

The RNA helicase DDX39B activates FOXP3 RNA splicing to control T regulatory cell fate

Minato Hirano^{1,2†}, Gaddiel Galarza-Muñoz^{1,3†}, Chloe Nagasawa^{1,4}, Geraldine Schott^{1†}, Liuyang Wang⁵, Alejandro L Antonia⁵, Vaibhav Jain⁶, Xiaoying Yu^{1,7}, Steven G Widen¹, Farren BS Briggs⁸, Simon G Gregory^{5,6,9}, Dennis C Ko^{5,10}, William S Fagg^{1,11}, Shelton Bradrick^{12*§}, Mariano A Garcia-Blanco^{1,13,14*}

¹Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, United States; ²National Research Center for the Control and Prevention of Infectious Disease, Nagasaki University, Nagasaki, Japan; ³Autoimmunity Biological Solutions, Galveston, United States; ⁴Human Pathophysiology and Translational Medicine Program, Institute for Translational Sciences, University of Texas Medical Branch, Galveston, United States; ⁵Department of Molecular Genetics and Microbiology, Duke University, Durham, United States; ⁶Duke Molecular Physiology Institute, Duke University, Durham, United States; ⁷Department of Preventive Medicine and Population Health, University of Texas Medical Branch, Galveston, United States; ⁸Department of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University, Cleveland, United States; ⁹Department of Neurology, Duke University School of Medicine, Durham, United States; ¹⁰Division of Infectious Diseases, Department of Medicine, Duke University, Durham, United States; ¹¹Transplant Division, Department of Surgery, University of Texas Medical Branch, Galveston, United States; ¹²Institute of Human Infections and Immunity, University of Texas Medical Branch, Galveston, United States; ¹³Department of Internal Medicine, University of Texas Medical Branch, Galveston, United States; ¹⁴Department of Microbiology, Immunology and Cancer Biology, University of Virginia, Charlottesville, United States

*For correspondence: ssbradri@utmb.edu (SB); maragarc@utmb.edu (MAG-B)

†These authors contributed equally to this work

Present address: [†]Quivira Trce, League, United States; [§]The Trudeau Institute, Algonquin Ave, New York, United States

Competing interest: See page 20

Funding: See page 20

Received: 10 January 2022

Preprinted: 27 January 2022

Accepted: 09 May 2023

Published: 01 June 2023

Reviewing Editor: Apurva Sarin, Institute for Stem Cell Science and Regenerative Medicine, India

© Copyright Hirano, Galarza-Muñoz et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract Genes associated with increased susceptibility to multiple sclerosis (MS) have been identified, but their functions are incompletely understood. One of these genes codes for the RNA helicase DEXD/H-Box Polypeptide 39B (DDX39B), which shows genetic and functional epistasis with interleukin-7 receptor- α gene (*IL7R*) in MS-risk. Based on evolutionary and functional arguments, we postulated that DDX39B enhances immune tolerance thereby decreasing MS risk. Consistent with such a role we show that DDX39B controls the expression of many MS susceptibility genes and important immune-related genes. Among these we identified *Forkhead Box P3* (*FOXP3*), which codes for the master transcriptional factor in CD4⁺/CD25⁺ T regulatory cells. DDX39B knockdown led to loss of immune-regulatory and gain of immune-effector expression signatures. Splicing of *FOXP3* introns, which belong to a previously unrecognized type of introns with C-rich polypyrimidine tracts, was exquisitely sensitive to DDX39B levels. Given the importance of *FOXP3* in autoimmunity, this work cements DDX39B as an important guardian of immune tolerance.

Editor's evaluation

DDX39B is a helicase with known functions in mRNA splicing and nuclear export. This important study provides convincing evidence that DDBX39B regulates Foxp3, a lineage marker for T-regulatory cells in the immune system. The work provides a detailed analysis of the post-transcriptional regulation of Foxp3, and also positions DDX39B more broadly, as an important player in the regulation of autoimmune responses. The work will be of interest to RNA biologists, immunologists, and those studying autoimmune disorders.

Introduction

Autoimmune diseases are caused by a combination of environmental and genetic factors. Genetic factors associated with increased susceptibility to multiple sclerosis (MS), an autoimmune disease of the central nervous system, have been identified, but their mechanisms of action are incompletely understood (*Briggs, 2019; Patsopoulos et al., 2019*). We established that the association between MS risk and the interleukin-7 receptor- α gene (*IL7R*) is mediated by alternative splicing of *IL7R* transcripts (*Gregory et al., 2007*). The disease associated allele of the single-nucleotide polymorphism (SNP) rs6897932 in exon 6 of *IL7R* strengthens an exonic splicing silencer (ESS) and increases skipping of exon 6 leading to increased production of soluble *IL7R* (s*IL7R*) (*Evsyukova et al., 2013; Gregory et al., 2007; Schott et al., 2021*). Elevated levels of s*IL7R* have been shown to exacerbate the inducible murine MS-like model experimental autoimmune encephalomyelitis (EAE) and are proposed to increase the bioavailability of *IL7* (*Lundström et al., 2013*). Given the importance of exon 6 splicing we identified RNA binding proteins that bind and impact its splicing, and among these was the RNA helicase DExD/H-Box Polypeptide 39B (DDX39B; *Evsyukova et al., 2013; Galarza-Muñoz et al., 2017*).

DDX39B, known to RNA biologists as U2AF65-Associated Protein 56 KDa (UAP56), plays roles in RNA splicing and nucleocytoplasmic transport (*Shen, 2009*), but was first discovered by immunologists, who named it HLA-B Associated Transcript 1 (BAT1) (*Spies et al., 1989*) and linked it to autoimmune diseases (*Degli-Esposti et al., 1992*). We connected the RNA biology and immunology roles of DDX39B showing that it activates splicing of *IL7R* exon 6 and represses production of s*IL7R* (*Galarza-Muñoz et al., 2017*). Furthermore, *DDX39B* shows genetic and functional epistasis with *IL7R* in enhancing MS risk (*Galarza-Muñoz et al., 2017*). Based on these studies, we proposed that DDX39B plays protective roles in MS and other autoimmune diseases (*Galarza-Muñoz et al., 2017*). The alternative splicing of *IL7R* exon 6 is conserved only in primates, while *DDX39B* is conserved and located in the major histocompatibility complex (MHC) class III region in all vertebrates (*Schott and Garcia-Blanco, 2021*), suggesting that DDX39B controls important immune modulators other than *IL7R*.

Here, we show that DDX39B controls expression of gene products involved in autoimmunity including Forkhead Box P3 (FOXP3), a master regulator of the development, maintenance and function of CD4⁺/CD25⁺ regulatory T cells (Tregs; *Georgiev et al., 2019; Hori, 2021; Josefowicz et al., 2012*), and a repressor of autoimmune diseases (*Bennett et al., 2001; Brunkow et al., 2001; Chatila et al., 2000; Wildin et al., 2001*). Splicing of FOXP3 introns, which have C-rich polypyrimidine tracts, is strongly dependent on DDX39B, making FOXP3 expression highly sensitive to the levels of this RNA helicase.

Results

DDX39B controls expression of genes associated with multiple sclerosis risk

We had previously knocked down DDX39B with two independent shRNAs (Sh3 and Sh5) in primary human CD4⁺ T cells from six healthy human donors (Donors 1–6) and shown that DDX39B depletion led to increased skipping of *IL7R* exon 6 (*Galarza-Muñoz et al., 2017*). To identify other RNAs affected by DDX39B in these immune relevant cells, we carried out RNA sequencing (RNAseq) of polyadenylated RNA from control and DDX39B depleted CD4⁺ T cells from two of these donors (donors 1 and 4). More transcript level changes were detected with Sh3 treatment than with Sh5

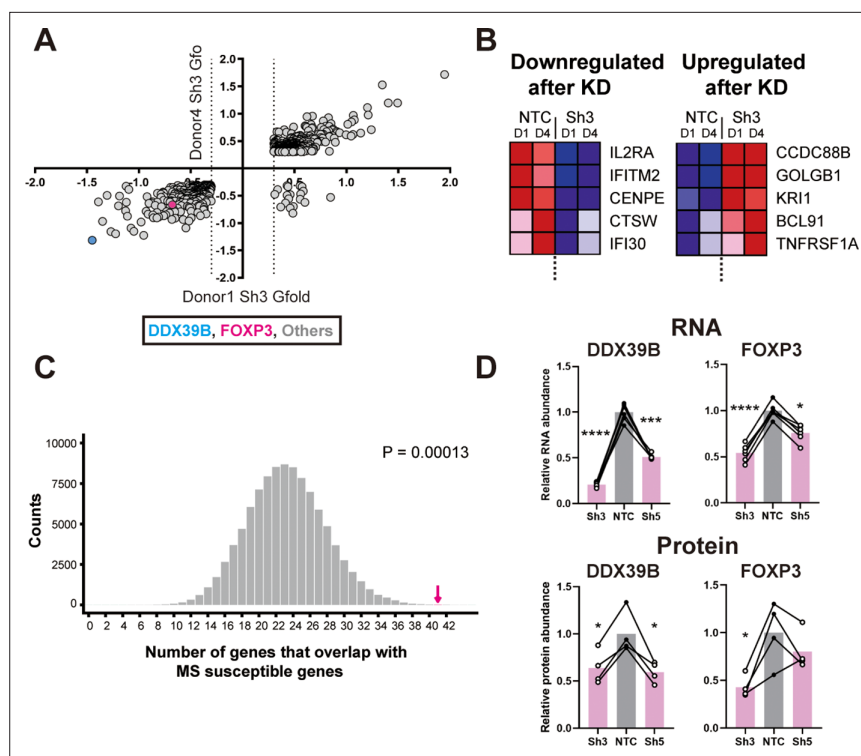


Figure 1. Loss of FOXP3 expression in DDX39B-depleted T cells. **(A)** RNA abundance changes between control (NTC) and DDX39B-depleted (Sh3) CD4⁺ T cells from two healthy individuals (Donor 1 and Donor 4) identified by RNAseq. Data points for DDX39B (cyan) and FOXP3 (magenta) are indicated. **(B)** Examples of MS susceptibility genes differentially expressed upon DDX39B knockdown. The heat map shows expression of five genes between control (NTC) and DDX39B depleted (Sh3) CD4⁺ T cells. **(C)** Enrichment analysis of MS susceptibility genes in DEGs following DDX39B depletion. Resampling 100,000 times resulted in a distribution of the number of genes that overlap by chance, with 23 being the most common result. The observed overlap of 41 (Magenta arrow) demonstrates substantial enrichment (empirical $p=0.00013$). **(D)** Levels of DDX39B RNA and FOXP3 RNA, normalized to EEF1A1 RNA levels, after DDX39B depletion in CD4⁺ T cells and levels of DDX39B and FOXP3 protein relative to Tubulin after DDX39B depletion in CD4⁺ T cells. Connected dots indicate samples from the same donor. In all figures the error bars indicate standard deviation. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$ and ****: $p<0.0001$.

The online version of this article includes the following source data and figure supplement(s) for figure 1:

Figure supplement 1. Transcript level changes upon DDX39B depletion in primary human CD4⁺ T cells.

Figure supplement 2. Schematic diagram of the enrichment analysis of MS-susceptibility genes among genes whose expression is altered by DDX39B knockdown.

Figure supplement 3. Enrichment analysis of the MS_Pathogenic or MS_Protective genes.

Figure supplement 4. Coverage tracks of reads mapping to the *FOXP3* gene in control and DDX39B-depleted CD4⁺ T cells.

Figure supplement 5. Western blot analysis of FOXP3 expression upon DDX39B depletion in CD4⁺ T cells.

Figure supplement 5—source data 1. Two va files have been uploaded for all gels/blots shown in this figure:

(1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labeled.

treatment (**Figure 1—figure supplement 1** and **Supplementary file 1**), which was consistent with Sh3 treatment reducing DDX39B RNA levels more profoundly (**Supplementary file 1**) and DDX39B protein levels more consistently (**Galarza-Muñoz et al., 2017**). Therefore, analysis of transcriptome alterations were first carried out by comparing CD4⁺ T cells depleted of DDX39B using Sh3 to cells treated with a control shRNA; however, all critical results were confirmed with both shRNAs and in samples from multiple human donors.

We evaluated the overlap between 762 transcripts that were significantly altered by Sh3-mediated DDX39B knockdown in CD4⁺ T cells from both donors 1 and 4 (**Figure 1A** and **Supplementary file 2**) and 558 putative MS susceptibility genes identified by the *Patsopoulos et al., 2019*. Expression of 421 of these MS susceptibility genes was detected in our RNAseq data and expression of 41 of these was significantly altered by DDX39B depletion (e.g. IL2RA, **Figure 1B**). This overlap meant that DDX39B differentially expressed genes (DEG) were highly enriched among MS susceptibility genes ($p=0.00013$; **Figure 1C** and **Figure 1—figure supplement 2**).

Based on eQTL information of the disease associated alleles (*GTEx Consortium, 2017*), we classified 250 of the 558 MS susceptibility genes as potentially 'pathogenic' since the disease-associated allele also associates with higher gene expression in lymphoblastoid cells, and 262 genes as potentially 'protective' since the disease-associated allele associates with lower expression of the gene (**Figure 1—figure supplement 3** and **Supplementary file 3**). We found significant overlap between MS pathogenic genes and genes upregulated upon DDX39B knockdown ($p=0.00015$, **Figure 1—figure supplement 3**), and also between MS protective genes and genes downregulated upon DDX39B knockdown, ($p=0.03$; **Figure 1—figure supplement 3**). In contrast, there was no significant overlap when the directional pairing was flipped to compare MS pathogenic genes with those downregulated by DDX39B knockdown ($p=0.7$), or MS protective genes with those upregulated by DDX39B knockdown ($p=0.1$; **Figure 1—figure supplement 3**). These data strongly suggested a shift of gene expression signature from MS-protective to MS-pathogenic upon DDX39B depletion, and provide support for a protective role for DDX39B in MS risk.

DDX39B controls expression of FOXP3

Among the 762 transcripts that were significantly altered by Sh3-mediated DDX39B knockdown in CD4⁺ T cells from both donors (**Figure 1A**, **Supplementary file 2**) we identified Forkhead Box P3 (FOXP3) transcripts. Indeed, FOXP3 transcripts were among 122 high confidence targets of DDX39B significantly reduced in CD4⁺ T cells from both donors 1 and 4 with both Sh3 and Sh5 (**Figure 1—figure supplement 4**, **Supplementary file 1**). Given the importance of FOXP3 in the development, maintenance and function of Tregs (*Georgiev et al., 2019*; *Josefowicz et al., 2012*), and its strong association with autoimmune diseases (*Dominguez-Villar and Hafler, 2018*) we investigated this further.

