Published online 1 July 2016

Nucleic Acids Research, 2016, Vol. 44, No. 14 ***6817–6829***
doi: 10.1093/nar/gkw591

Global intron retention mediated gene regulation during CD4+ T cell activation

Ting Ni[1,*,†], Wenjing Yang[2,†], Miao Han[1,†], Yubo Zhang[2], Ting Shen[1], Hongbo Nie[1], Zhihui Zhou[3], Yalei Dai[3], Yanqin Yang[2], Poching Liu[2], Kairong Cui[2], Zhouhao Zeng[4], Yi Tian[4,5], Bin Zhou[1], Gang Wei[1], Keji Zhao[2], Weiqun Peng[4,*] and Jun Zhu[2,*]

[1]State Key Laboratory of Genetic Engineering & MOE Key Laboratory of Contemporary Anthropology, Collaborative Innovation Center of Genetics and Development, School of Life Sciences, Fudan University, Shanghai 200438, P.R. China, [2]Systems Biology Center, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA, [3]Department of Immunology, Tongji University School of Medicine, Shanghai 200092, P.R. China, [4]Department of Physics, George Washington University, Washington, DC 20052, USA and [5]Institute of Immunology, PLA, Third Military Medical University, Chongqing 400038, P.R. China

Received November 05, 2015; Revised June 17, 2016; Accepted June 17, 2016

ABSTRACT

T cell activation is a well-established model for studying cellular responses to exogenous stimulation. Using strand-specific RNA-seq, we observed that intron retention is prevalent in polyadenylated transcripts in resting CD4+ T cells and is significantly reduced upon T cell activation. Several lines of evidence suggest that intron-retained transcripts are less stable than fully spliced transcripts. Strikingly, the decrease in intron retention (IR) levels correlate with the increase in steady-state mRNA levels. Further, the majority of the genes upregulated in activated T cells are accompanied by a significant reduction in IR. Of these 1583 genes, 185 genes are predominantly regulated at the IR level, and highly enriched in the proteasome pathway, which is essential for proper T cell proliferation and cytokine release. These observations were corroborated in both human and mouse CD4+ T cells. Our study revealed a novel post-transcriptional regulatory mechanism that may potentially contribute to coordinated and/or quick cellular responses to extracellular stimuli such as an acute infection.

INTRODUCTION

T cell activation is an essential step in immune response (1,2). Upon receiving appropriate signals, resting T cells transform from a relatively quiescent state to an active proliferation state, producing cytokines (2). The activation process involves a coordinated program of gene expression regulation. Transcriptional regulation has been known to play an important role in the activation process. Activator protein 1 (AP-1), nuclear factor of activated T-cells (NFAT) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) are among the essential transcriptional regulators responsible for the induction of a battery of genes including key cytokines such as interlukin-2 (IL2) (3,4).

The role of chromatin in T cell biology especially the activation process has attracted considerable interests. In fact, T cell was one of the first model systems that was used to depict a comprehensive picture of the epigenome had been depicted using ChIP-seq (chromatin immunoprecipitation followed by deep sequencing) (5–7). These studies along with studies in other systems have established epigenetic landscape as a powerful tool for annotating functional elements involved in the regulation and execution of transcriptional events (7,8). For example, expressed genes are enriched with trimethylation of histone H3 at lysine 4 (H3K4me3), histone acetylations and RNA polymerase II (Pol II) at the promoter; and trimethylation of histone H3 at lysine 36 (H3K36me3) in the transcribed regions. Beyond simple absence and presence, the levels of many histone modifications correlate well with their gene expression levels (9). As a result, it is expected that up-regulated (down-regulated) genes on average would exhibit a concomitant increase (decrease) of active histone modification marks that has indeed been observed in many systems (6,9,10).

Interestingly, recent genome-wide studies (11,12) in T cell activation had led to a number of unexpected observations. First of all, many inducible genes that are silent in resting cells, have already been primed with higher levels of active

*To whom correspondence should be addressed. Tel: +86 21 51630627; Fax: +86 21 51630627; Email: tingni@fudan.edu.cn
Correspondence may also be addressed to Jun Zhu. Tel: +1 301 443 7927; Fax: +1 301 443 7927; Email: jun.zhu@nih.gov
Correspondence may also be addressed to Weiqun Peng. Tel: +1 202 994 0129; Fax: +1 202 994 3001; Email: wpeng@gwu.edu
†These authors contributed equally to this work as the first authors.

chromatin marks (e.g. H3K4me3, H2A.Z and H3K9ac), as well as higher levels of Pol II at the promoter. Secondly, following T cell activation, little or no change in chromatin states was detected upon gene induction (11,12). These observations have led to the hypothesis that those genes are 'poised' for rapid transcription upon stimulation. However, the discordance between histone modification levels and gene expression levels, as well as the lack of dynamical changes in the histone modification levels upon gene induction, is not limited to silent genes. Similar observations have also been made in a number of rapidly induced primary response genes that have low basal expressions in resting cells (13). These results suggest that a transcription-independent regulatory mechanism might play a significant role during T cell activation.

IR is one of the major forms of alternative splicing in eukaryotes (14–17). Recent studies show that IR or 'detained' introns could be found in diverse human and mouse tissues, and progressively accumulate during development and terminal erythropoiesis (18–20). The functional involvement of IR in gene regulation has drawn attention only until recently (19,21–24). In yeast, IR has been reported to coordinate the expression of ribosomal proteins and genes related to meiotic processes (25,26). In higher eukaryotes, gene regulation mediated by intron retention has been observed during neurogenesis (23), granulocyte differentiation (21) and terminal erythropoiesis (20). However, it is unknown whether IR is regulated during cell stimulation such as T cell activation.

In this study, we found that IR is prevalent in resting $CD4^+$ T cells and dramatically decrease upon cell activation. We provide evidence that transcripts with retained introns are less stable, implying the potential coupling between IR and RNA degradation in dialing up/down gene expression. The reduction of IR has a striking correlation with increased expression of their corresponding genes upon T cell activation. Furthermore, we identified a subset of genes whose expressions are mainly regulated via IR. These genes are highly enriched in proteasome genes that are known to be involved in T cell proliferation and cytokine release. Regulated IR was also observed in primary mouse $CD4^+$ T cells upon activation. Taken together, we propose that IR, in combination with RNA degradation, serves as a novel post-transcriptional mechanism broadly involved in regulating gene expression during $CD4^+$ T cell activation.

MATERIALS AND METHODS

$CD4^+$ T cell isolation and activation

Human $CD4^+$ T cells were isolated from peripheral blood monocytes using the Dynabeads Untouched Human CD4 T Cells kit (Invitrogen) as previously described (5). $CD4^+$ T cell activation was performed with the Dynabeads Human T-Activator CD3/CD28 for T-Cell Expansion and Activation kit (Invitrogen) by incubating them at 37°C for 18 hours (h). The cell morphology was observed under the microscope, and CD25, a cell surface marker, was used to confirm T cell activation by fluorescence activated cell sorting (FACS). Mouse $CD4^+$ T cells were isolated from the lymph nodes, and activated with the CD3/CD28 antibody for a period of 8, 12 or 24 h. Total RNA was isolated using TRIzol reagent (Invitrogen). The resulting RNA was further purified with the RNeasy mini kit (Qiagen) and an on-column DNase I treatment to eliminate DNA contamination.

Library construction

The human $CD4^+$ T cell RNA-seq libraries were constructed according to the DeLi-seq protocol, a strand-specific RNA-seq procedure (27). All RNA samples were subjected to two rounds of polyA+ RNA selection with oligo$(dT)_{25}$ Dynabeads (Invitrogen) before library construction. ChIP-seq libraries were constructed as previously described (5,7). ChIP grade antibodies for Pol II and H3K36me3 were purchased from Abcam with cat. no. ab26721 and ab9050, respectively. All libraries were sequenced using an Illumina HiSeq 2000 sequencer, and the data sets used in this study are summarized in Supplementary Table S6.

Sequence alignment

For the human $CD4^+$ T cell RNA-seq and ChIP-seq libraries, we performed single-end mapping using BWA (28) with default parameters against human genome hg19. For the mouse $CD4^+$ T cell RNA-seq libraries, we performed paired-end mapping using BWA with default parameters against mouse genome mm9. Only uniquely mapped reads were used for downstream analyses.

Compilation of gene model

We utilized an annotation system where each gene is represented by its Entrez ID and the associated isoforms are represented by their RefSeq IDs. To reliably calculate the IR levels, it is critical to minimize potential biases due to alternative designation of introns (or exons) in multiple isoforms. To this end, transcript isoforms were consolidated and only the shared intronic (exonic) regions in all annotated isoforms are designated as intronic (exonic) regions (Supplementary Figure S3). In addition, Entrez genes were removed from the analysis if they had one of the following issues: (i) contained one or more NR transcripts (non-coding transcripts in the RefSeq annotation system); or (ii) contained two or more transcripts that do not overlap or on different strands; or (iii) overlapped with another gene on the same strand; or (iv) had no intron.

