nature cell biology

Article

https://doi.org/10.1038/s41556-023-01118-8

Arg-tRNA synthetase links inflammatory metabolism to RNA splicing and nuclear trafficking via SRRM2

Received: 28 March 2022

Accepted: 27 February 2023

Published online: 14 April 2023

Check for updates

Haissi Cui[1], Jolene K. Diedrich[1], Douglas C. Wu[2], Justin J. Lim[3,4], Ryan M. Nottingham[2], James J. Moresco[1,5], John R. Yates III[1], Benjamin J. Blencowe[3,4], Alan M. Lambowitz[2] ✉ & Paul Schimmel[1] ✉

Cells respond to perturbations such as inflammation by sensing changes in metabolite levels. Especially prominent is arginine, which has known connections to the inflammatory response. Aminoacyl-tRNA synthetases, enzymes that catalyse the first step of protein synthesis, can also mediate cell signalling. Here we show that depletion of arginine during inflammation decreased levels of nuclear-localized arginyl-tRNA synthetase (ArgRS). Surprisingly, we found that nuclear ArgRS interacts and co-localizes with serine/arginine repetitive matrix protein 2 (SRRM2), a spliceosomal and nuclear speckle protein, and that decreased levels of nuclear ArgRS correlated with changes in condensate-like nuclear trafficking of SRRM2 and splice-site usage in certain genes. These splice-site usage changes cumulated in the synthesis of different protein isoforms that altered cellular metabolism and peptide presentation to immune cells. Our findings uncover a mechanism whereby an aminoacyl-tRNA synthetase cognate to a key amino acid that is metabolically controlled during inflammation modulates the splicing machinery.

Inflammation accompanies most diseases, including cancer, neurodegeneration and auto-immune disorders[1]. One of the metabolic hallmarks of inflammation is a decrease in systemic arginine levels due to its increased consumption by pro- and anti-inflammatory enzymes, such as iNOS and Arginase-1, with plasma arginine recovering when the inflammatory insult is resolved[2,3]. Consequently, arginine and signalling molecules derived from arginine (for example, nitric oxide and ornithine) are key regulators of cellular signalling pathways and necessary for immune cell activation during inflammation[4,5].

The function of aminoacyl-tRNA synthetases (aaRSs) in cell signal transduction has come into focus in recent years. Aminoacyl-tRNA synthetases are highly conserved proteins that are essential for messenger RNA (mRNA) translation, where they catalyse the attachment of amino acids to the 3′ end of their cognate transfer RNAs (tRNAs)[6]. Aside from tRNAs, aaRSs can also aminoacylate proteins[7,8]. Eight aaRSs and three adaptor proteins form the multi-synthetase complex (MSC), which is crucial for the intra-cellular distribution of certain aaRSs and specifically for their nuclear localization[9]. The MSC is dynamic and acts as a reservoir to release aaRSs for extra-translational functions[10,11]. Although nuclear-localized aaRSs were historically seen as quality control agents for the correct processing of nascent tRNAs before their export to the cytoplasm[12], they have also been linked to other functions[13], including small molecule sensing[14] and regulation of gene expression[15,16].

The cell nucleus is compartmentalized by membraneless organelles, some of which display condensate-like behaviour, with the

[1]Department of Molecular Medicine, The Scripps Research Institute, La Jolla, CA, USA. [2]Institute for Cellular and Molecular Biology and Departments of Molecular Biosciences and Oncology, University of Texas at Austin, Austin, TX, USA. [3]The Donnelly Centre, University of Toronto, Toronto, Ontario, Canada. [4]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [5]Center for the Genetics of Host Defense, UT Southwestern Medical Center, Dallas, TX, USA. ✉e-mail: lambowitz@austin.utexas.edu; schimmel@scripps.edu

characteristics of phase-separated liquids[17,18]. Despite their lack of a membrane and constant exchange with the surrounding nucleoplasm, these subnuclear domains display structural integrity[19]. In the nucleus, DNA-sparse interchromatin granule clusters (termed nuclear speckles) sequester proteins with repetitive sequences of low complexity and high disorder, among them factors necessary for mRNA maturation and mRNA splicing[20,21]. RNA polymerases and splicing factors are in steady exchange between these structures and the surrounding nucleoplasm[22], raising the possibility that nuclear protein trafficking regulates their functions.

Different protein isoforms can be derived from the same coding sequence by alternative splicing, potentially resulting in different functions[23]. Alternative splicing events leading to intron retention can post-transcriptionally regulate mRNA levels by impeding transport to the cytoplasm and increasing RNA degradation[24]. Twenty per cent of alternatively spliced cassette exons (where splice sites are variably recognized, leading to an exon being either retained or spliced out) are evolutionarily conserved between humans and mice, suggesting functional consequences[25].

Nuclear speckles store mRNA processing factors at their periphery and/or are hubs for mRNA transcription and processing[26,27]. Serine/arginine repetitive matrix protein 2, (SRRM2, also known as SRm300, gene name *SRRM2*), a component of spliceosomal complexes[28,29] and SON, an mRNA binding protein and splicing factor, are required for nuclear speckle formation[30]. During mRNA splicing, the conserved N-terminal domain of SRRM2 engages with exon sequences and regulates splice-site selection[31,32]. In line with its importance, SRRM2 mutations have been identified in developmental disorders[33], cancer[34] and Parkinson's disease[35].

Given the role of arginine in inflammation[2], we were interested in investigating how arginine-metabolizing enzymes, particularly arginyl-tRNA synthetase (ArgRS, gene name *RARS*) might respond to fluctuations in external arginine. In this article, by examining the effects of arginine depletion and ArgRS knockdown, we delineated a pathway through which cells respond to arginine starvation during inflammation via an interaction between nuclear ArgRS and SRRM2. Our results suggest a mechanism in which an aaRS interacts with a splicing factor to regulate nuclear trafficking and alternative mRNA splicing in response to a metabolic signal.

Results

Arginine regulates nuclear ArgRS

Previous work showed that ArgRS localizes to the nucleus as part of the MSC[36]. To begin, we re-investigated the nuclear localization of ArgRS using four model systems: (1) human hepatic carcinoma cell line HepG2, (2) human embryonic kidney cell line 293T, (3) murine embryonic fibroblasts and (4) mouse liver. We confirmed the presence of ArgRS in the nuclear fraction of all four samples (Fig. 1a and Extended Data Fig. 1a). As arginine levels regulate central cellular signalling pathways[2], we tested whether nuclear ArgRS levels were affected by reduction of extracellular arginine. We starved 293T cells of arginine for 6 h by incubation in arginine-free medium and found a decrease in the level of nuclear ArgRS, which could be reversed by adding arginine (Fig. 1b). Cytoplasmic ArgRS levels were unchanged by these treatments (Fig. 1b). Nuclear localization of ArgRS after re-addition of arginine was reduced but not abrogated by a mutation that abolishes ArgRS catalytic activity or by mutating a putative nuclear localization sequence (Extended Data Fig. 1b). This is consistent with previous findings showing that nuclear localization of ArgRS is governed mainly through the MSC[9].

Inflammatory arginine depletion decreases nuclear ArgRS

Next, we verified the effect of arginine starvation on nuclear ArgRS levels in a more physiological model. We reduced systemic arginine in mice by intra-peritoneal injection of recombinant murine Arginase-1. Plasma arginine in arginine-depleted mice was reduced to 30–50% of vehicle controls (Extended Data Fig. 1c), while nuclear ArgRS in the spleen, which consists mostly of immune cells, was reduced to <40% of vehicle controls (Fig. 1c, $P = 0.03$).

As plasma arginine concentration decreases during inflammation[3], we tested whether systemic inflammation induced by the liver toxin carbon tetrachloride (CCl_4) (ref. 37) would similarly affect nuclear ArgRS in mice. Plasma arginine levels dropped to 50% of vehicle controls (Extended Data Fig. 1c), and nuclear ArgRS levels in the spleen were reduced by ~60% (Fig. 1d). These findings indicate that systemic arginine depletion itself or as a result of systemic inflammation decreased nuclear ArgRS levels, and we focused our attention on the function of nuclear ArgRS.

The ArgRS interactome contains nuclear proteins

To identify interaction partners of nuclear ArgRS, which could offer clues to its function, we enriched ArgRS from the nuclear fraction of HepG2 cells using polyclonal ArgRS antibodies and characterized its protein interactome by mass spectrometry[38]. We used label-free quantification (LFQ) to compare intensities against an isotype-matched control antibody (Supplementary Tables 1–4). Proteins with a P value <0.05 were identified as significantly enriched (Fig. 2a). All MSC-bound aaRSs, except leucyl-tRNA synthetase (LeuRS, gene name *LARS*), were retrieved from the HepG2 nuclear fraction (Fig. 2a, orange, and Supplementary Table 1), suggesting that LeuRS may be absent or substoichiometric in the nuclear MSC. The adaptor protein AIMP-3 was enriched, albeit not significantly ($P > 0.05$). Thus, by enriching only for ArgRS, we captured nearly all interactions within the MSC. These results suggest that the MSC is mostly intact in the nucleus of HepG2 cells but may be more dynamic than in the cytoplasm. They also confirm we can identify ArgRS interaction partners, even if the interactions are indirect and mediated through other proteins.

Besides MSC components, we found several proteins with predominantly nuclear localization in the nuclear ArgRS interactome, including the nuclear speckle and spliceosome component SRRM2 (Fig. 2a,b, yellow). To avoid artefacts due to the isolation of nuclei, we repeated the experiment using whole cell lysates and confirmed specific and reproducible binding ($P < 0.05$) to nuclear proteins (Fig. 2a,b and Supplementary Table 2), despite ArgRS being a predominantly cytoplasmic protein. Consistent with the known cytoplasmic composition of the MSC, LeuRS was retrieved by ArgRS immunoprecipitation from the whole cell lysates of HepG2 cells (Fig. 2a).

To test if these interactions are limited to the liver carcinoma cell line HepG2, we used the same approach on 293T cells (Fig. 2a,b and Supplementary Table 3). Affinity enrichment of ArgRS retrieved the MSC and several proteins with predominately nuclear localization (Fig. 2a,b). Comparison between the ArgRS interactomes in the two cell types showed significant overlap but also distinct binding partners (hypergeometric test, Extended Data Fig. 2a). In contrast, the same experimental procedure in 293T cell lysates with a polyclonal antibody against another MSC-bound aaRS, methionyl-tRNA synthetase (MetRS, gene name *MARS*)[39], yielded exclusively MSC proteins as significant interactors ($P < 0.05$, Fig. 2a,b and Supplementary Table 4). Among the ArgRS binding partners, SRRM2 was highly and reproducibly enriched in both 293T and HepG2 cells (Fig. 2a,b, upper right and lower left, and Supplementary Table 5).

Reciprocal co-immunoprecipitation of ArgRS and SRRM2

We focused on SRRM2 because of its critical roles in both RNA splicing and nuclear compartmentalization. We confirmed, with different antibodies, in reciprocal co-immunoprecipitations and western blotting in both HepG2 and 293T cell lysates, that pulldown of ArgRS enriched for SRRM2 and vice versa (Extended Data Fig. 2b–f). Domain mapping with recombinant, biotinylated C- and N-terminally truncated ArgRS confirmed that the catalytic domain of ArgRS is necessary for interaction

https://doi.org/10.1038/s41556-023-01118-8

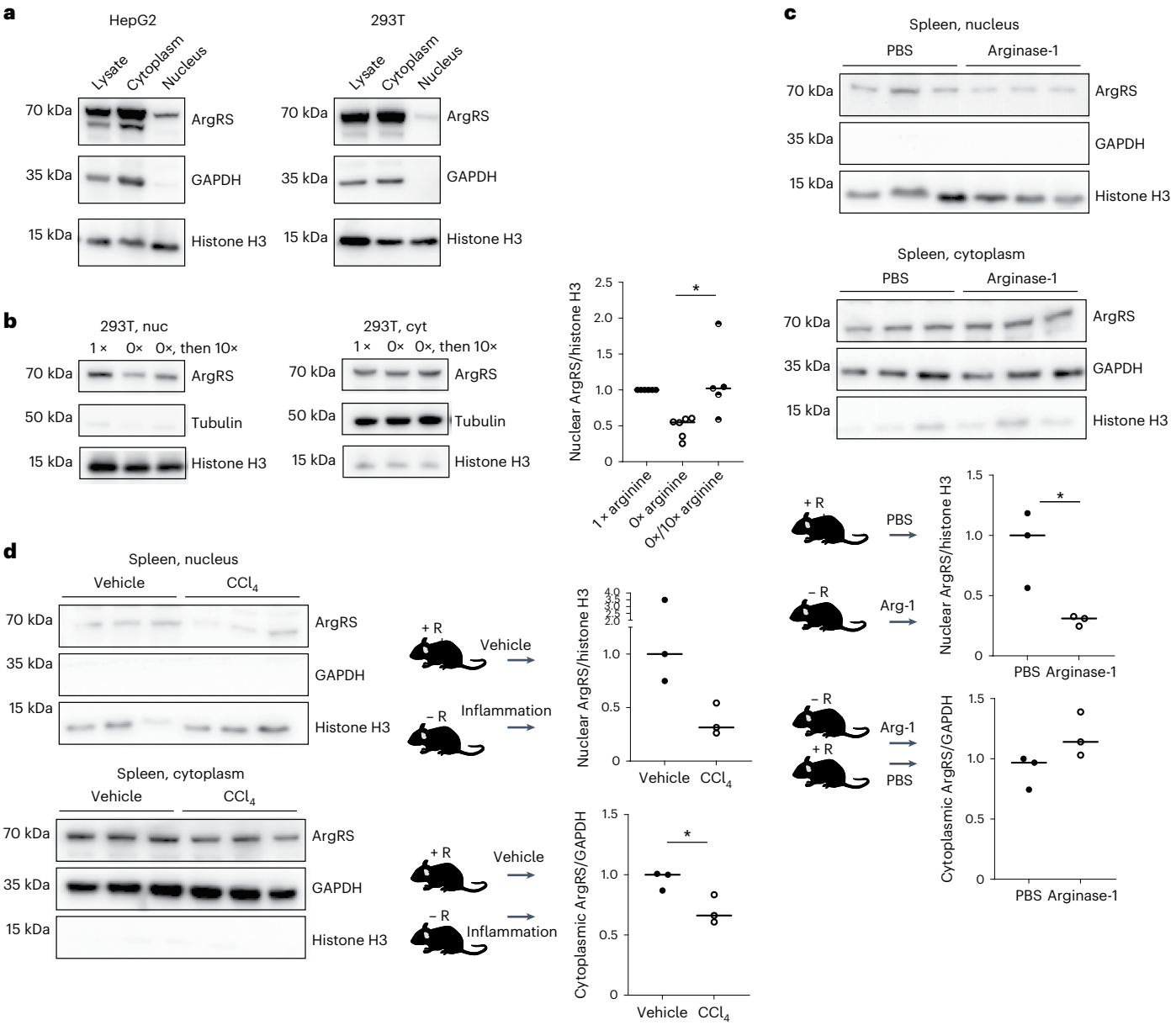


Fig. 1 | ArgRS localized to the nucleus in response to external arginine levels. **a**, ArgRS in the cytoplasmic and nuclear fractions of human hepatocellular carcinoma HepG2 or HEK 293T cells. Experiments were repeated at least three times. **b**, ArgRS in the nucleus (nuc, left) and cytoplasm (cyt, right) following cell fractionation. Nuclear ArgRS could be restored by addition of external arginine after starvation (0×/10×). Arginine was re-added at 10× to DMEM to replenish lysosomal arginine storage and ensure sufficient arginine to rescue ArgRS nuclear localization. Arginine concentrations are relative to DMEM. $n = 6$ and 6, with 5 different cell passages, $^*P = 0.02$. Unpaired, two-tailed t-test. **c**, Injection of recombinant arginase-1-reduced nuclear ArgRS in the spleen of mice. $n = 3$, with 3 biologically independent animals, $^*P = 0.03$, $P = 0.1$. Unpaired, two-tailed t-test. **d**, Systemic inflammation by CCl_4-reduced nuclear ArgRS in the spleen of mice. $n = 3$, with 3 biologically independent animals, $^*P = 0.03$. R, arginine. Unpaired, two-tailed t-test. **a**–**d** show western blots and densitometric quantification. GAPDH, cytoplasmic loading control; tubulin, cytoplasmic loading control; histone H3, nuclear loading control.

with SRRM2 (Extended Data Fig. 2g) and showed that catalytically inactive ArgRS also retrieved SRRM2 from cell lysates (Extended Data Fig. 2h). Deletion of the ArgRS leucine zipper (ΔLZ), which integrates ArgRS into the MSC and thereby regulates its nuclear import[9], did not abolish the interaction with SRRM2, suggesting that ArgRS and SRRM2 can interact directly in cell lysates (Extended Data Fig. 2i). In reciprocal experiments, the N-terminal 276 amino acids of SRRM2 but not a C-terminal arginine-rich stretch (amino acids 445–800) pulled down ArgRS (Extended Data Fig. 2j). By contrast, MetRS, another component of the nuclear MSC, was not enriched by SRRM2 pulldown (Extended Data Fig. 2k). These findings, together with the lack of SRRM2 enrichment in the MetRS interactome identified by mass spectrometry (Fig. 2b, lower right), indicated that ArgRS interacts with SRRM2 outside of the complete MSC.

ArgRS and SRRM2 co-localize in the nucleus

To localize where ArgRS and SRRM2 might interact in intact cells, we used immunofluorescence (IF) staining of HepG2 cells followed by Airyscan super-resolution microscopy. SRRM2 accumulated in chromatin-low nuclear structures (Fig. 3, Extended Data Fig. 3a and antibody specificity in Extended Data Fig. 3h). Eighty per cent of SON, which together with SRRM2 is crucial for nuclear speckle formation[30], co-localized with SRRM2 (Extended Data Fig. 3a), suggesting that these structures were indeed nuclear speckles.

https://doi.org/10.1038/s41556-023-01118-8

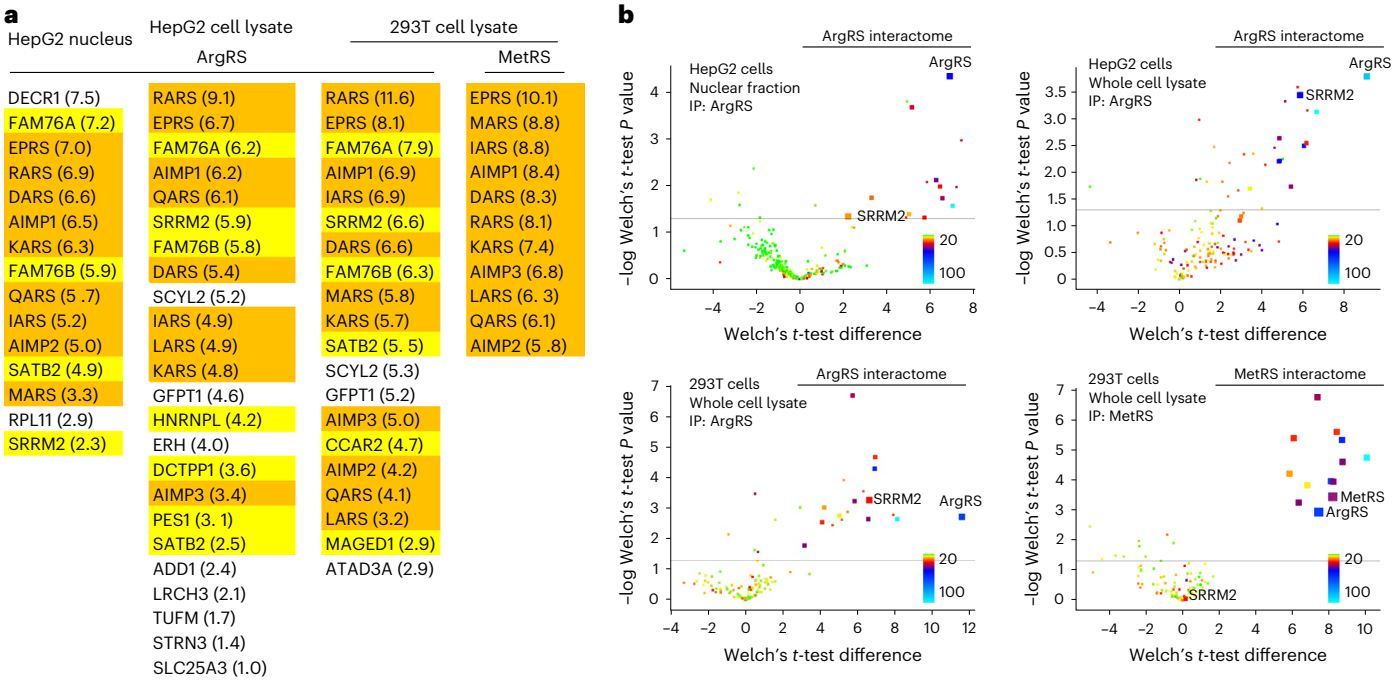


Fig. 2 | Characterization of the ArgRS interactome revealed nuclear proteins as interaction partners. a, Comparison among the ArgRS interactome in nucleus, HepG2 whole cell lysate, 293T whole cell lysate and the interactome of MetRS, another aminoacyl-tRNA synthetase. MSC components, orange; nuclear proteins, yellow. Proteins are identified by gene names and listed in order of enrichment over a non-targeted antibody of the same isotype, with fold change noted in parentheses. **b**, Volcano plots of ArgRS or MetRS interactomes. MSC proteins are highlighted by larger symbol sizes. Axes denote enrichment (Welch's *t*-test enrichment, *x* axis) and reproducibility between three repeats (Welch's *t*-test significance, *y* axis, $n = 3$ different cell passages). Colour: number of unique peptides detected; line: $P = 0.05$.

We next assessed where ArgRS and SRRM2 co-localized. Immunofluorescence co-staining of ArgRS visualized ArgRS in and at the periphery of SRRM2-rich areas (Fig. 3a, Supplementary Video 2, additional images in Extended Data Fig. 3e,f and antibody specificity and staining pattern in Extended Data Fig. 3b,g,h). The calculated Manders overlap coefficient, which denotes the correlation of absolute intensities, was >0.6 for SRRM2 protein in ArgRS/SRRM2 co-staining (Fig. 3a and Extended Data Fig. 3c). If the Costes method[40] was used for thresholding, co-localization increased to 0.8 (Extended Data Fig. 3d). The Pearson correlation coefficient, another measure of co-localization that denotes the correlation of relative intensities, was 0.45 (Extended Data Fig. 4c). About 60% of SRRM2 was calculated to be co-localized with ArgRS (Fig. 3a), and similar results were obtained with alternative ArgRS/SRRM2 antibodies (Extended Data Fig. 3c,e,f).

Nuclear ArgRS was also found outside of SRRM2-rich areas and partially co-localized with paraspeckle protein SFPQ (Manders coefficient <0.3, Extended Data Fig. 4a), which was found throughout the nucleus, but not with coilin, which exclusively located to Cajal bodies (Manders coefficient near 0, Extended Data Fig. 4b). Taken together, these results suggest ArgRS and SRRM2 specifically co-localize in the periphery of SRRM2-rich, low-chromatin regions identified as nuclear speckles. Arginine starvation significantly decreased the Pearson coefficient between ArgRS and SRRM2 from 0.6 to 0.4 ($P = 0.01$, Extended Data Fig. 4c), in agreement with our finding that nuclear ArgRS levels were decreased by arginine starvation.