Consistent with the measurement from the RNAseq, RT-qPCR showed FOXP3 RNA was reduced in DDX39B-depleted primary CD4⁺ T cells from the two aforementioned donors and from four additional healthy donors (**Figure 1D**). Importantly, FOXP3 protein was decreased in CD4⁺ T cells lysates that had measurable levels of total protein from four of these donors (**Figure 1D** and **Figure 1—figure supplement 5**).

In blood-derived CD4⁺ T cells FOXP3 is predominantly expressed in CD4⁺/CD25⁺ Tregs (*Josefowicz et al., 2012*), which are a small fraction of CD4⁺ T cells. To address the effect of DDX39B knockdown in Tregs we depleted DDX39B in the MT-2 Treg-like human cell line (*Hamano et al., 2015*). As observed in CD4⁺ T cells, knockdown of DDX39B led to markedly decreased expression of FOXP3 RNA and protein in MT-2 cells (**Figure 2A–C**). All FOXP3 isoforms detected in western blots were downregulated upon DDX39B knockdown (**Figure 2C**). Circulating primary human Tregs were isolated from PBMCs from Donors 7 and 8, or induced from PBMCs (*Liu et al., 2006*) from Donors 9 and 10 (iTregs), and depleted of DDX39B by expression of Sh3 and/or Sh5 (**Figure 2D**). DDX39B depletion led to a decrease in FOXP3 RNA in all Tregs from all four donors (**Figure 2D**) and a decrease in FOXP3 protein in Tregs from the two donors (7 and 8) where we had protein lysates (**Figure 2E**). All the above data indicate that robust FOXP3 expression in Tregs requires DDX39B.

DDX39B depletion disrupts Treg-specific gene expression

Given that DDX39B knockdown led to low FOXP3 levels we predicted important alterations in immune relevant gene expression networks and tested this prediction using RNAseq data from primary CD4⁺ T cells depleted of DDX39B. Using Gene Set Enrichment Analysis (GSEA) (*Subramanian et al., 2005*), we identified gene sets enriched in genes differentially expressed upon DDX39B knockdown in CD4⁺ T cells (**Supplementary file 4**). Among 648 gene sets enriched for genes downregulated upon DDX39B knockdown (nominal p -value ≤ 0.05) we found many related to immune function, including several gene sets of FOXP3 targets (*Gavin et al., 2007*; **Figure 3A** and **Figure 3—figure supplement**

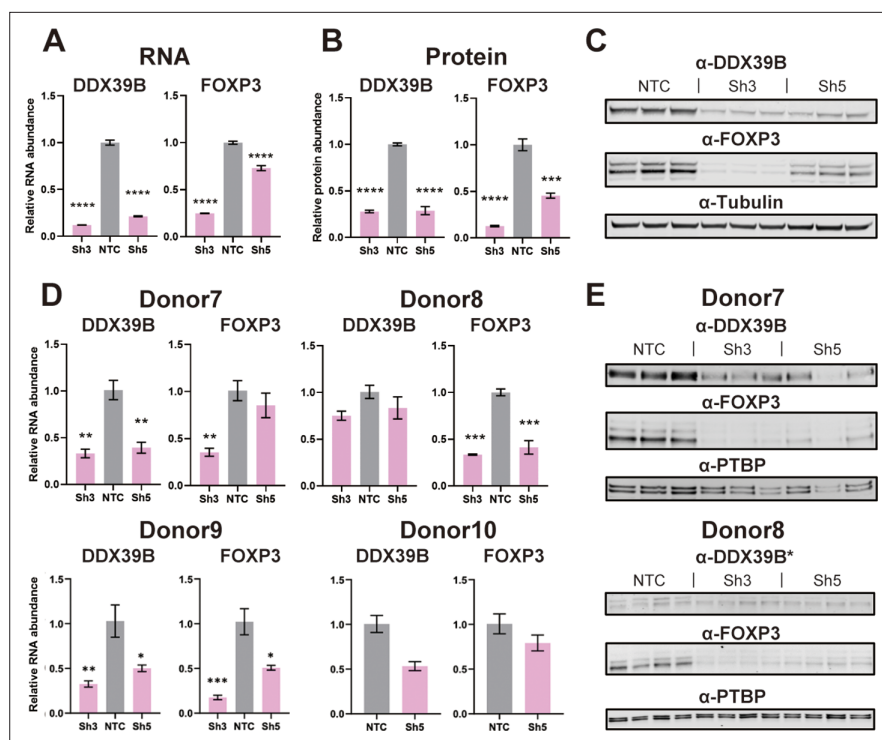


Figure 2. Effect of DDX39B depletion on FOXP3 RNA and protein expression. **(A)** Quantification of DDX39B and FOXP3 RNA between control (NTC) and DDX39B-depleted (Sh3 and Sh5) MT-2 cells. DDX39B and FOXP3 RNA abundance was measured by RT-qPCR and normalized to *EEF1A1* RNA expression. **(B)** Quantification of DDX39B and FOXP3 protein expression between control (NTC) and DDX39B-depleted (Sh3 and Sh5) MT-2 cells from western blot in **(C)**. DDX39B and FOXP3 protein abundance was normalized to alpha-tubulin. **(D)** DDX39B and FOXP3 expression upon DDX39B depletion in primary or induced Tregs. DDX39B was depleted in primary Tregs (Donors 7 and 8) and induced Tregs (Donors 9 and 10) via lentivirus transduction with either non-targeting control (NTC) or DDX39B targeting (Sh3 and/or Sh5) shRNAs. DDX39B and FOXP3 RNA abundance was measured by RT-qPCR and normalized to *EEF1A1* RNA expression. **(E)** Western blot of DDX39B and FOXP3 expression between control (NTC) and DDX39B-depleted (Sh3 and Sh5) primary Tregs. PTBP was used as a loading control. Error bars indicate SEM except for NTC in Donor 10 where error bars represent range due to loss of one replicate sample; *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$.

The online version of this article includes the following source data for figure 2:

Source data 1. Two versions of source data files have been uploaded for all gels/blots shown in this figure: (1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labeled.

1). To confirm regulation of FOXP3-driven gene expression networks by DDX39B in an independent data set, we explored the ConnectivityMap (CMap) database provided by The Broad Institute (*Subramanian et al., 2017*). CMap is a large-scale catalog of transcriptional responses of human cells to chemical and genetic perturbations that enables the identification of conditions causing similar transcriptional responses. We found that the transcriptional profile of DDX39B knockdown (ID: CGS001-7919) showed significant overlap with that of FOXP3 knockdown (ID: CGS001-50943, *Supplementary file 5*). The GSEA and CMap results suggested that the gene expression networks of DDX39B and FOXP3 are tightly linked, which is consistent with robust FOXP3 expression requiring DDX39B.

We evaluated the effect of DDX39B knockdown on well-characterized FOXP3 targets, and established that some targets (e.g. *IL2RA* (CD25)) were significantly downregulated in DDX39B-depleted CD4⁺ T cells and primary Tregs (*Figure 3B*). Other DDX39B targets behaved differently in these two cell populations, for instance *EBI3*, *ICAM1*, and *TNFRSF9* were consistently downregulated by DDX39B depletion in CD4⁺ T cells but not in Tregs (*Figure 3C, Figure 3—figure supplements 2–3*). These observations are consistent with the existence of FOXP3-dependent and FOXP3-independent pathways altered by DDX39B knockdown.

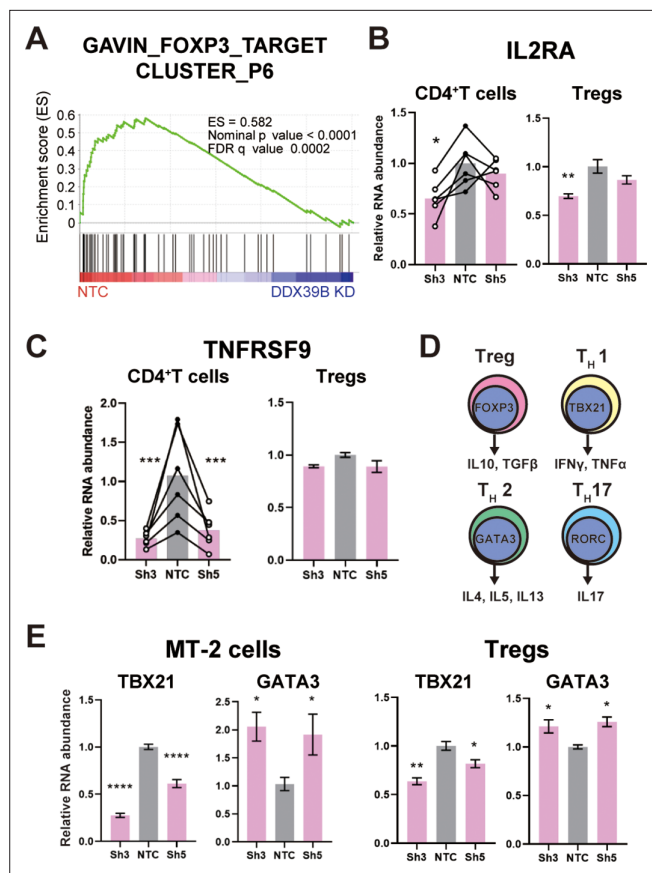


Figure 3. DDX39B depletion causes loss of Treg gene expression signature. **(A)** GSEA results of GAVIN_FOXP3_TARGET_CLUSTER_P6 gene set enriched in normal over DDX39B depleted CD4⁺ T cells **(B)** RNA abundance of IL2RA (CD25) relative to EEF1A1 in DDX39B-depleted CD4⁺ T cells (Left) and primary Tregs (Right). **(C)** RNA abundance of TNFRSF9 relative to EEF1A1 in DDX39B-depleted CD4⁺ T cells (Left) and primary Tregs (Right). **(D)** Schematic of genes expressed in different T cell lineages: Tregs, T_H17, T_H1, and T_H2 effector cells. **(E)** RNA abundance of transcription factors regulating T cell differentiation: TBX21 (T_H1), GATA3 (T_H2) in MT-2 cells (Left) and primary Tregs (Right). *: p<0.05, **: p<0.01, ***: p<0.001 and ****: p<0.0001.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. GSEA results of FOXP3-related gene sets enriched in normal over DDX39B-depleted cells.

Figure supplement 2. Effect of DDX39B depletion on expression of genes in CD4⁺ T cells.

Figure supplement 3. Effect of DDX39B depletion on expression of Treg genes.

We explored the effect of DDX39B knockdown on the expression of important transcriptional regulators of different T cell lineages (**Figure 3D**). DDX39B knockdown in MT-2 cells and primary Tregs decreased expression of TBX21, a transcriptional regulator of T_H1 cells, but increased expression of GATA3, a transcriptional regulator of T_H2 cells (**Figure 3E**). Given the complex composition of T cell populations (*Josefowicz et al., 2012*) these results have to be carefully interpreted; nonetheless, they suggest DDX39B depletion led to gene expression patterns associated with T_H2 cells.

Low DDX39B exacerbates retention of FOXP3 introns

We investigated the mechanistic basis for the dependence of FOXP3 RNA accumulation on DDX39B. Since DDX39B plays important roles in constitutive splicing (*Fleckner et al., 1997; Kistler and Guthrie, 2001; Shen et al., 2008*), alternative splicing (*Galarza-Muñoz et al., 2017; Nakata et al., 2017*), and nucleo-cytoplasmic transport (*Gatfield et al., 2001; Huang et al., 2018; Jensen et al., 2001; Luo et al., 2001*) we investigated which of these roles was affected in FOXP3 transcripts. We depleted DDX39B in the Treg-like MT-2 cells and biochemically fractionated cytoplasm, nucleoplasm and chromatin (**Figure 4A**) to determine which cellular compartment(s) showed reduced FOXP3 transcripts. In

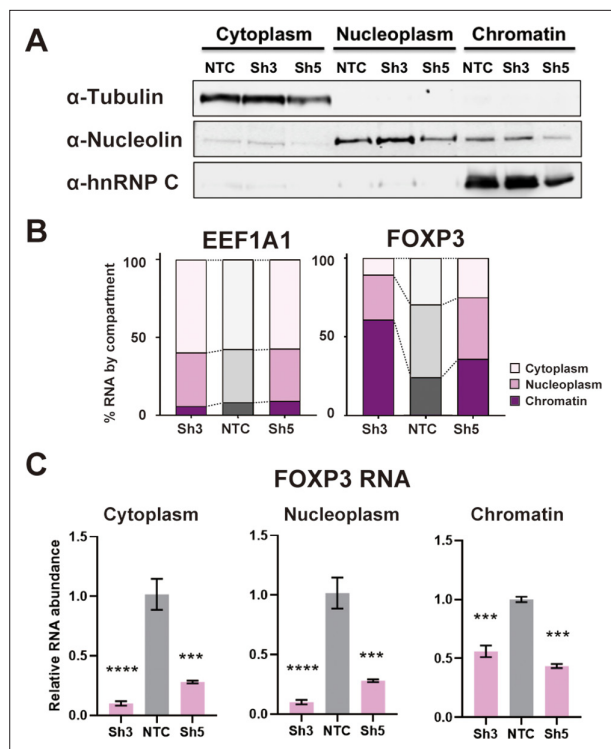


Figure 4. DDX39B depletion disturbs an early step in FOXP3 RNA biogenesis. (A) Protein abundance of subcellular compartment markers in fractionated control (NTC) or DDX39B-depleted (Sh3 and Sh5) MT-2 cells. (B) Percent EEF1A1 or FOXP3 RNA in subcellular compartments. (C) FOXP3 RNA abundance relative to EEF1A1 in subcellular compartments upon DDX39B depletion.

The online version of this article includes the following source data for figure 4:

Source data 1. Two versions of source data files have been uploaded for all gels/blots shown in this figure: (1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labelled.

control conditions, FOXP3 transcripts were disproportionately found to be associated with chromatin or in the nucleoplasm relative to transcripts coding for the translation elongation factor EEF1A1, and this association was increased for FOXP3 but not EEF1A1 RNAs upon DDX39B knockdown (**Figure 4B**). A significant reduction of FOXP3 RNAs was clearly observed in all three cellular fractions and the dramatic reduction in both nucleoplasm and cytoplasm argues against an effect on nucleocytoplasmic transport (**Figure 4C**). These data indicate that the effect of DDX39B knockdown on FOXP3 transcripts occurred early during their biogenesis.

