Electronic Supplementary Information (ESI) for

Substoichiometric ribose methylations in spliceosomal snRNAs

Nicolai Krogh, Martin Kongsbak-Wismann, Carsten Geisler and Henrik Nielsen

The ESI contains:

Materials and Methods

RiboMeth-seq analysis. RiboMeth-seq analysis was performed in triplicates according to a previously described protocol.[1, 2] Briefly, the RNA was degraded by alkaline into short fragments and the 20-40 nt fraction purified from gels. Then, adaptors were ligated to the library fragments using a modified *Arabidopsis* tRNA ligase. Finally, the library was sequenced on the Ion Proton sequencing platform, mapped, and scored for read-end counts. For internal positions, the score expresses the fraction methylated at a given position based on comparison of the read-end counts at the queried position compared to a weighted average of six neighboring positions at either side.[1] In the case of cap proximal nucleotides, there are no upstream neighbors and we adapted the score to use pos. 3-8 as a reference for calculating the RiboMeth-seq score (RMS) at the queried positions 1 and 2. In a few cases (indicated by (*) in Table S2), a barcode correction was applied. We have noted that commercial RNA oligonucleotides can be heterogeneous in length. If an oligo missing a nucleotide at its 5' end is ligated to a library fragment with an identical nucleotide at its 3' end, the library nucleotide will be removed during the sequence processing step intended to remove adapter sequence. Since all analyses are made in triplicates using adapters with different barcodes, such errors are easily detected and dealt with by excluding the 3' end data set from the calculation at the affected site. BAM files with sequence reads mapped to snRNAs analysed in this study and an excel file with the complete set of read-end data, calculated RMS, statistical analyses, and read-end count analysis are deposited to NCBI Gene Expression Omnibus (GEO) and accessible at GSE104947.

RNA from solid tissues. Whole cell RNA extracted from solid tissues pooled from 5 adult donors was purchased from BioChain (Cat nos. brain (R1234035-P), skeletal muscle (R1234171-P), skin (R1234218-P), pancreas (R1234188-P), liver (R1234149-P).

Primary $CD4^+$ T cells. Primary $CD4^+$ T cells were purified from buffy coats. Buffy coats were obtained from anonymous healthy blood donors. Written informed consent was obtained from blood donors at the Department of Clinical Immunology, University Hospital Rigshospitalet, Copenhagen and used without the possibility to identify case specific information. Use of these buffy coats for research was approved by the ethical committee, Region H, The Capital Region of Denmark. Mononuclear cells were isolated by Lymphoprep (Axis-Shield, Oslo, Norway) density gradient centrifugation using SepMate™ tubes (#86460, Stemcell Technologies, Grenoble, France). From the mononuclear cells, naïve $CD4^+$ T cells were isolated and cultured as previously described[3], using EasySep Human Naïve $CD4^+$ T cell Enrichment Kit (#19155, Stemcell Technologies). The resulting cell population contained 95-98% $CD4^+$ T cells of which >96% were $CD45RA^+$. The purified naïve $CD4^+$ T cells were subsequently cultured in serum-free X-VIVO 15 medium (#BE02-060F, Lonza, Verviers, Belgium) at 37°C, 5% CO_2 at a cell concentration of 1 x 10^6 cells/ml in flat-bottomed 24-well tissue culture plates (#142485, NuncTM, ThermoFisher Scientific, MA, USA). For generation of activated $CD4^+$ T cells, the cell cultures were supplemented with Dynabeads Human T-Activator CD3/CD28 (#111.32D, Thermo Fisher Scientific) at a cell to bead ratio of 5:2 for up to 120 hours.

Cultivation of Jurkat cells. The Jurkat cell line (E6-1) was from American Type Culture Collection. After thawing, the cells were transferred to cell culture flasks in a concentration of 0.5 x 10^6 cells/ml. Cell culture medium for Jurkat experiments was RPMI 1640 (#11875093, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum, 1% L-glutamine, 0.5 IU/l penicillin and 500 mg/l streptomycin. When reaching a concentration of 2.5 x 10^6 cells/ml, the cells were split and transferred to new culture flasks.