In accordance with our mass spectrometry results (Fig. 2), the Manders coefficient for SRRM2/MetRS was only 0.1, as opposed to 0.6 for SRRM2/ArgRS (Extended Data Figs. 3c and 4d and Supplementary Video 3, $P < 0.0001$). The majority of nuclear ArgRS did not co-stain with nuclear MetRS, while cytoplasmic ArgRS and MetRS showed high co-localization (Manders coefficients 0.1 versus 0.7, respectively, Extended Data Figs. 3c and 4e and Supplementary Video 4), suggesting that the majority of nuclear ArgRS is not associated with the MSC. These results point to an ArgRS-specific, MSC-independent association with the nuclear speckle protein SRRM2 within or at the periphery of SRRM2-rich structures (Extended Data Fig. 4f).

Arginine or ArgRS depletion increases SRRM2 mobility

As ArgRS localized to the periphery of SRRM2-rich signals, we next wondered whether ArgRS might thereby influence SRRM2 mobility and trafficking in the nucleus. To address this question, we used fluorescence recovery after photobleaching (FRAP) to monitor SRRM2 flux (scheme in Fig. 3b). For FRAP measurements, we fluorescently labelled endogenous SRRM2 in 293T cells by introducing an mVenus-tag near its C-terminus (Extended Data Fig. 5a). As expected, SRRM2-mVenus localized to discrete nuclear regions similar to those detected by antibody staining (Extended Data Fig. 5b).

For FRAP experiments, regions of interest (ROIs) were chosen that contain one nuclear speckle including its surrounding nucleoplasm (Fig. 3b). Bleaching of SRRM2-mVenus in ROIs was followed by a time-dependent recovery of fluorescence that plateaued within 60 s (Extended Data Fig. 5c) through influx of SRRM2 from the surrounding SRRM2-rich compartments and nucleoplasm (Fig. 4a and Supplementary Videos 5–7). About 50% of SRRM2 in the ROIs was immobile (Extended Data Fig. 5d). A one-exponential curve fit of background-subtracted, normalized fluorescent recordings estimated a recovery half time of 17 s in arginine-rich medium (Fig. 3c). Arginine starvation, which decreased the amount of nuclear ArgRS (Fig. 1), increased SRRM2 mobility as indicated by the shorter half time of recovery of SRRM2 (11 s versus 18 s) (Fig. 3c and Extended Data Fig. 5e).

The above findings suggested that nuclear ArgRS decreases SRRM2 mobility. To further investigate, we stably introduced short hairpin RNAs (shRNAs) directed against the 3′ untranslated region (shRARS_1) or 5′ end (shRARS_2) of *RARS* and thereby reduced ArgRS

https://doi.org/10.1038/s41556-023-01118-8

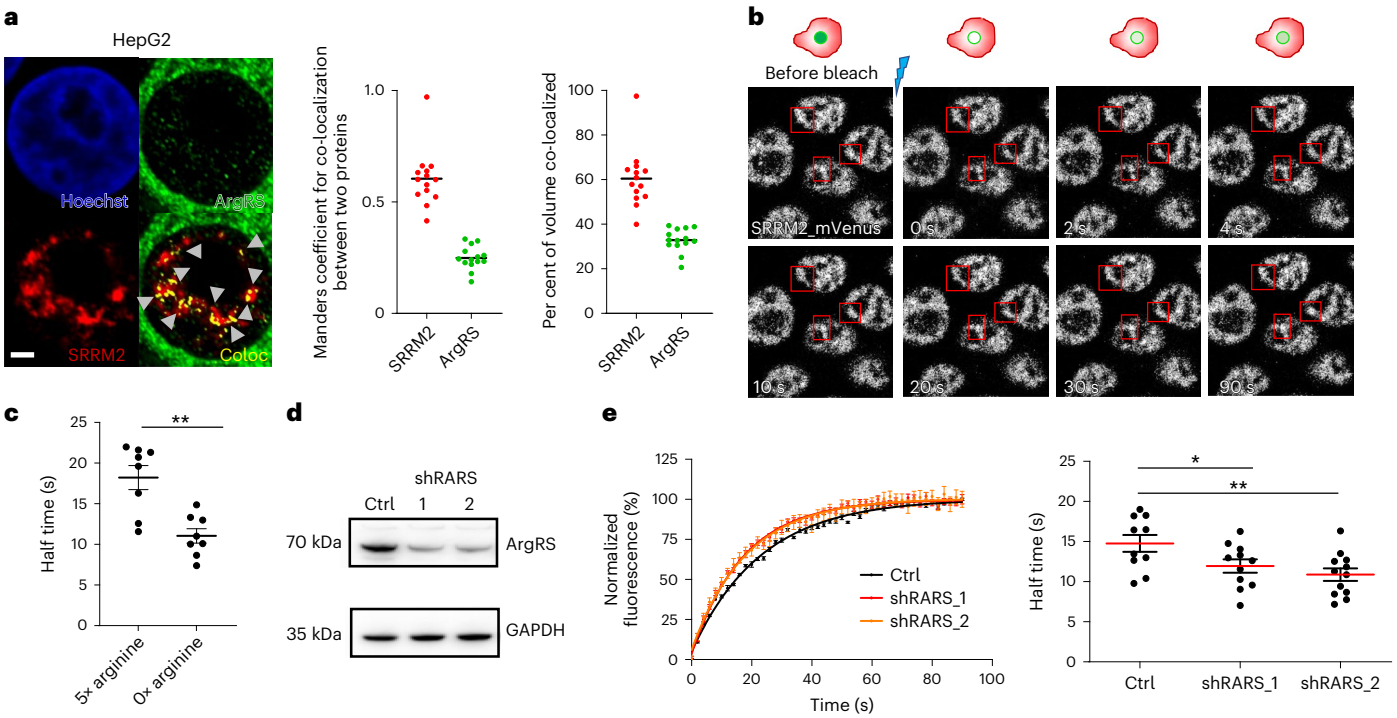


Fig. 3 | ArgRS knockdown increased SRRM2 mobility. a, Confocal microscopy following immunofluorescent staining of ArgRS/SRRM2. A single slice of a series of *z* stacks is shown. Co-localized areas in yellow. Scale bar, 2 µm. Experiments were repeated at least three times and with different antibodies (Extended Data Fig. 3e,f). **b**, Fluorescence recovery after photobleaching (FRAP). SRRM2-mVenus-dense areas were bleached, and recovery was monitored over 90 s. SRRM2-mVenus is displayed as greyscale. These images illustrate the data collection process leading to quantification shown in **b**–**d**. **c**, Arginine starvation reduced SRRM2-mVenus FRAP half time $t_{1/2}$ to 60% as compared with high arginine. Arginine concentrations relative to DMEM, $n = 8$, $^{**}P = 0.001$. Replicates are averages of 16 or more regions of interest (ROIs). **d**, Stable knockdown of ArgRS in SRRM2-mVenus cell lines with two different shRNAs (shRARS_1 and shRARS_2) reduced ArgRS protein levels to 40% of a non-targeting shRNA control (ctrl), shown by western blot. Relative *RARS* mRNA levels quantified by qRT–PCR: shSCR, 1.00 ± 0.17; shRARS_1, 0.20 ± 0.17; shRARS_2, 0.07 ± 0.06. A representative of more than three repeats is shown. **e**, Left: recovery of fluorescence after photobleaching. Means with s.e.m. are shown. $n = 5$, 5 and 7. Right: $t_{1/2}$ of SRRM2-mVenus recovery time after bleaching was reduced to 70% by stable ArgRS knockdown. shSCR, $n = 10$; shRARS_1, $n = 11$; shRARS_2, $n = 12$ averages of 16 or more ROIs, $P = {}^{*}0.048$, $^{**}0.007$. In **c** and **e**, means with s.e.m. are shown. We used an unpaired two-tailed *t*-test.

protein levels in SRRM2-mVenus 293T cells to ~40% compared with a non-targeting shRNA control (Fig. 3d). This level of ArgRS knockdown did not decrease cell viability or proliferation (Extended Data Fig. 5f), in agreement with findings for knockdowns of other aminoacyl-tRNA synthetases[41,42] and in a haplo-insufficient mouse[43]. SRRM2 protein expression levels were not reduced by the ArgRS knockdowns (Extended Data Fig. 5g). In both ArgRS knockdown cell lines, the half time of recovery for SRRM2-mVenus after bleaching decreased to ~70% of the control ($P = 0.048$ and 0.007) (Fig. 3e and Supplementary Videos 5–7), consistent with the increased mobility of SRRM2 seen for arginine starvation (Fig. 3c). Collectively, these findings indicate that ArgRS slowed SRRM2 trafficking, and that this interaction is modulated by external arginine concentrations, which in turn control nuclear ArgRS levels.

ArgRS affects processing of mRNAs but not small ncRNAs

As SRRM2 is both an organizer of nuclear compartments and a spliceosome component, we hypothesized that arginine-mediated changes in nuclear ArgRS levels, which impact SRRM2 mobility, could affect RNA processing. As manipulation of arginine triggers several cross-talking cellular signalling pathways[4], we used ArgRS and SRRM2 knockdowns as surrogates to assess how the interaction of SRRM2 with ArgRS might affect gene expression. We stably knocked down ArgRS in HepG2 cells (to ~30% on the protein level and ~40% on the RNA level, Extended Data Fig. 6a,b) and reset the cellular response to arginine by 6 h of arginine starvation followed by 2 h of arginine supplementation. Previously, we lowered total cellular ArgRS by 90% without affecting tRNA aminoacylation or mRNA translation[9]. Consequently, the milder knockdown of ArgRS did not decrease rates of cell proliferation or protein synthesis as monitored by puromycin incorporation assay[44] (Extended Data Fig. 6c,d).

For a comprehensive view of the effect of ArgRS depletion on cellular RNA synthesis, we used thermostable group II intron reverse transcriptase sequencing (TGIRT-seq) of ribosomal RNA-depleted cellular RNAs, which enables profiling of coding and non-coding RNAs (ncRNAs) simultaneously (Fig. 4a; ref. 45). Comparing 21,540 different coding and ncRNAs showed relatively little difference in the proportion of different RNA biotypes between the ArgRS shRNA knockdown and control shRNA (Fig. 4b). Most differentially expressed genes were protein coding (Fig. 4c,d). The ArgRS knockdown also resulted in numerous differences in exon usage for these genes (Extended Data Fig. 6e). In contrast, we found only minor differences in the levels of small ncRNAs (Fig. 4e–h), suggesting that ArgRS knockdown did not grossly perturb cellular RNA metabolism or induce cell stress responses[46]. Only a small change in the relative abundance of one tRNA ($tRNA^{Asn}$-GTT, $P = 0.01$) and no significant changes in $tRNA^{Arg}$ isodecoders were observed (Fig. 4h and Supplementary Table 6). Collectively, these findings indicate that the ArgRS knockdown primarily affected the expression and/or alternative splicing of protein-coding genes.

ArgRS knockdown alters mRNA processing and protein isoforms

Due to the higher representation of small ncRNAs, the sequencing depth we achieved with TGIRT-seq for mRNA exon–exon junctions was

https://doi.org/10.1038/s41556-023-01118-8

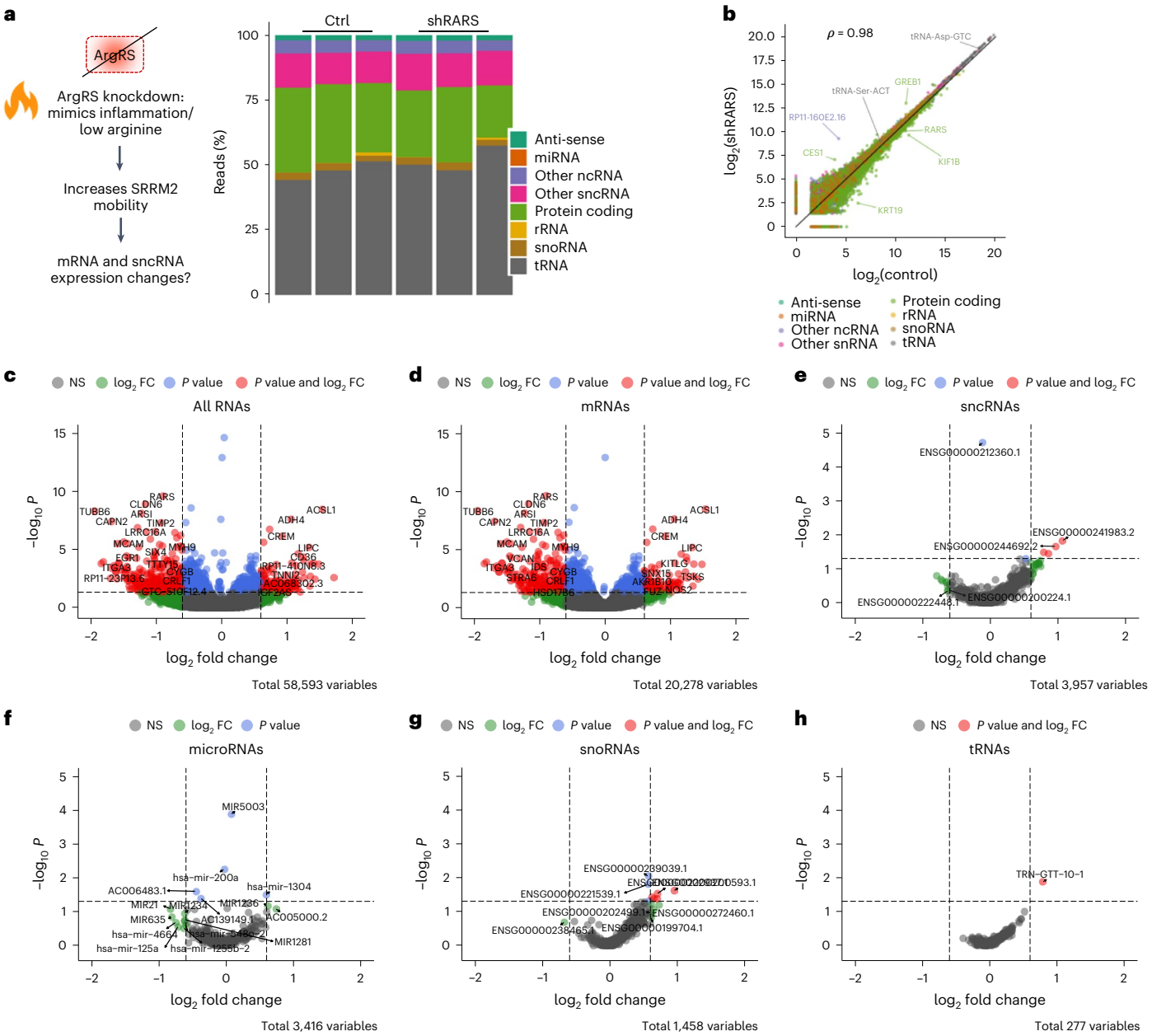


Fig. 4 | ArgRS knockdown primarily affected mRNAs but not ncRNAs. **a**, Stacked bar graphs showing RNA biotype distributions. **b**, Scatter plot comparing RNAs. ρ is the Spearman's correlation coefficient. **c**–**h**, Volcano plots comparing all and different RNA biotypes in ArgRS knockdown versus control HepG2 cells. Vertical dashed lines depict a 1.5-fold change (FC) and horizontal dashed lines indicate padj <0.05. Differential expression of all RNAs (**c**), mRNAs (**d**), small non-coding RNAs (**e**), microRNAs (**f**), small nucleolar RNAs (**g**) and tRNAs (**h**). In **a**–**h**, ctrl, HepG2 shSCR control; shRARS, HepG2 ArgRS knockdown. Three replicate TGIRT-seq datasets derived from different cell passages; *P* values were calculated with DESeq2. miRNA, microRNA; sncRNA, small non-coding RNA; rRNA, ribosomal RNA; snoRNA, small nucleolar RNA.

insufficient to draw conclusions about splice junction usage. Thus, we prepared and sequenced libraries from the same RNA preparations after pre-selection for poly(A)-containing mRNAs and used vast-tools[47] to evaluate previously annotated splice junctions on the basis of reads spanning exons (referred to as 'alternative splicing events'). We first identified alternative splicing events that were affected by ArgRS knockdown compared with the shSCR control. For this purpose, we used the minimum value of the difference (denoted MV) as a measure of significance, where MV >0 means a ≥0.95 probability that splice junction usage changed between the ArgRS knockdown and the control[48] (Methods). These events were then further classified by |ΔPSI|, defined as the per cent difference spliced in. We identified 393 events in 345 genes with alternative splicing changes between the ArgRS knockdown and control with MV >0 (Fig. 5a shows events with |ΔPSI| >0.1, and Supplementary Table 7). These alternative splicing events corresponded predominantly to the inclusion or skipping of cassette exons and increased retained introns, while inclusion or skipping of microexons (3–27 nt) or use of alternative 5′ or 3′ splice sites were relatively unaffected by the ArgRS knockdown (Fig. 5a).

We also analysed differential transcript usage with DEXseq[49], which compares expression levels of individual exons to the rest of the gene as opposed to quantifying splice junction spanning reads. By using this complementary method, we identified 1,004 exons in 533 genes that were differentially expressed upon ArgRS knockdown compared

https://doi.org/10.1038/s41556-023-01118-8

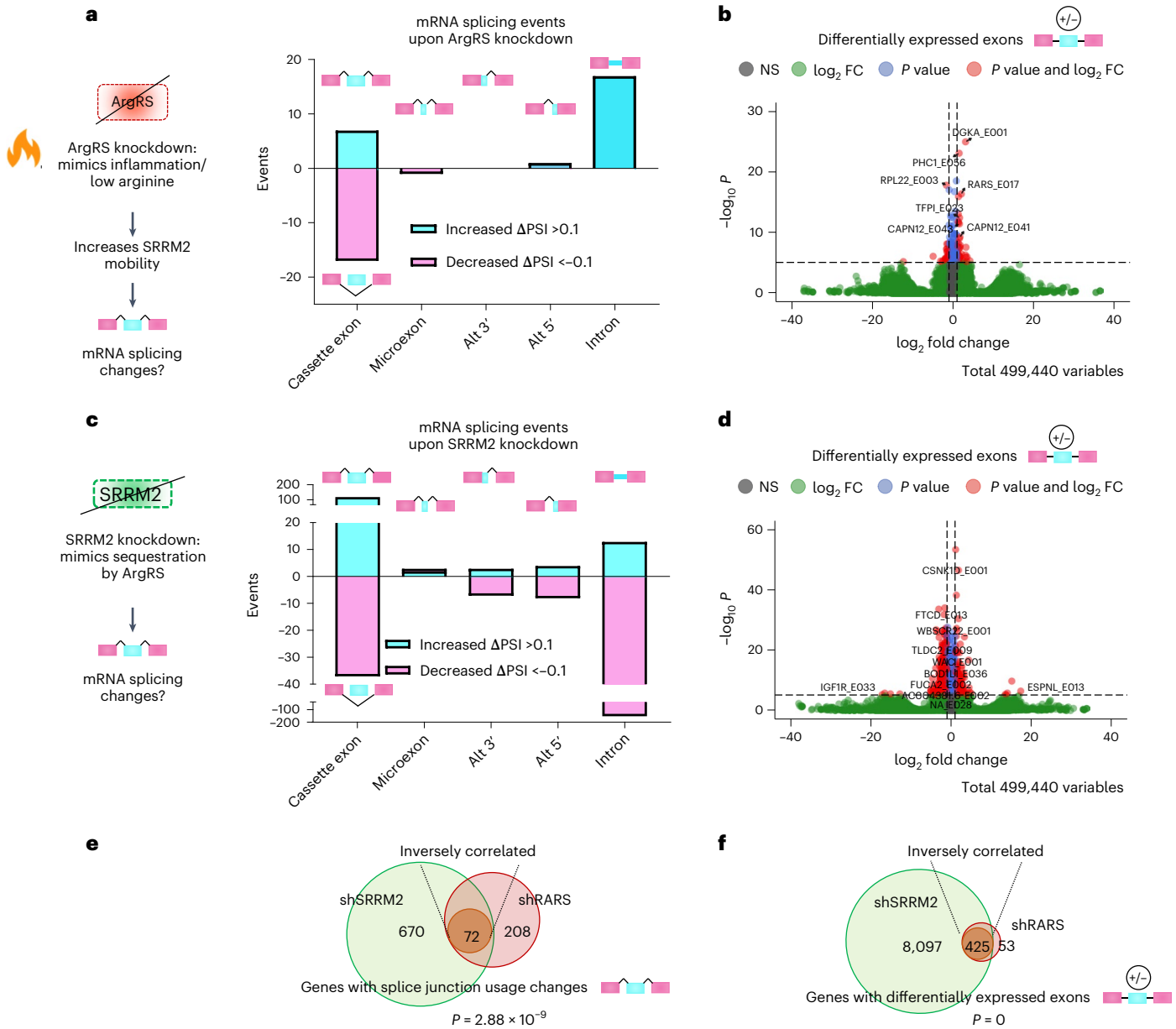


Fig. 5 | Differential mRNA processing upon ArgRS knockdown. a, Splicing event changes upon ArgRS knockdown (shRARS). **b**, Volcano plot of differentially expressed exons upon ArgRS knockdown (shRARS). **c**, Splicing event changes upon SRRM2 knockdown. **d**, Volcano plot of differential exon usage upon SRRM2 knockdown. In **a** and **c**, splicing events are divided by event type relative to shSCR control, as detected by RNA-seq. Events with a |ΔPSI| >0.1 and MV >0 are shown. Increased event frequency shown in cyan, events with decreased frequency in pink. MV >0: significant splicing changes; cassette: cassette exons/exon skipping; microexons: 3–15 nucleotide exons; alt 3: alternative 3′ end splice site; alt 5′: alternative 5′ end splice site; intron: intron retention. A total of 115,814 (shRARS) and 111,346 (shSRRM2) events were detected. In **b** and **d**, dashed vertical lines: twofold change; dashed horizontal line: *P* value 1×10^{-6}; NS, not significant. *P* values were calculated with DEXSeq. **e**,**f**, Venn diagram of genes with shared alternative splicing events or differentially expressed exons upon ArgRS and SRRM2 knockdown. Genes with inversely regulated splicing events (**e**) (ΔPSI >0 for shRARS and ΔPSI <0 for shSRRM2 or vice versa) or differentially expressed exons (**f**) ($\log_2$ fold change >0 for shRARS or <0 for shSRRM2 or vice versa) in the brown circle. Hypergeometric test for enrichment of inversely correlated, ArgRS-regulated splicing events was calculated in R. In **a**–**f**, data were derived from sequencing of three replicates (different cell passages). All changes relative to shSCR control.

with the control (adjusted *P* value (padj) <0.05, Fig. 5b and Supplementary Table 8).

We directly confirmed alternative splicing events upon ArgRS knockdown for *FGFR3* and *PLEKHG5* in HepG2 cells with reverse transcription polymerase chain reaction (RT–PCR) (Extended Data Fig. 7a,b). These genes displayed the same regulation in 293T cells upon ArgRS knockdown (Extended Data Fig. 7a,b). Rescue of ArgRS knockdown by overexpression of a *RARS* transgene resulted in a lower ratio of *FGFR3* splice variants similar to that seen in control cells (Extended Data Fig. 7c,d), but overexpression of catalytically inactive ArgRS did not (Extended Data Fig. 7e). This suggests that the catalytic activity or a specific conformation of ArgRS, which could be induced by arginine binding, was required for splicing regulation.