Given this early effect, we analyzed changes in FOXP3 RNA splicing. The RNAseq data suggested that DDX39B knockdown in CD4⁺ T cells led to increased retention of FOXP3 introns (**Figure 5A**). The DDX39B-sensitive retention of multiple FOXP3 introns was confirmed using FOXP3 intron-specific RT-qPCR from total primary CD4⁺ T cell RNA (**Figure 5B**), total and chromatin-associated MT-2 RNA (**Figure 5C** and **Figure 5—figure supplement 1**), and total Treg RNA (**Figure 5—figure supplement 2**).

Retention of FOXP3 introns 2, 4, 6, 7, 9, and 11 in DDX39B depleted MT-2 cells was rescued by expression of recombinant DDX39B (**Figure 5D and E**). FOXP3 protein expression was also rescued although to a lesser extent than intron retention (**Figure 5D**). The lack of strong rescue of FOXP3 protein expression is likely due to transcriptional downregulation of FOXP3 secondary to the sustained knockdown of FOXP3 preceding the expression of DDX39B. Nonetheless, given the robust rescue of the splicing phenotype we concluded that DDX39B is required for efficient intron removal for several FOXP3 introns and this results in a dramatic decrease in overall levels of FOXP3 RNA and protein.

To ascertain if changes in intron retention were seen for other DDX39B targets, we examined global changes in alternative splicing upon DDX39B knockdown in our CD4⁺ T cell RNAseq data.

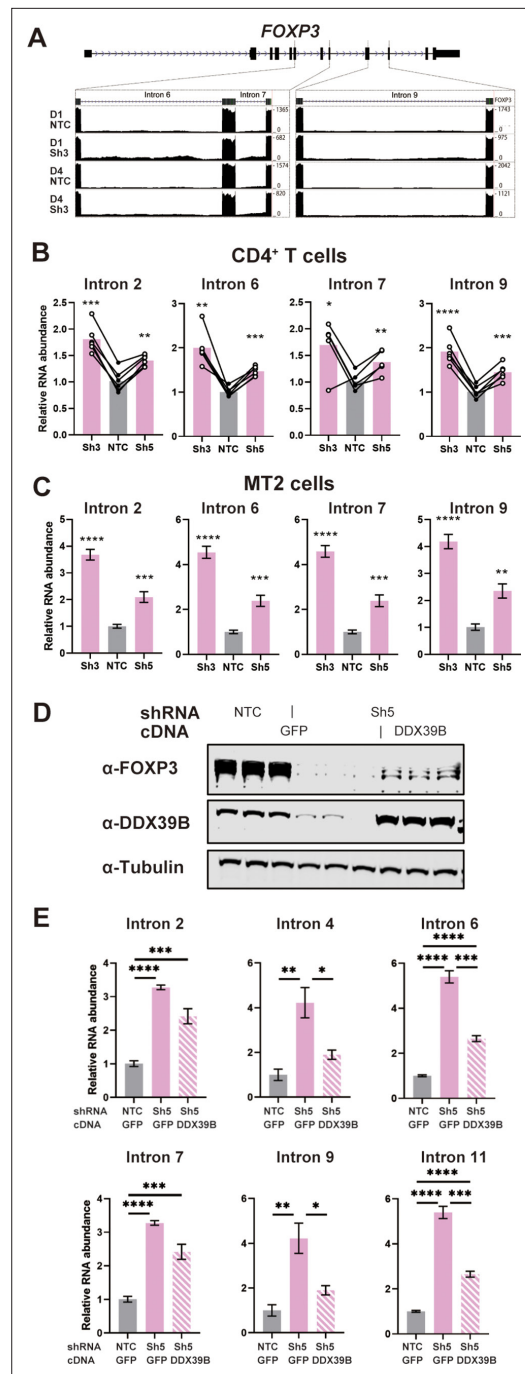


Figure 5. DDX39B depletion triggers the retention of FOXP3 introns. (A) RNAseq reads mapping to the FOXP3 genomic region in control (NTC) or DDX39B depleted (Sh3) CD4⁺ T cells from Donors 1 and 4. Read counts for FOXP3 introns 6, 7, and 9 are shown on the Y-axis in the two insets. (B, C) Abundance of FOXP3 RNA introns relative to total FOXP3 RNA after DDX39B depletion in CD4⁺ T cells from six donors (B) or MT-2 cells (C). (D–E) Rescue of the DDX39B depletion (Sh3) by exogenous expression (GFP or DDX39B) in MT-2 cells. The abundance of FOXP3 and DDX39B relative to Tubulin (D) or FOXP3 RNA introns relative to total

Figure 5 continued on next page

Figure 5 continued

FOXP3 RNA (E) are shown. *: p<0.05, **: p<0.01, ***: p<0.001 and ****: p<0.0001.

The online version of this article includes the following source data and figure supplement(s) for figure 5:

Source data 1. Two versions of source data files have been uploaded for all gels/blots shown in this figure: (1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labeled.

Figure supplement 1. Quantification of FOXP3 retained introns in the chromatin fraction of MT-2 cells.

Figure supplement 2. Quantification of FOXP3 retained introns in primary and induced Tregs.

Figure supplement 3. Alternative splicing events observed upon DDX39B depletion in CD4⁺ T cells.

First and foremost, we noted that only a limited number of splicing events detected by RNAseq were significantly altered by DDX39B depletion (Figure 5—figure supplement 3). As with transcript level changes, we noted more alternative splicing events changing with Sh3 than Sh5, and more with Donor 1 than Donor 4 (Figure 5—figure supplement 3). We detected very few events in common for all conditions, perhaps because of the inefficient knockdown of DDX39B observed in the libraries from cells treated with Sh5 and the well-documented heterogeneity of alternative splicing between individuals (Wang et al., 2008). Therefore, most of our analysis focused on events changed by Sh3 in CD4⁺ T cells from Donors 1 and 4 (Supplementary file 6). Retained introns were the most frequently observed type of RNA splicing event changed (n=784) with increased retention in 397 introns and decreased retention in 387 (Figure 5—figure supplement 3). The second most common changes were in cassette exon use (n=488), followed by alternative use of 5' and 3' splice sites (Figure 5—figure supplement 3). We conclude that splicing of a subset of genes is sensitive to DDX39B levels and this is primarily due to modulation of intron retention as observed for FOXP3 transcripts.

FOXP3 introns have C-rich py tracts

Since removal of FOXP3 introns was sensitive to DDX39B levels and since FOXP3 is on the X chromosome, we compared FOXP3 splice sites to those found in introns of other protein coding genes on the X chromosome. 5' splice sites in FOXP3 introns were only modestly weaker than those found in the average gene, whereas 3'

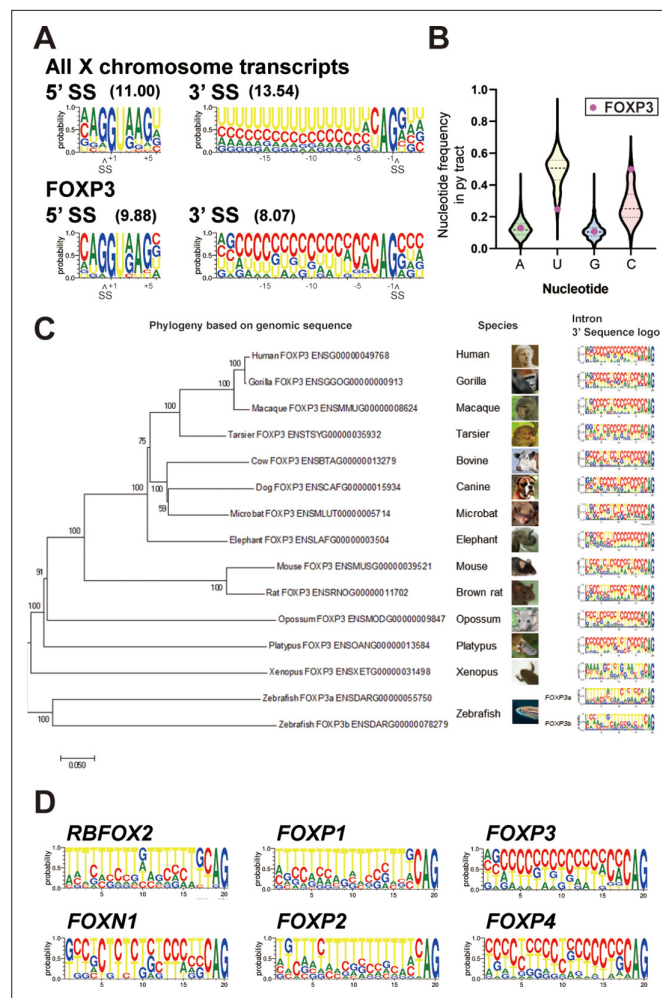


Figure 6. *FOXP3* introns have conserved C-rich py tracts. **(A)** Sequence logos of the 5' and 3' splice sites (SS) of *FOXP3* versus all other chromosome X genes. 5' or 3' MaxEntScore based on the highest-probability sequence are shown. **(B)** Distribution of the py tract nucleotide frequency among chromosome X genes. Magenta dots indicate *FOXP3*. **(C)** Phylogenetic tree and 3' SS sequence logos of the *FOXP3* among 14 species. Zebrafish has two *FOXP3* genes (*FOXP3a* and *FOXP3b*). **(D)** 3' SS sequence logos of FOX family genes. *RBFOX2* is shown as an outgroup.

splice sites were considerably weaker (**Figure 6A**). This correlated with an altered composition of the polypyrimidine tract (py tract): *FOXP3* introns contain uracil (U)-poor and cytosine (C)-rich py tracts, which on the average are more likely to have a C at each position, except positions -19 and -20 (measured relative to the guanine (G) in the 3' splice site AG marked as position -1) (**Figure 6A and B** and **Supplementary file 7**). This contrasts with our analysis of py tracts in other genes on the X chromosome (**Figure 6A and B**) and previous analysis for all human introns spliced by the major (U2 snRNP dependent) spliceosome (**Yeo and Burge, 2004**), where the average position in the py tract is most likely to be occupied by U.

C-rich py tracts in *FOXP3* introns are conserved in mammals from both monotreme and theriiformes subclasses, which diverged over 200 million years ago (**Tarver et al., 2016; Figure 6C** and **Supplementary file 7**). Furthermore, the amphibian *Xenopus laevis* *FOXP3* has a U-poor py tract, although in this case the Us are replaced by all three other nucleotides (**Figure 6C**). The zebrafish *Danio rerio* has two *FOXP3* paralogs and py tracts in one of these (*FOXP3a*) trend towards the C-rich py tracts of mammals, while the other (*FOXP3b*) has U-rich tracts (**Figure 6C**). An examination of other FOX family genes revealed that close relatives of *FOXP3* diverge in their py tracts, *FOXP4* shares C-rich py tracts, while *FOXP1* and *FOXP2* have U-rich tracts (**Figure 6D**). *RBFOX2*, which encodes for a RNA binding protein and not a member of the FOX family of transcription factors, is shown as a phylogenetic outlier with U-rich py tracts (**Figure 6D**). U-poor and C-rich py tracts appear to be

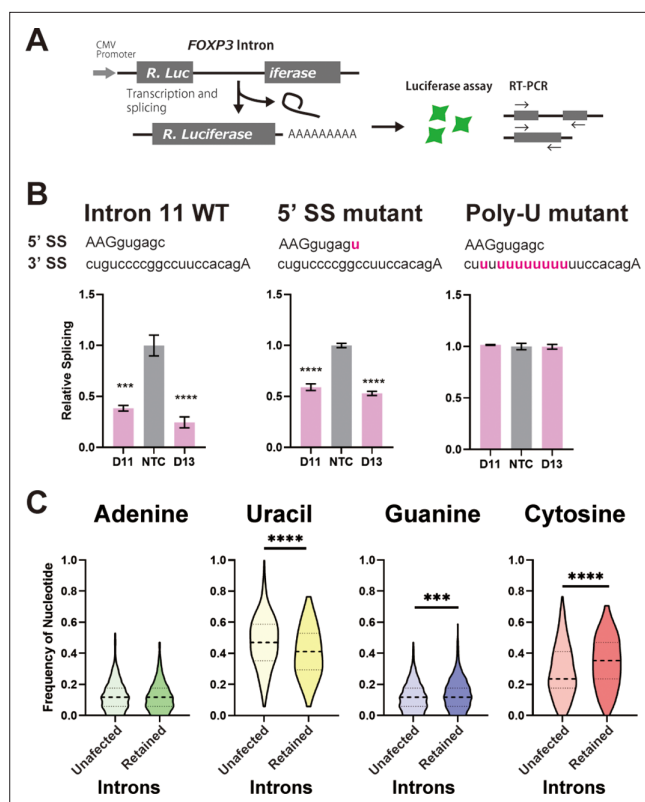


Figure 7. C-rich py tracts in *FOXP3* introns determine DDX39B sensitivity. **(A)** Schematic diagram of the splicing reporter. **(B)** Relative splicing efficiency of the wild-type (WT) or mutant (5' SS mutant and Poly-U mutant) *FOXP3* intron 11 splicing reporters after DDX39B depletion with D11 and D13 siRNAs. The sequence of 5' splice sites (5' SS) for these introns is shown indicating the three upstream exonic residues in capital letters and the six downstream intronic residues in minuscule letters. The sequence of the py tracts and 3' splice sites (3' SS) for these introns is shown indicating the intronic residues in minuscule letters and a single exonic residue in capital letters. **(C)** Nucleotide frequency in the py tract of introns that are insensitive (unaffected) or sensitive (retained) to DDX39B depletion. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$ and ****: $p < 0.0001$.

The online version of this article includes the following source data and figure supplement(s) for figure 7:

Figure supplement 1. Luciferase activity of the splicing reporters.

Figure supplement 2. RT-PCR analysis of splicing efficiency of wild-type or mutant *FOXP3* Intron 11 RLuc reporters.

Figure supplement 2—source data 1. Two versions of source data files have been uploaded for all gels/blots shown in this figure: (1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labeled.

Figure supplement 3. RT-PCR analysis of splicing efficiency of wild-type or mutant *FOXP3* Intron 7 RLuc reporters.

Figure supplement 3—source data 1. Two versions of source data files have been uploaded for all gels/blots shown in this figure: (1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labeled.

Figure supplement 4. Comparison of Max entropy splice sites (SS) score between unaffected and DDX39B-sensitive introns.

