after TNF stimulation, as represented by the temporal change of CoSI values, or the heights of the coverage plot relative to neighboring exons. This means the splicing rate or half-life of an intron is not necessarily constant but may be time-dependent, at least in the case of TNF stimulation. Shouldn't the half-life measurements be designed in a way to measure the half-life at multiple time points after TNF stimulation? And maybe the measured half-lives of some introns will show as time-dependent?

(5) In Supplemental Figure 6, the interpretation is a little confusing to me: If delayed splicing is causing delayed expression of the corresponding gene, shouldn't the non-immediate gene groups (early/intermediate/Late) have low CoSI beginning from the early time points (e.g. 4 minutes)? Why does the slowdown of splicing seem to peak at a later time point? Does it mean immediately after TNF stimulation, there's a different mechanism in delaying the expression of the non-immediate gene groups? Maybe it's better to have more explanation or use a different visualization to show what non-immediate gene groups are experiencing at very early time points.

(6) On the fine-tuning of the deep sequence model: it's a little unclear whether the input and output are time-dependent. It's stated that expression at multiple time points is used for training, but it's unclear whether the model outputs time-dependent expression patterns and whether the time information is used as input.

https://doi.org/10.7554/eLife.109726.1.sa1

Reviewer #3 (Public review):

Summary:

The manuscript by Dearborn et al investigates the kinetics of intron splicing in inflammation-associated transcripts after TNF-stimulation of macrophages, using targeted sequencing of chromatin-associated RNA to obtain high coverage across a focused set of induced genes. The authors' main conclusion is that splicing kinetics are heterogeneous across these transcripts, and that delayed introns (which they term "bottleneck introns") are associated with weak donor sequences. Using a deep learning approach, they have also identified additional sequence features that might contribute to intron splicing kinetics.

Overall, I think the findings in the manuscript are very intriguing and will be of interest to readers working on RNA biology. The changes the authors have made to the manuscript in response to some very valid comments from reviewers have strengthened the manuscript. While the existing data might not be sufficient to directly address some of the broader mechanistic claims made by the authors, I think the findings are nonetheless very interesting and should contribute towards a better understanding of the post-transcriptional regulation of gene expression.

Strengths:

A strength of the manuscript is the experimental design. The targeted capture approach is innovative and well-suited to the goal of measuring intron-specific splicing behaviour across

time. The inclusion of experimental validation in minigene assays of some of the computational predictions also strengthens the claims made by the authors.

The authors have made a constructive effort to address some of the concerns raised in a previous round of review. The revised manuscript reads as a balanced text.

Weaknesses:

The study still does not fully resolve the downstream consequences of delayed splicing. In particular, it remains unclear whether the bottleneck introns lead primarily to delayed production of mature transcripts, reduced productive transcript output, or some combination of the two.

On a related point, the minigene reporter assays measure a steady-state level of the transcript and don't provide insights into the kinetics directly.

Lastly, given that the detailed analyses were performed on a selected subset of (inflammation-induced) transcripts, a broader evolutionary interpretation needs to be restrained given the current data.

https://doi.org/10.7554/eLife.109726.1.sa0

Author response:

We thank the Reviewing Editor and reviewers for their thoughtful and constructive evaluation of our manuscript, *Programmed Delayed Splicing: A Mechanism for Timed Inflammatory Gene Expression*. We are encouraged that the reviewers found the study valuable, the experimental design strong for the core findings. We appreciate the reviewers' careful attention to the limits of inference in several parts of the manuscript, and will address these points in a revised version. We especially want to acknowledge that this paper has benefited from the abiding interest in splicing regulation by the editors and reviewers who have meticulously improved nearly every aspect of this multifaceted work in its present state.

Our planned revisions will focus on five areas. **First**, we will more carefully evaluate and discuss the extent to which the hybrid-capture strategy may impose position-dependent constraints on apparent splicing behavior, particularly across 5' and 3' introns. **Second**, we will clarify the use of the term "bottleneck introns," distinguishing descriptive use in the main text from the ranked subsets used in downstream analyses. **Third**, we will revise the framing of the reporter assays to make explicit that these measure steady-state reporter output and do not, on their own, resolve all downstream kinetic consequences of delayed splicing. **Fourth**, we will clarify the interpretation of the actinomycin D experiments as providing estimates of intron excision behavior under transcriptional arrest rather than a complete time-resolved model of splicing during TNF induction. **Fifth**, we will substantially revise the scope and stated limitations of the deep learning-aided interpretations of data in this work.

Reviewer #1

We thank Reviewer #1 for the positive assessment of the hybrid-capture strategy, the splice-site reporter experiments, and the potential value of the neural-network-based analysis. We appreciate the reviewer's view that these approaches help extend a well-established system for studying temporal gene expression in TNF-stimulated macrophages. We address the main concerns raised in the public review below.

(1) While evidence is provided that these introns are distinct from previously published splicing kinetics studies, "bottleneck" introns are not adequately placed in context for

assessment of how they are similar or different.