Calculation of intron retention index

The intron retention index (IRI) is devised to quantify the extent of IR based on the gene model defined above. Specifically, the IRI of a gene is defined as the ratio of read density of shared intronic regions and that of shared exonic regions. Since the RefSeq and Entrez annotation system is unlikely to include all the genes and isoforms, we removed genes whose IRI > 1 in either resting or activated states from further analysis to reduce the negative impact due to missing annotations. In genes with low expression levels, the IRI values are more susceptible to statistical errors. Therefore, in downstream analysis that involves genes characterized by IRI, only genes expressed in resting state (RPKM > 1) were

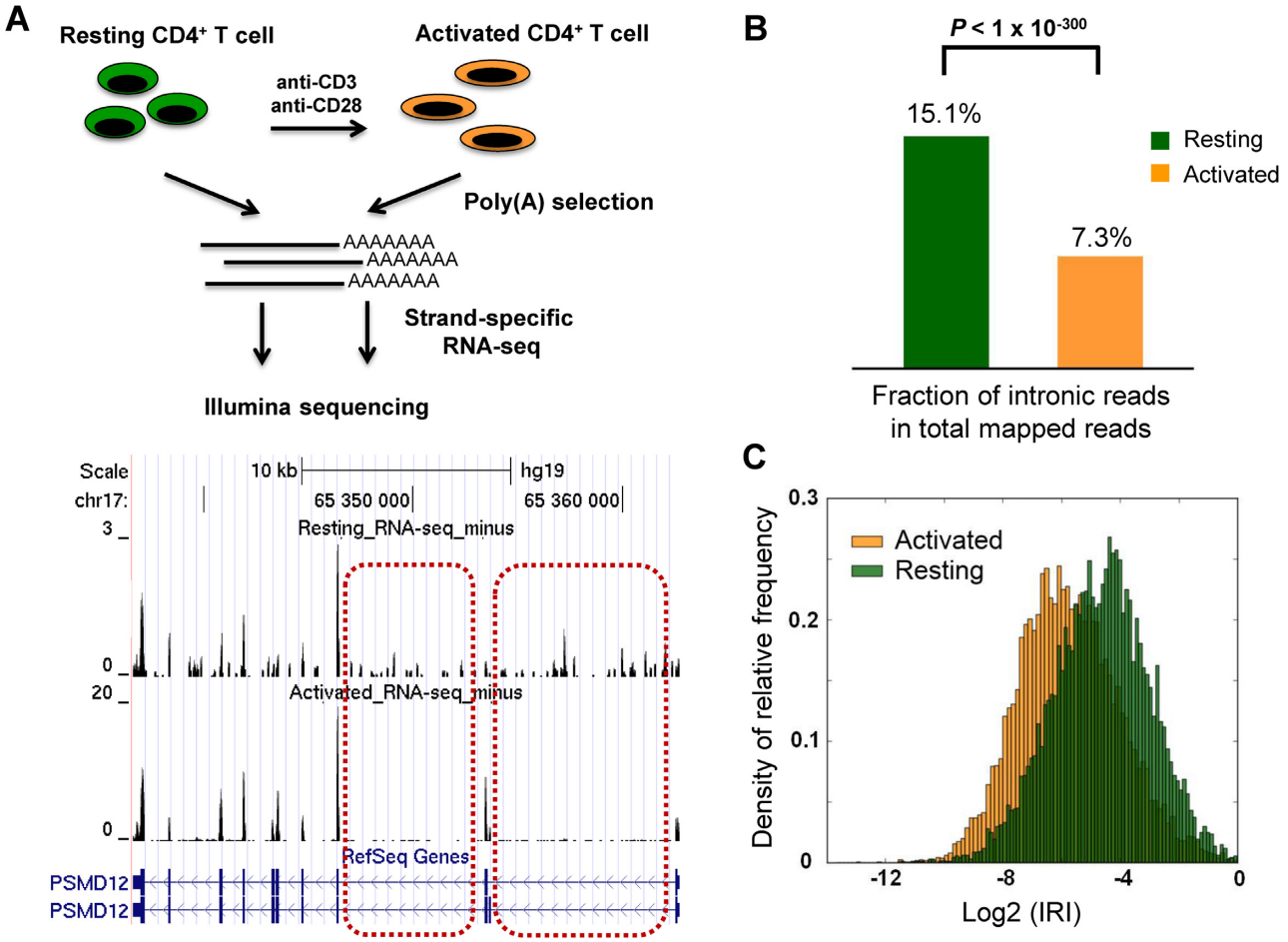

Figure 1. Intron retention (IR) is prevalent in resting human CD4$^+$ T cells and dramatically reduced upon activation. (**A**) The transcriptomes of human resting and activated CD4$^+$ T cells were profiled by RNA-seq. Oligo(dT) was used to enrich polyA+ RNAs, which were sequenced using DeLi-seq, a strand-specific RNA-seq procedure. The normalized expression profiles of a candidate gene, *PSMD12*, are shown in the bottom panel. Two representative introns with differential IR between resting and activated T cells are shown in dashed rectangle. (**B**) The fraction of intronic reads in total mapped reads are shown for resting and activated T cells. Compared to resting T cells, activated T cells show a significantly reduced fraction of intronic reads (*P*-value $< 1 \times 10^{-300}$, X^2 test). (**C**) The frequency density distribution of IR Index (IRI) of individual genes in resting and activated human T cells. The integral of the area under the vertical bars is 1 for both resting and activated T cells. For each gene locus, IRI is defined as the ratio of normalized tag density between shared introns and shared exons. The mean and median IRI for the resting condition is 7.1% and 4.0%, respectively, whereas the mean and median IRI for the activated condition is 3.8% and 1.8%, respectively.

used. Furthermore, in the IRI histograms shown in Figure 1C and Supplementary Figure S5B, only genes expressed (shared exonic RPKM ≥ 1) in resting state and those with an IRI > 0 in both states are included.

Calculation of IR percentage

An alternative strategy to characterize IR is to use reads covering splice junctions. We define a constitutive junction as an annotated splice junction that connects a shared intronic region and a shared exonic region. To remove ambiguity, we only examine annotated constitutive junctions. We consider a read to be across-junction if it covers the junction point and overlaps both sides for at least 8 bp. We consider a read to be spliced if its coverage jumps over the junction point and overlaps the shared exonic region for at least 8 bp. Genome-wide, we calculate the intron retention percentage (IRP) as the fraction of all junction reads (i.e. across-junction + spliced) that are across-junction. For each gene, intron retention ratio is defined in a similar manner, with the constitutive junctions limited to those belonging to this gene. For other reads, only those overlapping with upstream exon for at least 8 bp were included in the IRP calculation. In Supplementary Figure S4B, only genes that are used in Figure 1C are included. Furthermore, genes with zero exon–intron reads or zero exon–exon reads in either resting or activated state are excluded.

RNA-seq analysis

Differentially expressed genes were selected by requiring a fold change (FC) > 2 and a reads per kilobase per million (RPKM) in shared exons > 1 in the state with higher expression.

ChIP-seq analysis

To eliminate background noise in ChIP-seq data sets, SICER was used to determine ChIP-enriched regions for each sample (29). Only reads for ChIP-enriched regions were used for downstream analysis. For all ChIP-seq libraries, the window size was chosen to be 200 bp. For histone modifications H3K36me3 and H4K20me1, the gap size was chosen to be 600 bp. For histone modification H3K4me3, the gap size was chosen to be 200 bp. For RNA Pol II, the gap size was chosen to be 400 bp. In all cases, the false discovery rate (FDR) was chosen to be 0.01.

Gene ontology analysis

Gene ontology analysis was carried out using the GeneRank program with default settings in the Genomatix Genome Analyzer (GGA) package.

RT-PCR validation of retained introns

We performed reverse transcription followed by polymerase chain reaction (RT-PCR) to validate the intron retention ratio change upon T cell activation. Primer pair in the upstream exon and downstream exon was used for detection of gene expression. Primer pair in the upstream exon and downstream intron was used to reflect the IR signal. Primers were designed by Primer3 (version 0.4.0) and synthesized by IDT. Reverse transcription of 1 μg DNA-free total RNA was performed with SuperScript II reverse transcriptase (Invitrogen) in a 20 μl reaction, containing 1x first strand synthesis buffer (Invitrogen), 5 μM dNTP (Bioline), 0.5 μM oligo(dT) primer (5′-TTTTTTTTTTTTTTTTTTTTVN-3′; V = A/C/G and N = A/C/G/T), 10 mM DTT, 40 units of RNasin (Promega) and 200 units of SuperScript II Reverse Transcriptase (Invitrogen). The RT reaction was incubated at 42°C for 2 minitues (min) before adding reverse transcriptase. We then incubated the reaction at 42°C for 60 min and 75°C for 15 min. The 20 μl PCR reaction contains 1 μl of RT template, 1x PCR buffer (Qiagen), 0.4 nmol dNTP, 0.2 μM of the forward primer and two reverse primers (one is for exon signal and the other is for intron signal) and 1 μl of Taq DNA Polymerase (Qiagen). Thermal cycling was carried out as follows: 94°C for 30 seconds (s); 30–40 cycles of 94°C for 30 s, 52–58°C for 30 s and 72°C for 30 s; 72°C for 10 min; hold at 4°C. A total of 3 μl of PCR product was taken out and run in a 2% agarose gel (Invitrogen EX E-gel). The PCR cycle number is adjusted according to the gene expression and the annealing temperature is adjusted according to the gene-specific primers used.

Quantitative real-time PCR validation of retained introns

The reverse transcription reaction for regular RT-PCR was also used for quantitative PCR (qPCR). The 20 μl qPCR reaction contains 0.5 μl RT reaction (cDNA template), 0.2 μM of the forward primer and the reverse primer and 10 μl 2x SYBR Green Power Mix (ABI). The qPCR reaction was prepared in triplicates. Thermal cycling was carried out as follows: 95°C for 10 min; 40 cycles of 95°C for 15 s, 52–58°C for 30 s and 72°C for 30 s. The annealing temperature was adjusted based on the gene-specific primers used. Melting curve module was included in the StepOnePlus RT-PCR system (Applied Biosystems).

ChIP-PCR validation of candidate genes

ChIP coupled with PCR assays were performed as described previously (30). Monoclonal anti-Pol II (ab76123, Abcam) antibody was used for immunoprecipitation. Primers for *PSMD7* and *STAT1* genes are listed in Supplementary Table S5.