Figure supplement 5. Comparison of Max entropy splice site (SS) scores (top panel) and nucleotide composition of py tracts (bottom panel) of unaffected introns or introns that were more retained in CD4⁺ T cells from either donor 1 or 4 depleted of DDX39B by treatment with Sh3 or Sh5.

Figure supplement 6. Detection of intron retention events in X chromosome genes with C-rich or U-rich py tracts.

Figure supplement 7. Retention of *FOXP1* introns is not increased upon DDX39B depletion in MT-2 cells.

Figure supplement 8. RT-PCR analysis of splicing efficiency of wild-type or mutant *FOXP1* Intron 19 RLuc reporters.

Figure 7 continued on next page

Figure 7 continued

Figure supplement 8—source data 1. Two versions of source data files have been uploaded for all gels/blots shown in this figure: (1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labeled.

Figure supplement 9. Comparison of py tract composition between unaffected and less-retained introns.

the ancestral form in the *FOX* family as py tracts in the more divergent *FOXN1* are more like those in *FOXP3* and *FOXP4* (Figure 6D). The evolutionary conservation of C-rich py tracts in *FOXP3* and other *FOX* family genes suggests important regulatory function.

C-rich py tracts are required for the sensitivity of *FOXP3* introns to DDX39B depletion

Since we had previously shown that replacement of U with C in intronic py tracts inhibits splicing of model pre-mRNA substrates in vitro (Roscigno et al., 1993), we posited that the C-rich py tracts in *FOXP3* introns would make these inefficient and highly sensitive to DDX39B levels. To test this, we made reporter constructs where *FOXP3* introns interrupted a Renilla luciferase ORF, and splicing efficiency could be inferred by luciferase activity or measured directly using RT-PCR (Figure 7A). Importantly, splicing reporters containing *FOXP3* introns were markedly dependent on DDX39B (Figure 7B and Figure 7—figure supplements 1–3). While conversion of the 5' splice site in *FOXP3* introns 7 and 11 to consensus only modestly relieved DDX39B dependency, replacement of their py tracts with U-rich tracts made these introns insensitive to DDX39B knockdown (Figure 7B and Figure 7—figure supplements 2–3). These findings indicated that weak C-rich py tracts are necessary for the strong DDX39B dependency of *FOXP3* introns.

To determine whether or not other DDX39B-sensitive intron retention events shared sequence features with *FOXP3* introns, we analyzed our RNAseq data from DDX39B-depleted CD4⁺ T cells. When 500 randomly selected unaffected introns were compared to 397 introns with increased retention upon DDX39B knockdown with Sh3 in either donor 1 or 4 we found no difference in the 5' splice site maximum entropy score and a modest but statistically significant decrease in the 3' splice site maximum entropy score (Figure 7—figure supplement 4). Since the lower 3' splice site scores for

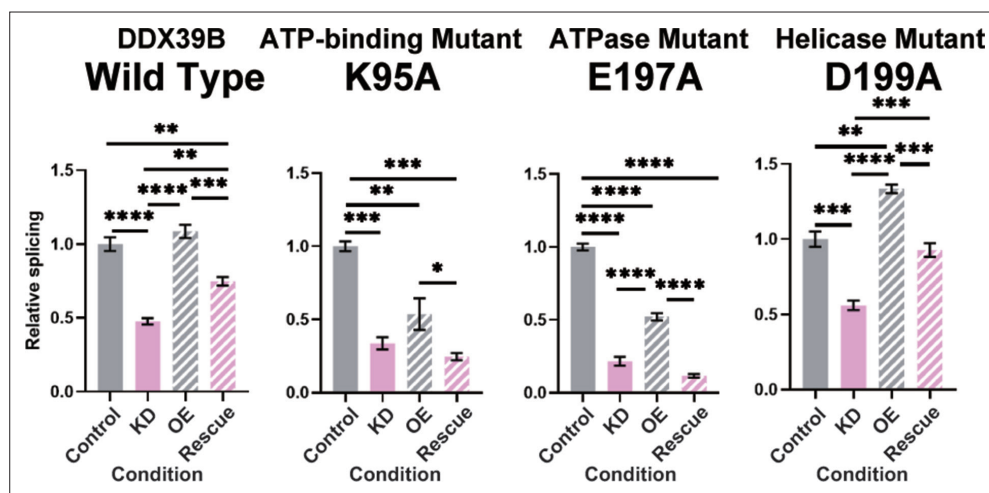


Figure 8. *FOXP3* intron 11 splicing requires DDX39B ATPase activity but not its helicase activity. Rescue of the splicing of *FOXP3* intron 11 reporter by induced expression of wild type (WT) or mutant (K95A, E197A, and D199A) DDX39B. KD and OE indicate knockdown and over-expression, respectively. *: p < 0.05, **: p < 0.01, ***: p < 0.001 and ****: p < 0.0001.

The online version of this article includes the following source data and figure supplement(s) for figure 8:

Figure supplement 1. Expression of WT or mutant DDX39B in stable cell lines.

Figure supplement 1—source data 1. Two versions of source data files have been uploaded for all gels/blots shown in this figure: (1) the original files of the full raw unedited gels or blots and; (2) figures with the uncropped gels or blots with the relevant bands clearly labeled.

FOXP3 introns were driven by the C-rich py tract, we examined the composition of the py tract in DDX39B-sensitive introns. We found that the frequency of U residues in the py tracts of DDX39B sensitive introns was significantly lower than in unaffected ones, and the frequency of C was significantly higher (**Figure 7C**). There was a statistically significant but very small increase in G residues in these U-poor, C-rich py tracts (**Figure 7C**). An equivalent analysis interrogating all intron retention events altered with DDX39B knockdown with both Sh3 and Sh5 led to same conclusion: DDX39B sensitive introns are more likely to have U-poor, C-rich py tracts than unaffected introns (**Figure 7—figure supplement 5**).

To further test our conclusions on the role of C-rich py tracts, we interrogated the behavior of introns in other X chromosome genes selected only because they contained either C-rich or U-rich py tracts. Of twelve introns with C-rich py tracts, in *FAM3A*, *PORCN*, *RBM10*, *RENBP*, *CFP*, and *G6PD*, all but two, both in the *CFP* gene, were more retained upon DDX39B knockdown in MT-2 cells (**Figure 7—figure supplement 6**). In contrast, four U-rich py tract introns in *FMR1* and *DDX3X* were tested and were not more retained upon DDX39B depletion (**Figure 7—figure supplement 6**), thereby supporting the requirement of C-rich py tracts for DDX39B-dependency.

The results above on *CFP* introns suggested that a U-poor, C-rich py tract was not sufficient to confer DDX39B sensitivity. To address this, we analyzed splicing of introns 11, 14, and 19 in the *FOXP1* gene (**Figure 7—figure supplement 7**), which encodes a homologue of *FOXP3* that contains introns with U-rich py tracts (**Figure 6D** and **Figure 7—figure supplement 7**). Retention of *FOXP1* introns 11, 14, and 19 was not increased with DDX39B knockdown in MT-2 cells (**Figure 7—figure supplement 7**; contrast to *FOXP3* introns in **Figure 5C**). We modified the *FOXP3* intron 11 luciferase reporter by replacing the intron with the first and last 75 nucleotides of *FOXP1* intron 19. As expected the *FOXP1* intron 19 reporter was not sensitive to DDX39B knockdown (**Figure 7—figure supplement 8**). We replaced the *FOXP1* intron 19 py tract with that of *FOXP3* intron 11 and noted that this C-rich tract did not confer DDX39B sensitivity on *FOXP1* intron 19 (**Figure 7—figure supplement 8**) indicating that the C-rich tract in *FOXP3* intron 11 was not sufficient for this sensitivity. Consistent with this, *en masse* analysis of the 387 introns that were less retained upon DDX39B depletion revealed a modest increase in C frequency in py tracts suggesting that this is not sufficient to confer DDX39B-dependency (**Figure 7—figure supplement 9**). Therefore, a C-rich py tract was required but not sufficient to impart DDX39B-dependency on an intron.

Collectively, the data presented in this section show that C-rich py tracts are required for the strong DDX39B-dependency of *FOXP3* introns and suggest the existence of a type of introns where the C-rich py tract is required for said dependency.

Splicing of *FOXP3* introns requires the ATPase activity of DDX39B but not its helicase activity

The py tract is recognized early in the splicing reaction by the U2 small nuclear RNA auxiliary factor (U2AF) composed of U2AF1 (U2AF³⁵) and U2AF2 (U2AF⁶⁵) (**Ruskin et al., 1988; Zamore and Green, 1989**). U2AF2 binds preferentially to U-rich sequences in the py tract *in vitro* (**Singh et al., 1995; Zamore et al., 1992**) and this is recapitulated by crosslinking-immunoprecipitation experiments *in vivo* (**Wu and Fu, 2015**). U2AF2 binds DDX39B, which licenses the U2AF-mediated recruitment of U2 snRNP to the neighboring branchpoint sequence to form the pre-spliceosome (**Shen et al., 2008**). We conjectured that the exquisite dependency of C-rich py tract *FOXP3* introns on DDX39B was mediated by its ability to promote pre-spliceosome formation. Pre-spliceosome formation requires the ATPase activity of DDX39B, but not its helicase activity, which is required for later steps in spliceosome assembly (**Shen et al., 2008**). While expression of wildtype (WT) DDX39B rescued splicing of *FOXP3* intron 11 in DDX39B depleted cells, two ATPase defective mutants, K95A and E197A (**Shen et al., 2007**), did not rescue splicing (**Figure 8** and **Figure 8—figure supplement 1**). On the other hand, the helicase defective DDX39B mutant D199A that retains ATPase activity and supports pre-spliceosome formation (**Shen et al., 2008; Shen et al., 2007**) rescued *FOXP3* intron 11 splicing (**Figure 8** and **Figure 8—figure supplement 1**). Of note, the data in **Figure 8** indicate that the DDX39B D199A mutant rescued *FOXP3* intron 11 splicing even better than the WT DDX39B. This result could be explained by the fact that while the D199A lacks helicase activity, it has higher ATPase activity than WT DDX39B (**Shen et al., 2007**). These results indicate that the C-rich py tract of *FOXP3* intron 11 is highly dependent on the DDX39B ATPase activity and this unique dependency may distinguish its

spliceosome assembly pathway from more conventional introns. Nine of eleven FOXP3 introns belong to the C-rich py tract type, and are all likely to share this requirement for the DDX39B ATPase activity, which explains why overall FOXP3 expression is exquisitely sensitive to DDX39B levels.

Discussion

DDX39B (also known as BAT1) was proposed to be an anti-inflammatory factor (Allcock et al., 2001) and we have shown that *DDX39B* is genetically associated with MS, an autoimmune disease (Galarza-Muñoz et al., 2017). Previously we discovered a molecular mechanism that partially explained these phenomena. DDX39B is a potent activator of IL7R exon 6 splicing and a repressor of the production of sIL7R (Galarza-Muñoz et al., 2017), which exacerbates the MS-like EAE (Lundström et al., 2013). Complementing this molecular analysis, we demonstrated epistasis between alleles in DDX39B and IL7R and MS risk, with individuals homozygous for the risk alleles at both loci having an approximately 3-fold higher chance of developing MS (Galarza-Muñoz et al., 2017). While this connection between DDX39B and autoimmunity was compelling, evolutionary arguments suggested that this RNA helicase would play other roles in immunity. Our exploration of this suggestion led to the discovery, described above, that DDX39B is exquisitely required for the splicing of FOXP3 transcripts. Low levels of DDX39B lead to low levels of FOXP3 and by inference low numbers of functional Tregs.

Our work highlights the importance of post-transcriptional mechanisms in the regulation of FOXP3. These have not been noted before and indeed CRISPR screens to discover regulators of FOXP3 did not discover DDX39B (Cortez et al., 2020; Loo et al., 2020; Schumann et al., 2020), but this is likely because knockout of DDX39B results in lethality.

The connection between FOXP3 and autoimmunity is well established: mutations in FOXP3 cause a systemic and severe autoimmune syndrome in humans called immunodysregulation polyendocrinopathy enteropathy X-linked (IPEX) syndrome (Bennett et al., 2001; Chatila et al., 2000; Wildin et al., 2001) and the equivalent scurfy in mice (Brunkow et al., 2001). While there is evidence supporting a prominent role for FOXP3 in MS (Fletcher et al., 2009; Sambucci et al., 2018), genetic variation within FOXP3 has yet to be associated with MS in large genome wide associations studies largely due to the fact that the X chromosome has been understudied in human genetic association studies. The role of Tregs, which depend on FOXP3 transcriptional control (Hori et al., 2003), on autoimmunity has been carefully documented (Dominguez-Villar and Hafler, 2018). Furthermore, it is clear that high expression of FOXP3 is required for full Treg function and there is a FOXP3 dose-dependent effect on immune suppressive activity (Wan and Flavell, 2007). Therefore, given the strong dependence of FOXP3 expression on DDX39B levels, it is likely that humans with low levels of DDX39B will have reduced FOXP3 expression and by extension low levels of Treg cell development, maintenance and suppressive function. Furthermore, given our data that DDX39B knockdown promotes GATA3 expression it is possible that low levels of DDX39B could promote the conversion of Tregs to T_H2-like self-reactive effector cells (Komatsu et al., 2014; Noval Rivas et al., 2015; Zhou et al., 2009). The DDX39B control of FOXP3 expression cements its role in immune regulation.

There are other intriguing connections between DDX39B and immunity. First, we show that putative MS susceptibility genes are regulated by DDX39B and some are likely regulated independently of effects on FOXP3. Second, DDX39B regulates the nucleocytoplasmic transport of human circular RNAs (circRNAs) larger than 1200 nucleotides (Huang et al., 2018). There is an emerging realization that circRNAs play important, although heretofore incompletely understood roles, in immunity and autoimmune disease (Zhou et al., 2019). Third, *DDX39B* resides in the class III region of the MHC and many of its neighbors code for proteins involved in RNA transactions (e.g. *DXO* and *SKIV2L*) (Lehner et al., 2004; Schott and Garcia-Blanco, 2021). This overrepresentation of genes involved in RNA metabolism within the class III region is found in all vertebrates we have examined, suggesting an important immune function for these RNA binding proteins. These observations, first revealed by studying DDX39B, suggest new and interesting connections between RNA and immunity.