We appreciate this point and agree that the current manuscript does not yet place these introns in sufficiently clear context relative to prior literature. Our study builds on foundational work describing regulated changes in splicing kinetics, widespread intron retention, and detained introns as biologically meaningful modes of gene regulation, including transcript-specific regulation of splicing in response to stress (Pleiss, *Mol Cell.*, 2007), widespread functional intron retention in mammals (Braunschweig, *Genome Res.*, 2014), and the definition of detained introns as a distinct class of post-transcriptionally spliced introns (Boutz, *Genes Dev.*, 2015). In revision, we will expand the comparison to previously described classes of delayed or retained introns and clarify more explicitly how the introns studied here are defined in the setting of inducible inflammatory transcripts and their temporal resolution over the course of stimulation. We will also revise the relevant Results and Discussion text so that the distinction is made directly in the manuscript rather than relying on inference from the broader presentation.

(2) Splicing reporters are a good approach, but the complexities of post-transcriptional gene expression regulation are not adequately addressed.

We agree that the interpretive limits of the reporter assays should be stated more clearly and consistently. In revision, we will revise the presentation of the minigene experiments to make explicit that these are steady-state reporter assays and therefore do not, on their own, resolve all downstream kinetic consequences of delayed splicing in the endogenous context. At the same time, we believe the assay remains informative because it provides a controlled system in which the contribution of splice donor sequence can be tested directly in matched reporter constructs. In that sense, the reporter experiments are valuable as a reductionist test of whether weak donor sequences are sufficient to alter reporter output, even if they do not fully recapitulate the broader endogenous post-transcriptional environment. We will emphasize that these data support an association between weak donor sites and altered reporter output, while moderating any broader mechanistic claims that extend beyond what the assay directly measures.

(3) Deep learning models are a potentially powerful tool for identifying novel regulatory sequences; however, their use here is underdeveloped.

We appreciate this concern and agree that the deep-learning section should be revised substantially. In a revised manuscript, we will clarify the training setup, the definition of the slow-intron subsets used in downstream analyses, and the interpretation of the attribution and motif analyses. Alongside, we believe the assay remains informative because it provides a controlled system in which the contribution of splice donor sequence can be tested directly in matched reporter constructs. In that respect, the reporter experiments are valuable as a reductionist test of whether weak donor sequences are sufficient to alter reporter output, even if they do not fully recapitulate the broader endogenous post-transcriptional environment. We will revise the framing of these results so that they are presented more explicitly as identifying candidate sequence features associated with delayed splicing, rather than as direct evidence of specific causal regulatory mechanisms.

Reviewer #2

We thank Reviewer #2 for the thoughtful and detailed comments, and for recognizing the strengths of the measurement strategy and the clarity of the manuscript. We appreciate the reviewer's view that the study will be of interest to a broad audience, and we agree that several conclusions will be strengthened by additional analysis and clearer explanation. We address the main concerns raised in the public review below.

(1) Concern regarding possible bias of the hybrid-capture strategy toward introns closer to the 3' end, and whether 5' introns should be treated separately in some analyses.

We thank the reviewer for this careful and important point. We agree that this is a potential limitation of the approach and that it should be addressed more explicitly in the manuscript. Our assay begins with poly(A)-selected RNA and then enriches transcripts of interest through terminal-exon capture, so the molecules analyzed are completed, polyadenylated transcripts rather than nascent partial transcripts. This feature is important for reducing ambiguity arising from incomplete transcription, particularly in the chromatin-associated fraction. At the same time, we agree that for introns near the 5' end, the assay may have limited power to distinguish very rapid splicing from moderately rapid splicing if excision is largely complete by the time the transcript is fully synthesized and polyadenylated.

In revision, we will address this concern directly in two ways. First, we will revise the Results and Discussion to clarify that the assay provides a population-level measure of splice completion in completed transcripts and that interpretation is strongest for introns whose excision is not already fully resolved before transcript completion. Second, we will more systematically evaluate whether apparent slow splicing covaries with transcript position, distance from the 3' end, and intron length, and we will perform sensitivity analyses with and without the most 5' introns to determine which conclusions are robust to these positional constraints. We will also examine transcript coverage patterns in greater detail to better assess the extent to which library construction and cDNA generation may contribute to apparent positional bias. Our preliminary inspection suggests that transcript position is not the sole determinant of the observed heterogeneity, but we agree that a more explicit treatment of this issue is warranted in the revised manuscript.

(2) Request for more detailed discussion of alternative library-construction choices.

We appreciate this suggestion and agree that the revised manuscript would benefit from a fuller discussion of the strengths and limitations of the current enrichment strategy. We chose poly(A) selection followed by terminal-exon capture because this design enriches completed transcripts of interest and reduces ambiguity from nascent partial transcripts, which is particularly important in the chromatin-associated fraction. This approach also provides greater read depth over the selected inflammatory transcripts, enabling more informative intron-level comparisons within the targeted dataset. In revision, we will clarify this rationale more explicitly in the manuscript. We will also discuss the tradeoffs of this design relative to alternative exon-targeting strategies and how those alternatives might provide different, but complementary, views of splicing kinetics.