To investigate whether the splicing changes upon ArgRS knockdown led to an altered proteome, we assessed splice variants using mass spectrometry. As proteins derived from alternative splicing can

only be identified from individual peptides[50], we sought to increase the sensitivity of detection. Therefore, we focused on events where alternative splicing would give rise to distinct N-termini and enriched N-terminal peptides by removing all other (tryptic) peptides[51]. We identified 167 peptides from 135 proteins (Supplementary Table 9) that were detected at higher or lower intensity upon ArgRS knockdown (|pdiff| >1). Ten proteins were encoded by genes with differential exon usage and seven by genes with alternative splicing events upon ArgRS knockdown, including splicing changes that affected peptides in the proximity of potential alternative translation start sites (Extended Data Fig. 7f). For the splicing factor PTBP1, we saw an increase of the N-terminal peptide for one isoform, while the N-terminal peptide of another isoform was simultaneously reduced (Extended Data Fig. 7g). Differences in PTBP1 isoform usage were confirmed by western blots (Extended Data Fig. 7h). Collectively, these findings indicate that ArgRS knockdown at levels that do not affect cell proliferation results in alternative splicing changes that lead to different protein isoforms.

Identification of SRRM2-dependent splicing changes

Nuclear ArgRS interacts with and impedes SRRM2 mobility in the nucleus (Fig. 3) and could thereby reduce the levels of SRRM2 available for RNA splicing. To identify alternative splicing events that are sensitive to decreased levels of SRRM2, we knocked down SRRM2 by 50% (RNA and protein level, Extended Data Fig. 8a). This knockdown led to 1,033 alternative splicing events in 807 genes with MV >0 (Fig. 5c shows events with |ΔPSI| >0.1, Supplementary Table 10) and 42,140 differential exon usage events in 8,563 genes that were changed relative to the control shRNA (Fig. 5d and Supplementary Table 11, padj <0.05). The numbers of changed alternative splicing and differential exon usage events were much higher than those for the ArgRS knockdown (393 and 1,004, respectively) (Fig. 5a,b), as expected for a spliceosome component, whose activity might be subject to multiple modes of regulation. Cassette exons were changed in both directions upon SRRM2 knockdown, while introns were increasingly spliced out (Fig. 5c). To exclude that the alternative splicing changes were due to off-target effects of the SRRM2 knockdown with a single shRNA, we transfected 293T cells with a pool of small interfering RNAs (siRNAs) against SRRM2 (Extended Data Fig. 8b). Alternative splicing of *FGFR3* and *CDK5RAP3* genes could also be observed upon the reduction of SRRM2 with an heterogeneous pool of siRNAs, suggesting that this was a direct consequence of lowered SRRM2 availability (Extended Data Fig. 8b).

Inverse regulation of splicing by ArgRS and SRRM2

After obtaining lists of genes that were sensitive to the reduction of either SRRM2 or ArgRS, we next set out to identify genes that were shared between both. To control for indirect effects due to the knockdown of an aaRSs, we knocked down MetRS to 10% at the RNA and protein level (Extended Data Fig. 9a and Supplementary Table 12). We identified 602 alternative spicing events in 493 genes (|MV| >0, events with |ΔPSI| >0.1 shown in Extended Data Fig. 9b) and 50,896 differentially expressed exons in 7,668 genes that changed upon MetRS knockdown (Extended Data Fig. 9b).

A total of 282 alternative splicing events overlapped between the ArgRS, SRRM2 and MetRS knockdowns (Extended Data Fig. 9c, MV >0, no criterion for |ΔPSI| to depict all events regardless of magnitude). Of these, 70 splicing events were shared between the SRRM2 and ArgRS knockdowns (Extended Data Fig. 9c) and their ΔPSIs were negatively correlated ($R = -0.29$, $P = 0.037$). We then compared the directionality of the overlapping splicing events between ArgRS and SRRM2 knockdowns. This identified 27 of the 70 shared alternative splicing events (38.6%) that were inversely regulated by ArgRS and SRRM2, meaning that the knockdown of ArgRS led to a ΔPSI >0, while knockdown of SRRM2 led to a ΔPSI <0 or vice versa (Extended Data Fig. 9d, MV >0, hypergeometric test $P = 0.01$). In contrast, only 12 of 185 splicing events (6.2%) were inversely regulated by MetRS and SRRM2 (Extended Data Fig. 9d, MV >0, hypergeometric test $P = 1$). We confirmed splicing changes by RT–PCR for prominently changed events (*CDK5RAP3*, *IL32*, *RAD52* and *STX16*, all Extended Data Fig. 9e).

If focusing on genes that were alternatively spliced and shared between the ArgRS and SRRM2 knockdowns (instead of individual splicing events within a shared gene), almost 500 genes had splicing events or differentially expressed exons that were inversely regulated by ArgRS and SRRM2 (72 and 475, respectively, Fig. 5e,f). These findings suggested that the observed inverse regulation of alternative splicing by ArgRS and SRRM2 might lead to the production of different protein isoforms that contribute to the cellular response to inflammation.

Inverse mRNA splicing alters metabolism and communication

To explore the possible biological consequences of the ArgRS/SRRM2 interaction, we examined the enrichment of pathways and Gene Ontology (GO) terms in the genes changed in opposite directions in the ArgRS and SRRM2 knockdowns (Fig. 6a,b and Extended Data Fig. 10a–e). Among these were genes that might benefit the cellular response to inflammation such as the pro-inflammatory cytokine IL32 (ref. 52). Over-represented pathways included mTORC1 signalling, which is regulated by arginine (Fig. 6b).

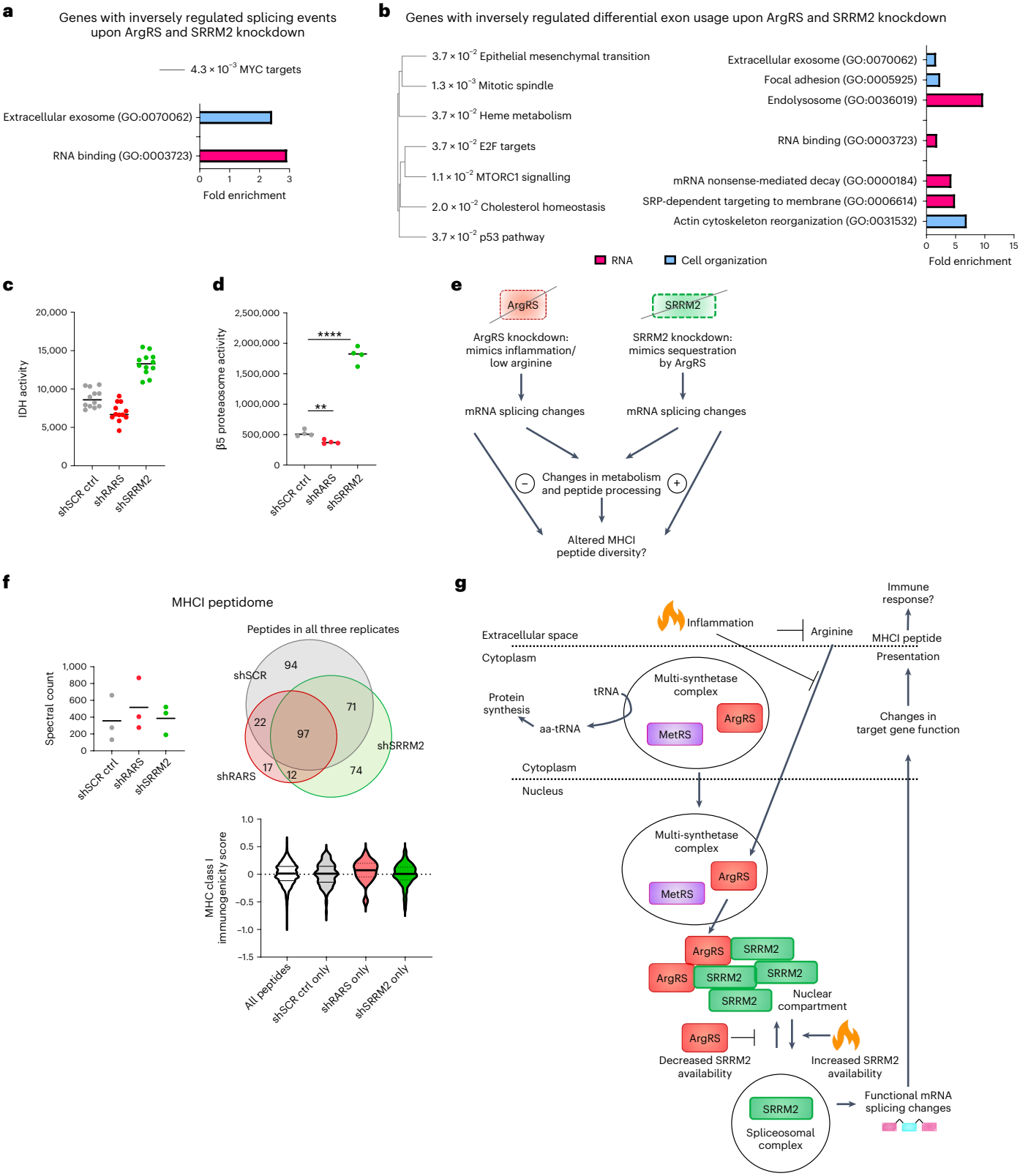
To investigate the functional consequences of differential exon usage that might be regulated by ArgRS/SRRM2, we chose several mTORC1-associated genes (Supplementary Table 13) and compared the enzymatic activity of their encoded proteins after ArgRS or SRRM2 knockdown. The enzyme Isocitrate dehydrogenase 1 (IDH1) responds to cell stress by regulating intra-cellular NADP/NADPH levels[53]. Differential exon usage in *IDH1* correlated with an increase in IDH enzymatic activity in cell lysates upon SRRM2 knockdown, while ArgRS knockdown led to decreased activity (Fig. 6c). Differential exon usage was also found for genes encoding components of the ubiquitin/proteasome system (UPS), including the 19S proteasomal subunit de-ubiquitinating enzyme UCHL5 and the catalytically active proteasomal subunit PSMB5 (β5 subunit) (Extended Data Fig. 10f). These correlated with increased β5 activity upon SRRM2 knockdown and decreased β5 activity upon ArgRS knockdown (Fig. 6d).

Fig. 6 | SRRM2- and ArgRS-regulated mRNA processing led to metabolic changes and an altered MHCI peptidome. a, Molecular Signatures Database (MSigDB) and GO analysis of genes with inversely regulated splicing events upon ArgRS and SRRM2 knockdown. **b**, MSigDB and GO analysis of genes with inversely regulated differentially expressed exons upon ArgRS and SRRM2 knockdown. In **a** and **b**, GO terms associated with RNA-interacting proteins are shown in fuchsia and GO terms associated with cellular organization and organelles are shown in blue. **c**, IDH1 enzyme activity. The IDH activity assay is representative of three repeats with 12 technical replicates (one representative repeat is shown, repeats are different passages). Deletion of ENSE00000934683 leads to an in-frame deletion of 126 amino acids. **d**, PSMB5 differential exon usage, β5 proteasomal subunit (PSMB5 protein) activity. Quantification of differentially used exon: β5 activity is representative of three repeats (different cell passages) with four replicates, $^{**}P = 0.006$, $^{****}P < 0.0001$; one-way analysis of variance and Tukey's multiple comparisons test. A representative repeat is shown. **e**, A model of ArgRS and SRRM2 knockdown, their connection to inflammation and consequences on their splice targets. **f**, Total spectral counts (upper left) correspond to peptide quantity. $n = 3$, 3 and 3 replicates, which are different cell passages. $P = 0.53$ and $P = 0.88$. Venn diagrams of identified MHCI peptides by mass spectrometry following isolation of MHCI complex (upper right) found in all replicates upon ArgRS or SRRM2 knockdown. Bottom: violin plot of MHCI immunogenicity score for peptides found in all samples, shSCR ctrl, ArgRS knockdown and SRRM2 knockdown only. Median MHCI immunogenicity score is indicated by a black line. In **a**–**f**, shSCR ctrl, HepG2 shSCR control; shRARS, HepG2 ArgRS knockdown; shSRRM2, HepG2 shSRRM2 knockdown. **g**, A model of cellular response to inflammation by arginine/ArgRS/SRRM2.

https://doi.org/10.1038/s41556-023-01118-8

The UPS provides peptides that are presented on the cell surface by the major histocompatibility complex I (MHCI) and alternative splicing of UPS components could result in different peptides being presented to immune cells. Additionally, splicing changes that result in different protein isoforms could yield alternative peptides directly through differences in amino acid sequence and differing protease cleavage sites. These MHCI-bound peptides communicate the status of the presenting cell to immune cells[54].

To investigate whether communication to immune cells through MHCI might be changed by ArgRS and SRRM2 knockdown (Fig. 6e), we compared their MHCI-bound peptidomes using mass spectrometry. We retrieved 1,601 unique MHCI-bound peptides from three



independent biological replicates of each of the three conditions (shSCR control, shRARS and shSRRM2, Supplementary Table 16). These peptides were 8–14 amino acids long and matched previously described MHCI-bound peptides in length and preference for Tyr and Val (ref. 55) (Fig. 6f and Extended Data Fig. 10g). Overall, MHCI protein levels and the spectral counts for the measured peptides were comparable between all treatments (Extended Data Fig. 10g). However, upon sequencing of the peptides presented by MHCI by mass spectrometry, we found that ArgRS knockdown substantially reduced the peptide diversity displayed on MHCI relative to the shSCR control (148 versus 284, found in all three replicates, Fig. 6f and Extended Data Fig. 10h). In contrast, SRRM2 knockdown led to the presentation of a comparable number of different peptides by MHCI as the shSCR control (284 versus 254 in all three replicates, Fig. 6f and Extended Data Fig. 10h). Seventy-four of these peptides were unique for the SRRM2 knockdown cell line and 94 were unique for the shSCR control, suggesting that SRRM2 knockdown led to the presentation of an altered set of MHCI-bound peptides (Fig. 6f and Extended Data Fig. 10h).

To predict how these MHCI peptide pools might influence communication with immune cells, we calculated MHCI immunogenicity scores (Fig. 6f). A higher score predicts that peptides would more likely be recognized by T cells and therefore better at eliciting an immune response[56]. Consistent with our hypothesis that ArgRS knockdown mimics inflammation, MHCI peptides presented only by ArgRS knockdown cells had a higher median score than peptides that were presented only by SRRM2 knockdown cells (Fig. 6f). Collectively, these findings support a model in which alternative splicing events that are inversely regulated by ArgRS and SRRM2 could contribute to changes in cellular metabolism and communication as part of a response to inflammation (Fig. 6g).

Discussion

Here we found that a fraction of ArgRS localizes to the cell nucleus and that nuclear ArgRS levels fluctuate in an arginine-dependent manner (Fig. 1). We identified SRRM2 in the ArgRS interactome by mass spectrometry (Fig. 2). Further, SRRM2 and ArgRS localized in proximity to each other in the nucleus and reduction of nuclear ArgRS increased SRRM2 mobility (Fig. 3), suggesting that the interaction impedes SRRM2 trafficking. ArgRS knockdown correlated with changes in SRRM2-dependent alternative splicing, while ncRNAs were largely unaffected (Figs. 4 and 5). SRRM2 and ArgRS modulated alternative splicing of a subset of genes in opposite directions, suggesting that ArgRS could affect SRRM2's function (Fig. 5), contributing to altered biochemical activities of the encoded proteins and MHCI peptide presentation (Fig. 6).

Arginine levels decrease during inflammation[2], and eventually, the depletion of arginine leads to the resolution of inflammation due to a restraining effect on immune cells, especially T cells[3,57]. As manipulation of arginine triggers several cross-talking cellular signalling pathways, including mTOR[4], we used ArgRS knockdowns to investigate the effects of ArgRS' interaction with SRRM2. The observed downregulation of IDH1 activity and thereby NADPH production by ArgRS knockdown would support the metabolic shutdown initiated by mTOR inhibition upon arginine depletion.

We showed previously that the presence of ArgRS in the nucleus is dependent on the MSC as nuclear ArgRS was strongly reduced upon exclusion of ArgRS from the MSC[9]. Here, we found that ArgRS—once in the nucleus—interacted with SRRM2, independent of the MSC. These findings suggest that in addition to tRNA quality control[12], the nuclear MSC might act as a reservoir of nuclear aaRSs with regulatory functions, similar to the cytoplasmic MSC[11]. ArgRS, with changing arginine levels as a cue, thus connects two distinct protein organizational structures with central roles in transcription and translation. Due to their ability to sense and respond to changes in amino acids levels and interact with a variety of other proteins, aaRSs are uniquely poised to function in metabolite sensing and signal transduction. In addition to sequestration, ArgRS could potentially add post-translational modifications to SRRM2 by arginylating lysine side chains[7].

Approximately 50% of ArgRS was bound by SRRM2 and/or associated with the MSC. Our interactome study identified both direct and indirect interaction partners of ArgRS, as shown by the retrieval of MSC components that are not in direct contact with ArgRS. Thus, we cannot exclude that additional factors might be needed to mediate the interaction between ArgRS and SRRM2 or to effect the observed changes in SRRM2 mobility and alternative splicing. Other nuclear speckle proteins could also alter SRRM2 distribution and trafficking within the nucleus. Lastly, we cannot exclude that ArgRS and SRRM2 also interact outside of discrete nuclear structures, which would be more challenging to visualize by microscopy.

SRRM2 is both essential for nuclear speckle formation[30] and a component of active spliceosomes[58]. Importantly, RNA transcription and splicing are reported to take place at or in close proximity to nuclear speckles[26]. The interaction between ArgRS and SRRM2 at the interface of SRRM2-rich areas could therefore impact alternative mRNA splicing by disrupting the function of SRRM2 within the spliceosome or its cellular trafficking. The N-terminal 1% of SRRM2 is immobilized sufficiently in the spliceosome to be resolved by cryo-electron microscopy, while the remaining 99% (including the low-complexity Ser/Arg-rich tail) are disordered[58]. SRRM2 would therefore be an ideal candidate to anchor and organize spliceosomes within or on the periphery of nuclear speckles, and provides ample interaction surface for other regulatory proteins.

A previous study in a murine liver cell line found that depletion of SRRM2 to lower levels than in our study resulted in the upregulation of immune response gene transcription and processing[59]. Here we found that modulation of SRRM2 levels impacted cellular metabolism and MCHI peptide processing (Fig. 6). MHCI presentation allows cells to communicate with the immune system[54]. In pathologies such as viral infection or cancer, an altered MHCI peptidome could induce different immune responses. While the expression of specialized proteasome subunits strongly influences T-cell maturation[60,61], effects on the MHCI repertoire have also been described[62]. We found that ArgRS knockdown, which mimics arginine starvation by decreasing ArgRS in the nucleus, reduced peptide diversity and increased the immunogenicity of peptides presented by MHCI (Fig. 6f). Under inflammatory conditions, this could enable more efficient communication with the immune system. Recently, the MHCI peptidome has been modulated with splicing inhibitors to increase sensitivity to immune therapy in cancer[63], and our findings suggest that a similar mechanism could occur physiologically during inflammation. The interplay found here among arginine, ArgRS and SRRM2 would enable cells to detect ongoing extracellular inflammation by sensing metabolic cues and respond adequately by adapting their proteome and communication with immune cells.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-023-01118-8.

References

1. Netea, M. G. et al. A guiding map for inflammation. *Nat. Immunol.* **18**, 826–831 (2017).
2. Bronte, V. & Zanovello, P. Regulation of immune responses by L-arginine metabolism. *Nat. Rev. Immunol.* **5**, 641–654 (2005).
3. Murray, P. J. Amino acid auxotrophy as a system of immunological control nodes. *Nat. Immunol.* **17**, 132–139 (2016).