DDX39B, which is also known as UAP56, was characterized as a factor required for formation of splicing complexes (Fleckner et al., 1997; Shen et al., 2008). Here we describe a type of introns that have C-rich py tracts and depend on high levels of DDX39B for splicing. Although these py tracts are poor binding sites for U2AF2 (Singh et al., 1995), we propose that DDX39B-driven U2 snRNP binding to the branchpoint sequence stabilizes U2AF2 at these suboptimal sites. We suggest that for these introns the U2 snRNP complex may be the first splicing commitment complex, which

would be distinct from U2 snRNP-independent commitment complexes in introns with U-rich py tracts (Jamison *et al.*, 1992; Li *et al.*, 2019; Seraphin and Rosbash, 1989). Our studies indicate that these introns are exceedingly sensitive to DDX39B ATPase activity, but much less so to its helicase activity. Shen *et al.* showed that DDX39B ATPase activity, but not helicase activity, is required for binding of U2AF and recruitment of U2 snRNP to the branchpoint sequence of the 3' splice site to form the pre-spliceosome (Shen *et al.*, 2008). This suggests that once bound to the branchpoint sequence and py tract in these C-rich introns, the U2 snRNP-U2AF complex does not require further remodeling by DDX39B. In contrast to this, activation of IL7R exon 6 requires DDX39B helicase activity (Galarza-Muñoz *et al.*, 2017). Our data thus indicate distinct functions for DDX39B in different splicing events, which implies multiple pathways to spliceosome assembly as has been suggested before (Kistler and Guthrie, 2001; Newnham and Query, 2001). An intriguing possibility is that FOXP3 introns, and perhaps many introns with C-rich py tracts, are actually prone to intron 'detention' rather than retention (Boutz *et al.*, 2015) and that certain stimuli may free them from DDX39B dependency.

Materials and methods

Cells

MT-2 cells were kindly provided by Bryan R. Cullen (Duke University) and grown at 37 °C in Roswell Park Memorial Institute (RPMI) 1640 medium (Thermo Fisher Scientific) with 10% (v/v) fetal bovine serum (FBS) (Genesee Scientific) and 1% (v/v) penicillin-streptomycin (Thermo Fisher Scientific). Human embryonic kidney 293T/17 (HEK 293T/17; ATCC CRL-11268) and HeLa (ATCC CCL-2) cell lines were obtained from the Duke University Cell Culture Facility. The HeLa Flp-In T-Rex cell line (HeLa-Flp-In) was kindly provided by Dr. E. Dobrikova (Duke University). HEK293T/17, HeLa, and HeLa-Flp-In cells were cultured at 37 °C in Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher Scientific) with 10% (v/v) FBS, 1% (v/v) penicillin-streptomycin and 2.5 µg/ml Plasmocin (InvivoGen). HeLa-Flp-In cell lines expressing wild-type or mutant versions of DDX39B (K95A, E197A and D199A) were grown as described above except media was supplemented with 2.5 µg/ml blasticidin (InvivoGen) and 200 µg/ml Hygromycin B (Thermo Fisher Scientific). All cells were free of Mycoplasma contamination as confirmed by routine testing using the MycoAlert Mycoplasma Detection Kit (Lonza).

Primary human CD4⁺ T cells were isolated and cultured as described previously (Galarza-Muñoz *et al.*, 2017). In brief, peripheral blood mononuclear cells (PBMCs) from healthy donors were isolated from whole blood or buffy coats using the Ficoll-Plaque Gradient method. Blood samples were collected at Duke University following the institutional IRB protocol (# Pro00070584) and buffy coats were purchased from the New York Blood Center. CD4⁺ T cells were further isolated using CD4⁺ T cell isolation kit (Miltenyi Biotec). The isolated cells were cultured in RPMI medium supplemented with 20% (v/v) FBS and 100 ng/mL human recombinant IL-2 (PeproTech). Two days prior to transduction, CD4⁺ T cells or MT-2 cells were activated with media containing 50 ng/mL anti-CD3 (eBioscience) and 100 ng/mL anti-CD28 (BD Biosciences) antibodies.

Primary human T regulatory cells (Tregs) were isolated from PBMCs using CD4⁺ CD25⁺ CD127^{dim/-} Regulatory T Cell Isolation Kit (Miltenyi Biotec). The isolated Treg cells were expanded in culture for 14 days with TexMACS medium (Miltenyi Biotec) supplemented with 5% (v/v) Human AB Serum (Sigma), 500 IU/mL recombinant human IL2 (PeproTech) and Treg Expansion kit (Miltenyi Biotec) which contains beads pre-loaded with anti-CD3 and anti-CD28 antibodies. Induced T regulatory cells (iTregs) were purchased from IQ Biosciences (IQB-Hu1-iTr-1) and were expanded in culture for 6 days using the same media and Treg Expansion kit as the primary human Tregs.

Plasmids

shRNA pLKO.1 plasmids expressing non-targeting control shRNA (NTC: SHC002) or anti-DDX39B shRNAs (Sh3: TRCN0000286976; and Sh5: TRCN0000294383) were purchased from Millipore Sigma. The construct pLCE-DDX39B was generated by cloning the coding sequence of *DDX39B* into the pLCE vector. Splicing reporter plasmids (pcDNA3.1-RLuc_FOXP3 Intron 7, pcDNA3.1-RLuc_FOXP3 Intron 9, pcDNA3.1-RLuc_FOXP3 Intron 11, and pcDNA3.1-RLuc_HGB1 Intron 2) were constructed by inserting introns 7, 9 or 11 of FOXP3 (NC_000023), or intron 2 of HGB1 (NC_000011) within the open reading frame of *Renilla Luciferase* (Rluc) at nucleotide positions 599, 601, 408, or 253, respectively, in the pcDNA3.1 vector (pcDNA3.1-Rluc). Splice sites (SS) mutations were introduced in the

corresponding parental plasmid using In-Fusion HD Cloning kit (Takara Bio) to generate pcDNA3.1-Rluc_FOXP3 Intron 7 5'-SS G-U, pcDNA3.1-Rluc_FOXP3 Intron 7 3'-SS PolyU, pcDNA3.1-Rluc_FOXP3 Intron 11 5'-SS G-U, and pcDNA3.1-Rluc_FOXP3 Intron 11 3'-SS PolyU. To construct pcDNA3.1-Rluc_FOXP1 Intron 19 WT and 3' C-rich py tract mutant, synthesized DNA fragments with the FOXP1 intron sequences (NG_028243.1) were purchased from Integrated DNA Technologies and were subcloned into the pcDNA3.1-Rluc plasmid. Point mutants DDX39B K95A, E197A and D199A were introduced in the parental pcDNA5/FRT/TO-DDX39B plasmid (*Galarza-Muñoz et al., 2017*) using QuickChange Lightning Mutagenesis Kit (Agilent Technologies). All constructs and mutations were confirmed by Sanger sequencing. Primers used to generate these constructs are shown in **Supplementary file 8**.

Establishment of stable cell lines

Generation of the inducible HeLa cell line expressing wild-type DDX39B or DDX39B mutant D199A was described previously (*Galarza-Muñoz et al., 2017*). HeLa cell lines stably expressing DDX39B mutants (K95A and E197A) were constructed similarly. In brief, HeLa Flp-In T-Rex cells were co-transfected with these constructs in the pcDNA5/FRT/TO plasmid and pOG44 plasmid, which encodes Flp recombinase, using Lipofectamine 2000 (Thermo Fisher Scientific). Transfected cells were selected with 2.5 µg/ml blasticidin and 200 µg/ml Hygromycin B for 15 days. Expression of the transgene was induced by addition of 1 µg/mL doxycycline in the culturing media.

Lentiviral packaging

Lentiviral packaging of shRNA constructs was conducted in 293T/17 cells using Lipofectamine 2000 (Thermo Fisher Scientific). In brief, 1.0×10^7 293 T/17 cells were seeded in 15 cm dishes in DMEM media and cultured overnight (three 15 cm dishes per construct). Cells were co-transfected with 17 µg of the corresponding shRNA pLKO.1 vector, pLCE-DDX39B or pLCE-GFP, 17 µg of packaging plasmid (pCMVR8.74) and 7 µg of VSV-G envelope plasmid (pMD2.G) in serum-free media and 160 µL of Lipofectamine 2000. The media was replaced with 20 mL fresh DMEM media 18 hr post-transfection. After 72 hr, supernatants were collected, filtered through 0.45 mm filters and concentrated to 6 mL in Amicon Ultra 100 K centrifugal filter units (Millipore Sigma). Concentrated lentiviral particles were used immediately or stored at -80°C .

Lentiviral transduction of human primary CD4⁺ T cells, MT-2 cells and primary Tregs

Transduction of human primary CD4⁺ T cells with either NTC or anti-DDX39B shRNAs was described previously (*Galarza-Muñoz et al., 2017*). In brief, 4.0×10^6 activated primary CD4⁺ T cells from each donor were transduced in T25 flasks with lentiviruses encoding control (NTC) or DDX39B-targeting (Sh3 or Sh5) shRNAs for 3 days, and transduced cells were selected under medium supplemented with 1.5 µg/mL puromycin for 4 days. Cells were then cultured for 24 hr in the absence of puromycin and collected for functional analyses (8 days after initial transduction). Transduction of activated MT-2 cells (1.0×10^7 cells) was carried out as for primary CD4⁺ T cells. Activated primary Tregs (1.0×10^5 cells) were transduced in 96-well plates and cultured for 7 days with half media exchange every 2 days (without puromycin selection). Induced Tregs (1.0×10^5 cells) were transduced in 96 well plates with Sh3 and/or Sh5 for two days then subjected to puromycin selection for two days. Cells were then cultured for an additional 24 hr in the absence of puromycin and harvested for functional analyses. In all cell types, depletion of DDX39B protein was confirmed by western blot using anti-UAP56 (DDX39B) antibody (ab18106, Abcam) and α -Tubulin (AB_1904178, Cell Signaling Technology), PTBP1 (anti-PTB rabbit serum *Wagner and Garcia-Blanco, 2002*), or α -beta-actin (sc-47778, Santa Cruz) as loading control.

DDX39B rescue in MT-2 cells

To rescue DDX39B expression in MT-2 cells, 1.0×10^7 cells were co-transduced with the following combination of shRNAs and expression plasmids for 3 days: (i) NTC shRNA +pLCE GFP (control), (ii) Sh5 shRNA +pLCE GFP (knockdown), and (iii) Sh5 shRNA +pLCE-DDX39B (rescue). The cells were then selected with RPMI medium supplemented with 1.5 µg/mL puromycin for 2 days, followed by 24 hr culture in the absence of puromycin, and collection of cell lysates for RNA and protein analyses (6 days after initial transduction). We used the DDX39B shRNA Sh5 since it targets the 3' untranslated region of DDX39B mRNA, and thus it depletes the endogenous DDX39B mRNAs but not the

trans-gene DDX39B mRNAs. Depletion or rescue of DDX39B protein was confirmed by western blot using anti-UAP56 (DDX39B) antibody (ab18106, Abcam) and α -Tubulin (AB_1904178, Cell Signaling Technology) as loading control.

RNA-seq library preparation, sequencing and analyses

Total RNA was isolated from control (NTC) or DDX39B-depleted (Sh3 or Sh5) CD4⁺ T cells using ReliaPrep RNA Cell Miniprep System (Promega), and treated in-column with DNase I following the manufacturer's recommendations. Poly-A⁺ RNA was enriched from 1 μ g of total RNA and used as template to generate libraries using the Illumina TruSeq platform as recommended by the manufacturer. Libraries were sequenced on a 2x100 paired-end format on an Illumina Hi-Seq 1500. Reads were aligned to the human GRCh38 reference with program STAR version 2.5.2b, using the parameters recommended for the ENCODE consortium (*ENCODE Project Consortium, 2011; Dobin et al., 2013*). The STAR genome index was built from the GENCODE primary genome assembly and the corresponding primary annotation file. Gfold software, version 1.1.4, with default parameters was used to count reads per gene and estimate expression differences between treatments. The GENCODE V24 basic annotation file was used for the gene counts. In order to determine significantly changed transcript abundance, an initial cutoff was used to discard transcripts averaging an RPKM of less than 2, then with a secondary cutoff of $|\text{Gfold value}| > 0.3$ or > 0.1 (for Sh3 or Sh5, respectively, compared to NTC from the same donor). A lower cutoff was used for Sh5 since it was less effective than Sh3 at knocking down DDX39B. Finally, only transcripts fulfilling these parameters in both Donor 1 and Donor 4 were included as significantly changed (see *Supplementary file 1*).

Splicing analysis was carried out using Vast-tools program version 0.2.1 (*Irimia et al., 2014*) by aligning the paired-end reads to the Vast-tools human database (vastdb.hsa.7.3.14) using the default parameters. The Sh3-treated sample from each donor were compared using the differential function of vast-tools to determine changes in splicing relative to NTC in the same donor. Subsequently, we used the minimum value of change percent spliced in at 95% confidence interval value (MV|dPSI|; <https://github.com/vastgroup/vast-tools>, copy archived at [swh:1:rev:fb87ceb4379cd9fbd36fc402e4a3429567021548](https://www.swh.io/rev/fb87ceb4379cd9fbd36fc402e4a3429567021548); *Irimia, 2023*) to filter for significantly changed events using 0.15 as cutoff (*Supplementary file 6*).

MS susceptibility gene enrichment analysis

The study by the International Multiple Sclerosis Genetics (IMSG) Consortium (*Patsopoulos et al., 2019*) identified single nucleotide polymorphisms (SNPs) genetically and/or functionally associated with increased MS risk, from which a map of MS susceptibility genes was determined. We created a functional classification of the MS susceptibility genes (the genes with exonic or intronic SNPs, or cis-eQTL, MS_Susceptibility: 558 genes; see supplemental tables 7 & 19 in *Patsopoulos et al., 2019*), in lymphocytes based on expression quantitative trait loci (eQTL). Gene expression profiles of the MS susceptibility genes in EBV-transformed lymphocytes were obtained from GTExPortal (*GTEx Consortium, 2017*) and merged with data of odds ratios (OR) of MS risk to generate risk SNP and susceptibility gene pairs. Pairs of SNPs and genes with missing expression profiles were excluded from the analysis. A list of 539 pairs of SNPs and genes was generated (*Supplementary file 3*). If the MS risk allele (OR >1) showed higher expression of the corresponding susceptibility gene (Normalized expression: NES <0) in the lymphoblastoid cells, that gene was defined as a pathogenic MS gene (MS_Pathogenic: 250 genes). If the MS risk allele (OR >1) showed lower expression of the corresponding susceptibility gene (NES >0), that gene was defined as a protective MS gene (MS_Protective: 262 genes).