RESULTS

Prevalent IR in human resting T cells compared with activated T cells

We employed DeLi-seq, a strand-specific RNA-seq procedure based on directional ligation of two adaptors to the double-stranded cDNA (27), to characterize transcriptome profiles of human $CD4^+$ T cells before and after activation. Consistent with previous findings (11,12), we did not find concordant changes of active mark H3K4me3 in genes that were up- or down-regulated upon T cell activation (Supplementary Figure S1). This is in sharp contrast to genes differentially expressed during hematopoietic stem cell differentiation (Supplementary Figure S1A and B), suggesting that mechanism(s) other than transcriptional regulation might be involved in controlling gene expression during T cell activation.

A close examination of the RNA-seq data, however, revealed that a significant number of reads reside in intronic regions. This is unexpected because polyA+ transcripts were used for constructing RNA-seq libraries (Figure 1). Strikingly, the phenomenon is more pronounced in resting T cells; the proportion of intronic reads is significantly reduced after T cell activation (15.1% versus 7.3%, *P*-value < 1E-300, chi-squared test, Figure 1B). Because strand-specific RNA-seq was used to produce the transcriptome profile, one can largely rule out that these intronic reads result from antisense transcripts (Supplementary Figure S2) (27). To characterize IR at the individual gene level, we devised an IRI that computes the ratio between normalized read counts at intronic and exonic regions based on a consolidated gene model (Supplementary Figure S3, see Materials and Methods for details). The resulting IRI distributions show that IR is a prevalent phenomenon in resting T cells, and is dramatically reduced upon T cell activation (Figure 1C, *P*-value < 1E-300, Mann–Whitney U test). The median value of the IRI is 4.0% in resting T cells and decreases to 1.8% in activated T cells (Figure 1C). In resting cells, the median IRI values for the top and bottom 1000 intron-retained genes are 20.0% and 0.62%, respectively. An alternative strategy, which characterizes IR using exon–intron and exon–exon junction reads (19), confirmed these findings (Supplementary Figures S4A and B). To ascertain the reproducibility of the stimulation-induced reduction in IR, we carried out DeLi-seq with resting and activated $CD4^+$ T cells isolated from another blood donor. In agreement with our initial observation, the results showed that the percentage of intronic reads is significantly reduced

after T cell activation (19.8% versus 8.5%, *P*-value < 1E-300, chi-squared test, Supplementary Figure S5A). In addition, the IRI distribution also showed a significant shift to a lower level (Supplementary Figure S5B), suggesting a global reduction of IR upon activation.

IR is associated with transcript instability

It has been proposed that retained introns may render the corresponding transcripts unstable that are possibly degraded by the nuclear or cytoplasmic RNA surveillance complex (21,23). We reasoned that if intron-retained transcripts are unstable, they have to be transcribed at a higher level to reach a steady-state mRNA level comparable to those transcripts without IR (Supplementary Figure S6). To test this hypothesis, we examined the transcriptional activities of genes with different levels of IR using Pol II ChIP-seq signal. Two groups of genes were compiled based on their IRI status: one group consists of top 1000 genes with the highest IRI in resting T cells (IRI-high group), while the other group consisted of the bottom 1000 genes with the lowest IRI (IRI-low group). The IRI-high gene group had significantly higher RNA Pol II occupancy than the IRI-low group (Figure 2A), suggesting that genes in the IRI-high group are transcribed at a higher rate. The observation is further confirmed by an independent ChIP-seq experiment (Supplementary Figure S7). To rule out the possibility that the expression levels between these two gene groups may affect the analysis, we further stratified the IRI-high genes and IRI-low genes into 6 sub-groups to ensure matching expression levels (Figure 2D and E). In every subgroup, although the steady-state mRNA levels are comparable, Pol II occupancy is significantly higher comparing IRI-high genes with IRI-low genes (*P*-value < 1E-45, Mann–Whitney U test) (Figure 2F). The results were further corroborated by the occupancies of H3K36me3 (Figure 2B and G) and H4K20me1 (Figure 2C), which are two histone variants known to spread over the promoter as well as the gene body of transcribed loci (5,7). Taken together, the higher transcriptional activity and comparable steady-state mRNA level strongly suggest that the intron-retained transcripts tend to be less stable than those transcripts with little or no intron-retention. Therefore, coupling of IR and RNA degradation may regulate gene expression at the post-transcriptional level.

IR is broadly involved in regulated gene expression during human T cell activation

Since IR is significantly reduced upon T cell activation, we then examined the potential link between IR and differential gene expression between resting and activated human T cells. Intriguingly, we found that the reduction in IR is correlated with increased gene expression in a linear fashion (Figure 3A–F), thereby suggesting that regulated IR might play an important role in controlling gene expression. To further explore this notion, we focused on gene loci that were upregulated in activated T cells, reasoning that the increase in the steady-state mRNA level might result from transcriptional activation and/or post-transcriptional regulation (reduced IR and increased RNA stability). Of the 2090 upregulated genes, one fourth (507 genes or 24.3%) do not show significant reduction in IR (group I genes; Figure 4A left panel), suggesting that they are likely to be predominately regulated at the transcriptional level (Figure 4A). This is further confirmed by ChIP-seq results that show a significantly higher Pol II level in activated T cells than in resting T cells (Figure 4A, left panel, Figure 4B, Supplementary Figure S8). In contrast, the majority of upregulated genes (1583 genes or 75.7%) showed a >2-fold reduction in intron retention (Figure 4B), suggesting that regulated IR might serve as a prevalent post-transcriptional gene regulatory mechanism. We further divided the 1583 genes into two groups based on whether there is a detectable change of the Pol II level (group II, Figure 4A, middle panel) or not (group III, Figure 4A, right panel). Even for group II genes, the extent of Pol II increase is less pronounced than group I genes (Figure 4A, comparing left and middle panel, Supplementary Figure S8), suggesting IR contributes to expression changes, although transcription activation and reduced IR might not be mutually exclusive.

Proteasome genes are predominantly regulated by IR

To further pinpoint the biological significance of regulated IR in human T cell activation, we focused on a subset of up-regulated genes with reduced IRI, but no change in Pol II levels, upon T cell activation. Thus, increase in gene expression is presumably due to more efficient splicing (or decreased intron retention) rather than elevated transcription. A total of 185 genes were identified (group III genes in Figure 4A, Supplementary Table S1) and they are highly enriched in the proteasome pathway (*P*-value < 1E-7, Figure 4C), which is known to play a critical role in T cell proliferation and immune function (31,32). In contrast, group I up-regulated genes (Figure 4A) are enriched in the IL-12 and IFNγ signaling pathways (Supplementary Table S2), consistent with transcription mediated gene up-regulation. Inhibition of proteasome function in $CD4^+$ T cells can result in cell cycle arrest, apoptosis and reduced cell proliferation (33). Proteasome inhibitors also lead to a decreased production of CD25 and IFNγ, two important markers for T cell activation (34,35). Taken together, IR might be critical for T cell proliferation or activation by regulating the steady-state mRNA level of proteasome related genes.

A subset of candidate proteasome genes was then randomly selected for validation. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) was used to determine the expression level of the spliced and intron-retained transcripts of each individual locus (Supplementary Figure S9, Supplementary Table S3 and S4). For all the loci tested, IR was significantly reduced while the spliced variant increased (Supplementary Figure S9B and S9C).

To further corroborate the findings, ChIP-PCR was used to analyze the Pol II level at the promoter of *PSMD7*, one of the proteasomal genes that is predominately regulated at the post-transcriptional level based on the RNA-seq and ChIP-seq results (Figure 5A). As expected, the Pol II level of *PSMD7* remained relatively constant between resting and activated T cells (Figure 5A,F), although the level of spliced mRNA variant and the protein level are significantly increased (Figure 5B–E, Supplementary Table S5). As a pos-