RNA extraction from T cells and Jurkat cells. Whole cell RNA from T cells and Jurkat cells were extracted using TRIzol™ Reagent (Thermo Fisher Scientific) in accordance with the manufacture's protocol.

Statistics. Data is shown as mean ± standard deviations, significance is calculated using unpaired Student's t-test for 2'-O-Me sites having a difference in mean RMS >0.05. * indicates $p \leq 0.05$, **$p \leq 0.01$ and ***$p \leq 0.001$.

Table S1. 2'-O-Me sites in human snRNAs and their putative scaRNA guides. Note that several of the positions for which the guide RNA is currently unknown are neighbours to sites with an assigned scaRNA. Generally, these assignments are based on sequence inspection only. It has been documented in several cases that box C/D RNAs can guide methylations at two neighbouring positions, and this may account for methylations at U2-G12, U6-G54, and U6-C63. Predictions of scaRNAs are based on the snoRNA-LBME database,[4] Karijolich *et al*.,[5] and Machyna *et al*.[6]

snRNA	Cap-proximal	Internal	Guide RNA for internal 2'-O-Me
U1	A1	A70	SCARNA7
	U2		
U2	A1	G11	SCARNA2
	U2	G12	Not known
		G19	SCARNA9
		G25	SCARNA2
		A30	SCARNA9
		C40	Not known
		U47	SCARNA28
		C61	SCARNA2
U4	A1	C8	SCARNA17
	G2	A65	SCARNA5
U5	A1	G37	Not known
	U2	U41	SCARNA5/6
		C45	SCARNA10
U6	None	A47	SNORD7
		A53	SNORD8/9
		G54	Not known
		C60	SNORD67
		C62	SNORD94
		C63	Not known
		A70	Not known
		C77	SNORD10

Table S2. RMS and standard deviations at all snRNA 2'-O-Me sites identified in solid tissues, T cells and Jurkat cells. *: Sites where barcode correction was applied.

Solid tissues – RMS (fraction methylated).