In order to estimate enrichment of differentially expressed genes (DEGs) in pre-determined MS gene sets, we used a simulation method. We determined DEGs in two donors with the shRNA that effectively knocked down DDX39B, Sh3, based on the Gfold method (*Feng et al., 2012*) with a cutoff of Gfold value change >0.3. The observed number of DEGs found in the MS gene set, k , was calculated by intersecting DEGs with each MS gene set (MS_Susceptibility, MS_Pathogenic, or MS_Protective). We then calculated enrichment using a resampling method. We constructed a sampling distribution for the null hypothesis (X) by randomly resampling n_1 genes from N, and counting the number of genes overlapping with each MS gene list n_2 over 100,000 permutations. An empirical p

value was calculated based on the fraction of instances that simulated (X) was greater than or equal to the actual observed value of k for each MS gene set.

RT-qPCR analysis of RNA expression

Total RNA was isolated from control or DDX39B-depleted cells using ReliaPrep RNA Cell Mini-prep System (Promega) or Direct-zol RNA kit (Zymo Research), and treated in-column with DNase I following the manufacturer's recommendations. Reverse transcription was conducted with random primers using the High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). DDX39B and FOXP3 RNA levels were measured by real-time quantitative PCR (RT-qPCR) using PowerUP SYBR Green Master Mix (Thermo Fisher Scientific) in a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific). *EEF1A1* was used as normalization control since its expression was not affected by DDX39B depletion.

Genes whose expression was found to be differentially expressed upon DDX39B depletion in the RNAseq dataset in primary CD4⁺ T cells were analyzed by RT-qPCR in control (NTC) or DDX39B-depleted (Sh3 and Sh5) primary CD4⁺ T cells (from 6 donors), MT-2 cells and primary Tregs (from 2 donors). Target genes and primers are shown in **Supplementary file 8**. The data were normalized to *EEF1A1* expression. Differentially expressed genes were determined by comparing levels in DDX39B-depleted cells versus control cells.

Western blot analysis of DDX39B and FOXP3 protein expression

Whole cell lysates from control (NTC) or DDX39B-depleted (Sh3 and Sh5) cells were collected in 1 X RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, and 50 mM Tris-HCl at pH 7.5) freshly supplemented with 1 X protease inhibitors (Roche). The cell lysates were quantified by Bradford and equal amount of total protein was loaded per lane on NuPAGE 4–12% Bis-Tris pre-cast gels (Life Technologies), transferred to nitrocellulose membranes (Whatman), and blotted using standard protocols with antibodies against DDX39B (ab18106, Abcam), FOXP3 (AB_467554, Thermo Fisher Scientific), and either PTB, α -Tubulin (AB_1904178, Cell Signaling Technology) or α -beta-actin (sc-47778, Santa Cruz) as loading control.

Gene-set enrichment analysis (GSEA)

To conduct GSEA analysis with our RNAseq data set in primary CD4⁺ T cells, we generated a gene list from the control condition (NTC: Donor1-NTC, Donor4-NTC) and a gene list from the DDX39B knock-down condition (KD: Donor1-Sh3, Donor4-Sh3). All the genes that passed the initial cutoff (RPKM of ≥ 2 in both NTC libraries) were used. GSEA was conducted with the C2 curated genesets collections of Molecular Signatures Database (**Subramanian et al., 2005**). Among 5501 genesets in the C2 collection, 730 genesets passed the cutoff of the nominal p-value ($P < 0.05$). Of these, 648 and 82 genesets showed enrichment to NTC and KD groups, respectively (**Supplementary file 4**).

Subcellular fractionation of MT-2 cells

Subcellular fractionation of MT-2 cells was conducted following a protocol described previously (**Bhatt et al., 2012**) with minor modification. Control (NTC) or DDX39B-depleted (Sh3 and Sh5) MT-2 cells were washed with Phosphate-buffered saline (PBS), then lysed in NP40 cytoplasmic lysis buffer [0.075% (v/v) NP-40, 20 mM Tris-HCl pH7.5, 150 mM NaCl, 1 mM DTT, and 1 X protease inhibitor] for 2.5 minutes on ice. Cell lysates were then layered on top of a sucrose cushion [24% (w/w) sucrose, 20 mM Tris-HCl pH7.5, 150 mM NaCl, 1 mM DTT, and 1 X protease inhibitor] and centrifuged at 14,000 rpm for 10 min. For further separation of nucleoplasm and chromatin, the pelleted nuclei were treated with glycerol nucleoplasm lysis buffer [50% (v/v) Glycerol, 20 mM Tris-HCl pH7.5, 75 mM NaCl, 0.5 mM EDTA, 1 mM DTT and 1 X protease inhibitor], mixed gently, and treated with urea nuclei lysis buffer [1% (v/v) NP-40, 1 M Urea, 20 mM HEPES pH7.5, 1 mM DTT, 7.5 mM MgCl₂, 0.2 mM EDTA and 1 X protease inhibitor] for 2 min on ice, followed by centrifugation at 14,000 rpm for 2 min. Supernatants were collected as nucleoplasm fraction and the pellets were collected as the chromatin fraction. Successful separation of subcellular compartments was confirmed by western blot with antibodies against markers of the different compartments: cytoplasmic α -Tubulin (AB_1904178, Cell Signaling Technology), nucleoplasmic Nucleolin (AB_533406, Bethyl Laboratories), and chromatin-associated hnRNP C (AB_627731, Santa Cruz Biotechnology). RNA was isolated from each fraction

using Direct-zol RNA kit (Zymo Research), and 200 ng of RNA from each compartment was used for reverse transcription with random primers as above. The abundance of FOXP3 and EEF1A1 RNAs were quantified in each fraction by RT-qPCR and normalized to total RNA as above. The percentage of EEF1A1 and FOXP3 RNA in each compartment was calculated by dividing by the corresponding total signal from the three compartments.

RT-qPCR analysis of retained introns in *FOXP3*, *FOXP1* and other RNAs

To quantify intron retention events, we designed RT-qPCR primers to amplify total FOXP3 transcripts (constitutive exon 11), spliced FOXP3 transcripts (spanning the exon-exon junction between constitutive exons 10 and 11) and retained intron-containing FOXP3 transcripts: intron 2 (spanning the intron 2/exon 3 junction), intron 4 (spanning the intron 4/exon 5 junction), intron 6 (spanning the exon 6/intron 6 junction), intron 7 (spanning the intron 7/exon 8 junction), intron 9 (spanning the exon 9/intron 9 junction), or intron 11 (spanning the intron 11/exon 12 junction) (**Supplementary file 8**). Likewise, we designed primers to amplify total FOXP1 transcripts (within constitutive exon 21) and FOXP1 transcripts with retained introns: intron 11 (spanning intron 11/exon 12 junction), intron 14 (spanning intron 14/exon 15 junction), or intron 19 (spanning intron 19/exon 20 junction). Retained intron values were normalized to the corresponding total FOXP3 or FOXP1 RNA, and the data for DDX39B-depleted (Sh3 or Sh5) primary CD4⁺ T cells, MT-2 cells or primary Tregs is presented as fold-change over control (NTC). For the rescue experiments of endogenous DDX39B-depletion in MT-2 cells, cells transduced with NTC and GFP lentiviruses were used as controls. Detections of retained introns in FAM3A, PORCN, RBM10, RENBP, CFP and G6PD (C-rich py tracts) or FMR1 and DDX3X (U-rich py tracts) were conducted in the same manner. All primers used for this analysis are shown in **Supplementary file 8**.

Phylogenetic analysis of the *FOXP3* gene

The evolutionary history of the *FOXP3* gene was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (*Tamura and Nei, 1993*), and a phylogenetic tree was generated with MEGA7 software (*Kumar et al., 2016*). The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 15 genomic sequences of *FOXP3*, *FOXP3a* and *FOXP3b* gene from different vertebrate organisms: Human_ENSG00000049768, Gorilla_ENSGGOG00000000913, Macaque_ENSMUG00000008624, Tarsier_ENSTSYG00000035932, Microbat_ENSMLUT00000005714.2, Mouse_ENSMUSG00000039521, Rat_ENSRNOG00000011702, Cow_ENSBTAG00000013279, Dog_ENSCAFG00000015934, Elephant_ENSLAFG00000003504, Opossum_ENSMODG00000009847, Platypus_ENSOANG00000013584, *Xenopus*_ENSXETG00000031498, Zebrafish_FOXP3a_ENSDARG00000055750, and Zebrafish_FOXP3b_ENSDARG00000078279.

Likewise, the evolutionary history of the FOX family gene was inferred by using the Maximum Likelihood method as described above. The analysis involved 12 protein-coding sequences of human FOX family genes (FOXA1: ENST00000250448.3, FOXD1: ENST00000615637.3, FOXE1: ENST00000375123.4, FOXF1: ENST00000262426.6, FOXH1: ENST00000377317.4, FOXN1: ENST00000226247.2, FOXO1: ENST00000379561.6, FOXP1: ENST00000318789.9, FOXP2: ENST00000350908.9, FOXP3: ENSG0000004976, FOXP4: ENST00000373063.7, and RBFOX2: ENST00000449924.6).

FOXP3 splicing reporter assays

A total of 2.0×10^5 HeLa cells were transfected in 12-well plates with control (NTC: all-stars non-targeting control, Qiagen) or DDX39B-targeting (D11: Hs_BAT1_11, and D13: Hs_BAT1_13, Qiagen) siRNAs using Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher Scientific). After two days of culture, the Renilla luciferase (RLuc) splicing reporter plasmids were co-transfected with a Firefly luciferase plasmid (pGL3 control-FLuc, Promega) as transfection control. After 24 hr, cell lysates were collected for measurements of luciferase activity and RNA isolation. RLuc and FLuc activities

were measured by Dual-Luciferase Reporter Assay System (Promega), and the data are presented as RLuc/FLuc.

To directly quantify the splicing efficiency of the reporters, RNA was isolated and used for reverse transcription as described above. Spliced and unspliced reporter transcripts were measured by endpoint PCR (Platinum Taq DNA polymerase, Thermo Fisher Scientific) with primers specific to the RLuc coding sequence. PCR amplicons were detected by electrophoresis on 6% non-denaturing polyacrylamide/TBE gels with SYBR Gold Nucleic Acid Gel Stain (Thermo Fisher Scientific). Splicing efficiency was measured by densitometry analysis of the PCR amplicons in image J (*Schindelin et al., 2012; Schneider et al., 2012*).

For the splicing-rescue assay, HeLa-Flp-In cells stably expressing either wild-type or mutant (K95A, E197A or D199A) DDX39B *trans*-genes were transfected with control siRNA (NTC) or DDX39B-targeting siRNA D13, which targets the 3' untranslated region (3'UTR) of DDX39B mRNA. After two days of culture, expression of DDX39B *trans*-genes in NTC- or D13-treated cells was either induced by 1 μ g/mL doxycycline or not (doxycycline withheld) to generate the following four conditions: NTC –doxycycline (control), NTC +doxycycline (overexpression), D13 –doxycycline (knockdown) and D13 +doxycycline (rescue). The cells were then transfected with pcDNA3.1-RLuc_FOXP3 Intron 11 and harvested 24 hr after for analyses. DDX39B protein expression was confirmed by western blot, and the splicing efficiency of the reporters was evaluated by endpoint PCR as described above.

Py tract sequence analysis of retained introns

5' SS (from –3 to +6 nt at the 5' Exon-Intron junction) and 3' SS (from –20 to +3 nt at the 3' Intron-Exon junction) sequences of *FOXP3*, *FOX* family genes and X chromosome genes (2155 genes) were obtained from Human genome assembly (GRCh38.p13) using the UCSC table browser (*Karolchik et al., 2004; Kent et al., 2002*) and Galaxy platform (*Afgan et al., 2018*). Sequence logos of the splice sites of *FOX* family genes and X chromosomal genes (total 2155 genes) were generated by WebLogo3 (*Crooks et al., 2004*). Sequences of 3' SS (from –20 to –3 nt) were defined as poly-pyrimidine tract (Py tract) sequence, and nucleotide probability at each of these positions was determined per gene. The sequence logo for a given gene shows the average at each position of the Py tract from all introns for that particular gene.

For comparison of Py tracts of introns that were DDX39B-sensitive (retained in the knockdown) versus DDX39B-insensitive (unchanged in the knockdown), we compared the 402 introns that were significantly retained versus 500 events randomly selected from the total unchanged introns. 5' SS and 3' SS sequences for these events were obtained using the Galaxy platform, and MaxEntScore for these were determined by MaxEntScan (*Yeo and Burge, 2004*). The nucleotide probability at each position of the Py tract region was determined as above.

Statistical analysis

In all figures, error bars represent standard deviation (S.D.) unless otherwise noted. Asterisks denote level of statistical significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$. Statistical analyses were conducted in Prism8 (GraphPad Software) or SAS (SAS Institute). For the statistical analysis of experiments using CD4⁺ T cells from individual donors, One-way repeated measures ANOVA was used. To specifically compare each experimental (Sh3 and Sh5) group to control (NTC) after a significant Omnibus *f* test, Dunnett's test was performed to control family-wise type I error rate due to multiple comparisons. For the experiments other than CD4⁺ T cells, One-way ANOVA model followed by aforementioned Dunnett's tests were conducted to compare between experimental (Sh3 and Sh5) and control (NTC). For experiments done using iTregs from Donor 10, no statistical methods were used due to the difference in number of replicates for each experimental condition (NTC: $n=2$; Sh5: $n=3$). For the rescue assay of splicing, One-way ANOVA models followed by Tukey's tests were used to perform pairwise comparisons among all tested conditions. For the sequence analysis of retained introns, outliers of Max entropy scores were identified and removed (ROUT, $Q=0.5\%$) (*Motulsky and Brown, 2006*). Then, Mann-Whitney tests or Kruskal-Wallis tests were conducted.

Statistical analysis of RT-qPCR data in CD4⁺ T cell were carried out using a linear mixed model on the delta CT values (Δ CT). For each gene, the linear mixed model includes treatment (Control, DDX39B Knockdown1 [Sh3], DDX39B Knockdown2 [Sh5]) as fixed effect, and random intercept to account for the heterogeneity of the individuals and correlations among repeated measures from

the same individual. The fold-change differences (delta-delta CT [ddCT] values) and their standard error (SE) between treatment groups versus the control were estimated from the model. The p-values were adjusted by False Discovery Rate (FDR) method, which controls the false discovery rate due to multiplicity in hypothesis testing. To facilitate the interpretation using relative expression, the point estimate of the ddCT values were converted to 2^{-ddCT} .