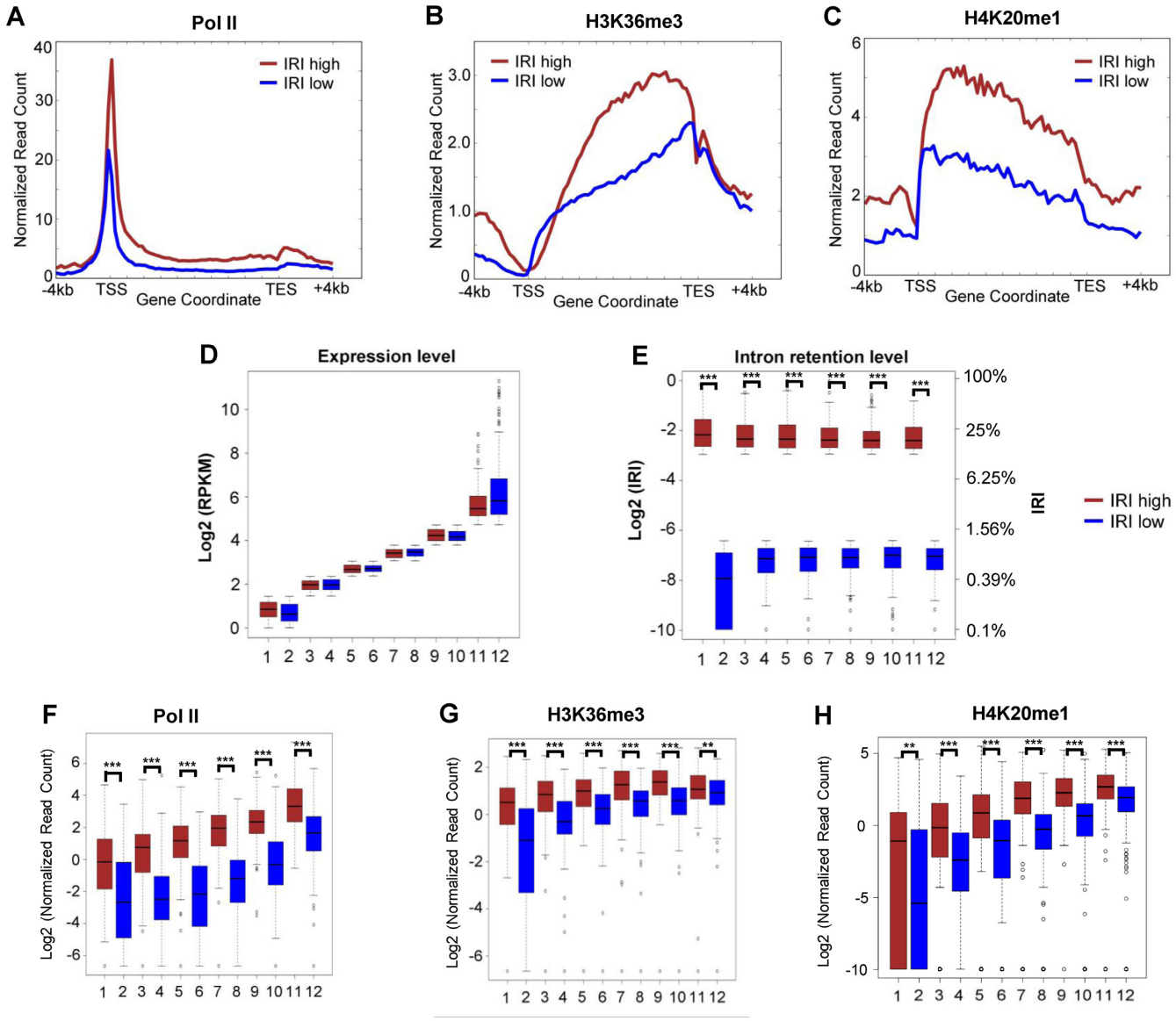

Figure 2. IRI-high transcripts are less stable than IRI-low transcripts. (**A**), (**B**) and (**C**) show the RNA Pol II, H3K36me3 and H4K20me1 density profile, respectively. The top 1000 genes (brown) and the bottom 1000 genes (blue) were selected based on their relative IRI ranks in resting T cells. The read counts were normalized to uniquely mapped reads. Six pairs of gene groups were compiled with matched (**D**) expression profiles, the corresponding levels of (**E**) intron retention, (**F**) Pol II, (**G**) H3K36me3 and (**H**) H4K20me1 are shown. In resting T cells, the median IRI values of IRI-high and IRI-low genes (E, IRI values on the right-hand side) are 20.0% and 0.6%, respectively. ** and *** represents *P*-value < 0.01 and < 0.001, respectively.

itive control, the Pol II level of *STAT1*, a gene whose expression is up-regulated at the transcriptional level (36), is increased upon activation (Figure 5F, Supplementary Table S5). By blocking a new round of transcription, expression of intron-retained and fully-spliced human *PSMD7* transcript was assayed as a function of time. We found that intron-retained transcript has a higher degradation rate than its spliced counterpart (Figure 5G). To gain insight into where IR occurs, cytoplasmic, nuclear and chromatin-associated fractions were isolated from resting human T cells. qRT-PCR results showed a significantly higher level of IR in nucleus than cytoplasm (Figure 5H). Interestingly, chromatin-associated fraction has the highest level of retained introns (Figure 5H). This result indicates that IR occurs co-transcriptionally and detained in nucleus, at least in the case of the *PSMD7* gene. Together, these data suggest that proteasome-related genes might be predominantly regulated at the intron retention level during T cell activation.

Regulated IR is conserved in mouse CD4$^+$ T cells

We next asked whether the observed IR-mediated gene regulation is conserved during evolution. To address this question, we analyzed IR using published mouse CD4$^+$ T cell transcriptome data (37). Similar to human T cells, IRI decreased in activated mouse T cells compared to naïve T cells (Figure 6). The mouse *PSMD7* gene was used for further validation and regulated IR can be confirmed, in mouse T cells (Figure 6A and C). Interestingly, we observed that

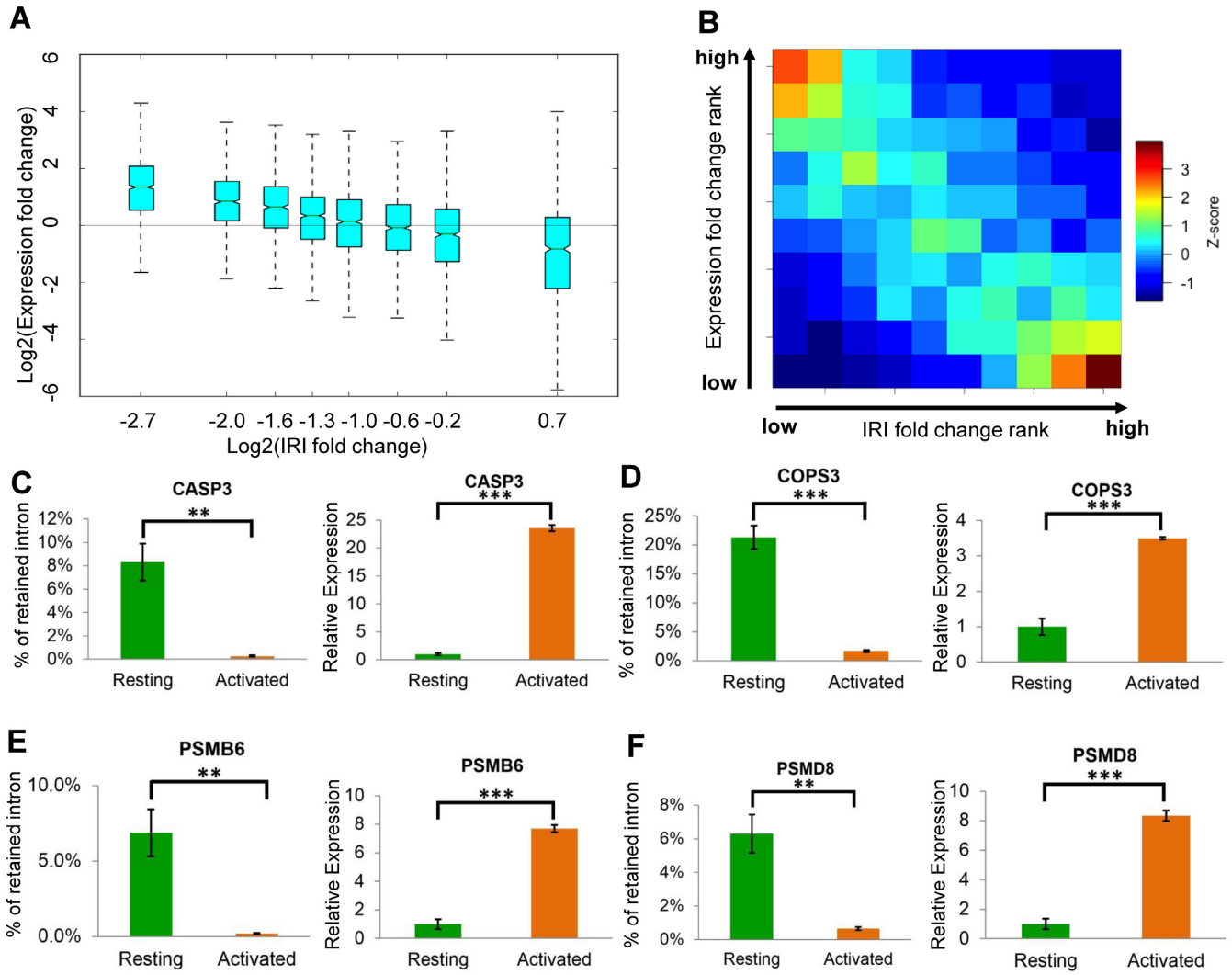

Figure 3. Changes in IR are negatively correlated with expression changes. (**A**) The fold changes in the IRI and expression levels were computed for each expressed Entrez gene as the result of human T cell activation. Entrez genes were then evenly divided into 8 bins based on IRI fold changes (X axis). The changes in expression levels for each gene in each group (Y axis) are shown as a box plot. (**B**) Both IRI fold-change rank (X axis) and expression fold change rank (Y axis) of individual genes were evenly divided into 10 bins. Each Entrez gene was subsequently placed into one of the 10 × 10 grids. The enrichment scores (Z score) were computed based on the observed gene density in each grid compared to a random distribution, and collectively shown as a heat map. Spearmen's rank correlation value is -0.43, with a *P*-value < 1E-300. (**C–F**) show four gene loci with a decreased IR level (left panel) and an increased transcript abundance (right panel) upon T cell activation.

the reduction of IR ratio (and the increase in gene expression) takes place in a time-dependent manner, and that the changes in IR level became apparent 8 h after cell activation (Figure 6B). These data strongly suggested that gene expression regulation via IR is evolutionary conserved.

DISCUSSION

In this study, we showed that IR is prevalent in resting human CD4$^+$ T cells and the level of retained introns is dramatically reduced upon activation (Figure 1). Majority (75%) of up-regulated genes in activated T cells exhibit a decrease in IR. The rest of the up-regulated genes, for which intron retention is not reduced, do show a more pronounced increase of Pol II levels at the promoter region (Figure 4A). These results explain at least in part the lack of concordant changes of Pol II occupancy and histone modifications for differentially expressed loci (Supplementary Figure S1). Extending the previous reports (19,23), we provided evidence that intron-retained transcripts are unstable and subject to RNA degradation (Figures 2 and 5G). Therefore, regulated IR coupled with RNA instability might serve as a novel post-transcriptional mechanism to control gene expression duering T cell activation (Figure 7).