4. Bar-Peled, L. & Sabatini, D. M. Regulation of mTORC1 by amino acids. *Trends Cell Biol.* **24**, 400–406 (2014).
5. Weichhart, T., Hengstschläger, M. & Linke, M. Regulation of innate immune cell function by mTOR. *Nat. Rev. Immunol.* **15**, 599–614 (2015).
6. Schimmel, P. R. & Söll, D. Aminoacyl-tRNA synthetases: general features and recognition of transfer RNAs. *Annu. Rev. Biochem.* **48**, 601–648 (1979).
7. He, X.-D. et al. Sensing and transmitting intracellular amino acid signals through reversible lysine aminoacylations. *Cell Metab.* **27**, 151–166.e6 (2018).
8. Vo, M.-N. et al. ANKRD16 prevents neuron loss caused by an editing-defective tRNA synthetase. *Nature* **557**, 510–515 (2018).
9. Cui, H. et al. Regulation of ex-translational activities is the primary function of the multi-tRNA synthetase complex. *Nucleic Acids Res.* **49**, 3603–3616 (2021).
10. Lee, S. W., Cho, B. H., Park, S. G. & Kim, S. Aminoacyl-tRNA synthetase complexes: beyond translation. *J. Cell Sci.* **117**, 3725–3734 (2004).
11. Ray, P. S., Arif, A. & Fox, P. L. Macromolecular complexes as depots for releasable regulatory proteins. *Trends Biochem. Sci.* **32**, 158–164 (2007).
12. Lund, E. & Dahlberg, J. E. Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science* **282**, 2082–2085 (1998).
13. Guo, M. & Schimmel, P. Essential nontranslational functions of tRNA synthetases. *Nat. Chem. Biol.* **9**, 145–153 (2013).
14. Sajish, M. & Schimmel, P. A human tRNA synthetase is a potent PARP1-activating effector target for resveratrol. *Nature* **519**, 370–373 (2015).
15. Yannay-Cohen, N. et al. LysRS serves as a key signaling molecule in the immune response by regulating gene expression. *Mol. Cell* **34**, 603–611 (2009).
16. Shi, Y. et al. tRNA synthetase counteracts c-Myc to develop functional vasculature. *eLife* **3**, e02349 (2014).
17. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
18. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
19. Zhu, L. & Brangwynne, C. P. Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.* **34**, 23–30 (2015).
20. Spector, D. L. & Lamond, A. I. Nuclear speckles. *Cold Spring Harb. Perspect. Biol.* **3**, a000646 (2011).
21. Galganski, L., Urbanek, M. O. & Krzyzosiak, W. J. Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res.* **45**, 10350–10368 (2017).
22. Guo, Y. E. et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019).
23. Galarza-Muñoz, G. et al. Human epistatic interaction controls IL7R splicing and increases multiple sclerosis risk. *Cell* **169**, 72–84.e13 (2017).
24. Braunschweig, U. et al. Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res.* **24**, 1774–1786 (2014).
25. Pan, Q. et al. Revealing global regulatory features of mammalian alternative splicing using a quantitative microarray platform. *Mol. Cell* **16**, 929–941 (2004).
26. Chen, Y. & Belmont, A. S. Genome organization around nuclear speckles. *Curr. Opin. Genet. Dev.* **55**, 91–99 (2019).
27. Smith, K. P., Hall, L. L. & Lawrence, J. B. Nuclear hubs built on RNAs and clustered organization of the genome. *Curr. Opin. Cell Biol.* **64**, 67–76 (2020).
28. Blencowe, B. J., Issner, R., Nickerson, J. A. & Sharp, P. A. A coactivator of pre-mRNA splicing. *Genes Dev.* **12**, 996–1009 (1998).
29. Blencowe, B. J. et al. The SRm160/300 splicing coactivator subunits. *RNA* **6**, 111–120 (2000).
30. Ilik, İ. A. et al. SON and SRRM2 are essential for nuclear speckle formation. *eLife* **9**, e60579 (2020).
31. Zhang, X. et al. An atomic structure of the human spliceosome. *Cell* **169**, 918–929.e14 (2017).
32. Gautam, A., Grainger, R. J., Vilardell, J., Barrass, J. D. & Beggs, J. D. Cwc21p promotes the second step conformation of the spliceosome and modulates 3′ splice site selection. *Nucleic Acids Res.* **43**, 3309–3317 (2015).
33. Kaplanis, J. et al. Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature* **586**, 757–762 (2020).
34. Tomsic, J. et al. A germline mutation in SRRM2, a splicing factor gene, is implicated in papillary thyroid carcinoma predisposition. *Sci. Rep.* **5**, 10566 (2015).
35. Shehadeh, L. A. et al. SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. *PLoS ONE* **5**, e9104 (2010).
36. Nathanson, L. & Deutscher, M. P. Active aminoacyl-tRNA synthetases are present in nuclei as a high molecular weight multienzyme complex. *J. Biol. Chem.* **275**, 31559–31562 (2000).
37. Horiguchi, N. et al. Dissociation between liver inflammation and hepatocellular damage induced by carbon tetrachloride in myeloid cell-specific signal transducer and activator of transcription 3 gene knockout mice. *Hepatology* **51**, 1724–1734 (2010).
38. Keilhauer, E. C., Hein, M. Y. & Mann, M. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE–MS) rather than affinity purification mass spectrometry (AP–MS). *Mol. Cell. Proteom.* **14**, 120–135 (2015).
39. Kim, M. H. & Kim, S. Structures and functions of multi-tRNA synthetase complexes. *Enzymes* **48**, 149–173 (2020).
40. Costes, S. V. et al. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys. J.* **86**, 3993–4003 (2004).
41. Ofir-Birin, Y. et al. Structural switch of lysyl-tRNA synthetase between translation and transcription. *Mol. Cell* **49**, 30–42 (2013).
42. Hayano, M., Yang, W. S., Corn, C. K., Pagano, N. C. & Stockwell, B. R. Loss of cysteinyl-tRNA synthetase (CARS) induces the transsulfuration pathway and inhibits ferroptosis induced by cystine deprivation. *Cell Death Differ.* **23**, 270–278 (2016).
43. Seburn, K. L., Nangle, L. A., Cox, G. A., Schimmel, P. & Burgess, R. W. An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a charcot-marie-tooth 2D mouse model. *Neuron* **51**, 715–726 (2006).
44. Schmidt, E. K., Clavarino, G., Ceppi, M. & Pierre, P. SUnSET, a nonradioactive method to monitor protein synthesis. *Nat. Methods* **6**, 275–277 (2009).
45. Boivin, V. et al. Simultaneous sequencing of coding and noncoding RNA reveals a human transcriptome dominated by a small number of highly expressed noncoding genes. *RNA* **24**, 950–965 (2018).
46. Watkins, C. P., Zhang, W., Wylder, A. C., Katanski, C. D. & Pan, T. A multiplex platform for small RNA sequencing elucidates multifaceted tRNA stress response and translational regulation. *Nat. Commun.* **13**, 2491 (2022).
47. Tapial, J. et al. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res.* **27**, 1759–1768 (2017).

48. Han, H. et al. Multilayered control of alternative splicing regulatory networks by transcription factors. *Mol. Cell* **65**, 539–553.e7 (2017).
49. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).
50. Blencowe, B. J. The relationship between alternative splicing and proteomic complexity. *Trends Biochem. Sci.* **42**, 407–408 (2017).
51. Kleifeld, O. et al. Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates. *Nat. Protoc.* **6**, 1578–1611 (2011).
52. Aass, K. R., Kastnes, M. H. & Standal, T. Molecular interactions and functions of IL-32. *J. Leukoc. Biol.* **109**, 143–159 (2021).
53. Itsumi, M. et al. Idh1 protects murine hepatocytes from endotoxin-induced oxidative stress by regulating the intracellular $NADP^+$/NADPH ratio. *Cell Death Differ.* **22**, 1837–1845 (2015).
54. Rock, K. L., Reits, E. & Neefjes, J. Present yourself! By MHC Class I and MHC Class II molecules. *Trends Immunol.* **37**, 724–737 (2016).
55. Bassani-Sternberg, M., Pletscher-Frankild, S., Jensen, L. J. & Mann, M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol. Cell. Proteom.* **14**, 658–673 (2015).
56. Calis, J. J. A. et al. Properties of MHC class I presented peptides that enhance immunogenicity. *PLoS Comput. Biol.* **9**, e1003266 (2013).
57. Geiger, R. et al. L-Arginine modulates T cell metabolism and enhances survival and anti-tumor activity. *Cell* **167**, 829–842.e13 (2016).
58. Bertram, K. et al. Structural insights into the roles of metazoan-specific splicing factors in the human Step 1 spliceosome. *Mol. Cell* **80**, 127–139.e6 (2020).
59. Hu, S., Lv, P., Yan, Z. & Wen, B. Disruption of nuclear speckles reduces chromatin interactions in active compartments. *Epigenetics Chromatin* **12**, 43 (2019).
60. Kincaid, E. Z., Murata, S., Tanaka, K. & Rock, K. L. Specialized proteasome subunits have an essential role in the thymic selection of $CD8^+$ T cells. *Nat. Immunol.* **17**, 938–945 (2016).
61. van den Eshof, B. L., Medfai, L., Nolfi, E., Wawrzyniuk, M. & Sijts, A. J. A. M. The function of immunoproteasomes—an immunologists' perspective. *Cells* **10**, 3360 (2021).
62. Kincaid, E. Z. et al. Mice completely lacking immunoproteasomes show major changes in antigen presentation. *Nat. Immunol.* **13**, 129–135 (2011).
63. Lu, S. X. et al. Pharmacologic modulation of RNA splicing enhances anti-tumor immunity. *Cell* **184**, 4032–4047.e31 (2021).

Methods

Our research complies with all relevant ethical regulations. Specifically, all animals were humanely euthanized in accordance with animal protocol 19-0023, which was approved by the Scripps Research IACUC Department. Reagents are listed in Supplementary Table 17.

Mouse model

Staff and researchers were wearing scrubs, facemasks and gloves while handling animals. The rooms were kept between 69 °F and 78 °F, on a 12 h dark/light cycle and without regulated humidity. Animals were housed in an individually ventilated rack caging system, consisting of a cage bottom, wire bar lid and filter top, which was sanitized in a cage wash. Feed was commercially available laboratory-grade rodent diet. Purified water through reverse osmosis was provided through an automatic watering system. Cages were changed approximately every 10–14 days.

Six-to-8-week-old, female C57BL6/J mice from the Scripps Research breeding facility were used. Animals with the same birthdate were randomly assigned to experimental groups. All animals were humanely euthanized in accordance with animal protocol 19-0023, which was approved by the Scripps Research IACUC Department. Groups consisted of mice injected with (1) 100 µl 35 mg ml^{-1} murine Arginase, (2) 100 µl phosphate-buffered saline (PBS) control, (3) 50 µl 20% CCl4 in olive oil or (4) 50 µl olive oil.

Cell lines

Both cell lines were regularly tested for mycoplasma contamination with a PCR-based, in-house assay and were consistently negative. HepG2 and 293T cells were verified by the ATCC Cell Line Authentication service using STR Profiling Results. Both cell lines were found to be a 93% match with the reference cell line profile. HepG2 cells were originally derived from a male donor and human embryonic kidney (HEK) 293T cells from a female embryo. Murine embryonic fibroblasts (MEFs) were a kind gift from Yao Tong (Scripps Research). MEFs were generated from C57Bl/6J embryos around embryonic day E13.5 and used at cell passage 3. The sex of the embryos that were used for MEF generation was not determined. HepG2 human hepatocellular carcinoma cells and HEK 293T cells were cultured in high-glucose Dulbecco's modified Eagle medium (DMEM; Gibco) supplemented with 10% foetal bovine serum (FBS, Omega Scientific) and penicillin–streptomycin (Gibco).

Arginase-1 activity assay

Purified Arginase-1 was mixed with a reaction buffer containing 320 mM arginine, 100 mM HEPES pH 8.5 and 10 mM $CoCl_2$. The reaction was incubated for 45 min at 37 °C in a thermocycler. Urea was detected by a colourimetric assay monitoring the conversion of α-isonitrosopropiophenone. α-isonitrosopropiophenone was dissolved in ethanol (2 g 50 ml^{-1}) and freshly diluted 1:10 in 1:1:3 water:sulfuric acid:phosphoric acid to make urea detection reagent. Urea detection reagent was mixed with the Arginase-1 reaction product or a urea standard 5:1 and incubated in a thermocycler for 60 min at 100 °C followed by 15 min at 25 °C.

Arginine depletion in mice

Power analysis was used to pre-determine sample size. To enzymatically deplete systemic arginine, mice were intra-peritoneally injected with 35 mg kg^{-1} recombinant, lipopolysaccharide-free (6.15 endotoxin units (EU) mg^{-1}) murine Arginase-1 purified from clear *Escherichia coli* (Lucigen) and humanely euthanized 5–7 h later. Arginase-1 activity was tested by measuring conversion of arginine to urea as described above.

CCl_4 (Sigma-Aldrich) was used to elicit systemic inflammation. Mice were intra-peritoneally injected with 50 µl 20% CCl_4 in olive oil, or olive oil only, and humanely euthanized 30 h later. In all experiments, blood was collected from the inferior vena cava with EDTA-coated syringes and plasma was separated by centrifugation at 500*g* for 3 min at 4 °C.

Arginine determination in plasma

As an internal control, 5 µl 250 µM $13C_6$-arginine (Pierce) was spiked in 20 µl plasma. Plasma was precipitated with 200 µl 80% ice-cold methanol and dried in a SpeedVac. Metabolites were resuspended in 40 µl water and incubated with 20 µl freshly made 1% Marfey's reagent in acetone (wt/vol) and 4 µl 1 M bicarbonate for 1 h at 40 °C in a thermocycler. Derivatized amino acids were dried in a SpeedVac and resuspended in 40 µl 40% methanol. Samples were measured in the Scripps Research open-access mass spectrometry core using an Agilent 6125 single quadrupole mass spectrometer coupled to an Agilent 1260 LC stack. The column used was an Agilent SB-C8 5 µm 300 A 4.6 × 50 mm running at a flow rate of 0.5 ml min^{-1}. Mobile phases consisted of A (H_2O/0.1% formic acid) and B (ACN/0.1% formic acid). Spray chamber conditions were dry gas flow rate of 12 l min^{-1} at 350 C, nebulizer pressure of 35 psi and capillary voltage of 4,000 V for positive mode and 3,500 V for negative mode. Sample (20 µl) was injected and eluted with a 15 ml 10–35% A to B gradient. Derivatized arginine was eluted at 9.3 ml. Spectra were collected in a mass range of 150–600 for both positive and negative mode and additionally in the 420–430 Da and 430–440 Da range.

Modulation of arginine levels

For the modulation of arginine in tissue culture, SILAC DMEM with dialysed FBS and substituted with lysine was used as a base medium. Arginine concentrations are given as relative to 1× DMEM (0.084 g l^{-1} L-arginine monohydrochloride, dissolved in PBS). For 0×, arginine concentrations were under the detection limit of the assay described above. Cells were deprived of arginine for 4–6 h, and arginine was added back at the indicated concentrations for 2 h.

Cell fractionation

Cultured cells were seeded at a density of 5×10^6 cells per 10 cm tissue culture dish the day before (293T, MEF) or at 2.5×10^6 per 10 cm tissue culture dish 2 days before (HepG2) collection in 500 µl cell fractionation buffer (20 mM HEPES, pH 7.5, 10 mM KCl, 2 mM $MgCl_2$, 2 mM EDTA, 1 mM dithiothreitol (DTT) and protease inhibitor). Spleens of arginine-depleted or control mice were homogenized by passing through a 40 µm cell sieve. Mouse livers were collected from 4-week-old C57Bl/6J mice and homogenized using a 40 µm cell sieve. The cell suspension was left on ice for 20 min for osmotic lysis and 100 µl was taken as whole cell lysate. The remaining cells were passed ten times through a 27-gauge needle. Nuclei and cytoplasm were separated by centrifugation at 750*g* for 5 min. The supernatant was taken as cytoplasmic fraction and the nuclear pellet washed three times with cell fractionation buffer. Radioimmunoprecipitation assay buffer (10×: 0.5 M Tris–HCl, pH 7.4, 1.5 M NaCl, 2.5% deoxycholate, 10% IGE-PAL CA-630 and 10 mM EDTA) was added to 2× final concentration to all fractions. Complete lysis was achieved by freezing and thawing. Lysates were spun down at 14,000*g* for 20 min to remove insoluble components and sodium dodecyl sulfate (SDS) loading buffer was added before western blot analysis.

Affinity enrichment/co-immunoprecipitation

Two days before the experiment, 1×10^7 293T or 5×10^6 HepG2 cells were seeded on a 15 cm tissue culture dish. Cells were lysed for 20 min under mild conditions in 1% IGE-PAL CA-630 (Sigma-Aldrich) in Tris-buffered saline (TBS) with protease (Pierce Protein Biology) and phosphatase inhibitors (Pierce Protein Biology). Insoluble components were pelleted by centrifugation at 14,000*g* for 20 min. Protein A/G agarose beads (30 µl, Santa Cruz Biotechnology) were equilibrated in TBS. Cell lysate was added to the beads together with 2 µl antibody against ArgRS (Biorbyt, RARS-5; ref. 64). For immunoprecipitation of SRRM2, antibody against SRRM2 (Santa Cruz Biotechnology) was conjugated to beads using the Co-Immunoprecipitation kit (Pierce Protein Biology) according to the manufacturer's instructions or used as described for

ArgRS (Life Technologies antibody). For pull downs using biotinylated recombinant protein as bait, streptavidin sepharose beads were used (Biovision). Immunoprecipitation was performed for 3 h (ArgRS) or overnight (SRRM2, biotinylated proteins). After incubation, the beads were washed twice with TBS containing 0.1% IGE-PAL CA-630 and twice with TBS only. Precipitated proteins were either eluted using 0.1 M glycine, pH 3, for western blot or by tryptic digestion for interactome analysis.

Western blot

Cell lysate from different cellular fractions or immunoprecipitated proteins were run on a 4–15% gradient gel (Bolt, ThermoFisher Scientific) and subsequently blotted on a polyvinylidene difluoride membrane using an iBlot or iBlot2 device (ThermoFisher Scientific). For SRRM2 western blots, maximum voltage and transfer time were used to account for the high molecular weight of SRRM2. Membranes were blocked for at least 1 h with 5% bovine serum albumin (BSA) or 1% dry milk in TBS-T. Membranes were then cut horizontally and incubated in the corresponding antibody. Primary antibody was incubated overnight at 4 °C at a 1:5,000 (ArgRS: Biorbyt, GAPDH, Tubulin, Histone H3, Lamin A/C: Cell Signaling Technology, MetRS: Abcam, GFP, PTBP1: Proteintech Group) or a 1:100 (SRRM2, Santa Cruz Biotechnology) dilution in 5% BSA/TBS-T. The next day, blots were washed three times for at least 10 min per wash in TBS-T and HRP-labelled secondary antibody (Invitrogen) was added at a 1:10,000 (anti-rabbit) or 1:5,000 (anti-mouse) dilution in BSA/TBS-T or milk/TBS-T for at least 90 min. For murine ArgRS detection, secondary antibody was incubated overnight. After three additional washes, membranes were developed with ProSignal Dura ECl (Genesee Scientific) and imaged with a FluorChem M (Proteinsimple). All sections of one blot were developed together, but exposure time was optimized for each antibody.

ArgRS interactome

Proteomes were eluted by incubation with 25 µl 2 M urea, 5 ng µl^{-1} trypsin (Pierce Protein Biology), 1 mM DTT in 50 mM Tris–HCl, pH 7.5. After 30 min at room temperature, 5 mM chloroacetamide (Sigma-Aldrich) was added for alkylation of cysteine residues to a total volume of 125 µl. Samples were digested overnight at 37 °C and quenched with 0.5% trifluoroacetic acid (Sigma-Aldrich) the next morning. Stage tip desalting was done using C18 spin tips with a 100 µl bed (Pierce Protein Biology) according to the manufacturer's instructions. The digested samples were analysed on a Q Exactive mass spectrometer (ThermoFisher Scientific). Approximately 250 ng of digest was injected directly onto a 20 cm, 100 µm inner diameter column packed with Aqua 3 µm C18 resin (Phenomenex). Samples were separated at a flow rate of 400 nl min^{-1} on an Easy nLCII (ThermoFisher Scientific). Buffers A and B were 0.1% formic acid in 5% acetonitrile and 0.1% formic acid in 80% acetonitrile, respectively. A gradient of 1–35% B over 80 min, an increase to 80% B over 25 min and holding at 80% B for 5 min before returning to 1% B was used for 120 min total run time. The column was re-equilibrated with buffer A before the injection of sample. Peptides were eluted directly from the tip of the column and nanosprayed directly into the mass spectrometer by application of 2.5 kV voltage at the back of the column. The Q Exactive was operated in a data-dependent mode. Full MS^1 scans were collected in the Orbitrap at 70 K resolution with a mass range of 400–1,800 *m/z*. The ten most abundant ions per cycle were selected for tandem mass spectrometry and dynamic exclusion was used with exclusion duration of 15 s.

Three biological replicates were measured sequentially. Mass spectrometry data were processed with MaxQuant 1.5.7.0 (ref. 65) and Perseus 1.6.2.3 (ref. 66) as described previously[38] to identify ArgRS interaction partners. In brief, raw files were searched against *Homo sapiens* reference proteome UP000005640_9606 (ref. 67) (Uniprot) with the Andromeda search engine integrated in MaxQuant. Default settings were used for LFQ. Phosphorylated peptides (STY) were included when searching for nuclear interaction partners of ArgRS. The resulting protein groups file was loaded into Perseus and filtered for 'reverse', 'potential contaminants' and 'only identified by site'. The $\log_2$ values of LFQ intensities were calculated, and all proteins with less than two valid values/group were discarded. Missing values were replaced from a normal distribution (width 0.3 and downshift 1.8), and Welch's *t*-test was used to calculate *t*-test significance and difference.

Generation of SRRM2-mVenus cell line

A guide RNA (CCAUGAGACACCGCUCCUCC) targeted to the second last exon of SRRM2 (exon 14) was inserted into pSpCas9(BB)-2A-GFP. pSpCas9(BB)-2A-GFP (PX458) was a gift from Feng Zhang, Addgene plasmid no. 48138; http://n2t.net/addgene:48138; RRID:Addgene_48138 (ref. 68)). mVenus (mVenus C1) was a gift from Steven Vogel, Addgene plasmid no. 27794; http://n2t.net/addgene:27794; RRID:Addgene_27794 (ref. 69). The donor vector for insertion of mVenus was based on pcDNA6. The CMV promoter and multiple cloning site of pcDNA6 were substituted for mVenus flanked by 800 bp upstream and downstream of the Cas9 cleavage site. Cells were transfected with both plasmids at a 1:2 ratio gRNA/Cas9:donor vector and sorted for GFP expression after 2 days (successful expression of pSpCas9-GFP) by the Flow Cytometry Core at Scripps Research. Single cells were expanded into monoclonal cell lines and evaluated for mVenus expression after expansion by flow cytometry. About one-third of all tested clones showed mVenus fluorescence. Correct genomic integration was assessed by PCR on isolated genomic DNA (DNeasy Blood and Tissue Kit, Qiagen, Extended Data Fig. 5a). PCR primers: GTGGTGCCTGAGGTGGTGCC/CCACTCCCAAAATGGGGCCG.

Successful labelling of SRRM2 only was confirmed by the reduction of mVenus by two individual shRNAs directed against SRRM2 (Extended Data Fig. 6a).

Flow cytometry and fluorescence-activated cell sorting

Flow cytometry was performed on instruments maintained by the Flow Cytometry core at Scripps Research. The 293T cells were detached using trypsin/EDTA and resuspended in DMEM with 10% FBS. The cell suspension was mixed 1:1 with sorting buffer (2.5 mM EDTA, 25 mM HEPES, pH 7.0, 1% FBS and 1% penicillin–streptomycin in PBS) and analysed on an LSR II analytical flow cytometer (BD Bioscience). mVenus fluorescence was detected in the FITC channel (filter 525/50) after excitation by a 488 nm laser. Flow cytometry results were analysed with FlowJo version 10.06 or higher. Fluorescence-activated cell sorting was performed by members of the Scripps Research Flow Cytometry Core on a MoFlo Astrios EQ jet-in-air sorting flow cytometer (Beckman Coulter). Two days after transfection, cells were detached using Accutase (Stemcell Technologies), resuspended in DMEM with FBS and spun down to remove serum to prevent cells from adhering. Following a wash with PBS, cells were diluted in sorting buffer to 1×10^6 cells ml^{-1}. Single cells were sorted into a 96-well plate pre-filled with 100 µl medium and expanded to monoclonal cell lines. For quantification of MHCI expression, cells were detached with EDTA and stained with 1 µg ml^{-1} HB-95 W6/32 for 20 min on ice followed by Alexa 647-conjugated anti-mouse secondary antibody for 20 min on ice, both in sorting buffer.

ArgRS, MetRS and SRRM2 knockdown

shRNAs against ArgRS (shRARS_1: gtggacacaagcataagtaaa, shRARS_2: ggagcagttacaagaagaaaa, shRARS_3: ccacactctctgtgattatat, shRARS_4: ccttcactagaatcaggtcta) were cloned into a pLKO.1 vector (pLKO.1–TRC cloning vector was a gift from David Root, Addgene plasmid no. 10878; http://n2t.net/addgene:10878; RRID:Addgene_10878; ref. 70). A non-targeting control shRNA pLKO.1 shSCR was a gift from Sheila Stewart (Addgene plasmid no. 17920; http://n2t.net/addgene:17920; RRID:Addgene_17920; ref. 71). shRNAs against *SRRM2* and *MARS* were

acquired from The RNAi Consortium via Sigma-Aldrich (shSRRM2_1: cgccacctaaacagaaatc, shSRRM2_2: gttgggactggaggttgta, shMARS_1: caaggaaacattgtccgagaac, shMARS_2: tcgacatggcaaccaatatatc). Plasmids were co-transfected into 293T cells together with lentiviral packaging plasmids pRSV-rev, pMDLg/pRRE and pMD2.G (plasmids were a gift from Didier Trono (Addgene plasmid nos. 12253 and 12251; http://n2t.net/addgene:12253 and http://n2t.net/addgene:12251; RRID:Addgene_12253 and Addgene_12251; ref. 72) and a gift from the Torbett lab at Scripps Research) using Lipofectamine 2000 (Invitrogen). Medium was changed after overnight transfection and collected 2 days later. Supernatant containing viral particles was filtered through a 0.45 µm syringe filter, and either used directly or stored at −80 °C. For viral transduction, SRRM2-mVenus-tagged 293T or HepG2 cells were seeded a day prior on 6 cm tissue culture dishes and incubated with 1 ml undiluted viral particles in the presence of 8 µg polybrene (EMD Millipore) for 2 h. Afterwards, fresh medium was added, and the infection left to proceed for 2 days. Successfully transduced cells were selected with 10 µg ml^{-1} (293T) or 50 µg ml−1 (HepG2) puromycin (AdipoGen Life Science). ArgRS knockdown was confirmed via western blot and quantitative real-time PCR (qRT–PCR).