	pos.	Brain	Liver	Skin	Skeletal muscle	Pancreas
U1	A1	0.99 ±0.01	0.95 ±0.02	0.98 ±0.03	0.99 ±0.01	0.97 ±0.01
	U2	0.99 ±0.02	0.99 ±0.01	0.96 ±0.03	0.95 ±0.05	0.99 ±0.01
	A70	0.95 ±0.03	0.99 ±0.00	0.96 ±0.02	0.96 ±0.02	0.94 ±0.02
U2	A1	0.99 ±0.00	0.99 ±0.01	0.99 ±0.01	0.99 ±0.01	0.99 ±0.00
	U2	0.98 ±0.00	0.98 ±0.02	0.97 ±0.01	0.93 ±0.02	0.99 ±0.00
	G11	0.94 ±0.02	0.97 ±0.00	0.94 ±0.02	0.83 ±0.07	0.97 ±0.01
	G12	0.95 ±0.03	0.93 ±0.01	0.81 ±0.08	0.77 ±0.06	0.91 ±0.03
	G19	0.93 ±0.02	0.94 ±0.01	0.90 ±0.02	0.86 ±0.05	0.96 ±0.02
	G25	0.95 ±0.01	0.96 ±0.02	0.96 ±0.00	0.89 ±0.01	0.95 ±0.01
	A30	0.97 ±0.01	0.95 ±0.02	0.93 ±0.02	0.85 ±0.04	0.98 ±0.01
	C40	0.90 ±0.05	0.93 ±0.03*	0.87 ±0.03	0.83 ±0.02*	0.96 ±0.00
	U47	0.97 ±0.02	0.95 ±0.01	0.97 ±0.01	0.87 ±0.04	0.96 ±0.01
	C61	0.92 ±0.03	0.89 ±0.04	0.90 ±0.00	0.84 ±0.05	0.88 ±0.03
U4	A1	0.95 ±0.03	0.94 ±0.08	0.92 ±0.01	0.99 ±0.02	0.97 ±0.02
	G2	0.94 ±0.02	0.96 ±0.03	0.87 ±0.03	0.94 ±0.07	0.95 ±0.02
	C8	0.91 ±0.06	0.91 ±0.02	0.83 ±0.03	0.91 ±0.03	0.89 ±0.03
	A65	0.93 ±0.08	0.97 ±0.03	0.97 ±0.05	0.98 ±0.03	0.93 ±0.05
U5	A1	0.98 ±0.01	0.96 ±0.02	0.97 ±0.02	0.97 ±0.02	0.97 ±0.01
	U2	0.93 ±0.00	0.99 ±0.02	0.94 ±0.02	0.96 ±0.03	0.96 ±0.00
	G37	0.96 ±0.05	0.94 ±0.03	0.96 ±0.03	0.90 ±0.02	0.96 ±0.02
	U41	0.88 ±0.06	0.90 ±0.03	0.93 ±0.04	0.88 ±0.03	0.82 ±0.08
	C45	0.92 ±0.04	0.95 ±0.03	0.94 ±0.01	0.92 ±0.04	0.93 ±0.01
U6	A47	0.90 ±0.02	0.85 ±0.09	0.90 ±0.01	0.94 ±0.08	0.97 ±0.02
	A53	0.96 ±0.01	0.93 ±0.08	0.97 ±0.01	0.95 ±0.01	0.97 ±0.01
	G54	0.94 ±0.05*	0.96 ±0.05	0.96 ±0.04*	0.93 ±0.02	0.95 ±0.02
	C60	0.91 ±0.03	0.97 ±0.02	0.93 ±0.07	0.99 ±0.01	0.92 ±0.06*
U6	C62	0.92 ±0.07	0.95 ±0.03	0.99 ±0.01	0.97 ±0.02	0.94 ±0.02
	C63	0.89 ±0.04	0.83 ±0.07	0.94 ±0.05	0.86 ±0.10	0.98 ±0.03*
	A70	0.94 ±0.02	0.97 ±0.02	1.00 ±0.01	0.98 ±0.02	0.95 ±0.04
	C77	0.93 ±0.03*	0.88 ±0.03	0.95 ±0.02*	0.98 ±0.01	0.97 ±0.04*

Table 2 continues next page

T cells and Jurkat cells – RMS (fraction methylated).