The hallmark of an adaptive immune system is its ability to rapidly tune gene expression levels and cellular functions in response to specific stimuli. Transcriptional regulation has been extensively studied in T cell activation (38–40). Interestingly, we identified 185 up-regulated genes that show a concomitant decrease in IR without significant alterations of Pol II levels (Figure 4). These data highlighted that IR may drive differential gene expression in a transcription-

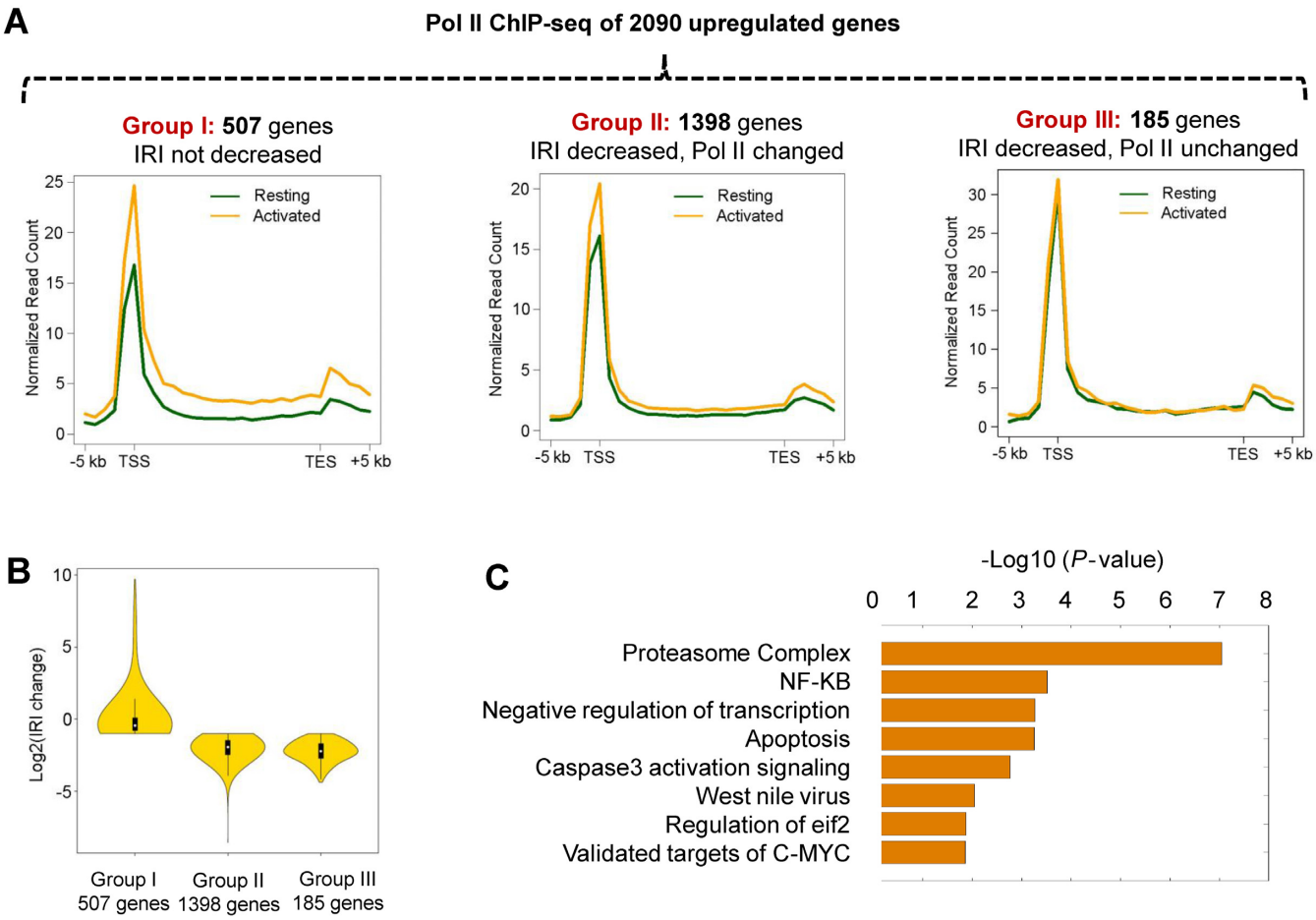

Figure 4. Upregulated expressions in activated T cells are mediated by transcriptional activation and/or intron retention. (**A**) A total of 2090 up-regulated genes identified by RNA-seq were initially divided into 2 groups based on their IRI changes. Group I consists of 507 genes that did not exhibit a decrease in intron retention. The rest of genes (1583 genes) were further divided into two groups (groups II and III) based on the Pol II ChIP-seq data obtained from resting and activated $CD4^+$ T cells. Group III genes (185 genes) show little or no change in Pol II occupancy (<30% change) and expression fold change > 3, while group II contains the remaining 1398 genes. Aggregated Pol II profiles of each group including 5 kb upstream of the transcription start sites (TSS) and 5 kb downstream of the transcription end sites (TES) are shown. (**B**) The distribution of the IRI fold-changes of three gene groups between resting and activated T cells. (**C**) Gene Ontology (GO) analysis of group III genes and the proteasome complex emerged as the most significant pathway (P-value $< 1 \times 10^{-7}$).

independent manner, although the two mechanisms are not necessarily mutually exclusive. Of these 185 genes, transcripts encoding subunits of the proteasome complex are highly enriched that is critical for T cell proliferation and cytokine release (33,34,41). We speculate that IR may provide an alternative way to achieve coordinated gene expression, similar to previous reports on fission yeast (42), budding yeast (26) and mammals (18–21,23).

Emerging studies have shown that IR is a widespread phenomenon in diverse tissues and cell types, and that the level of IR is often progressively increased during development, cell differentiation and terminal erythropoiesis (19–21,23). It has been proposed that IR may function to further repress low abundant transcripts to prevent leaky expression during cell or tissue specification (19). Extending these observations, our study demonstrated that IR is actively regulated upon acute T cell activation in both human and mice. Therefore, native T cells and possibly other terminally differentiated cells (e.g. neurons) may reach an 'active quiescent' state, and IR in a subset of genes may better prepare the cells for future encountering with pathogens or other stimuli. Compared to transcription activation, elevated gene expression from reduced IR levels does not necessarily require new rounds of transcription, thereby shortening the response time to stimuli. This can potentially be achieved by directly converting intron-retained transcripts into productive mRNAs by splicing machinery (43,44). It is well established that the activities of splicing factors (SR proteins and hnRNPs) are subject to the regulation of signaling cascade and may subsequently facilitate such splicing fate transition (45,46).

It is worth noting that the percentage of IR detected by RNA-seq or RT-PCR is at steady-state. Liu lab provided evidence that histone modifications such as H3K36me3, H3K4me3 and H3K27me3 can reflect transcriptional activity and are thus able to predict mRNA stability (47), supporting our hypothesis that mRNAs have higher H3K36me3, Pol II and H4K20me1 signals but comparably steady-state expression levels are less stable than mRNAs that have lower ChIP-seq signal (Figure 2B–D).