To rescue ArgRS expression with ArgRS active site or ΔNLS mutants, pcDNA6-based plasmids (ThermoFisher Scientific) were constructed that express ArgRS under the control of a CMV promoter. Site-specific mutations were introduced using PCR and clones were verified by Sanger sequencing. The expressed exogenous ArgRS contained a 14-amino acid V5-tag that can be detected with an antibody (ThermoFisher Scientific) to differentiate between endogenous and exogenous ArgRS. Cell lines with a stable knockdown of ArgRS (see above) were transfected with either Lipofectamine 2000, Lipofectamine 3000 (both ThermoFisher Scientific) or Calfectin (Signagen) according to the manufacturer's instructions. Medium was changed after overnight transfection in all cases, and cells were analysed after 24–48 h of recovery.

Cell viability

Cell viability and proliferation was assessed by measuring AlamarBlue (AbD Serotec) conversion. HepG2 or 293T cells were seeded in 96-well plates (10,000 cells per well) and 1:100 AlamarBlue was added 2 h before measurement on a Synergy H1 (Biotek, λ_{exc} = 530 nm, λ_{em} = 580 nm). Cell proliferation stagnates after confluency is reached. All measurements were done in technical quadruplicates.

Puromycin incorporation

Protein synthesis was monitored by puromycin incorporation as described previously[9,44]. Cells were seeded at 2×10^6 in 10 cm dishes 2 days before exposure to puromycin. Puromycin was directly added to the medium at 10 µg ml^{-1} for 30 min. Afterwards, cells were chased with medium for 1 h. Cells were then scraped into lysis buffer and analysed by western blot. Puromycin incorporation was quantified by densitometric analysis of western blots. Signal was quantified across the whole lane and normalized to the corresponding loading control.

Expression and purification of recombinant proteins

Variants of the ArgRS gene *RARS1* were subcloned from complementary DNA obtained from 293T HEK cells into a pET_SUMO vector (N-terminal His6-tag, N-terminal fusion protein with SUMO[73]) and expressed following induction with IPTG in RIPL *E. coli* (induction at OD_{600} 0.6–1.2, 0.5–1 mM IPTG, overnight at room temperature). Cells were lysed in buffer A (20 mM HEPES pH 7.5, 150 mM NaCl, 10 mM imidazole using a liquid fluidizer, centrifuged at 17,000*g* and loaded on 10 ml Ni-NTA resin (ThermoFisher Scientific or Qiagen)). Bound protein was washed with buffer A and eluted with 10 ml each of a stepwise gradient of buffer B (40 mM HEPES pH 8.0, 150 mM NaCl and 500 mM imidazole) at 10%, 20%, 40%, 60% and 100%. Fractions containing ArgRS were identified through SDS–PAGE, pooled and dialysed overnight against 20 mM HEPES, pH 7.5, 150 mM NaCl in the presence of SUMO protease (1:100) at 4 °C. Protein was run over Ni-NTA again to bind residual impurities and to separate ArgRS from SUMO protease and SUMO. ArgRS was then diluted to 50 mM NaCl, loaded on a 5 ml Q Sepharose column in an Akta FPLC system (both GE Healthcare), and eluted against a constant gradient to 500 mM NaCl to separate protein from RNA and for further purification. Finally, the fractions containing ArgRS were identified by OD_{260} as well as SDS–PAGE were pooled, concentrated and run over a Superdex 200 size exclusion column on a Akta FPLC system. Fractions containing ArgRS were concentrated and snap frozen in liquid nitrogen for storage at −80 °C. If protein was used for pulldown of interaction partners, it was biotinylated overnight using Biotin-dPEG 12-TFP ester (Quanta Bio) and 25 µl of 50 µM labelled protein was used as bait to pull out interaction partners from cell lysate.

Aminoacylation assay

Aminoacylation of tRNA with radiolabelled arginine was detected as described[74]. ArgRS was expressed and purified as described above. ArgRS was diluted in aminoacylation buffer (50 mM HEPES, pH 7.5, and 50 mM KCl) to 0.2 µM and then mixed with aminoacylation reaction mix (50 mM HEPES, pH 7.5, 50 mM KCl, 4 mM ATP, 2 mM DTT, 5 mM $MgCl_2$, 10 µM cold L-arginine, 0.004 mg ml^{-1} pyrophosphatase and ~1 µM radiolabelled arginine (arginine monohydrochloride L-[2,3,4-3H], PerkinElmer)) and incubated for 20 min at 37 °C in the presence of 1 mg ml^{-1} *E. coli* total tRNA. Timepoints were taken as indicated. Aminoacylation reactions were quenched in 300 mM NaOAc, pH 3.0, 100 mM EDTA and 0.5 mg ml^{-1} salmon sperm carrier DNA. tRNAs were precipitated in 20% ice-cold TCA. The filter plate was washed four times with 100 mM cold arginine in 5% TCA and once with ethanol. After drying, amino acids were hydrolysed with 0.1 M NaOH and eluted into Optiphase supermix scintillation cocktail (PerkinElmer). Counts per second (c.p.s.) were measured on a MicroBeta Trilux by averaging over 15 s.

Immunofluorescence

Glass cover slips (12 mm, 1.5 thickness, Neuvitro Corporation) were sterilized with 70% ethanol, coated with 0.01% Collagen I (MP Biomedicals Inc.) overnight, and stored in PBS until use. Two days before staining, 10,000 HepG2 or 293T cells were seeded on cover slips. Cells were washed, fixed with 4% PFA in PBS for 20 min, permeabilized in 0.25% Triton X-100 for 5 min, blocked with 1% FCS, 2% BSA in PBS (immunofluorescence (IF) buffer) for a minimum of 15 min and stained overnight with anti-SRRM2 antibody or anti-ArgRS antibody (both SCBT, 1:100, all dilutions in IF buffer) at 4 °C. Cells were washed twice with PBS and once with IF buffer between all steps. Anti-mouse AlexaFluor 488 secondary antibody (Abcam) was diluted 1:250 and added to the cells for 45 min at room temperature. Anti-methionine-tRNA synthetase (MetRS, Abcam) or ArgRS antibody (Biorbyt, made in-house) were added at a 1:100 dilution (ArgRS) or 1:200 dilution (MetRS, SON, SFPQ, and coilin) and incubated at room temperature for 3 h. Cross-adsorbed anti-rabbit Alexa Fluor 568 secondary antibody (Abcam) was added for 45 min at room temperature at a dilution of 1:250. Hoechst 33342 (AdipoGen Life Science) was added at 1:500 in PBS for 15 min. Cover slips were washed in PBS, dipped in water, mounted on soft-set Vectashield (Vectorlabs) and sealed with gel glue or nail polish. Slides were stored at 4 °C until imaging. Staining of SRRM2 with a monoclonal antibody raised against a C-terminal peptide of SRRM2 showed distinct nuclear structures, strongly reminiscent of an SRRM2 antibody staining deposited in the Human Protein Atlas with a different antibody (HPA066181). We speculate that the difference in staining patterns between the two SRRM2 antibodies used here is due to SRRM2 post-translational modifications.

Confocal microscopy

Microscopes were maintained by the Core Microscopy Facility at Scripps Research. For co-localization of SRRM2, ArgRS, SFPQ, coilin

and MetRS, a Zeiss LSM 880 Airyscan confocal laser scanning microscope with a 100× oil objective was used for imaging in Airyscan mode. Confocal images were recorded with a pixel dwell time of 1.5 µs in 16-bit depth with a scaling of 0.034 µm (x, y) and 0.145 µm (z). Z stacks covering the nucleus were taken with a motorized stage. Images were processed with 'Airyscan Processing' in Zen (Zeiss).

Co-localization

Estimation of co-localization was done with Imaris (Oxford Instruments). First, the nuclear volume of each cell was determined by using Hoechst DNA staining as a mask to delineate the cell nucleus, with an adequate threshold to filter spillover of Hoechst signal. Second, the nuclear signal of two antibody stains were co-localized by comparing nuclear volumes in individual cells and the percentage of co-localized volume was noted, as well as Manders overlap coefficients and Pearson correlation coefficient. Manders overlap coefficient calculates the overlap between two colours from their absolute intensities and results in individual coefficients for each colour (co-occurrence of colour1 with colour2 and vice versa). In contrast, Pearson correlation coefficient uses the deviation from the mean and results therefore in one coefficient for the correlation between both colours (relationship between the signal intensities). Background for the antibody staining was determined by comparison of signal intensity to a control in which cells were stained with secondary antibody only or by Costes thresholding in Imaris, if stated in the figure legend. Videos demonstrating co-localization (Supplementary Videos 1–4) were recorded with Imaris (Oxford Instruments). Extended Data Fig. 2k was taken on a Zeiss LSM 710 confocal laser scanning microscope with a 63× oil objective. Cells were fixed in 4% PFA/PBS for 10 min to retain mVenus fluorescence.

FRAP

SRRM2-mVenus 293T cells expressing either a non-targeting control shRNA (shSCR ctrl) or shRNA against ArgRS (shRARS) were seeded on collagen-I-coated (as described above for cover slips) glass-bottom 35 mm dishes with 1.5 thickness (World Precision Instrument) 2 days before imaging. Live cell imaging was performed on a Zeiss LSM 780 or a LSM 880 confocal laser scanning microscope at 37 °C, 5% CO_2 with a 63× water objective. Fluorescent bleaching was achieved by 20 cycles of 100% intensity at 514 nm. Pictures were taken every 2 s for 75 cycles with a pixel dwell time of 1.58 µs, and fluorescence in ROIs was recorded with ZEN (Zeiss). The resulting data were subtracted from background, divided by fluorescence intensity of the same area before bleaching (normalization) and the geometric mean of ROIs was fitted with a one-exponential model to calculate the half time of fluorescence recovery (GraphPad Prism 5). The mean of individual measurements was calculated and plotted together with standard error means. Pictures were exported with Fiji/ImageJ. Immobile fractions were calculated from dedicated measurements with images taken every 10 s (as opposed to every 2 s for the calculation of half time of recovery) to minimize photobleaching during imaging (at the cost of temporal resolution). FRAP recordings were excluded if sample drift was noted during the recording. Significance was tested with a two-tailed Student's t-test (GraphPad Prism 5). We compared SRRM2-mVenus fluorescence recovery half times with known literature values of SRSF1 (ref. 75) recovery. Secondary structure prediction of both proteins was carried out with Sable[76].

Reverse transcription and qRT–PCR

RNA was isolated using TRIzol (Invitrogen) according to the manufacturer's instructions. RNA (500 ng) was reverse transcribed using MultiScribe Reverse Transcriptase (Applied Biosystems) according to instructions in the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) with random hexamers (Qiagen) as primers. qRT–PCR was performed with Power SYBR Green Master Mix (ThermoFisher Scientific) on a StepOnePlus (Applied Biosystems). The following primers were used:

RARS: GAGAGTATAAGCCGCCTCTTTG/TCCTCCTCAGTATCAAACCTCT;
SRRM2: TGACAGCAAATCTCGACTATCC/GGTTCAGGAGAGGAATCAGAAC
18S: CTGAGAAACGGCTACCACATC/GCCTCGAAAGAGTCCTGTATTG

RNA isolation and TGIRT-seq library preparation

HepG2 cells (2.5×10^6) expressing either a non-targeting shRNA control (shSCR) or a shRNA directed against ArgRS (gene name *RARS*, shRARS) were seeded on 10 cm tissue culture dishes 2 days before RNA isolation. Cells were starved off arginine for 6 h and supplemented with 10× arginine 2 h before collection. RNA was isolated by using a mirVana miRNA isolation kit (Invitrogen) with phenol extraction following the manufacturer's instruction for total RNA isolation. RNA integrity was evaluated with a Bioanalyzer (Agilent). Extracted total RNA was treated with TURBO DNase (ThermoFisher) according to the manufacturer's instruction, and the DNase-treated RNA was then ribo-depleted by using a Ribo-Zero Gold (Human/Mouse/Rat) kit (Illumina). For each library, 50 ng of ribo-depleted RNA was fragmented by using an RNA Fragmentation Module kit (New England Biolabs) to a median size of 55–75 nt, and 3′ phosphates were removed by treatment with T4 polynucleotide kinase (Epicentre).

TGIRT-seq library preparation

TGIRT-seq libraries were prepared essentially as described[77] using a modified R2R adaptor that decreases adaptor dimer formation[78]. Reverse transcription reactions contained fragmented, dephosphorylated cellular RNA, 100 nM R2 RNA/R2R DNA starter duplex, and 1 µM TGIRT-III (InGex) in reaction buffer (20 mM Tris-HCl, pH 7.5, 450 mM NaCl, 5 mM $MgCl_2$ and 5 mM DTT). Reactions were pre-incubated at room temperature for 30 min and then initiated by addition of 1 mM dNTPs (an equimolar mix of 1 mM dATP, dCTP, dGTP and dTTP) and raising the temperature to 60 °C. Reverse transcription by TGIRT-III is initiated by template switching from a starting duplex consisting of a 35 nt DNA primer encoding the reverse complement of the Illumina Read 2 sequencing primer binding site (R2R) annealed to a 34 nt complementary RNA oligonucleotide (R2), leaving a single nucleotide 3′ DNA overhang composed of an equimolar mixture of A, G, C and T. The RNA oligonucleotide is blocked at its 3′ end with C3Sp (Integrated DNA Technologies) to inhibit template switching to itself. Reactions were incubated at 60 °C for 15 min and terminated by adding 5 N NaOH to a final concentration of 0.25 N and incubating at 95 °C for 3 min to degrade RNAs and denature protein. The reactions were then cooled to room temperature and neutralized with 5 N HCl. cDNAs were purified by using a Qiagen MinElute Reaction Cleanup Kit and then ligated at their 3′ ends to a DNA oligonucleotide encoding the reverse complement of the Illumina Read 1 primer binding site (R1R) using Thermostable 5′ AppDNA/RNA Ligase (New England Biolabs). Ligated cDNAs were re-purified with a MinElute Reaction Cleanup Kit and amplified by PCR for 12 cycles using Phusion DNA polymerase (ThermoFisher Scientific) with overlapping multiplex and index primers that add sequences necessary for Illumina sequencing. PCR products were purified with AMPure XP beads (Beckman Coulter) to remove unused PCR primers and adaptor dimers. Libraries were sequenced on a NextSeq 500 instrument (2 × 75 nt, paired-end reads) at the Genomic Sequencing and Analysis Facility at the University of Texas at Austin.

TGIRT-seq bioinformatic analysis

After sequencing, Fastq files were processed by using the TGIRT-map pipeline[79]. Briefly, the sequencing adaptors were removed with Atropos[80] using options trim -U 1–minimum-length = 15–threads = 24–no-cache-adapters–error-rate = 0.1 –b AAGATCGGAAGAGCACA

CGTCTGAACTCCAGTCAC -B GATCGTCGGACTGTAGAACTCTGAACGTG TAGA. Trimmed reads were then aligned against 5S rRNA (GenBank accession X12811.1), complete rRNA repeat unit (GenBank accession U13369.1) and hg19 Genomic tRNA database[81] using BOWTIE2 (ref. 82). The remaining unaligned reads were then sequentially aligned to the human genome (hg19) using a splice-aware aligner (HISAT2) (ref. 83) and a sensitive local aligner (BOWTIE2) (ref. 82). Finally, gene quantification was done on genomic loci, except for tRNA and rRNA, which required an additional step of re-aligning and counting (Supplementary Table 14).

Differential gene expression analysis was done using DESeq2 (ref. 84). Exon quantification was done by FeatureCount[85] and differential exon usage was quantified using DEXSeq[49]. DEXSeq uses a negative binomial model to test for differential exon usage between two samples while normalizing for the differences in gene expression[49]. Gene names were assigned using biomaRt[86].

Poly(A) RNA-seq library preparation and bioinformatic analysis

Poly(A) RNA-seq libraries were prepared and sequenced by the Genomic Sequencing and Analysis Facility at the University of Texas at Austin. RNAs were selected by using a Poly(A) Purist MAG Kit (ThermoFisher Scientific, AM1922) using the manufacturer's protocol followed by library preparation using a NEBNext Ultra II Directional RNA Library Kit for Illumina (New England Biolabs, E7760), also following the manufacturer's protocol. Samples were sequenced on a NovaSeq 6000 instrument using SP flowcells (2 × 150 nt, paired-end reads). Reads were trimmed with trimgalore and aligned to grch37 with STAR[87] (Supplementary Table 15). Counts were generated with Subread[85].

Differential exon usage analysis was performed with DEXSeq[49]. If exons were assigned to overlapping genes, gene symbols were assigned to the first gene and counted as one gene except in the direct comparison between DEXSeq results and N-terminal peptides, where all genes were used. Differentially expressed genes were identified with DESeq2 (ref. 84). Enrichment analysis was performed using Panther[88] and ShinyGO[89]. For Panther, all genes mapped to grch37 were used as a reference list. Vast -tools was used to identify differentially spliced events, including cassette exons, microexons, 3′ and 5′ alternative splice sites and retained splice sites[90]. Untrimmed reads were first aligned to the hg19 assembly and PSI values for all annotated splicing events were measured for each sample. Events were considered if, in at least of half of the profiled samples, they had ≥15 reads overlapping the included/excluded splice junctions and a ratio of reads mapping to the upstream and downstream junctions (a balance score) of <2 or 2–5. These events were than processed through the vast-tools diff module[48], which calculates the predicted ΔPSI between samples and assigns a minimum ΔPSI value (MV), with a ≥0.95 probability for each event. For example, MV >0 |ΔPSI| ≥0.1 indicates ≥0.95 probability of at least a 10% difference in splice junction usage in either direction.

Detection of N-terminal peptides

A modified protocol for the terminal amine isotopic labelling of substrate was used to enrich for N-terminal peptides[51]. Cells were lysed, proteins were extracted by a methanol/chloroform precipitation and N-termini were blocked overnight with formaldehyde. Proteins were extracted again with methanol/chloroform to avoid carry-over of reactive formaldehyde and digested with trypsin overnight. HPG-ALD polymers were acquired through Flintbox and added to the peptides at a 1:50 ratio to deplete peptides with unblocked reactive N-termini derived from the trypsin digest. The next day, the reaction was quenched, and free peptides separated from the polymer with a 10 kDa spin filter. Peptides were fractionated with Pierce High pH Reversed-Phase Peptide Fractionation kit (ThermoFisher Scientific) according to the manufacturer's instructions, all nine fractions were dried down, resuspended in 5% acetonitrile and 0.1% formic acid, and were analysed on the Q Exactive mass spectrometer as described above. Raw data were processed with MaxQuant[65] with dimethyllysine and dimethylated N-termini as a fixed modification. Peptides were searched against a reference proteome (Uniprot) with free N-termini and ArgC as protease and label-free quantification was used with default settings. The resulting intensities and LFQ values were filtered in Perseus[66] against a contaminant database. Reverse and peptides preceded by arginine, which are probably derived from trypsin cleavage, were excluded. MaxQuant (version 1.6.7.0) was run on the Scripps High Performance Computing core. Genes with overlapping differential exon usage (DEXSeq padj <0.05) and N-terminal peptides (Student's *t*-test difference >|Δ1|) were further manually analysed to identify the position of the differential exon in the gene and the resulting peptide sequence.

RT–PCR for verification of splicing events

RNA was reverse transcribed and amplified using OneStep RT–PCR (Qiagen) according to the manufacturer's instructions but in a reduced volume (10 µl final). HepG2 RNA (40 ng) or 293T RNA (100 ng) were transcribed per reaction. Primers were designed using VastDB[47] for splice junction changes and manually for differential exon usage. Amplified regions of different splice isoforms were separated on a 2.5% agarose/TAE gel and visualized after staining with SybrSafe on a Bio-Rad ChemiDoc Imager.

IDH and proteasome activity

HepG2 cells (10,000) were seeded in white 96-well plates 2 days before both assays. For IDH activity, cells were lysed in co-immunoprecipitation lysis buffer and mixed with an end concentration of 100 µM NADP, 2 mM $MnCl_2$ and 5 mM isocitrate. IDH activity was measured by monitoring fluorescence at 460 nm after excitation at 340 nm, once per minute for 10 min at 37 °C. Proteasomal subunit β5 activity was measured using the ProteasomeGlo assay (Promega) according to the manufacturer's instructions in the same format. Total protein concentration (IDH, with BCA assay) or cell count (β5 activity, with Alamar Blue assay) were used to normalize between treatment conditions.

MHCI peptidome

Cells (5×10^8) were seeded per biological replicate 2 days before collection and frozen at −80 °C until MHCI peptide enrichment. HB-95 W6/32 hybridoma were acquired by ATCC, cultured in DMEM with 10% FBS and diluted in HyClone CDM4MAb serum-free medium stepwise till 10% DMEM was reached. Cells were left to condition the serum-reduced medium for 5 days, then 300 ml medium was loaded on 3 ml of Protein A, washed with PBS and eluted with 100 mM glycine, pH 3.0, into a final concentration of 130 mM HEPES, pH 7.5 in five column volumes. The coupling of antibody to protein A was performed as described previously[91]. Cell lysis followed the instructions by Purcell et al., but centrifugation speed was changed to 40,000*g* for 30 min. Separation of MHCI from MHCI-bound peptides was performed as described previously[55]. Peptides were measured on a QE-HFX and searched against the human proteome with ProLuCID[92] without protease specificity and no requirements for tryptic peptides for protein identification. Peptides with 8–14 amino acids were analysed with NetMHCpan4.1 (ref. 93) for the alleles found in HepG2 (HLA-A A*02:01, 24:02; HLA-B B*35:14, 51:01; HLA-C C*04:01, 16:02). The MHCI immunogenicity score was calculated as described previously[56].

Statistics and reproducibility

If not mentioned otherwise, a Student's two-tailed *t*-test was used to determine significance between groups, except for mass spectrometry data for which Welch's *t*-test was used. Individual values are shown with lines indicating the median if not mentioned otherwise.

Repeats are different cell passages if not noted otherwise. For densitometric analysis, paired *t*-tests were used as repeats were run on individual gels. GraphPad Prism, versions 5–9 were used for testing, except for large datasets. Significance in large datasets was determined with Perseus for mass spectrometry results, DESeq2 for TGIRT-seq and RNA-seq, DEXSeq for exon analysis and Vast-Tools for splice junction usage. Correlation coefficients were calculated with Kendall Rank correlation or Pearson correlation and enrichment of gene sets by hypergeometric testing, both in R. If not stated otherwise, significance was defined as $*P < 0.05$, $**P < 0.01$ and $***P < 0.001$. No statistical method was used to pre-determine sample size if not mentioned otherwise. The experiments were not randomized. Data collection and analysis were not performed blind to the conditions of the experiments. No relevant data were excluded from the analyses unless obvious technical difficulties arose. Data distribution was assumed to be normal, but this was not formally tested.