Acknowledgements

We thank Drs. Paul Boutz (Rochester) and Harinder Singh (Pittsburgh) for excellent suggestions on an early draft of this manuscript. We acknowledge members of our laboratories for excellent discussions and comments to improve the science and the manuscript (MGB, SSB, SGG, DCK). WSF thanks John Paul Donohue (RNA Center for Molecular Biology, UC-Santa Cruz) for expert advice on the analysis of RNA sequencing data. Funding: We acknowledge support from the Uehara Foundation Fellowship and McLaughlin Postdoctoral Fund (MH), NIH F32 NS087899 (GGM), Duke Neurology startup and Stone family funds (SGG), KL2 TR001441-07 (WSF), R21AI133305 (DCK and LW), Duke MGM Start-up funds (DCK), R01 CA204806 (MGB), UTMB startup funds (MGB), and P01 AI150585 (MGB).

Additional information

Competing interests

Gaddiel Galarza-Muñoz, Mariano A Garcia-Blanco: I acknowledge that I have significant ownership in Autoimmunity Biologic Solutions, Inc (Galveston, TX), which is commercializing therapies that target the IL7R pathway in autoimmune diseases. While I do not believe this represents a conflict of interest it can lead to the perception of said conflict. The other authors declare that no competing interests exist.

Funding

Funder	Grant reference number	Author
Uehara Foundation		Minato Hirano
McLaughlin Family Foundation		Minato Hirano
National Institutes of Health	NIH F32 NS087899	Gaddiel Galarza-Muñoz
Stone Family Fund		Simon G Gregory
National Institutes of Health	KL2 TR001441-07	William S Fagg
National Institutes of Health	R21AI133305	Dennis C Ko
National Cancer Institute	R01 CA204806	Mariano A Garcia-Blanco
National Institutes of Health	P01 AI150585	Mariano A Garcia-Blanco

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Minato Hirano, Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review and editing; Gaddiel Galarza-Muñoz, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Writing – review and editing; Chloe Nagasawa, Formal analysis, Investigation, Writing – review and editing; Geraldine Schott, Investigation, Writing – review and editing; Liuyang Wang, Steven G Widen, Data curation, Formal analysis, Investigation; Alejandro L Antonia, Vaibhav Jain, Investigation; Xiaoying Yu, Data curation, Formal analysis, Methodology, Writing – review and editing; Farren BS Briggs, Formal analysis, Writing – review and editing; Simon G

Gregory, Conceptualization, Supervision, Writing – review and editing; Dennis C Ko, Conceptualization, Formal analysis, Supervision, Writing – review and editing; William S Fagg, Data curation, Formal analysis, Supervision, Investigation, Writing – review and editing; Shelton Bradrick, Conceptualization, Formal analysis, Supervision, Investigation, Writing – review and editing; Mariano A Garcia-Blanco, Conceptualization, Formal analysis, Supervision, Funding acquisition, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs

Steven G Widen  <http://orcid.org/0000-0002-6512-363X>

Dennis C Ko  <http://orcid.org/0000-0002-0113-5981>

Shelton Bradrick  <http://orcid.org/0000-0003-1566-9797>

Mariano A Garcia-Blanco  <http://orcid.org/0000-0001-8538-1997>

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.76927.sa1>

Author response <https://doi.org/10.7554/eLife.76927.sa2>

Additional files

Supplementary files

- Supplementary file 1. Differentially expressed genes after DDX39B-depletion observed with both shRNAs in both donors. Related to **Figure 1** and **Figure 1—figure supplement 1**.
- Supplementary file 2. Differentially expressed genes after DDX39B-depletion (Sh3) observed in both donors. Related to **Figure 1** and **Figure 1—figure supplement 1**.
- Supplementary file 3. List of MS Susceptibility SNPs and genes analyzed. Related to **Figure 1** and **Figure 1—figure supplements 2 and 3**.
- Supplementary file 4. Gene sets enriched in cells with normal levels of DDX39B (NTC). Related to **Figure 3** and **Figure 3—figure supplement 1**.
- Supplementary file 5. Connectivity Map analysis of the DDX39B KD. Related to **Figure 3** and **Figure 3—figure supplement 1**.
- Supplementary file 6. VastTool analysis of in CD4⁺T cells with DDX39B-depletion. Related to **Figure 5** and **Figure 5—figure supplement 3**.
- Supplementary file 7. Human and Mouse FOXP3 introns. Related to **Figure 6**.
- Supplementary file 8. Primer and Gblock sequences used in this study.
- Transparent reporting form

Data availability

There will be no restriction on any material described in this manuscript. The bulk RNA sequencing datasets have been uploaded to the Gene Expression Omnibus with accession number GSE145773. All data are available in the main text or supplementary materials.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Garcia-Blanco MA	2023	The RNA helicase DDX39B activates FOXP3 RNA splicing to control T regulatory cell fate	http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GGSE145773	NCBI Gene Expression Omnibus, GSE145773

References

Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltmann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 Update. *Nucleic Acids Research* **46**:W537–W544. DOI: <https://doi.org/10.1093/nar/gky379>, PMID: 29790989

- Allcock RJ**, Williams JH, Price P. 2001. The central MHC Gene, Bat1, may Encode a protein that down-regulates cytokine production. *Genes to Cells* **6**:487–494. DOI: <https://doi.org/10.1046/j.1365-2443.2001.00435.x>, PMID: 11380625
- Bennett CL**, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, Kelly TE, Saulsbury FT, Chance PF, Ochs HD. 2001. The immune dysregulation, Polyendocrinopathy, Enteropathy, X-linked syndrome (IPEX) is caused by mutations of Foxp3. *Nature Genetics* **27**:20–21. DOI: <https://doi.org/10.1038/83713>, PMID: 11137993
- Bhatt DM**, Pandya-Jones A, Tong AJ, Barozzi I, Lissner MM, Natoli G, Black DL, Smale ST. 2012. Transcript Dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150**:279–290. DOI: <https://doi.org/10.1016/j.cell.2012.05.043>, PMID: 22817891
- Boutz PL**, Bhutkar A, Sharp PA. 2015. Detained Introns are a novel, widespread class of post-Transcriptionally spliced Introns. *Genes & Development* **29**:63–80. DOI: <https://doi.org/10.1101/gad.247361.114>, PMID: 25561496
- Briggs F**. 2019. Unraveling susceptibility to multiple sclerosis. *Science* **365**:1383–1384. DOI: <https://doi.org/10.1126/science.aay1439>, PMID: 31604260
- Brunkow ME**, Jeffery EW, Hjerrild KA, Paepfer B, Clark LB, Yasayko SA, Wilkinson JE, Galas D, Ziegler SF, Ramsdell F. 2001. Disruption of a new Forkhead/winged-Helix protein, Scurfin, results in the fatal lymphoproliferative disorder of the Scurfy mouse. *Nature Genetics* **27**:68–73. DOI: <https://doi.org/10.1038/83784>, PMID: 11138001
- Chatila TA**, Blaeser F, Ho N, Lederman HM, Voulgaropoulos C, Helms C, Bowcock AM. 2000. Jm2, Encoding a fork head-related protein, is Mutated in X-linked Autoimmunity-allergic Disregulation syndrome. *The Journal of Clinical Investigation* **106**:R75–R81. DOI: <https://doi.org/10.1172/JCI11679>, PMID: 11120765
- Cortez JT**, Montauti E, Shifrut E, Gatchalian J, Zhang Y, Shaked O, Xu Y, Roth TL, Simeonov DR, Zhang Y, Chen S, Li Z, Woo JM, Ho J, Vogel IA, Prator GY, Zhang B, Lee Y, Sun Z, Ifergan I, et al. 2020. CRISPR screen in regulatory T cells reveals Modulators of Foxp3. *Nature* **582**:416–420. DOI: <https://doi.org/10.1038/s41586-020-2246-4>, PMID: 32499641
- Crooks GE**, Hon G, Chandonia JM, Brenner SE. 2004. Weblogo: A sequence logo generator. *Genome Research* **14**:1188–1190. DOI: <https://doi.org/10.1101/gr.849004>, PMID: 15173120
- Degli-Esposti MA**, Leelayuwat C, Dawkins RL. 1992. Ancestral Haplotypes carry Haplotypic and Haplospecific Polymorphisms of Bat1: Possible relevance to autoimmune disease. *European Journal of Immunogenetics* **19**:121–127. DOI: <https://doi.org/10.1111/j.1744-313x.1992.tb00051.x>, PMID: 1352699
- Dobin A**, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: Ultrafast universal RNA-Seq aligner. *Bioinformatics* **29**:15–21. DOI: <https://doi.org/10.1093/bioinformatics/bts635>, PMID: 23104886
- Dominguez-Villar M**, Hafler DA. 2018. Regulatory T cells in autoimmune disease. *Nature Immunology* **19**:665–673. DOI: <https://doi.org/10.1038/s41590-018-0120-4>
- ENCODE Project Consortium**. 2011. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLOS Biology* **9**:e1001046. DOI: <https://doi.org/10.1371/journal.pbio.1001046>, PMID: 21526222
- Evsyukova I**, Bradrick SS, Gregory SG, Garcia-Blanco MA. 2013. Cleavage and Polyadenylation specificity factor 1 (Cpsf1) regulates alternative splicing of interleukin 7 receptor (IL7R) Exon 6. *RNA* **19**:103–115. DOI: <https://doi.org/10.1261/rna.035410.112>
- Feng J**, Meyer CA, Wang Q, Liu JS, Shirley Liu X, Zhang Y. 2012. GFOLD: A generalized fold change for ranking Differentially expressed genes from RNA-Seq data. *Bioinformatics* **28**:2782–2788. DOI: <https://doi.org/10.1093/bioinformatics/bts515>
- Fleckner J**, Zhang M, Valcárcel J, Green MR. 1997. U2A65 recruits a novel human DEAD box protein required for the U2 snRNP-Branchpoint interaction. *Genes & Development* **11**:1864–1872. DOI: <https://doi.org/10.1101/gad.11.14.1864>, PMID: 9242493
- Fletcher JM**, Lonergan R, Costelloe L, Kinsella K, Moran B, O'Farrelly C, Tubridy N, Mills KHG. 2009. Cd39+Foxp3+ regulatory T cells suppress pathogenic Th17 cells and are impaired in multiple sclerosis. *The Journal of Immunology* **183**:7602–7610. DOI: <https://doi.org/10.4049/jimmunol.0901881>
- Galarza-Muñoz G**, Briggs FBS, Evsyukova I, Schott-Lerner G, Kennedy EM, Nyanhete T, Wang L, Bergamaschi L, Widen SG, Tomaras GD, Ko DC, Bradrick SS, Barcellos LF, Gregory SG, Garcia-Blanco MA. 2017. Human Epistatic interaction controls IL7R splicing and increases multiple sclerosis risk. *Cell* **169**:72–84. DOI: <https://doi.org/10.1016/j.cell.2017.03.007>, PMID: 28340352
- Gatfield D**, Le Hir H, Schmitt C, Braun IC, Köcher T, Wilm M, Izaurralde E. 2001. The Dexh/D box protein HEL/Uap56 is essential for mRNA nuclear export in *Drosophila*. *Current Biology* **11**:1716–1721. DOI: [https://doi.org/10.1016/s0960-9822\(01\)00532-2](https://doi.org/10.1016/s0960-9822(01)00532-2), PMID: 11696332
- Gavin MA**, Rasmussen JP, Fontenot JD, Vasta V, Manganiello VC, Beavo JA, Rudensky AY. 2007. Foxp3-dependent programme of regulatory T-cell differentiation. *Nature* **445**:771–775. DOI: <https://doi.org/10.1038/nature05543>
- Georgiev P**, Charbonnier LM, Chatila TA. 2019. Regulatory T cells: The many faces of Foxp3. *Journal of Clinical Immunology* **39**:623–640. DOI: <https://doi.org/10.1007/s10875-019-00684-7>, PMID: 31478130
- Gregory SG**, Schmidt S, Seth P, Oksenberg JR, Hart J, Prokop A, Caillier SJ, Ban M, Goris A, Barcellos LF, Lincoln R, McCauley JL, Sawcer SJ, Compston DAS, Dubois B, Hauser SL, Garcia-Blanco MA, Pericak-Vance MA, Haines JL, Multiple Sclerosis Genetics Group. 2007. Interleukin 7 receptor alpha chain (IL7R) shows Allelic and functional association with multiple sclerosis. *Nature Genetics* **39**:1083–1091. DOI: <https://doi.org/10.1038/ng2103>, PMID: 17660817