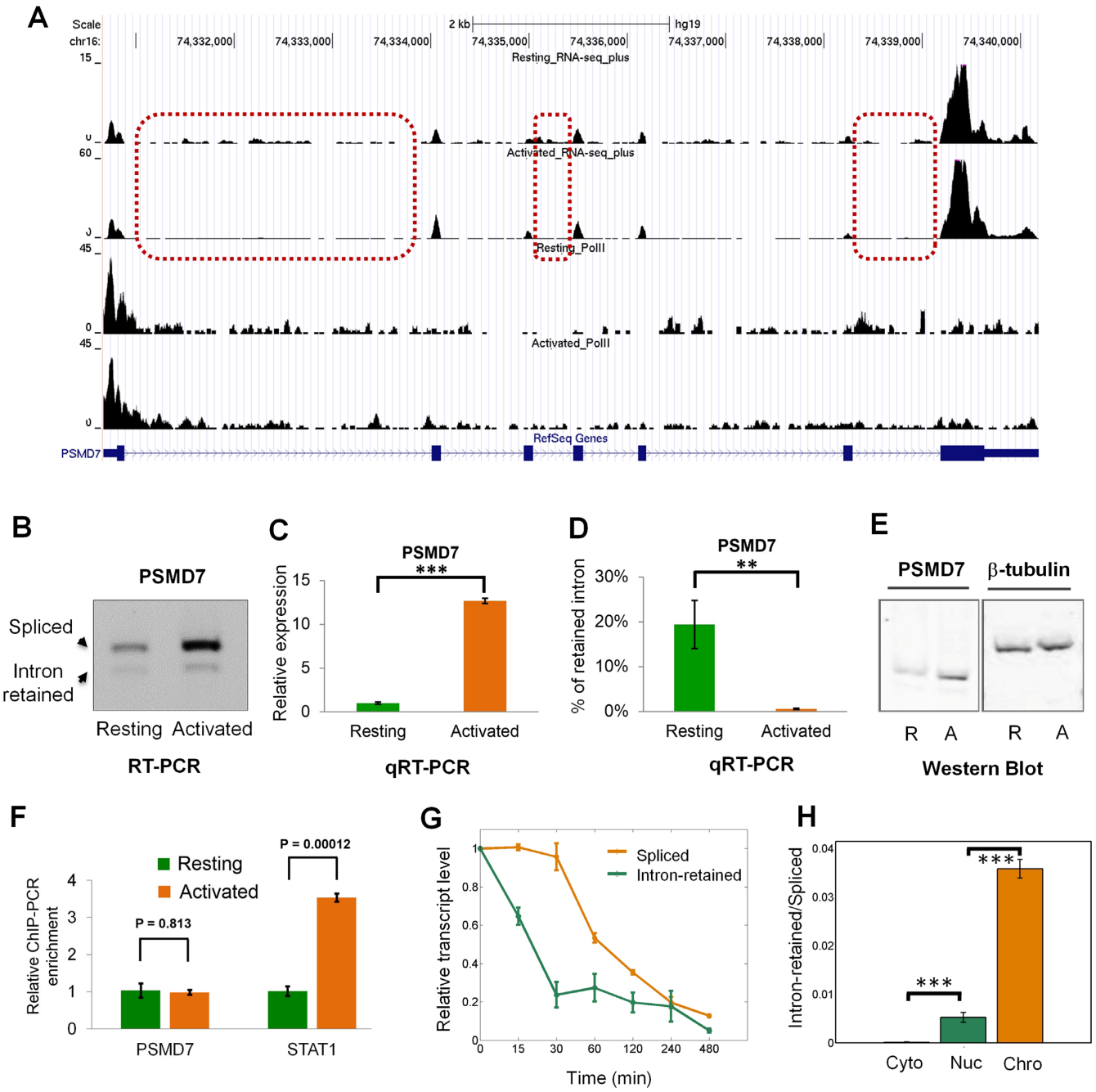

Figure 5. The expression of *PSMD7* gene is predominantly regulated by IR during human T cell activation. (**A**) The strand-specific RNA-seq (top two panels) and Pol II ChIP-seq (bottom two panels) tracks of the *PSMD7* gene. Three representative introns that have changed retention ratio between resting and activated human T cells are shown in dashed rectangles. (**B**) RT-PCR validation of regulated IR at the *PSMD7* locus. Total RNA isolated from resting and activated human $CD4^+$ T cells was reversely transcribed using oligo(dT) primer. The first-strand cDNAs were then amplified with three primers specific for *PSMD7* transcripts. The expected amplification products are shown as indicated. (**C**) The spliced transcript abundance and (**D**) IR level of human *PSMD7* gene were subsequently determined by quantitative RT-PCR (qRT-PCR) with primer pairs specific for spliced and intron-retained transcripts, respectively (mean ± SE, n = 3; two-tailed Student's *t*-test). *** and ** stands for *P*-value < 0.001 and < 0.01, respectively. (**E**) Western Blot for PSMD7 protein in resting (R) and activated (A) human $CD4^+$ T cells. β-tubulin serves as internal control. (**F**) ChIP-PCR analysis of Pol II occupancy at the promoter regions of *PSMD7* and *STAT1* genes, respectively (mean ± SE, n = 3; two-tailed Student's *t*-test). *STAT1*, a known transcriptionally regulated gene, was used as a positive control. (**G**) The degradation rate of spliced and intron retained transcript of *PSMD7* gene was quantified by qRT-PCR after inhibition of new round of transcription by adding Flavopiridol (FLV, 1 μM) in human resting T cells over the time course (min) shown in the figure. (**H**) The relative IR level (intron-retained versus spliced) of *PSMD7* gene was determined by qRT-PCR of RNA derived from cytoplasmic (Cyto), nuclear (Nuc) and chromatin-associated (Chro) fractions in human resting T cells. *** represents *P*-value < 0.001 (two-tailed Student's *t*-test).

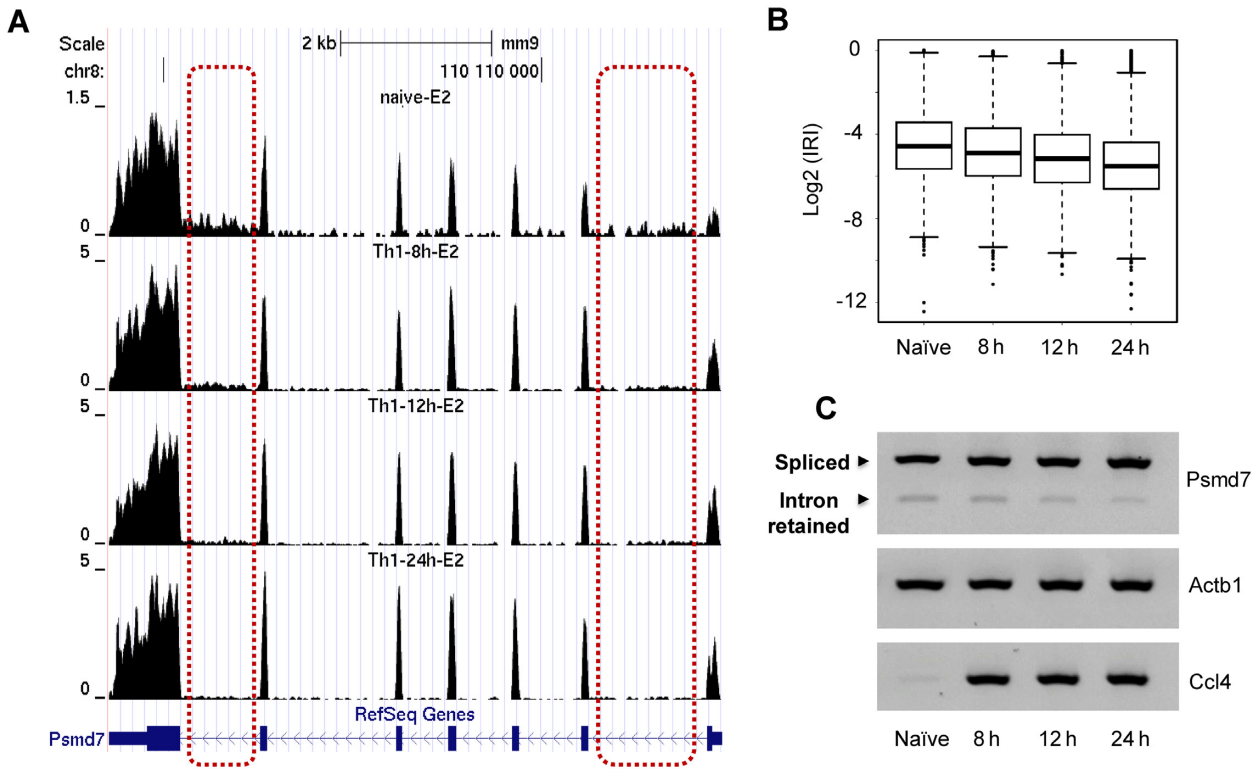

Figure 6. Conservation of IR-mediated gene regulation in mouse CD4$^+$ T cells. (**A**) The RNA-seq tracks of the *PSMD7* gene obtained from mouse naïve CD4$^+$ T cells, and Th1 cells 8 h, 12 h and 24 h post- activation. Two representative introns that have changed retention ratio are shown in dashed rectangles. (**B**) The overall IRI distribution of naïve and activated mouse Th1 cells. (**C**) RT-PCR validation of reduced IR at the mouse *PSMD7* locus along the course of Th1 cell activation. *Actb1* serves as the internal control, which didn't show significant expression changes for 4 different conditions. *Ccl4* gene serves as a positive control for mouse T cell activation.

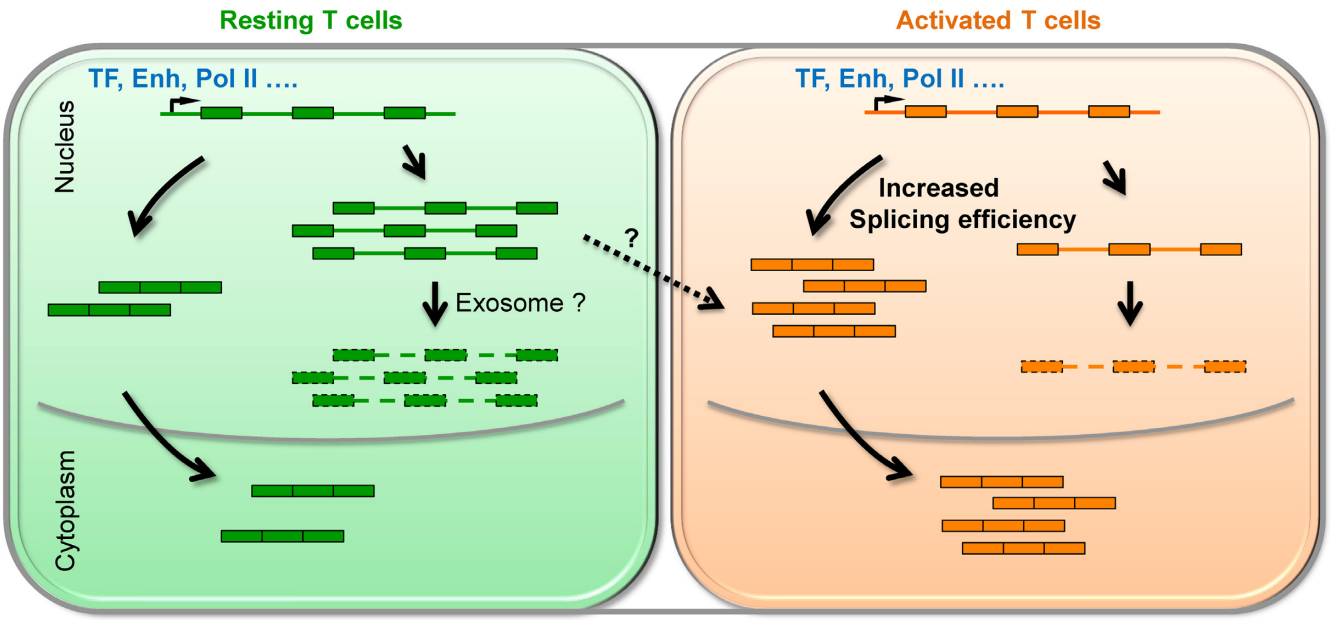

Figure 7. Proposed model for IR regulates steady-state mRNA level in both human and mouse CD4$^+$ T cells. In resting T cells, a certain number of genes (such as proteasome related genes) were not efficiently spliced that leads to high degree of degradation, possibly via the exosome-mediated pathway. Upon activation, splicing becomes more efficient and intron retained transcripts could be converted into productive mRNAs (Dashed line with an arrowhead). Therefore, an increase in steady-state mRNA level does not necessarily require transcriptional activation and could be solely achieved at the post-transcriptional level. 'TF' and 'Enh' is the abbreviation for 'transcription factor' and 'enhancer', respectively.