Materials availability

All unique reagents generated in this study are available from Prof. Paul Schimmel (schimmel@scripps.edu) or Haissi Cui (haissi.cui@utoronto.ca), with a completed Materials Transfer Agreement.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

RNA-seq and TGIRT-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE165513 and on NIH Bioproject with accession number PRJNA561913. Mass spectrometry data have been deposited in ProteomeXchange with the primary accession codes PXD015692, PXD024091 and PXD036232. All other data supporting the findings of this study are available from the corresponding author on reasonable request. Source data are provided with this paper.

Code availability

Custom code for the analysis of TGIRT-seq data can be found at https://github.com/wckdouglas/tgirt_map.

References

64. Preger, C. et al. Generation and validation of recombinant antibodies to study human aminoacyl-tRNA synthetases. *J. Biol. Chem.* **295**, 13981–13993 (2020).
65. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
66. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
67. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
68. Ran, F. A. et al. Genome engineering using the CRISPR–Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
69. Koushik, S. V., Chen, H., Thaler, C., Puhl, H. L. & Vogel, S. S. Cerulean, Venus, and VenusY67C FRET reference standards. *Biophys. J.* **91**, L99–L101 (2006).
70. Moffat, J. et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283–1298 (2006).
71. Saharia, A. et al. Flap endonuclease 1 contributes to telomere stability. *Curr. Biol.* **18**, 496–500 (2008).
72. Dull, T. et al. A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* **72**, 8463–8471 (1998).
73. Kanaji, T. et al. Tyrosyl-tRNA synthetase stimulates thrombopoietin-independent hematopoiesis accelerating recovery from thrombocytopenia. *Proc. Natl Acad. Sci. USA* **115**, E8228–E8235 (2018).
74. Beebe, K., Waas, W., Druzina, Z., Guo, M. & Schimmel, P. A universal plate format for increased throughput of assays that monitor multiple aminoacyl transfer RNA synthetase activities. *Anal. Biochem.* **368**, 111–121 (2007).
75. Phair, R. D. & Misteli, T. High mobility of proteins in the mammalian cell nucleus. *Nature* **404**, 604–609 (2000).
76. Adamczak, R., Porollo, A. & Meller, J. Accurate prediction of solvent accessibility using neural networks-based regression. *Proteins* **56**, 753–767 (2004).
77. Qin, Y. et al. High-throughput sequencing of human plasma RNA by using thermostable group II intron reverse transcriptases. *RNA* **22**, 111–128 (2016).
78. Xu, H., Yao, J., Wu, D. C. & Lambowitz, A. M. Improved TGIRT-seq methods for comprehensive transcriptome profiling with decreased adapter dimer formation and bias correction. *Sci. Rep.* **9**, 7953 (2019).
79. Wu, D. C., Yao, J., Ho, K. S., Lambowitz, A. M. & Wilke, C. O. Limitations of alignment-free tools in total RNA-seq quantification. *BMC Genomics* **19**, 510 (2018).
80. Didion, J. P., Martin, M. & Collins, F. S. Atropos: specific, sensitive, and speedy trimming of sequencing reads. *PeerJ* **5**, e3720 (2017).
81. Chan, P. P. & Lowe, T. M. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* **44**, D184–D189 (2016).
82. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
83. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
84. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
85. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
86. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
87. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
88. Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P. D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **47**, D419–D426 (2019).
89. Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).
90. Irimia, M. et al. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **159**, 1511–1523 (2014).
91. Purcell, A. W., Ramarathinam, S. H. & Ternette, N. Mass spectrometry-based identification of MHC-bound peptides for immunopeptidomics. *Nat. Protoc.* **14**, 1687–1707 (2019).
92. Xu, T. et al. ProLuCID: an improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J. Proteom.* **129**, 16–24 (2015).
93. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, W449–W454 (2020).

Acknowledgements
We thank the Center for Metabolomics, the Microscopy core, the High Performance Computing team and Flow Cytometry Core, all at Scripps Research, for training and instrument maintenance, and the staff at the Genomic Sequencing and Analysis Facility at the University of Texas at Austin for poly(A) library preparation and sequencing. Access to a plate reader and flow cytometry software was provided by J. Paulson and access to a ChemiDoc Imager by J. Kelly, both at Scripps Research. We thank A. Dutta for preliminary analysis of ArgRS- and MetRS-dependent signalling pathways. We also thank U. Braunschweig, R. Barutcu and M. Wu (all University of Toronto) for feedback on the experimental design (U.B.) and earlier versions of the manuscript (R.B. and M.W.). This work was supported by the National Foundation for Cancer Research to P.S., the DFG (327097878) and the HFSP (LT000207) to H.C., NIH grant P41 GM103533 to J.R.Y., NIH grant R35 GM136216 and Welch Foundation grant F-1607 to A.M.L. and a CIHR Foundation Grant to B.J.B. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions
H.C. planned and conducted experiments and analysed data. J.K.D., J.J.M. and J.R.Y. performed, analysed or supervised mass spectrometry experiments. D.C.W., R.M.N. and A.M.L. performed, analysed or supervised TGIRT-seq and RNA-seq library preparation and sequencing. J.J.L. and B.B. analysed RNA-seq results. P.S., A.M.L. and H.C. designed the research and wrote the manuscript with input and contributions from all authors.

Competing interests
Thermostable Group II Intron Reverse Transcriptase (TGIRT) enzymes and methods for their use are the subject of patents and patent applications that have been licensed by the University of Texas to InGex, LLC. A.M.L. and the University of Texas are minority equity holders in InGex, and A.M.L., some members of A.M.L.'s laboratory, and the University of Texas receive royalty payments from the sale of TGIRT enzymes. The remaining authors declare no competing interests.

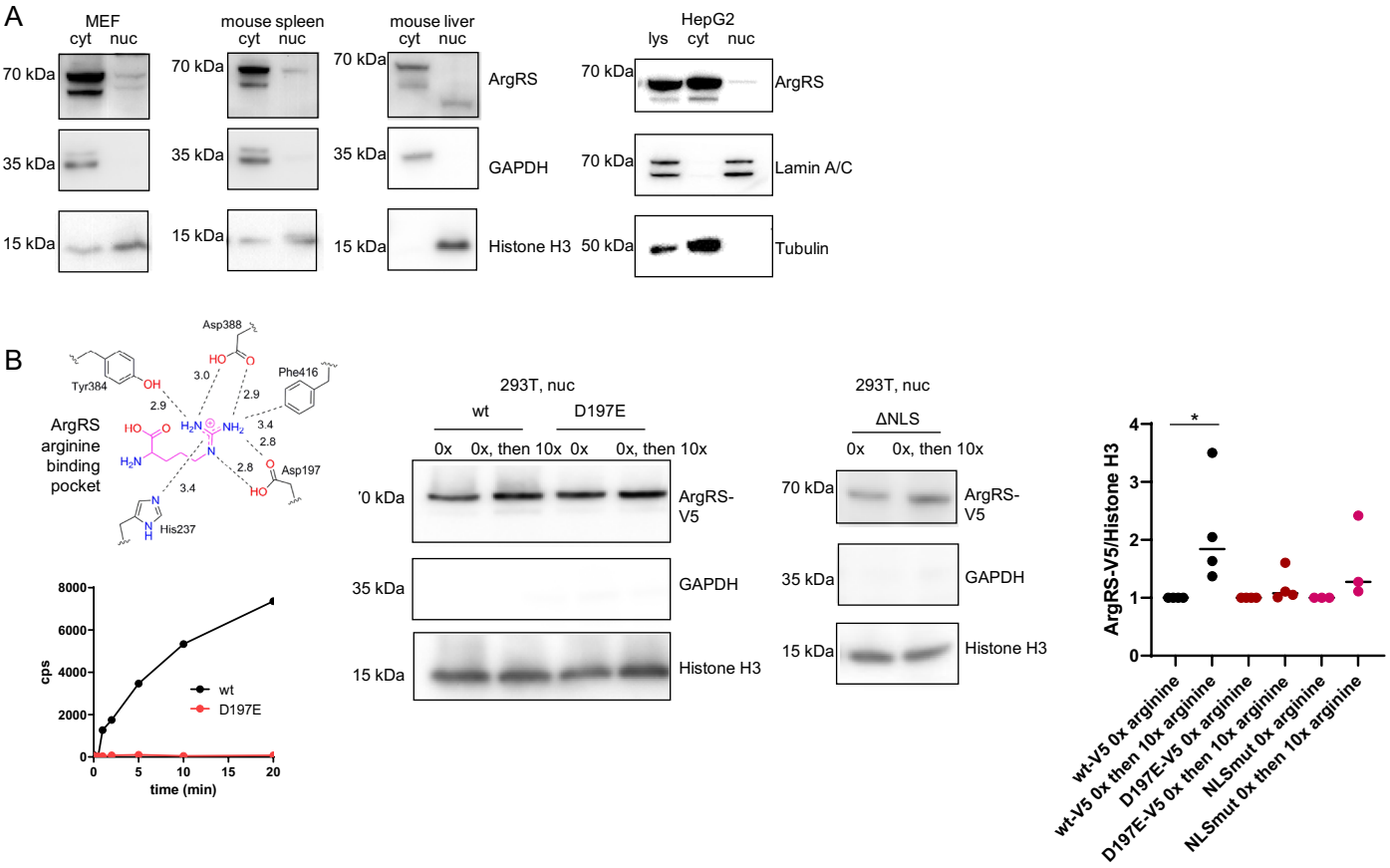
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41556-023-01118-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41556-023-01118-8.

Correspondence and requests for materials should be addressed to Alan M. Lambowitz or Paul Schimmel.

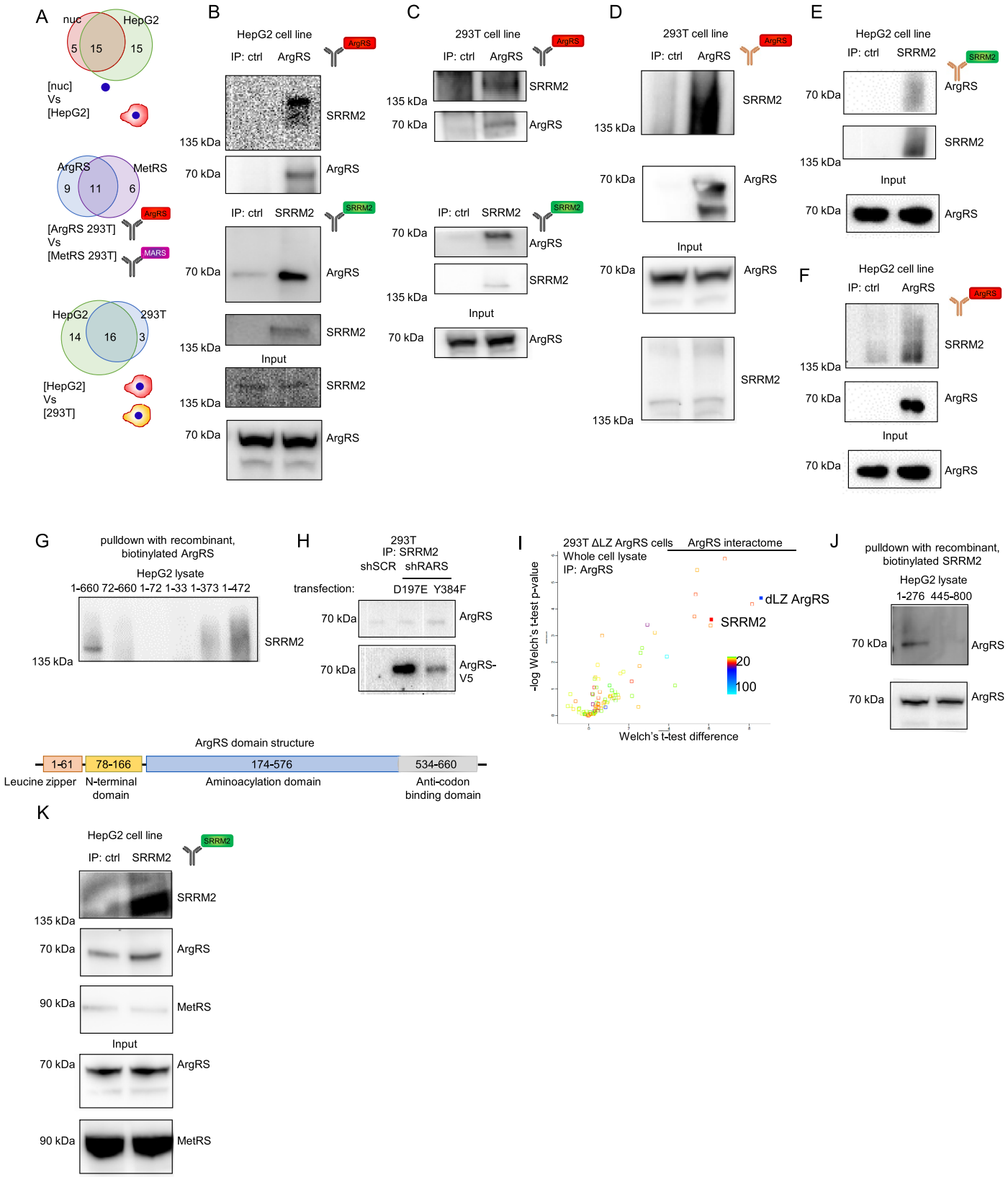
Peer review information *Nature Cell Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

https://doi.org/10.1038/s41556-023-01118-8



Extended Data Fig. 1 | ArgRS localized to the nucleus in murine cells and tissues and verification of decreased plasma arginine levels. (**a**) ArgRS in different cellular compartments of murine embryonic fibroblasts (MEF, left), mouse spleen (middle), and mouse liver (right) probed by western blot following cell fractionation. Far right: Cell fractionation of HepG2 cells shown with alternative markers for the cytoplasmic and nuclear fraction. GAPDH: cytoplasmic loading control, Histone H3: nuclear loading control. Lamin A/C: nuclear loading control, Tubulin: cytoplasmic loading control. lys: lysate, cyt: cytoplasm, nuc: nucleus. Experiments were repeated at least 3 times. (**b**) Active site and nuclear localization signal mutations in ArgRS do not abrogate nuclear import. Left panel: Schematic of the ArgRS active site with amino acids coordinating arginine highlighted and aminoacylation assays with purified recombinant wt and D197E mutant ArgRS proteins. Middle left panel: A mutation in the ArgRS active site reduced the increase of ArgRS in response to arginine but did not abrogate it completely. Western blot of the nuclear fraction of wt-V5 and D137E mutated ArgRS-V5. A C-terminal 14-amino acid V5-tag that can be recognized with an anti-V5 antibody was used to differentiate between endogenous and mutant ArgRS. Representative image of 4 replicates. Middle right panel: Mutations in the ArgRS nuclear localization sequence reduced but did not abrogate the increase of nuclear ArgRS in response to arginine. ΔNLS: K287N, K288N, K309N. Representative image of 3 replicates. Right panel: Densitometric quantification of ArgRS relative to the nuclear loading control Histone H3 and normalized to 0x arginine. n = 4, 4, 4, 4, 3, 3 biologically independent transfections, p = *0.048. One-way Anova, Tukey's multiple comparisons test. (**c**) LC/MS measurement of plasma arginine after (left panel) treatment with recombinant Arginase-1, an arginine-degrading enzyme, or (right panel) by induction of systemic inflammation with the liver toxin CCl4. n = 4, 3, 4, 4 independent mice. p = **0.002, p = ***0.0009. Unpaired, two-tailed t-test.

https://doi.org/10.1038/s41556-023-01118-8

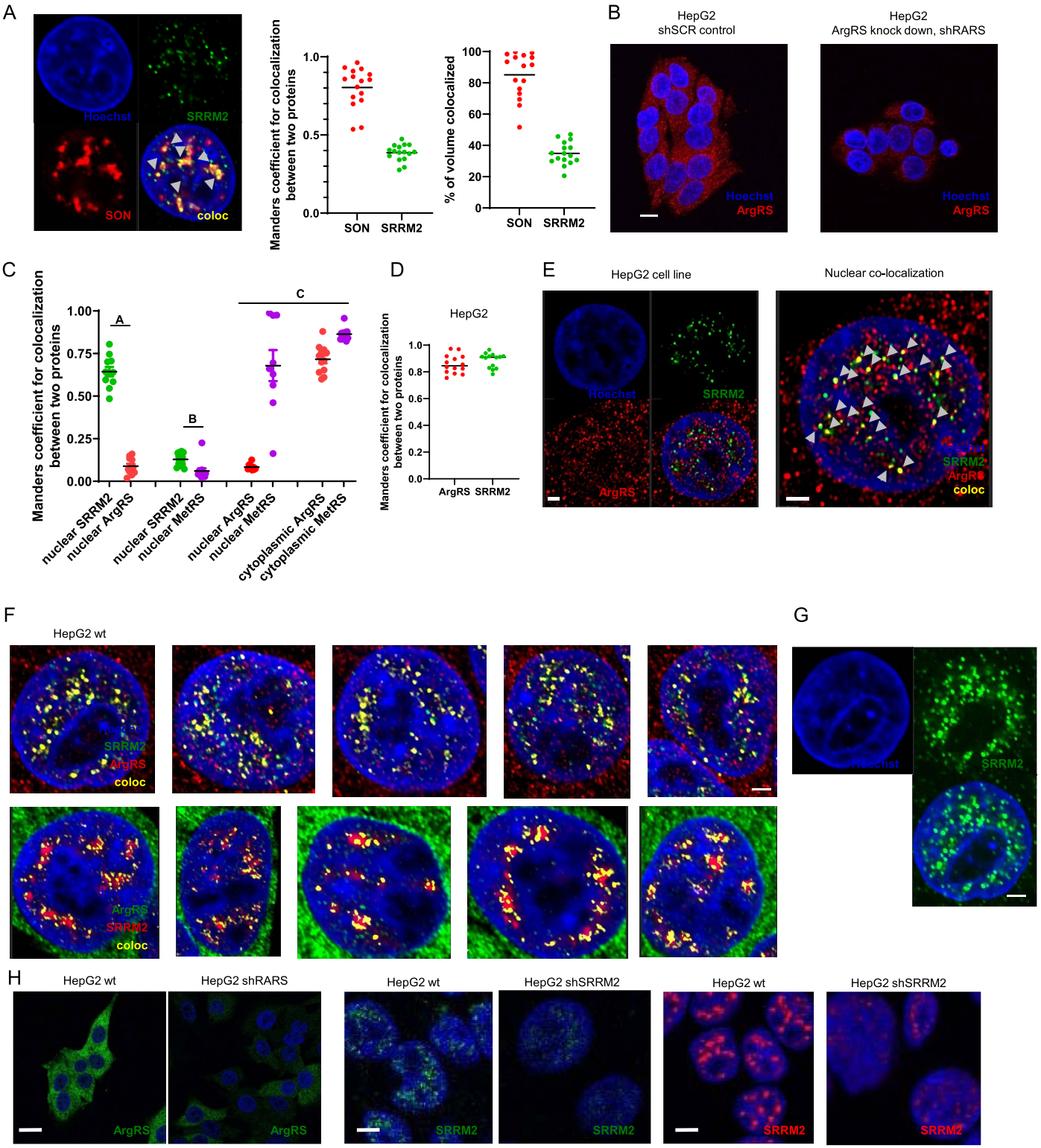


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41556-023-01118-8

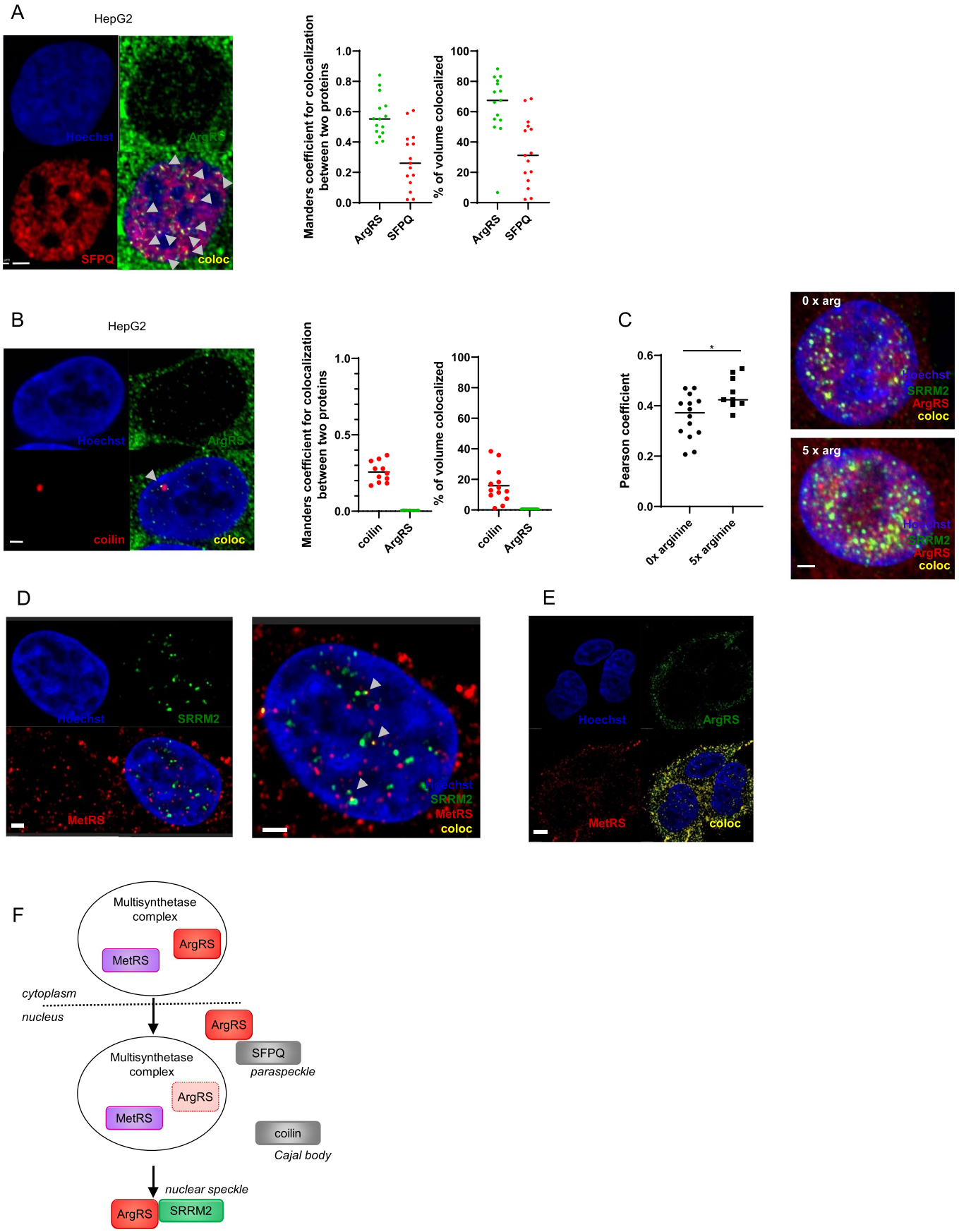
Extended Data Fig. 2 | Verification of the interaction between ArgRS and SRRM2 and domain mapping. (**a**) Venn diagrams comparing 3 different ArgRS and the MetRS interactomes. Statistical significance of overrepresentation by hypergeometric test: [nuc] vs [HepG2]: p = 7.6e-10, [HepG2] vs [293T]: p = 1.9e-10, [ArgRS 293T] vs [MetRS 293T]: p = 2.0e-7. [nuc]: HepG2 cells, nuclear fraction, IP: ArgRS, [HepG2]: HepG2 cells, whole cell lysate, IP: ArgRS, [293T]: 293T cells, whole cell lysate, IP: ArgRS, [MetRS]: 293T cells, whole cell lysate, IP: MetRS. Significance was calculated based on triplicate experiments (different cell passages). (**b-f**) Verification of ArgRS and SRRM2 interaction by immunoprecipitation (IP). Different antibodies are indicated by different colors of the antibody symbol. Input is indicated in the Figure. (**g-j**) Mapping of the interaction site between ArgRS and SRRM2. (**g**) Domain mapping of ArgRS to identify which domains are necessary for the interaction with SRRM2 using purified, biotinylated ArgRS fragments. Purified proteins were added to cell lysate and retrieved with streptavidin beads. N- and C-terminally truncated variants were used to retrieve SRRM2. Schematic of ArgRS domain architecture. (**h**) Immunoprecipitation of SRRM2 retrieved catalytically inactive ArgRS mutants as shown by western blot. A C-terminal V5-tag was used to differentiate between endogenous and mutant ArgRS. D197E and Y384F are mutations in the arginine binding site of ArgRS. See Extended Data Figs. 1b and 7e for confirmation of the loss of activity. (**b-h, j, k**) Experiments were repeated at least 3 times. (**i**) Volcano plot of the interactome of ArgRS with a deleted leucine zipper (remaining amino acids: 72–660). Axes denote enrichment (Welch's t-test enrichment, x-axis) and reproducibility (Welch's t-test significance, y-axis). Color: number of unique peptides detected (see scale at bottom right). Significance was calculated based on triplicate experiments (different cell passages). (**j**) Domain mapping of SRRM2 to identify which domains are necessary for interaction with ArgRS using purified, biotinylated fragments. The N-terminal coiled-coil domain (1–276) and an arginine-rich stretch of SRRM2 (445–800) were used to retrieve ArgRS. (**k**) SRRM2 immunoprecipitation did not enrich for MetRS but ArgRS. Detection of ArgRS, MetRS, and SRRM2 by western blot. (**b-g, k**) SRRM2 western blotting led to a smear-like appearance due to heavy post-translational modifications and inefficient transfer due to the high molecular weight of SRRM2.