(3) Questions regarding biological replicates, error bars, and statistical analysis in Figure 1C and other plots.

We agree that the replicate structure and intended interpretation of these plots should be clarified more explicitly. In revision, we will revise the figure legends and Methods to distinguish panels that display a single bulk RNA-seq time course (for example, Figure 1C) from panels that summarize distributions across many introns (for example, Figure 2 and Supplementary Figure 6). We will also add statistical comparisons where they are most appropriate and informative, such as in sequence-feature comparisons like Supplementary Figure 4C, while making clear that some CoSI panels are intended as descriptive summaries of intron-level heterogeneity rather than replicate-based inferential plots.

(4) Concern that intron half-lives may be time-dependent during TNF induction, and that the logic of the actinomycin D measurements is therefore unclear.

We appreciate this point and agree that the manuscript should distinguish more clearly between two related but non-identical quantities: the CoSI trajectories observed during

ongoing TNF induction, and the interruption-based half-life estimates derived from actinomycin D treatment. The actinomycin D experiments were performed using multiple post-treatment timepoints, but they were designed to estimate intron excision behavior after transcriptional arrest under a defined set of conditions, rather than to measure whether an individual intron's effective splicing rate changes across all phases of the TNF response. We agree that these estimates should therefore be interpreted as constrained measurements under the assay conditions used, rather than as a complete time-resolved model of splicing kinetics during induction. In revision, we will clarify this point in the Results, Methods, and Discussion, and we will more explicitly acknowledge that effective splicing behavior could vary across the induction time course.

(5) Concern that the interpretation of Supplementary Figure 6 is unclear, particularly why delayed splicing in non-immediate groups appears to peak later rather than at the earliest time points.

We appreciate this point and agree that the current presentation of Supplementary Figure 6 does not explain this behavior clearly enough. Our interpretation is not that delayed splicing is the sole determinant of early versus later induction classes. Rather, the earliest time points reflect a combination of transcriptional induction timing and RNA processing state. In this framework, the dip in CoSI shortly after stimulation reflects the appearance of newly induced, incompletely spliced transcripts, and the later kinetic groups appear to recover from this dip more slowly than the immediate-early group. Thus, the strongest signal of delayed splicing may become most apparent only after sufficient transcript accumulation, rather than necessarily at the very earliest time point. In revision, we will revise the text to make this logic clearer and will consider a more intuitive visualization of these group-specific CoSI trajectories.

(6) Concern that the deep-learning setup does not make clear whether the model input and output are time-dependent.

We appreciate this concern and agree that the current manuscript does not explain the model setup clearly enough. Briefly, we will clarify the role of the three TNF timepoints in model training, including the fact that these outputs were modeled jointly and that time itself was not provided as an explicit input to the model. We will also revise the Results and Methods so that the scope and interpretation of the resulting analyses are more explicit.

Reviewer #3

We thank Reviewer #3 for the positive assessment of the targeted capture design, the evaluation of overall interest of the findings, and the improvements in the current version. We appreciate the reviewer's view that the study is intriguing and that the manuscript has been strengthened in revision. We agree, however, that the manuscript should more clearly distinguish what is directly demonstrated from what remains mechanistically unresolved. We address the main concerns raised in the public review below.

(1) The study still does not fully resolve the downstream consequences of delayed splicing, including whether bottleneck introns lead primarily to delayed production of mature transcripts, reduced productive transcript output, or some combination of the two.

We agree with this assessment. The current data do not fully resolve whether delayed splicing primarily delays mature transcript production, reduces productive transcript output, or reflects some combination of the two. In revision, we will further moderate the framing of the downstream consequences of delayed splicing and will revise the Abstract, Results, and Discussion to make clear that the present data do not fully distinguish among delayed mature transcript production, reduced productive transcript output, or a combination of both. We

will ensure that the manuscript consistently presents these possibilities as alternatives not fully resolved by the current data.

(2) The minigene reporter assays measure a steady-state level of the transcript and do not provide direct insight into kinetics.

We agree and will revise the manuscript to make this limitation explicit throughout. In particular, we will ensure that the reporter assays are described consistently as steady-state reporter assays that support an association between splice donor strength and altered reporter output, while avoiding stronger claims that they directly resolve endogenous splicing kinetics or downstream transcript fate.

(3) Given that the detailed analyses were performed on a selected subset of inflammation-induced transcripts, a broader evolutionary interpretation should be restrained.

We agree that the broader evolutionary and mechanistic framing should be more carefully defined. In revision, we will restrain these interpretations so that they remain closely aligned with the inflammation-focused and targeted-transcript scope of the current study, and we will moderate language that extends beyond what is directly supported by the present dataset.

Closing Remarks

We again thank the reviewers for their constructive comments. We believe that the planned revisions will strengthen the manuscript by clarifying the scope of the mechanistic conclusions, sharpening the interpretation of the experimental approaches, and more carefully defining the role of the computational analyses. We appreciate the opportunity to revise the work and to provide this provisional response to accompany the Reviewed Preprint.

https://doi.org/10.7554/eLife.109726.1.sa4