Our data suggest that intron-retained transcripts have a higher turnover rate than intron-spliced isoforms. Therefore, it is plausible that the actual fraction of intron-retained transcripts produced by transcription is much higher than the observed fraction, especially in resting T cells. Consequently, changes of IRI could have significant impacts on steady-state gene expression by stabilizing otherwise degraded transcripts (Figure 5 and Supplementary Figure S7). Methods such as GRO-seq, Bru-seq and BruChase-seq (48,49), which measure the Pol II on-rate, nascent transcription and transcript turnover; respectively, are expected to help further elucidate the mechanism underlying IR-mediated gene regulation.

To address whether every intron along the transcript has similar retention levels, we computed the IRI for individual introns and grouped them into five categories based on their relative locations. Consistent with previous observations (50), IR levels gradually increase in a 5′ to 3′ orientation except the first intron, which on average has the highest retention level (Supplementary Figure S10). Furthermore, the fraction of intron retention of each category was significantly reduced upon T cell activation (Supplementary Figure S10), suggesting that there is no obvious location bias. To address whether specific intron(s) are selectively retained in individual genes, we focused on genes containing 6 or more shared introns, and determined the number of retained introns in each gene using different IRI cutoffs (Supplementary Figure S11). The results showed that genes with only one retained intron represent the most abundant IR class regardless of the IRI cut-offs used. As the number of retained introns required for each group increases, the group size gradually decreases (Supplementary Figure S11). Similar trends were observed in both resting and activated human T cells using all genes (Supplementary Figure S11A and B; based on shared introns and exons) and only the simple genes (or genes without any alternative isoform) (Supplementary Figure S11C and D). This observation suggests that transcripts with a single IR event may exist, although a considerable proportion of IR transcripts contain two or more introns. These results are consistent with recent studies that suggest specific introns might be preferentially retained (18,23,44).

It has been shown that retained introns are often associated with higher G/C content, reduced intron length and relatively weak splice site signals (19). Interestingly, we found that the overall gene length, compared to the intron length, may be more reliable in predicting IR in T cells (Supplementary Figure S12). In addition, the RNA Pol II elongation rate and pausing over exonic/intronic regions has been suggested to be associated with exon definition and splicing efficiency (19,51–55). We speculate that IR may reflect temporal competition between splicing and transcription during co-transcriptional splicing (50,56). Therefore, it may be necessary to include time as an important variable when investigating IR and other alternative splicing events. The situation is further confounded by chromatin states (e.g. histone modifications) that play direct or indirect roles in regulating transcription (57) and splicing (58,59). Further mechanistic characterizations are warranted for a better understanding of regulated IR as well as its potential coupling with transcriptional and post-transcriptional regulations at the systems level. Lynch lab and others showed that activation of T cells is characterized by a robust change in the alternative exon splicing program (60–62). These studies together with our results underscore that RNA processing is substantially regulated at both the exon and intron levels during T cell activation.

Recent studies have revealed a broad involvement of IR in development and cell differentiation. Our results further extend its role in regulating gene expression during CD4+ T cell activation. Our data highlight that IR, in combination with RNA degradation, serves as a novel post-transcriptional mechanism broadly involved in gene expression regulation in $CD4^+$ T cells and beyond.

ACCESSION NUMBERS

RNA-seq and ChIP-seq raw data can be found at the NCBI Sequence Read Archive (submission SRP058500). We also downloaded published data sets at the Gene Expression Omnibus with GEO number GSE48138 and NCBI Sequence Read Archive with SRA number SRA000206.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Yoshihisa Ishizuka, Julia Dovelkis, Lena Diaw and Ye Rin Koh for critical reading of the manuscript. We thank Genergy Biotech (Shanghai) Co., Ltd. for the deep sequencing service.

Authors' contributions: T.N., W.P. and J.Z. designed the experiments. T.N., Y.Z., T.S., H.N., Z.Z., Y.D., P.L., K.C., Y.T., B.Z. performed library construction and experimental validation for RNA-seq and ChIP-seq. W.Y., M.H., Y.Y., G.W., Z.Z. and W.P. performed bioinformatics analyses. K.Z. and J.Z. provided the deep sequencing service. T.N., W.P. and J.Z. wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

National Institute of Health (NIH) [Division of Intramural Research of National Heart Lung and Blood Institute to K. Z. and J. Z.; 1R21AI113806 to W. P.]; National Key Basic Research Program of China [973 programs: 2013CB530700 and 2015CB943000 to T. N., 2012CB518700 to Y. D.]; National Science Foundation of China [31271348 and 31471192 to T. N.]; Innovation Project of Shanghai Municipal Education Commission [14ZZ007 to T. N.]; Thousand Talents Plan of China and the 111 Project of China (B13016).

Conflict of interest statement. None declared.

REFERENCES

1. Chandok,M.R. and Farber,D.L. (2004) Signaling control of memory T cell generation and function. *Semin. Immunol.*, **16**, 285–293.
2. Smith-Garvin,J.E., Koretzky,G.A. and Jordan,M.S. (2009) T cell activation. *Annu. Rev. Immunol.*, **27**, 591–619.