https://doi.org/10.1038/s41556-023-01118-8



Extended Data Fig. 3 | Fluorescent microscopy of ArgRS and SRRM2. (**a**) Colocalization (yellow) of SRRM2 (green) and nuclear speckle protein SON (red) by immunofluorescent staining and imaging by confocal microscopy. 84% of SON colocalized with SRRM2 and 34% of SRRM2 colocalized with SON. n = 16, 16. Bar: 2 µm. (**b**) Maximum intensity projections of ArgRS immunofluorescent staining in HepG2 shSCR control and shRARS_1 cells to validate antibody specificity. Comparable experiments were repeated at least 3 times. Bar: 10 µm. (**c**) Manders coefficients of SRRM2, ArgRS, and MetRS. Hoechst staining was used to identify the cell nucleus. Manders coefficients were calculated for individual cell nuclei (n = 11, 12) or for an average for 2–3 cells (C, n = 8). (**d**) Manders coefficient of ArgRS and SRRM2 colocalization when Costes thresholding was used. n = 14, 14. (**e**) Colocalization (yellow) of ArgRS (green) and SRRM2 (red) with different antibodies than used in Fig. 3a. Bar: 2 µm. (**f**) Immunofluorescent staining of ArgRS and SRRM2 in representative nuclei. Bar: 2 µm. (**g**) SRRM2 staining without signal intensity threshold shows diffuse, connected areas where punctae accumulate. SRRM2 signal is stronger in DNA-low areas. (**h**) Verification of antibodies by ArgRS (shRARS) and SRRM2 (shSRRM2) knock down. Bar: wt and shRARS, 20 µm. wt and shSRRM2: 5 µm. Antibody validations for immunofluorescence were performed once.

https://doi.org/10.1038/s41556-023-01118-8

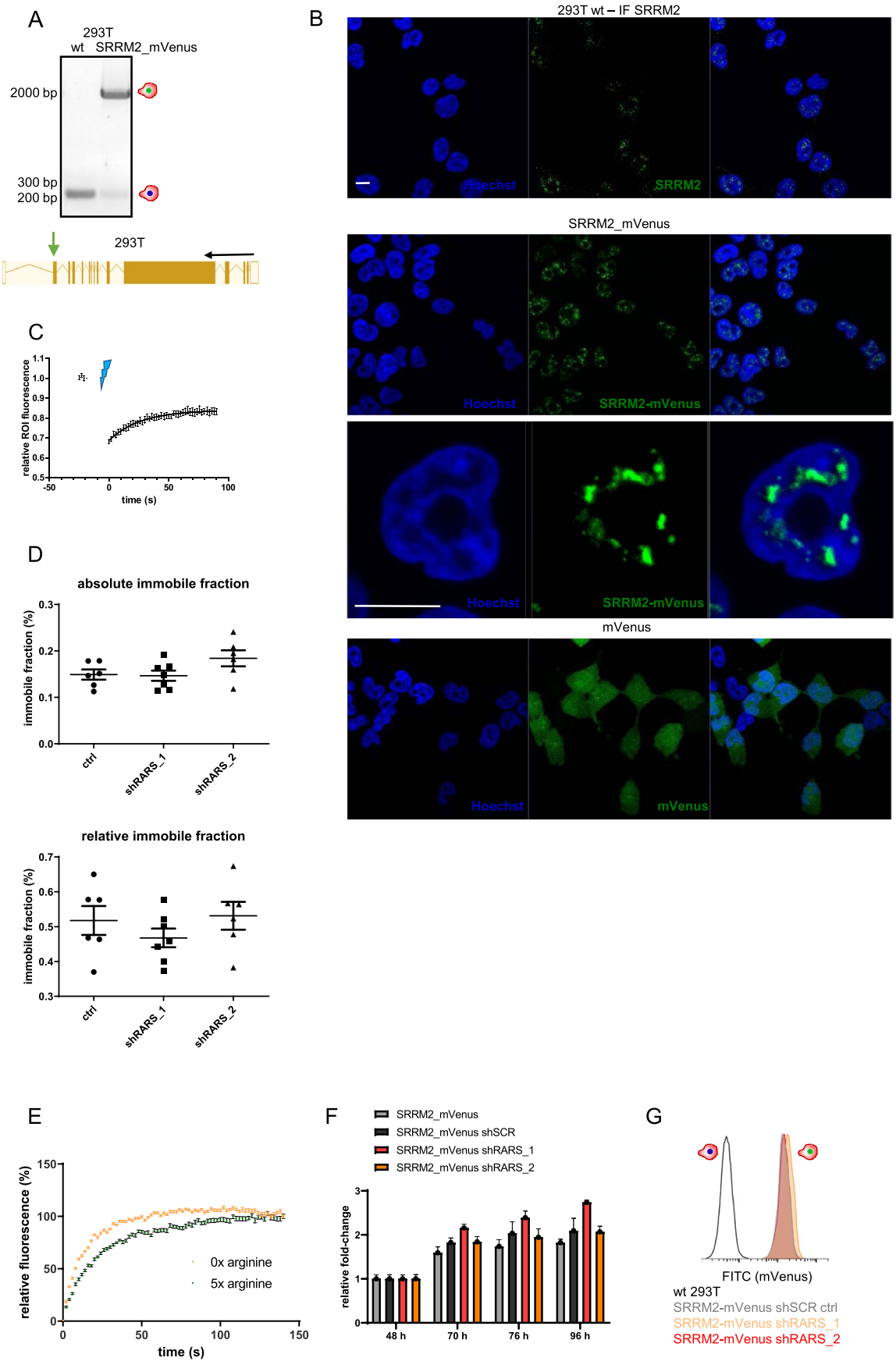


Extended Data Fig. 4 | See next page for caption.

https://doi.org/10.1038/s41556-023-01118-8

Extended Data Fig. 4 | Fluorescent microscopy to determine the nuclear ArgRS localization. (**a**) Colocalization of ArgRS (green) and paraspeckle protein SFPQ (red). Manders coefficient of colocalization after Costes thresholding. n = 15, 15. Bar: 2 µm. (**b**) Colocalization of ArgRS (green) and Cajal body protein Coilin (red). Manders coefficient of colocalization. n = 11, 11. Bar: 2 µm. (**c**) Pearson correlation coefficient of SRRM2 and ArgRS calculated from immunofluorescence staining during arginine starvation (0x arginine) and high arginine (5x arginine). Representative immunofluorescent staining of ArgRS and SRRM2. p = *0.01, n = 14, 9 independent cells. Unpaired, two-tailed t-test. Arginine concentrations are relative to DMEM. Bar: 2 µm. (**d**) Confocal microscopy following immunofluorescent staining of MetRS/SRRM2. A single slice of a series of z-stacks is shown. Colocalized areas in yellow. Bar: 2 µm. (**e**) Localization of MetRS and ArgRS in whole cells and in the nucleus. Bar: 5 µm. (**d, e**) Experiments were repeated at least 3 times. Quantification in Extended Data Fig. 3c. (**f**) Scheme of nuclear ArgRS localization relative to other MSC components and nuclear proteins.

https://doi.org/10.1038/s41556-023-01118-8

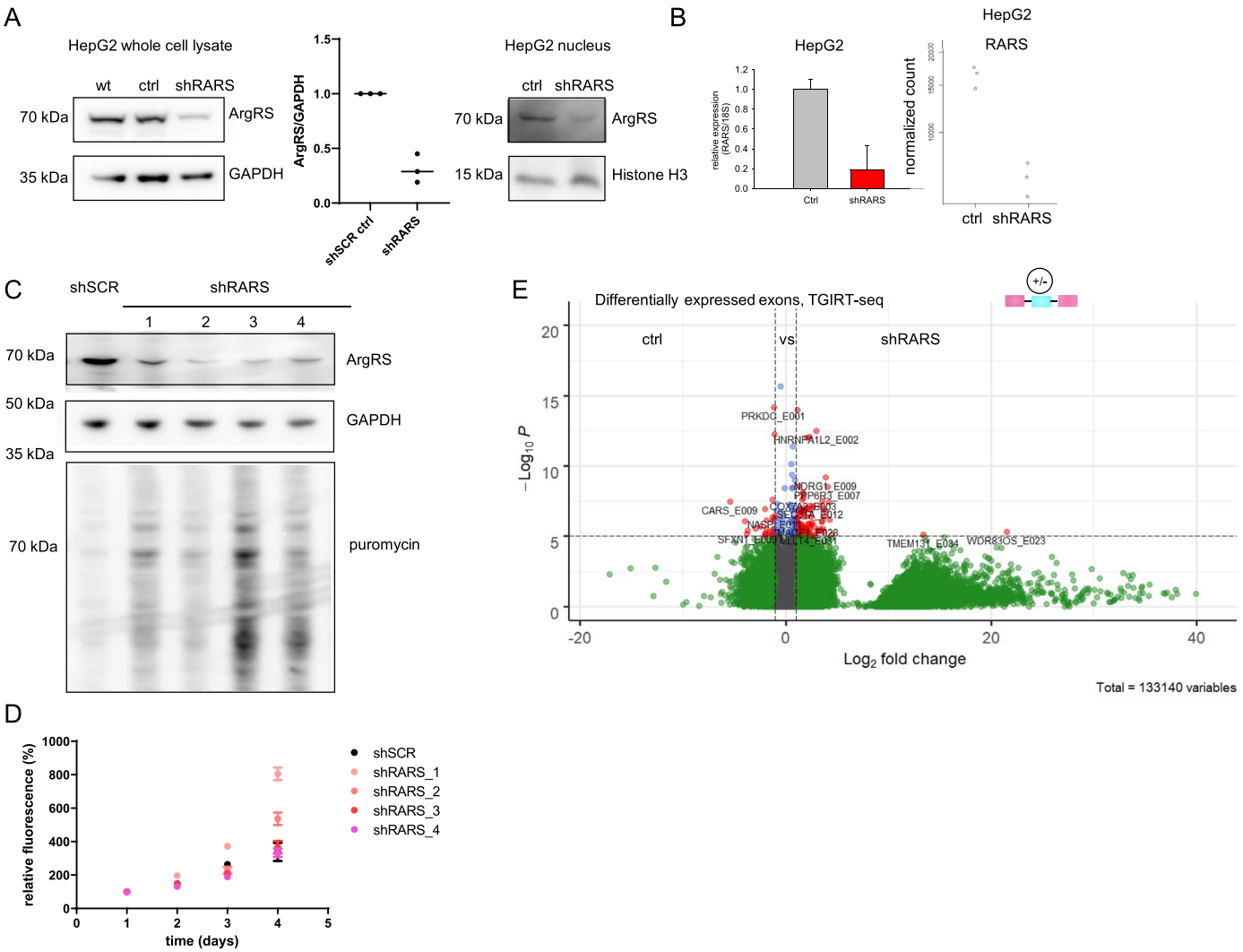


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41556-023-01118-8

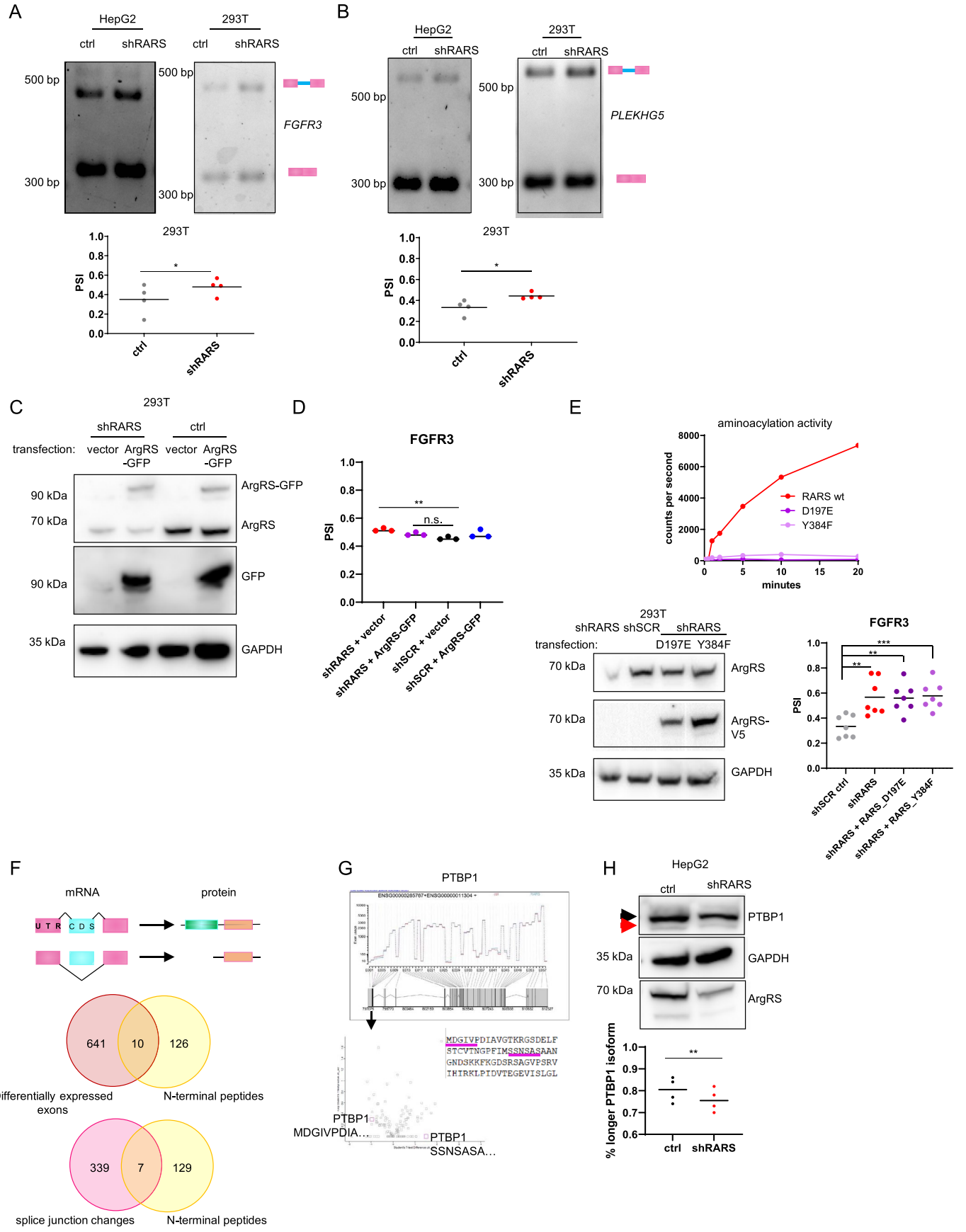
Extended Data Fig. 5 | Fluorescent microscopy to assess SRRM2 labeling and dynamics. (**a**) Labeling of endogenous SRRM2 by fusion with the fluorescent protein mVenus. mVenus was inserted four amino acids from the C-terminus of SRRM2. Verification of genomic mVenus insertion by PCR. Predicted amplification length of wildtype amplicon: 238 bp; predicted length of SRRM2-mVenus amplicon: 2071 bp. (**b**) Labeling with mVenus did not alter SRRM2 localization in 293T cells. Endogenous, unlabeled SRRM2 was visualized by immunofluorescence (upper panel). SRRM2-mVenus fluorescence displayed a similar pattern with fluorescence restricted to the cell nucleus and defined foci (middle panel). In comparison, stable expression of mVenus alone did not display a distinct nuclear localization (lower panel). Bar: 10 µm. (**c**) Fluorescence Recovery After Photobleaching (FRAP) measured over 90 s. Recovery plateaued after 60 s. (**d**) Absolute immobile SRRM2-mVenus fraction calculated from one-exponential curve fits of FRAP measurements and relative immobile fraction calculated by subtraction of residual fluorescence directly after bleaching. n = 6, 7, 6 with each derived from an average of > 16 regions of interest. Means with SEM are shown. (**e**) Recovery of fluorescent signal is faster in the absence of arginine. Fluorescent signal was corrected for background and normalized. Each measurement included 14 regions of interest and the geometric mean of 8 measurements is shown. Means with SEM are shown. (**f**) Cell viability upon ArgRS knock down in SRRM2-mVenus cells, measured with Alamar Blue. Technical quadruplets and means with SEM are shown. (**g**) Knock down of ArgRS did not affect SRRM2-mVenus protein levels measured by flow cytometry.

https://doi.org/10.1038/s41556-023-01118-8



Extended Data Fig. 6 | Characterization of ArgRS knock down. (**a**, **b**) ArgRS knock down was confirmed on protein level by western blot (A) and (B) on mRNA level by RT-qPCR (RQ and RQ max are shown) and RNA-seq. A representative of at least 3 repeats is shown for western blots and RT-qPCR. Replicates are different cell passages. (**c**) Protein synthesis was not inhibited by ArgRS knock down in HepG2 cells. Western blot verification of ArgRS knock down with 4 different shRNAs (shRARS_1–4). Western blot of puromycin incorporation. Puromycin is incorporated into newly synthesized proteins and detected with an anti-puromycin antibody to infer protein synthesis. A representatives of 6 (shRARS_1, 2) or 3 repeats (shRARS_3, 4) with different shRNAs are shown. Quantification of puromycin incorporation: shSCR: 1, shRARS_1: 1.5 ± 0.4, shRARS_2: 0.9 ± 0.3, shRARS_3: 2.6 ± 1.5, shRARS_4: 1.9 ± 1.1. (**d**) Alamar blue based cell viability assay, quadruplets of technical replicates with 4 different shRNAs are shown. Means with SEM are shown. (**e**) Volcano plot of differentially expressed exons upon ArgRS knock down as found by TGIRT-seq. Dashed vertical lines: 2-fold change, dashed horizontal line: p-value 1e-6. (**a-d**) ctrl: shSCR control, shRARS: ArgRS knock down. Significance was calculated with DEXSeq2.