- GTEx Consortium.** 2017. Genetic effects on gene expression across human tissues. *Nature* **550**:204–213. DOI: <https://doi.org/10.1038/nature24277>
- Hamano R, Wu X, Wang Y, Oppenheim JJ, Chen X.** 2015. Characterization of MT-2 cells as a human regulatory T cell-like cell line. *Cellular & Molecular Immunology* **12**:780–782. DOI: <https://doi.org/10.1038/cmi.2014.123>, PMID: 25544500
- Hori S, Nomura T, Sakaguchi S.** 2003. Control of regulatory T cell development by the transcription factor Foxp3. *Science* **299**:1057–1061. DOI: <https://doi.org/10.1126/science.1079490>, PMID: 12522256
- Hori S.** 2021. Foxp3 as a master regulator of Treg cells. *Nat Rev Immunol* **21**:618–619. DOI: <https://doi.org/10.1038/s41577-021-00598-9>
- Huang C, Liang D, Tatomer DC, Wilusz JE.** 2018. A length-dependent Evolutionarily conserved pathway controls nuclear export of circular RNAs. *Genes & Development* **32**:639–644. DOI: <https://doi.org/10.1101/gad.314856.118>, PMID: 29773557
- Irimia M, Weatheritt RJ, Ellis JD, Parikshak NN, Gonatopoulos-Pournatzis T, Babor M, Quesnel-Vallières M, Tapial J, Raj B, O'Hanlon D, Barrios-Rodiles M, Sternberg MJE, Cordes SP, Roth FP, Wrana JL, Geschwind DH, Blencowe BJ.** 2014. A highly conserved program of neuronal Microexons is Misregulated in autistic brains. *Cell* **159**:1511–1523. DOI: <https://doi.org/10.1016/j.cell.2014.11.035>, PMID: 25525873
- Irimia M.** 2023. Vast-tools. sw:1:rev:fb87ceb4379cd9fbd36fc402e4a3429567021548. Software Heritage. <https://archive.softwareheritage.org/sw:1:dir:707295b0b471e209527f4a1abf000b5fec5e9da9;origin=https://github.com/vastgroup/vast-tools;visit=sw:1:snp:1fc5b9d7de69f87d5af4f91c412e8e6520aba7cc;anchor=sw:1:rev:fb87ceb4379cd9fbd36fc402e4a3429567021548>
- Jamison SF, Crow A, Garcia-Blanco MA.** 1992. The Spliceosome assembly pathway in mammalian extracts. *Molecular and Cellular Biology* **12**:4279–4287. DOI: <https://doi.org/10.1128/mcb.12.10.4279-4287.1992>, PMID: 1383687
- Jensen TH, Boulay J, Rosbash M, Libri D.** 2001. The DECD box putative ATPase Sub2P is an early mRNA export factor. *Current Biology* **11**:1711–1715. DOI: [https://doi.org/10.1016/s0960-9822\(01\)00529-2](https://doi.org/10.1016/s0960-9822(01)00529-2), PMID: 11696331
- Josefowicz SZ, Lu LF, Rudensky AY.** 2012. Regulatory T cells: Mechanisms of differentiation and function. *Annual Review of Immunology* **30**:531–564. DOI: <https://doi.org/10.1146/annurev.immunol.25.022106.141623>, PMID: 22224781
- Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ.** 2004. The UCSC table Browser data retrieval tool. *Nucleic Acids Research* **32**:D493–D496. DOI: <https://doi.org/10.1093/nar/gkh103>, PMID: 14681465
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D.** 2002. The human genome Browser at UCSC. *Genome Research* **12**:996–1006. DOI: <https://doi.org/10.1101/gr.229102>, PMID: 12045153
- Kistler AL, Guthrie C.** 2001. Deletion of Mud2, the yeast Homolog of U2A65, can bypass the requirement for Sub2, an essential Spliceosomal ATPase. *Genes & Development* **15**:42–49. DOI: <https://doi.org/10.1101/gad.851601>
- Komatsu N, Okamoto K, Sawa S, Nakashima T, Oh-hora M, Kodama T, Tanaka S, Bluestone JA, Takayanagi H.** 2014. Pathogenic conversion of Foxp3+ T cells into Th17 cells in autoimmune arthritis. *Nature Medicine* **20**:62–68. DOI: <https://doi.org/10.1038/nm.3432>
- Kumar S, Stecher G, Tamura K.** 2016. Mega7: Molecular evolutionary Genetics analysis version 7.0 for bigger Datasets. *Molecular Biology and Evolution* **33**:1870–1874. DOI: <https://doi.org/10.1093/molbev/msw054>
- Lehner B, Semple JI, Brown SE, Counsell D, Campbell RD, Sanderson CM.** 2004. Analysis of a high-throughput yeast two-hybrid system and its use to predict the function of intracellular proteins encoded within the human MHC class III region. *Genomics* **83**:153–167. DOI: [https://doi.org/10.1016/s0888-7543\(03\)00235-0](https://doi.org/10.1016/s0888-7543(03)00235-0)
- Li X, Liu S, Zhang L, Issaian A, Hill RC, Espinosa S, Shi S, Cui Y, Kappel K, Das R, Hansen KC, Zhou ZH, Zhao R.** 2019. A unified mechanism for Intron and Exon definition and back-splicing. *Nature* **573**:375–380. DOI: <https://doi.org/10.1038/s41586-019-1523-6>
- Liu W, Putnam AL, Xu-Yu Z, Szot GL, Lee MR, Zhu S, Gottlieb PA, Kapranov P, Gingeras TR, Fazekas de St Groth B, Clayberger C, Soper DM, Ziegler SF, Bluestone JA.** 2006. Cd127 expression inversely correlates with Foxp3 and suppressive function of human Cd4+ T reg cells. *The Journal of Experimental Medicine* **203**:1701–1711. DOI: <https://doi.org/10.1084/jem.20060772>, PMID: 16818678
- Loo CS, Gatchalian J, Liang Y, Leblanc M, Xie M, Ho J, Venkatraghavan B, Hargreaves DC, Zheng Y.** 2020. A genome-wide CRISPR screen reveals a role for the non-Canonical Nucleosome-remodeling BAF complex in Foxp3 expression and regulatory T cell function. *Immunity* **53**:143–157. DOI: <https://doi.org/10.1016/j.immuni.2020.06.011>, PMID: 32640256
- Lundström W, Highfill S, Walsh STR, Beq S, Morse E, Kockum I, Alfredsson L, Olsson T, Hillert J, Mackall CL.** 2013. Soluble IL7Ralpha potentiates IL-7 bioactivity and promotes autoimmunity. *PNAS* **110**:E1761–E1770. DOI: <https://doi.org/10.1073/pnas.1222303110>, PMID: 23610432
- Luo ML, Zhou Z, Magni K, Christoforides C, Rappsilber J, Mann M, Reed R.** 2001. Pre-mRNA splicing and mRNA export linked by direct interactions between Uap56 and Aly. *Nature* **413**:644–647. DOI: <https://doi.org/10.1038/35098106>, PMID: 11675789
- Motulsky HJ, Brown RE.** 2006. Detecting Outliers when fitting data with Nonlinear regression - a new method based on robust Nonlinear regression and the false discovery rate. *BMC Bioinformatics* **7**:123. DOI: <https://doi.org/10.1186/1471-2105-7-123>, PMID: 16526949
- Nakata D, Nakao S, Nakayama K, Araki S, Nakayama Y, Aparicio S, Hara T, Nakanishi A.** 2017. The RNA Helicase Ddx39B and its Paralog Ddx39A regulate androgen receptor splice variant AR-V7 generation. *Biochemical and*

- Biophysical Research Communications* **483**:271–276. DOI: <https://doi.org/10.1016/j.bbrc.2016.12.153>, PMID: 28025139
- Newnham CM**, Query CC. 2001. The ATP requirement for U2 snRNP addition is linked to the pre-mRNA region 5' to the branch site. *RNA* **7**:1298–1309. DOI: <https://doi.org/10.1017/s1355838201010561>, PMID: 11565751
- Noval Rivas M**, Burton OT, Wise P, Charbonnier L-M, Georgiev P, Oettgen HC, Rachid R, Chatila TA. 2015. Regulatory T cell Reprogramming toward a Th2-cell-like lineage impairs oral tolerance and promotes food allergy. *Immunity* **42**:512–523. DOI: <https://doi.org/10.1016/j.immuni.2015.02.004>
- Patsopoulos NA**, Baranzini SE, Santaniello A, Shoostari P, Cotsapas C, Wong G, Beecham AH, James T, Replogle J, Vlachos IS, McCabe C, Pers TH, Brandes A, White C, Keenan B, Cimpean M, Winn P, Panteliadis IP, Robbins A, Andlauer TFM, et al. 2019. Multiple sclerosis Genomic map Implicates peripheral immune cells and Microglia in susceptibility. *Science* **365**:eaav7188. DOI: <https://doi.org/10.1126/science.aav7188>, PMID: 31604244
- Roscigno RF**, Weiner M, Garcia-Blanco MA. 1993. A mutational analysis of the Polypyrimidine tract of Introns. effects of sequence differences in Pyrimidine tracts on splicing. *The Journal of Biological Chemistry* **268**:11222–11229 PMID: 8496178.
- Ruskin B**, Zamore PD, Green MR. 1988. A factor, U2A_f, is required for U2 snRNP binding and splicing complex assembly. *Cell* **52**:207–219. DOI: [https://doi.org/10.1016/0092-8674\(88\)90509-0](https://doi.org/10.1016/0092-8674(88)90509-0)
- Sambucci M**, Gargano F, De Rosa V, De Bardi M, Picozza M, Placido R, Ruggieri S, Capone A, Gasperini C, Matarese G. 2018. Foxp3 Isoforms and PD-1 expression by T regulatory cells in multiple sclerosis. *Sci Rep* **8**:3674. DOI: <https://doi.org/10.1038/s41598-018-21861-5>
- Schindelin J**, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: An open-source platform for biological-image analysis. *Nature Methods* **9**:676–682. DOI: <https://doi.org/10.1038/nmeth.2019>, PMID: 22743772
- Schneider CA**, Rasband WS, Eliceiri KW. 2012. NIH image to Imagej: 25 Years of image analysis. *Nature Methods* **9**:671–675. DOI: <https://doi.org/10.1038/nmeth.2089>, PMID: 22930834
- Schott G**, Galarza-Muñoz G, Trevino N, Chen X, Weirauch M, Gregory SG, Bradrick SS, Garcia-Blanco MA. 2021. U2A_f binds IIR Exon 6 Ectopically and represses its inclusion. *RNA* **27**:571–583. DOI: <https://doi.org/10.1261/rna.078279.120>, PMID: 33568552
- Schott G**, Garcia-Blanco MA. 2021. MHC class III RNA binding proteins and immunity. *RNA Biology* **18**:640–646. DOI: <https://doi.org/10.1080/15476286.2020.1860388>, PMID: 33280511
- Schumann K**, Raju SS, Lauber M, Kolb S, Shifrut E, Cortez JT, Skartsis N, Nguyen VQ, Woo JM, Roth TL, Yu R, Nguyen MLT, Simeonov DR, Nguyen DN, Targ S, Gate RE, Tang Q, Bluestone JA, Spitzer MH, Ye CJ, et al. 2020. Functional CRISPR dissection of Gene networks controlling human regulatory T cell identity. *Nature Immunology* **21**:1456–1466. DOI: <https://doi.org/10.1038/s41590-020-0784-4>, PMID: 32989329
- Seraphin B**, Rosbash M. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to Spliceosome assembly and splicing. *Cell* **59**:349–358. DOI: [https://doi.org/10.1016/0092-8674\(89\)90296-1](https://doi.org/10.1016/0092-8674(89)90296-1)
- Shen J**, Zhang L, Zhao R. 2007. Biochemical characterization of the ATPase and Helicase activity of Uap56, an essential pre-mRNA splicing and mRNA export factor. *Journal of Biological Chemistry* **282**:22544–22550. DOI: <https://doi.org/10.1074/jbc.M702304200>
- Shen H**, Zheng X, Shen J, Zhang L, Zhao R, Green MR. 2008. Distinct activities of the Dexe/H-box splicing factor Huap56 facilitate stepwise assembly of the Spliceosome. *Genes & Development* **22**:1796–1803. DOI: <https://doi.org/10.1101/gad.1657308>
- Shen H-H**. 2009. Uap56- a key player with surprisingly diverse roles in pre-mRNA splicing and nuclear export. *BMB Reports* **42**:185–188. DOI: <https://doi.org/10.5483/BMBRep.2009.42.4.185>
- Singh R**, Valcárcel J, Green MR. 1995. Distinct binding Specificities and functions of higher Eukaryotic Polypyrimidine tract-binding proteins. *Science* **268**:1173–1176. DOI: <https://doi.org/10.1126/science.7761834>, PMID: 7761834
- Spies T**, Blanck G, Bresnahan M, Sands J, Strominger JL. 1989. A new cluster of genes within the human major Histocompatibility complex. *Science* **243**:214–217. DOI: <https://doi.org/10.1126/science.2911734>, PMID: 2911734
- Subramanian A.**, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* **102**:15545–15550. DOI: <https://doi.org/10.1073/pnas.0506580102>
- Subramanian A**, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, et al. 2017. A next generation Connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**:1437–1452. DOI: <https://doi.org/10.1016/j.cell.2017.10.049>, PMID: 29195078
- Tamura K**, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**:512–526. DOI: <https://doi.org/10.1093/oxfordjournals.molbev.a040023>, PMID: 8336541
- Tarver JE**, Dos Reis M, Mirarab S, Moran RJ, Parker S, O'Reilly JE, King BL, O'Connell MJ, Asher RJ, Warnow T, Peterson KJ, Donoghue PCJ, Pisani D. 2016. The interrelationships of Placental mammals and the limits of Phylogenetic inference. *Genome Biology and Evolution* **8**:330–344. DOI: <https://doi.org/10.1093/gbe/evv261>, PMID: 26733575
- Wagner EJ**, Garcia-Blanco MA. 2002. Rnai-mediated PTB depletion leads to enhanced Exon definition. *Molecular Cell* **10**:943–949. DOI: [https://doi.org/10.1016/s1097-2765\(02\)00645-7](https://doi.org/10.1016/s1097-2765(02)00645-7), PMID: 12419237

- Wan YY**, Flavell RA. 2007. Regulatory T-cell functions are subverted and converted owing to attenuated Foxp3 expression. *Nature* **445**:766–770. DOI: <https://doi.org/10.1038/nature05479>, PMID: 17220876
- Wang ET**, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative Isoform regulation in human tissue Transcriptomes. *Nature* **456**:470–476. DOI: <https://doi.org/10.1038/nature07509>, PMID: 18978772
- Wildin RS**, Ramsdell F, Peake J, Faravelli F, Casanova JL, Buist N, Levy-Lahad E, Mazzella M, Goulet O, Perroni L, Bricarelli FD, Byrne G, McEuen M, Proll S, Appleby M, Brunkow ME. 2001. X-linked neonatal diabetes mellitus, Enteropathy and Endocrinopathy syndrome is the human equivalent of mouse Scurfy. *Nature Genetics* **27**:18–20. DOI: <https://doi.org/10.1038/83707>, PMID: 11137992
- Wu T**, Fu XD. 2015. Genomic functions of U2Af in Constitutive and regulated splicing. *RNA Biology* **12**:479–485. DOI: <https://doi.org/10.1080/15476286.2015.1020272>, PMID: 25901584
- Yeo G**, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of Computational Biology* **11**:377–394. DOI: <https://doi.org/10.1089/1066527041410418>, PMID: 15285897
- Zamore PD**, Green MR. 1989. Identification, purification, and biochemical characterization of U2 small nuclear Ribonucleoprotein auxiliary factor. *PNAS* **86**:9243–9247. DOI: <https://doi.org/10.1073/pnas.86.23.9243>, PMID: 2531895
- Zamore PD**, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2Af. *Nature* **355**:609–614. DOI: <https://doi.org/10.1038/355609a0>, PMID: 1538748
- Zhou X**, Bailey-Bucktrout SL, Jeker LT, Penaranda C, Martínez-Llordella M, Ashby M, Nakayama M, Rosenthal W, Bluestone JA. 2009. Instability of the transcription factor Foxp3 leads to the generation of pathogenic memory T cells in vivo. *Nature Immunology* **10**:1000–1007. DOI: <https://doi.org/10.1038/ni.1774>
- Zhou Z**, Sun B, Huang S, Zhao L. 2019. Roles of circular Rnas in immune regulation and autoimmune diseases. *Cell Death & Disease* **10**:503. DOI: <https://doi.org/10.1038/s41419-019-1744-5>, PMID: 31243263