3. Macian,F., Lopez-Rodriguez,C. and Rao,A. (2001) Partners in transcription: NFAT and AP-1. *Oncogene*, **20**, 2476–2489.
4. Bendfeldt,H., Benary,M., Scheel,T., Steinbrink,K., Radbruch,A., Herzel,H. and Baumgrass,R. (2012) IL-2 Expression in Activated Human Memory FOXP3(+) Cells Critically Depends on the Cellular Levels of FOXP3 as Well as of Four Transcription Factors of T Cell Activation. *Front. Immunol.*, **3**, 264.
5. Barski,A., Cuddapah,S., Cui,K., Roh,T.Y., Schones,D.E., Wang,Z., Wei,G., Chepelev,I. and Zhao,K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.
6. Mikkelsen,T.S., Ku,M., Jaffe,D.B., Issac,B., Lieberman,E., Giannoukos,G., Alvarez,P., Brockman,W., Kim,T.K., Koche,R.P. *et al.* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553–560.
7. Wang,Z., Zang,C., Rosenfeld,J.A., Schones,D.E., Barski,A., Cuddapah,S., Cui,K., Roh,T.Y., Peng,W., Zhang,M.Q. *et al.* (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.*, **40**, 897–903.
8. Ernst,J. and Kellis,M. (2012) ChromHMM: automating chromatin-state discovery and characterization. *Nat. Methods*, **9**, 215–216.
9. Cui,K., Zang,C., Roh,T.Y., Schones,D.E., Childs,R.W., Peng,W. and Zhao,K. (2009) Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell*, **4**, 80–93.
10. Mohn,F., Weber,M., Rebhan,M., Roloff,T.C., Richter,J., Stadler,M.B., Bibel,M. and Schubeler,D. (2008) Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell*, **30**, 755–766.
11. Barski,A., Jothi,R., Cuddapah,S., Cui,K., Roh,T.Y., Schones,D.E. and Zhao,K. (2009) Chromatin poises miRNA- and protein-coding genes for expression. *Genome Res.*, **19**, 1742–1751.
12. Cuddapah,S., Barski,A. and Zhao,K. (2010) Epigenomics of T cell activation, differentiation, and memory. *Curr. Opin. Immunol.*, **22**, 341–347.
13. Lim,P.S., Hardy,K., Bunting,K.L., Ma,L., Peng,K., Chen,X. and Shannon,M.F. (2009) Defining the chromatin signature of inducible genes in T cells. *Genome Biol.*, **10**, R107.
14. Galante,P.A., Sakabe,N.J., Kirschbaum-Slager,N. and de Souza,S.J. (2004) Detection and evaluation of intron retention events in the human transcriptome. *RNA*, **10**, 757–765.
15. Wang,E.T., Sandberg,R., Luo,S., Khrebtukova,I., Zhang,L., Mayr,C., Kingsmore,S.F., Schroth,G.P. and Burge,C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470–476.
16. Barash,Y., Calarco,J.A., Gao,W., Pan,Q., Wang,X., Shai,O., Blencowe,B.J. and Frey,B.J. (2010) Deciphering the splicing code. *Nature*, **465**, 53–59.
17. Nilsen,T.W. and Graveley,B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**, 457–463.
18. Boutz,P.L., Bhutkar,A. and Sharp,P.A. (2015) Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev.*, **29**, 63–80.
19. Braunschweig,U., Barbosa-Morais,N.L., Pan,Q., Nachman,E.N., Alipanahi,B., Gonatopoulos-Pournatzis,T., Frey,B., Irimia,M. and Blencowe,B.J. (2014) Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res.*, **24**, 1774–1786.
20. Pimentel,H., Parra,M., Gee,S.L., Mohandas,N., Pachter,L. and Conboy,J.G. (2016) A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis. *Nucleic Acids Res.*, **44**, 838–851.
21. Wong,J.J., Ritchie,W., Ebner,O.A., Selbach,M., Wong,J.W., Huang,Y., Gao,D., Pinello,N., Gonzalez,M., Baidya,K. *et al.* (2013) Orchestrated intron retention regulates normal granulocyte differentiation. *Cell*, **154**, 583–595.
22. Cho,V., Mei,Y., Sanny,A., Chan,S., Enders,A., Bertram,E.M., Tan,A., Goodnow,C.C. and 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.
23. Yap,K., Lim,Z.Q., Khandelia,P., Friedman,B. and Makeyev,E.V. (2012) Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. *Genes Dev.*, **26**, 1209–1223.
24. Dvinge,H. and Bradley,R.K. (2015) Widespread intron retention diversifies most cancer transcriptomes. *Genome Med.*, **7**, 45.
25. Averbeck,N., Sunder,S., Sample,N., Wise,J.A. and Leatherwood,J. (2005) Negative control contributes to an extensive program of meiotic splicing in fission yeast. *Mol. Cell*, **18**, 491–498.
26. Parenteau,J., Durand,M., Morin,G., Gagnon,J., Lucier,J.F., Wellinger,R.J., Chabot,B. and Elela,S.A. (2011) Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. *Cell*, **147**, 320–331.
27. Ni,T., Tu,K., Wang,Z., Song,S., Wu,H., Xie,B., Scott,K.C., Grewal,S.I., Gao,Y. and Zhu,J. (2010) The prevalence and regulation of antisense transcripts in Schizosaccharomyces pombe. *PloS One*, **5**, e15271.
28. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
29. Zang,C.Z., Schones,D.E., Zeng,C., Cui,K.R., Zhao,K.J. and Peng,W.Q. (2009) A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics*, **25**, 1952–1958.
30. Maunakea,A.K., Chepelev,I., Cui,K.R. and Zhao,K.J. (2013) Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.*, **23**, 1256–1269.
31. Gibson,H.M., Mishra,A., Chan,D.V., Hake,T.S., Porcu,P. and Wong,H.K. (2012) Impaired proteasome function activates GATA3 in T cells and upregulates CTLA-4: relevance for Sezary syndrome. *J. Invest. Dermatol.*, **133**, 249–257.
32. Wang,X., Luo,H., Chen,H., Duguid,W. and Wu,J. (1998) Role of proteasomes in T cell activation and proliferation. *J. Immunol.*, **160**, 788–801.
33. Lu,M., Dou,Q.P., Kitson,R.P., Smith,D.M. and Goldfarb,R.H. (2006) Differential effects of proteasome inhibitors on cell cycle and apoptotic pathways in human YT and Jurkat cells. *J. Cell Biochem.*, **97**, 122–134.
34. Berges,C., Haberstock,H., Fuchs,D., Miltz,M., Sadeghi,M., Opelz,G., Daniel,V. and Naujokat,C. (2008) Proteasome inhibition suppresses essential immune functions of human CD4(+) T cells. *Immunology*, **124**, 234–246.
35. Yanaba,K., Yoshizaki,A., Muroi,E., Hara,T., Ogawa,F., Shimizu,K. and Sato,S. (2010) The proteasome inhibitor bortezomib inhibits T cell-dependent inflammatory responses. *J. Leukoc. Biol.*, **88**, 117–122.
36. Beadling,C., Guschin,D., Witthuhn,B.A., Ziemiecki,A., Ihle,J.N., Kerr,I.M. and Cantrell,D.A. (1994) Activation of JAK kinases and STAT proteins by interleukin-2 and interferon alpha, but not the T cell antigen receptor, in human T lymphocytes. *EMBO J.*, **13**, 5605–5615.
37. Hu,G.Q., Tang,Q.S., Sharma,S., Yu,F., Escobar,T.M., Muljo,S.A., Zhu,J.F. and Zhao,K.J. (2013) Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nat. Immunol.*, **14**, 1190–1198.
38. Wei,G., Wei,L., Zhu,J.F., Zang,C.Z., Hu-Li,J., Yao,Z.J., Cui,K.R., Kanno,Y., Roh,T.Y., Watford,W.T. *et al.* (2009) Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4(+) T Cells. *Immunity*, **30**, 155–167.
39. Christie,D. and Zhu,J.F. (2014) Transcriptional regulatory networks for CD4 T Cell differentiation. *Curr. Top Microbiol. Immunol.*, **381**, 125–172.
40. Zhu,J.F. and Paul,W.E. (2010) Peripheral CD4+T-cell differentiation regulated by networks of cytokines and transcription factors. *Immunol. Rev.*, **238**, 247–262.
41. Gibson,H.M., Mishra,A., Chan,D.V., Hake,T.S., Porcu,P. and Wong,H.K. (2013) Impaired proteasome function activates GATA3 in T cells and upregulates CTLA-4: relevance for Sezary syndrome. *J. Invest. Dermatol.*, **133**, 249–257.
42. Cremona,N., Potter,K. and Wise,J.A. (2011) A meiotic gene regulatory cascade driven by alternative fates for newly synthesized transcripts. *Mol. Biol. Cell*, **22**, 66–77.
43. Ninomiya,K., Kataoka,N. and Hagiwara,M. (2011) Stress-responsive maturation of Clk1/4 pre-mRNAs promotes phosphorylation of SR splicing factor. *J. Cell Biol.*, **195**, 27–40.
44. Boothby,T.C., Zipper,R.S., van der Weele,C.M. and Wolniak,S.M. (2013) Removal of retained introns regulates translation in the rapidly developing gametophyte of Marsilea vestita. *Dev. Cell*, **24**, 517–529.

45. David,C.J. and Manley,J.L. (2010) Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.*, **24**, 2343–2364.
46. Stamm,S. (2002) Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. *Hum. Mol. Genet.*, **11**, 2409–2416.
47. Wang,C., Tian,R., Zhao,Q., Xu,H., Meyer,C.A., Li,C., Zhang,Y. and Liu,X.S. (2012) Computational inference of mRNA stability from histone modification and transcriptome profiles. *Nucleic Acids Res.*, **40**, 6414–6423.
48. Core,L.J., Waterfall,J.J. and Lis,J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, **322**, 1845–1848.
49. Paulsen,M.T., Veloso,A., Prasad,J., Bedi,K., Ljungman,E.A., Tsan,Y.C., Chang,C.W., Tarrier,B., Washburn,J.G., Lyons,R. *et al.* (2013) Coordinated regulation of synthesis and stability of RNA during the acute TNF-induced proinflammatory response. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 2240–2245.
50. Tilgner,H., Knowles,D.G., Johnson,R., Davis,C.A., Chakraborty,S., Djebali,S., Curado,J., Snyder,M., Gingeras,T.R. and Guigo,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.
51. Brodsky,A.S., Meyer,C.A., Swinburne,I.A., Giles,H., Keenan,B.J., Liu,X.L.S., Fox,E.A. and Silver,P.A. (2005) Genomic mapping of RNA polymerase II reveals sites of co-transcriptional regulation in human cells. *Genome Biol.*, **6**, R64.
52. Alexander,R.D., Innocente,S.A., Barrass,J.D. and Beggs,J.D. (2010) Splicing-dependent RNA polymerase pausing in yeast. *Mol. Cell*, **40**, 582–593.
53. Chodavarapu,R.K., Feng,S.H., Bernatavichute,Y.V., Chen,P.Y., Stroud,H., Yu,Y.C., Hetzel,J.A., Kuo,F., Kim,J., Cokus,S.J. *et al.* (2010) Relationship between nucleosome positioning and DNA methylation. *Nature*, **466**, 388–392.
54. Kwak,H., Fuda,N.J., Core,L.J. and Lis,J.T. (2013) Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science*, **339**, 950–953.
55. Kornblihtt,A.R., Schor,I.E., Allo,M., Dujardin,G., Petrillo,E. and Munoz,M.J. (2013) Alternative splicing: A pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.*, **14**, 153–165.
56. Shukla,S. and Oberdoerffer,S. (2012) Co-transcriptional regulation of alternative pre-mRNA splicing. *Biochim. Biophys. Acta*, **1819**, 673–683.
57. Zhang,Y. and Reinberg,D. (2001) Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.*, **15**, 2343–2360.
58. Luco,R.F., Pan,Q., Tominaga,K., Blencowe,B.J., Pereira-Smith,O.M. and Misteli,T. (2010) Regulation of alternative splicing by histone modifications. *Science*, **327**, 996–1000.
59. Guo,R., Zheng,L., Park,J.W., Lv,R., Chen,H., Jiao,F., Xu,W., Mu,S., Wen,H., Qiu,J. *et al.* (2014) BS69/ZMYND11 reads and connects histone H3.3 lysine 36 trimethylation-decorated chromatin to regulated pre-mRNA processing. *Mol. Cell*, **56**, 298–310.
60. Martinez,N.M., Agosto,L., Qiu,J., Mallory,M.J., Gazzara,M.R., Barash,Y., Fu,X.D. and Lynch,K.W. (2015) Widespread JNK-dependent alternative splicing induces a positive feedback loop through CELF2-mediated regulation of MKK7 during T-cell activation. *Genes Dev.*, **29**, 2054–2066.
61. Cole,B.S., Tapescu,I., Allon,S.J., Mallory,M.J., Qiu,J., Lake,R.J., Fan,H.Y., Fu,X.D. and Lynch,K.W. (2015) Global analysis of physical and functional RNA targets of hnRNP L reveals distinct sequence and epigenetic features of repressed and enhanced exons. *RNA*, **21**, 2053–2066.
62. Thenoz,M., Vernin,C., Mortada,H., Karam,M., Pinatel,C., Gessain,A., Webb,T.R., Auboeuf,D., Wattel,E. and Mortreux,F. (2014) HTLV-1-infected CD4+ T-cells display alternative exon usages that culminate in adult T-cell leukemia. *Retrovirology*, **11**, 119.