https://doi.org/10.1038/s41556-023-01118-8

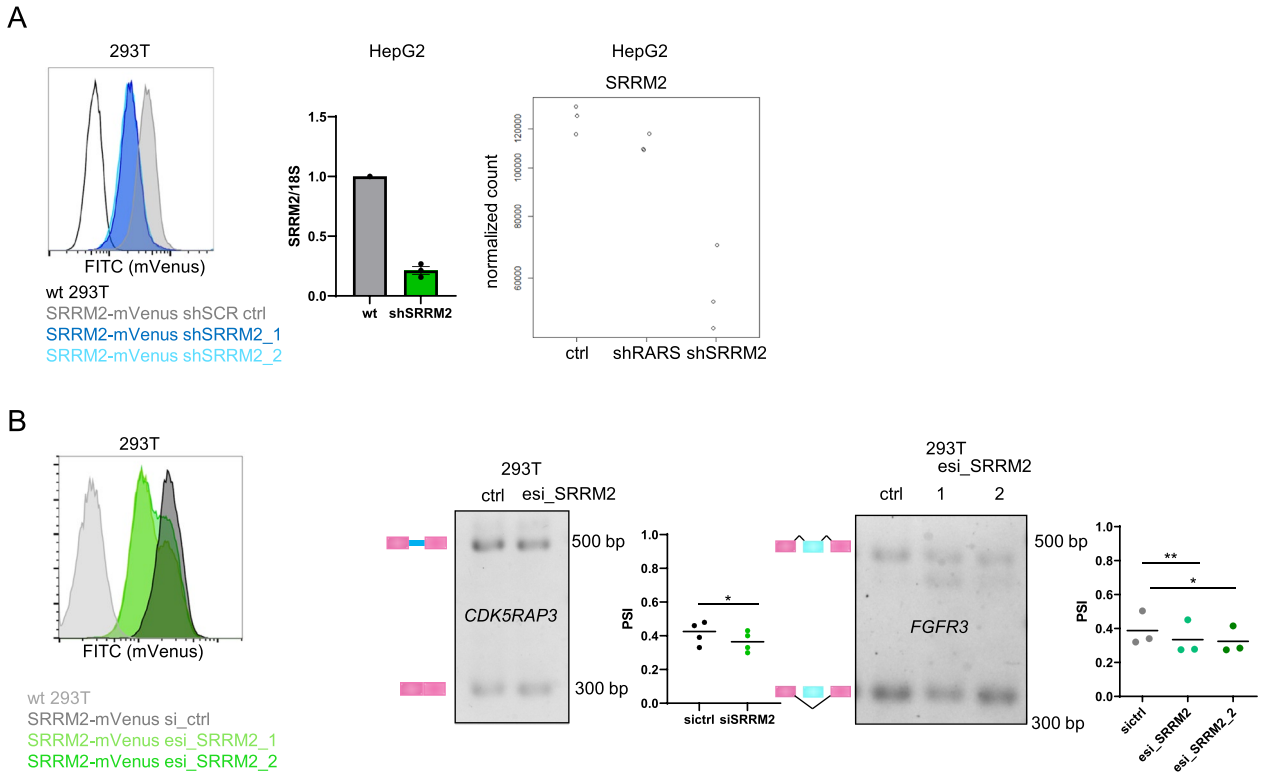


Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s41556-023-01118-8

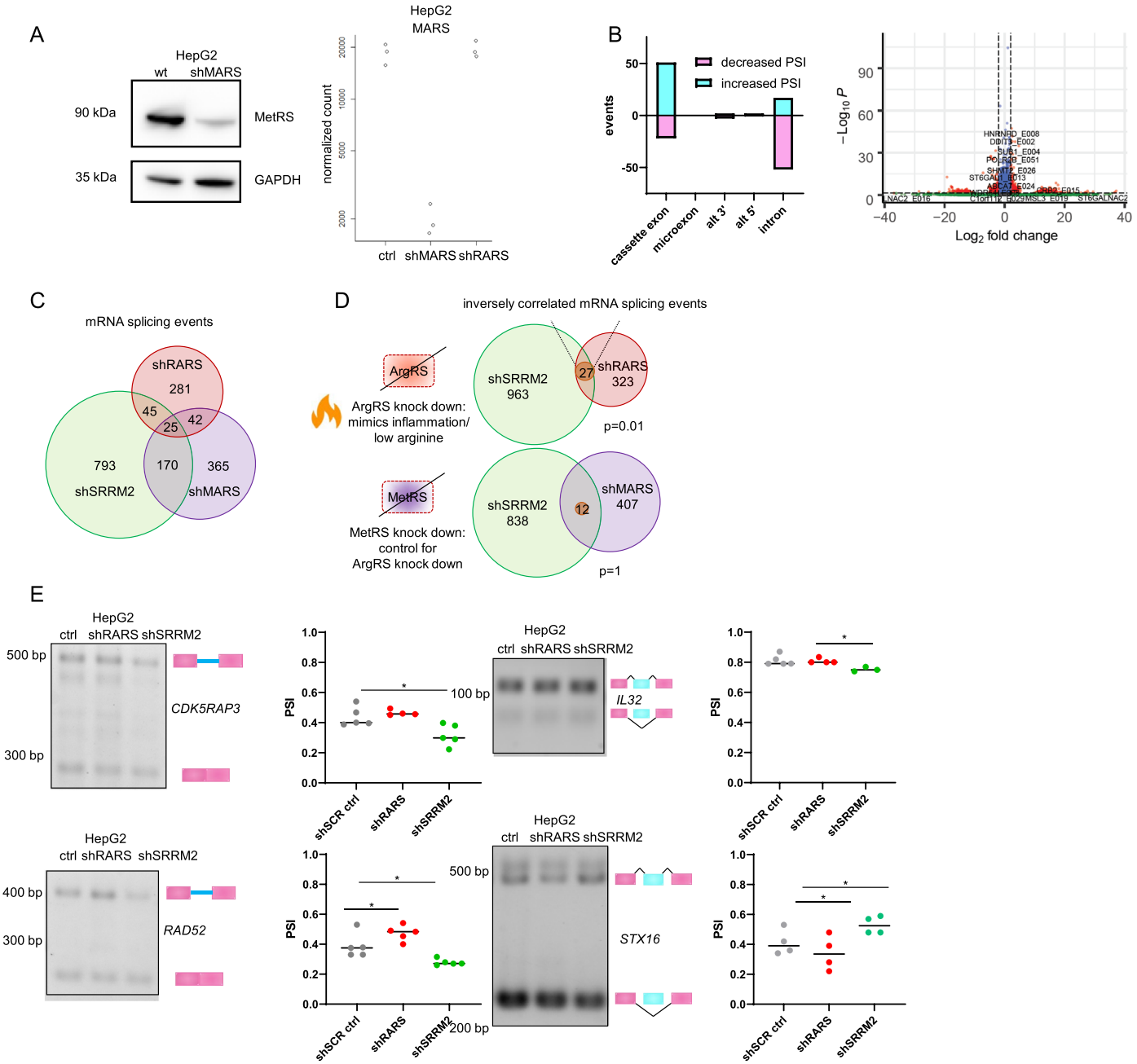
Extended Data Fig. 7 | ArgRS knock down affected mRNA processing and protein isoforms. (**a, b**) RT-PCR of splice isoforms of (A) FGFR3 and (B) PLEKHG5. Genes were detected as alternatively spliced in HepG2 cells by RNA-seq and confirmed for HEK 293T cells. A representative image from 3 or more replicates is shown as well as the densitometric quantification for 293T. PSI: percent spliced in. Ctrl: shSCR control. p = *0.04, n = 4, 4. p = *0.03, n = 4, 4 cell passages. Unpaired, two-tailed t-test. (**c**) Rescue of ArgRS knock down by overexpression of ArgRS-GFP by transient transfection in 293T cells, shown by western blot. (**d**) Rescue of FGFR3 splicing by overexpression of ArgRS-GFP. Densitometric analysis of 3 repeats is shown. p = *0.003, n = 3, 3, 3, 3 independent transfections. One-way Anova, Tukey's multiple comparisons test. (**e**) FGFR3 splicing change upon ArgRS knock down is not rescued by expression of catalytically inactive ArgRS mutants. Top: Aminoacylation assay with purified wild-type ArgRS and ArgRS mutants. Lower left: Western blot demonstrating restoration of ArgRS expression levels by overexpression of ArgRS-V5 mutants. Lower right: Densitometric analysis of RT-PCR of FGFR3 splice isoforms without or with overexpression of catalytically inactive ArgRS-V5 mutants. p = **0.003 (shSCR vs shRARS), **0.001 (shSCR vs shRARS + RARS_D197E), ***0.0006 (shSCR vs shRARS + RARS_Y384F), n = 7, 7, 7, 7, independent transfections. One-way Anova, Tukey's multiple comparisons test. (**f**) Genes and their corresponding proteins shared between differentially expressed exons or splice junction changes (RNA-seq) and N-terminal peptides (proteomics) in ArgRS knock down vs control cells. (**g**) N-terminal peptides found for PTBP1 in mass spectrometry. (**h**) Western blot verification of preference for a higher molecular weight form of PTBP1 after ArgRS knock down. p = *0.003, n = 4, 4. Paired, two-tailed t-test.

https://doi.org/10.1038/s41556-023-01118-8



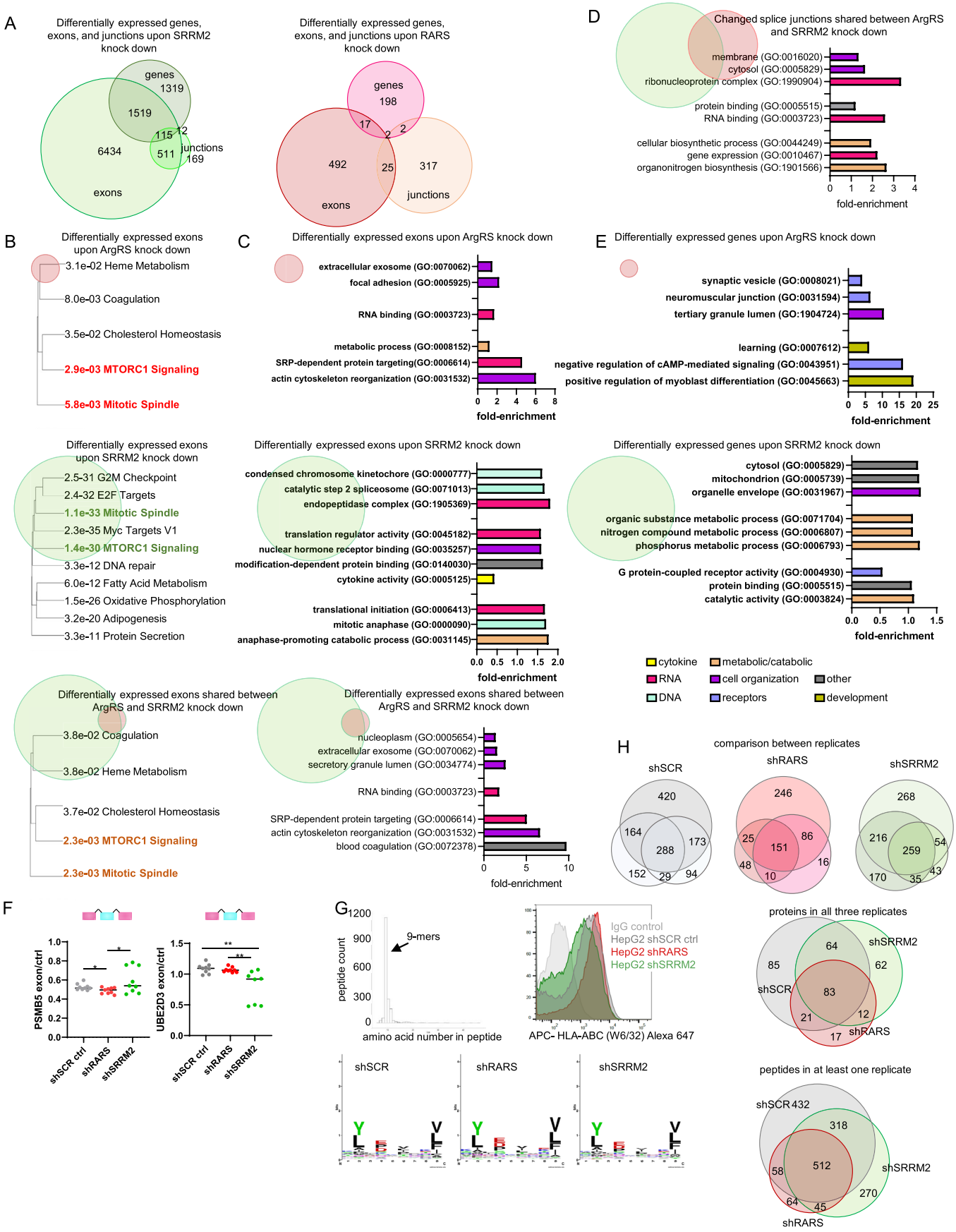
Extended Data Fig. 8 | Verification of stable and transient SRRM2 knock down and its effects on mRNA splicing. (**a**) Left to right: Knock down of SRRM2 in 293T_SRRM2_mVenus cells to verify shRNAs later used in HepG2 cells, assessed by flow cytometry. < 50% medium fluorescence intensity reduction upon SRRM2 knock down. Knock down of SRRM2 in HepG2 verified by qRT-PCR, three repeats, each measured in triplicates. Means with SEM are shown. Normalized SRRM2 counts in RNAseq. (**b**) Left: Verification of transient SRRM2 knock down by a heterogenous pool of siRNAs (esiRNA) in 293T_mVenus cells using flow cytometry. esi_SRRM2 1 and 2 are different batches of the same pool. ctrl: non-targeting pool of siRNAs. Right: RT-PCR of CDK5RAP3 and FGFR3 upon transient knock down of SRRM2 with a pool of esiRNA in 293T cells. CDK5RAP3: p = *0.03 paired, two-tailed t-test. n = 4, 4. FGFR3: p = **0.01, *0.04 paired, two-tailed t-test, n = 3, 3, 3. Replicates are independent transfections.

https://doi.org/10.1038/s41556-023-01118-8



Extended Data Fig. 9 | Characterization of MetRS knock down and verification of splicing events regulated by ArgRS and SRRM2 in opposite directions. (**a**) Verification of MetRS knock down (shMARS) by western blot and RNAseq. (**b**) Splicing events with |ΔPSI| > 0.1, MV > 0 and differential exon usage (padj < 0.05 DEXSeq). Dashed vertical lines: 2-fold change, dashed horizontal line: 1e-6. Significance was calculated with DEXSeq. (**c**) Venn diagram of shared splicing event changes upon knock down of ArgRS, SRRM2, or MetRS. All splicing event changes relative to shSCR control. MV > 0, no criterion for |ΔPSI|. (**d**) Inversely correlated alternative splicing events shared between ArgRS and SRRM2 or MetRS and SRRM2. Hypergeometric test for overrepresentation was calculated in R. (**e**) Knock down of ArgRS and SRRM2 had opposite effects on CDK5RAP3, RAD52, and STX16 processing as shown by RT-PCR. Representatives of 4 or 5 replicates shown with quantification by densitometry. CDK5RAP3: p = *0.02, n = 5, 4, 5. IL32: p = *0.03, n = 4, 4, 3. Unpaired, two-tailed t-test. RAD52: p = *0.02, *0.04, n = 5, 5, 5. STX16: p = *0.04, *0.04, n = 4, 4, 4. Paired, two-tailed t-test. Replicates are different cell passages.

https://doi.org/10.1038/s41556-023-01118-8



Extended Data Fig. 10 | See next page for caption.

https://doi.org/10.1038/s41556-023-01118-8

Extended Data Fig. 10 | SRRM2 and ArgRS induced changes on gene expression, exon and splice junction usage, and MHCI peptide presentation. (**a**) Comparison of genes with differential gene expression, differential exons, and splice junctions for shSRRM2 and shRARS. (**b**) Genes with differentially expressed exons enriched in MSigDB hallmark categories upon SRRM2 and ArgRS knock down (padj < 0.05) in HepG2. (**c**) Gene Ontology (GO) enrichment of genes with differentially expressed exons upon SRRM2 and ArgRS knock down (padj < 0.05) in HepG2. (**d**) GO enrichment of genes with changed splice junctions upon SRRM2 and ArgRS knock down (MV > 0) in HepG2. (**c, d**) All expressed genes were used as reference. (**e**) MSigDB hallmark categories and GO enrichment of genes with differential gene expression shared between SRRM2 and ArgRS knock down. (**b**-**e**) Circle size and Venn diagram visualize the proportion of shared exons, junctions, or genes. (**f**) Quantification of differentially expressed exons by qRT-PCR. PSMB5: p = *0.03, *0.04. UBE2D3: p = **0.006, **0.008. Unpaired, two-tailed t-test. n = 9, 9, 9, 9, 9, 8 different passages. PSMB5: Splicing in ENST00000493471 led to a frameshift, resulting in PSMB5 isoform 3. UBE2D3: Skipping of ENSE00003638329 leads to an alternative start site, resulting in UBE2D3 isoform 4. (**g**) (top left) Length distribution histogram of MHCI peptides confirmed that isolated peptides consisted predominantly of 9 amino acid peptides. (top right) Flow cytometry of MHCI on HepG2 shSCR ctrl, ArgRS knock down, and SRRM2 knock down cells after antibody staining. (bottom) Comparable overrepresentation of amino acids were found with WebLogo in 9 amino acid-long peptides. (**h**) Venn diagrams of MHCI peptides identified by mass spectrometry following the isolation of MHCI complex. Comparison between biological replicates (top panel), proteins corresponding to the detected peptides found in all three replicates (middle panel), and peptides found in at least one of three replicates (lower panel). All replicates are different cell passages.

nature portfolio

Corresponding author(s): Alan Lambowitz, Paul Schimmel

Last updated by author(s): Jan 22, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Zen (Zeiss) Instrument-associated software for Illumina, LSR II
Data analysis	Imaris 9 Bitplane https://imaris.oxinst.com/ Fiji Schindelin et al., 2012 https://imagej.net/Fiji Sable Adamczak et al., 2004 https://sable.cchmc.org/sable_doc.html Hisat2 Kim et al., 2015 http://daehwankimlab.github.io/hisat2/ Bowtie2 (Langmead and Salzberg, 2012 http://bowtie-bio.sourceforge.net/bowtie2/index.shtml Trimgalore Felix Krueger https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ Subread Liao et al., 2014 https://bioconductor.org/packages/release/bioc/html/Rsubread.html DESeq2 Love et al., 2014 http://bioconductor.org/packages/release/bioc/html/DESeq2.html DEXseq Anders et al., 2012 https://bioconductor.org/packages/release/bioc/html/DEXSeq.html Vast-Tools Irimia et al., 2014 https://github.com/vastgroup/vast-tools Vast Tools Diff module Han et al., 2017 https://github.com/vastgroup/vast-tools VastDB Tapial et al., 2017 http://vastdb.crg.eu/wiki/Main_Page Panther Mi et al., 2019 http://pantherdb.org/ ShinyGO Ge et al., 2020) http://bioinformatics.sdstate.edu/go/ Maxquant 1.5-2.2 Cox,J. and Mann,M., 2008 https://www.maxquant.org/ Perseus 1.5-1.6 Tyanova et al., 2016 https://maxquant.net/perseus/ GraphPad Prism 5,7-9 GraphPad Software LLC https://www.graphpad.com/ R Studio 3.4-4.2 R Studio https://rstudio.com/ Tidyverse 1.2-1.3 Tidyverse/RStudio https://www.tidyverse.org/ biomaRt Durinck S et al., 2009 https://bioconductor.org/packages/release/bioc/html/biomaRt.html EnhancedVolcano Blighe K, Rana S, Lewis M (2020). https://bioconductor.org/packages/release/bioc/html/EnhancedVolcano.html

TGIRTseq pipeline Douglas C. Wu https://github.com/wckdouglas/tgirt_map
ggpubr Alboukadel Kassambara https://CRAN.R-project.org/package=ggpubr
Subread to DEXSeq Vivek Bhardwaj https://github.com/vivekbhr/Subread_to_DEXSeq

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

TGIRT-seq data are deposited in the NCBI sequence read archive and accessible through the BioProject number PRJNA561913. Poly(A) RNA-seq data are deposited at GEO (GSE165513). Mass spectrometry search results and raw data are deposited in the PRIDE archive as PXD015692, PXD024091, and PXD027531.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Power analysis was used to predetermine sample numbers for animal experiments. Sample size was otherwise chosen based on previous preliminary experiments with the respective assays, standards in the field, or depending on the subsequent analysis. If no previous information was available, at least three replicates were used.
Data exclusions	FRAP data was excluded if sample drift during recording was noted. One animal was excluded as treatment with Arginase-1 did not show a reduction of plasma arginine. Experiments were excluded if obvious experimental errors took place.
Replication	PCA plots were used to estimate reproducibility between repeats for sequencing data sets. Replicates were determined to be successful if the primary manipulation was detectable (for example, knock down of ArgRS/SRRM2/MetRS). All replicates used for sequencing were successful.
Randomization	Samples were not randomized but compared based on treatment groups. However, samples were processed in a random order.
Blinding	Blinding was not implemented as samples were compared based on treatment groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	ArgRS Biorbyt orb247357

Antibodies used	ArgRS Abclonal A6307 ArgRS PMID 32817337 RARS-5 ArgRS Proteintech 66815-1-Ig SRRM2 SCBT sc-390315 SRRM2 Life Technologies PA566827 MetRS Abcam ab31541 MetRS Proteintech 14829-1-AP GFP Proteintech 10087-514 SON Sigma Aldrich HPA023535 Coilin CST 14168T SFPQ Proteintech 15585-1-AP PTBP1 Proteintech 12582-1-AP GAPDH CST 2118L Histone H3 CST 9715S Tubulin CST 2144S MHCI, HLA-A, B, C Purified in house from W6/32 hybridoma W6/32 Anti-Rabbit HRP Invitrogen 31464 Anti-Mouse HRP Invitrogen PI31432 Anti-Rabbit Alexa 568 Abcam ab175696-500ug Anti-Mouse Alexa 488 Abcam ab150113-500ug Anti-Mouse Alexa 647 Abcam ab150115-500ug
Validation	Manufacturer's website. shRNA mediated knock down (ArgRS, SRRM2, MetRS). GFP-antibody was validated by expression of GFP and comparison to an empty vector control. SON antibody was subjected to enhanced validation by the manufacturer: https://www.sigmaaldrich.com/CA/en/technical-documents/technical-article/protein-biology/immunohistochemistry/antibody-enhanced-validation CST antibodies are validated through 6 independent strategies: https://www.cellsignal.com/about-us/cst-antibody-validation-principles Proteintech antibodies had been validated by knock out or knock down according to the manufacturer. Secondary antibodies had been validated depending on application: https://www.abcam.com/primary-antibodies/how-we-validate-our-antibodies and https://www.thermofisher.com/ca/en/home/life-science/antibodies/secondary-antibodies.html

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	ATCC (HepG2, 293T). Yao Tong at Scripps Research (immortalized MEF cells).
Authentication	HepG2 and 293T cells were verified by the ATCC Cell Line Authentication service using STR Profiling Results. Both cell lines were found to be a 93% match with the reference cell line profile. HepG2 cells were originally derived from a male donor and HEK 293T cells from a female embryo. Murine embryonic fibroblasts (MEF) cells were a kind gift from Yao Tong (Scripps Research). MEFs were generated from C57Bl/6J embryos around E13.5 and used at cell passage 3. The sex of the embryos that were used for MEF generation was not determined.
Mycoplasma contamination	Cell lines were regularly tested for mycoplasma contamination with a PCR-based, in-house assay and were consistently negative.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Our research complies with all relevant ethical regulations. Specifically, all animals were sacrificed in accordance with animal protocol 19-0023 which was approved by the Scripps Research IACUC Department. 6-8-week-old, female C57BL6/J mice from the Scripps Research breeding facility were used. N-numbers can be found in the figure legends.
Wild animals	No wild animals were used.
Field-collected samples	No field collected samples were used.
Ethics oversight	All animals were sacrificed in accordance with animal protocol 19-0023 which was approved by the Scripps Research IACUC Department.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	293T cells were detached using Trypsin/EDTA and resuspended in DMEM with 10% FBS. The cell suspension was mixed 1:1 with sorting buffer (2.5 mM EDTA, 25 mM HEPES, pH 7.0, 1% FBS, 1% penicillin/streptomycin in PBS). Two days after transfection, cells were detached using Accutase (Stemcell Technologies), resuspended in DMEM with FBS, and spun down to remove serum to prevent cells from adhering. Following a wash with PBS, cells were diluted in sorting buffer to 1x106 cells/ml. For quantification of MHCI expression, cells were detached with EDTA and stained with 1 ug/ml HB-95 W6/32 for 20 minutes on ice followed by Alexa 647-cojugated anti-mouse secondary antibody for 20 minutes on ice, both in sorting buffer.
Instrument	Flow cytometry was performed on instruments maintained by the Flow Cytometry core at Scripps Research. Cells were analyzed on an LSR II analytical flow cytometer (BD Bioscience). mVenus fluorescence was detected in the FITC channel (filter 525/50) after excitation by a 488 nm laser. Fluorescence activated cell sorting was performed by members of the Scripps Research Flow Cytometry core on a MoFlo Astrios EQ jet-in-air sorting flow cytometer (Beckman Coulter).
Software	Flow cytometry results were analyzed with FlowJo version 10.06 or higher.
Cell population abundance	Fluorescence activated cell sorting was performed by members of the Scripps Research Flow Cytometry core on a MoFlo Astrios EQ jet-in-air sorting flow cytometer (Beckman Coulter). Two days after transfection, cells were detached using Accutase (Stemcell Technologies), resuspended in DMEM with FBS, and spun down to remove serum to prevent cells from adhering. Following a wash with PBS, cells were diluted in sorting buffer to 1x106 cells/ml. Single cells were sorted into a 96 well plate prefilled with 100 µL medium and expanded to monoclonal cell lines.
Gating strategy	Cells were gated to exclude debris and to identify single cells.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.