	pos.	T cells naive 1	T cells naive 2	T cells 24h 1	T cells 24h 2	T cells 120h 1	T cells 120h 2	Jurkat cells
U1	A1	0.98 ±0.01	0.97 ±0.02	1.00 ±0.00	0.98 ±0.02	0.97 ±0.01	0.98 ±0.03	0.86 ±0.05
	U2	0.95 ±0.02	0.93 ±0.01	0.97 ±0.01	0.90 ±0.03	0.95 ±0.01	0.93 ±0.04	0.91 ±0.03
	A70	0.86 ±0.02	0.81 ±0.10	0.87 ±0.04	0.92 ±0.03*	0.89 ±0.02	0.83 ±0.03	0.69 ±0.08
U2	A1	0.97 ±0.01	0.99 ±0.00	0.98 ±0.00	0.99 ±0.01	0.98 ±0.01	0.98 ±0.01	0.91 ±0.05
	U2	0.96 ±0.01	0.97 ±0.01	0.91 ±0.03	0.97 ±0.01	0.93 ±0.02	0.96 ±0.01	0.87 ±0.07
	G11	0.97 ±0.01	0.97 ±0.01	0.90 ±0.01	0.88 ±0.01	0.94 ±0.01	0.95 ±0.01	0.84 ±0.04
	G12	0.80 ±0.00	0.74 ±0.02	0.93 ±0.01	0.94 ±0.04	0.78 ±0.02	0.73 ±0.05	0.87 ±0.07
	G19	0.95 ±0.01	0.96 ±0.01	0.83 ±0.07	0.85 ±0.06	0.94 ±0.01	0.96 ±0.01	0.81 ±0.05
	G25	0.91 ±0.04	0.90 ±0.05	0.94 ±0.01	0.94 ±0.03	0.93 ±0.01	0.91 ±0.03	0.89 ±0.04
	A30	0.97 ±0.01	0.97 ±0.01	0.84 ±0.02	0.90 ±0.01	0.97 ±0.01	0.96 ±0.00	0.83 ±0.02
	C40	0.95 ±0.01	0.96 ±0.01	0.84 ±0.05	0.80 ±0.08	0.90 ±0.01	0.93 ±0.00	0.73 ±0.06
	U47	0.99 ±0.00	0.95 ±0.03	0.90 ±0.03	0.94 ±0.01	0.97 ±0.00	0.96 ±0.02	0.84 ±0.05
	C61	0.91 ±0.02	0.93 ±0.02	0.82 ±0.03	0.78 ±0.06	0.81 ±0.03	0.78 ±0.05	0.60 ±0.02
U4	A1	0.92 ±0.05	0.90 ±0.05	0.95 ±0.02	0.94 ±0.09	0.88 ±0.08	0.87 ±0.14	0.82 ±0.06
	G2	0.85 ±0.07	0.74 ±0.07	0.94 ±0.04	0.98 ±0.02	0.62 ±0.13	0.70 ±0.12	0.89 ±0.07
	C8	0.95 ±0.02	0.93 ±0.03	0.60 ±0.33	0.81 ±0.14	0.88 ±0.03	0.82 ±0.06	0.10 ±0.04
	A65	0.99 ±0.01	0.88 ±0.09	0.93 ±0.04	0.92 ±0.08	0.96 ±0.03	0.86 ±0.03	0.88 ±0.05
U5	A1	0.98 ±0.02	0.95 ±0.01	0.96 ±0.01	0.98 ±0.02	0.98 ±0.02	0.99 ±0.01	0.95 ±0.03
	U2	0.81 ±0.02	0.80 ±0.05	0.94 ±0.03	0.93 ±0.06	0.72 ±0.05	0.83 ±0.04	0.81 ±0.06
	G37	0.91 ±0.03	0.87 ±0.08	0.91 ±0.04	0.90 ±0.06	0.89 ±0.07	0.95 ±0.03	0.92 ±0.02
	U41	0.97 ±0.02	0.93 ±0.05	0.84 ±0.06	0.82 ±0.05	0.95 ±0.01	0.96 ±0.04	0.80 ±0.05
	C45	0.97 ±0.02	0.99 ±0.00	0.91 ±0.04	0.84 ±0.04	0.91 ±0.07	0.98 ±0.01	0.82 ±0.07
U6	A47	0.98 ±0.00	0.92 ±0.05	0.88 ±0.02	0.92 ±0.04*	0.97 ±0.01	0.92 ±0.00	0.88 ±0.01
	A53	0.98 ±0.01	0.97 ±0.02	0.90 ±0.03	0.92 ±0.02	0.97 ±0.01	0.96 ±0.01	0.76 ±0.03
	G54	0.93 ±0.05	0.93 ±0.09	0.84 ±0.10	0.93 ±0.02*	0.90 ±0.04	0.89 ±0.10	0.76 ±0.09
	C60	0.91 ±0.07	0.97 ±0.01	0.87 ±0.04	0.92 ±0.01	0.92 ±0.03	0.94 ±0.03	0.80 ±0.09
	C62	0.95 ±0.01	0.94 ±0.05	0.94 ±0.02	0.81 ±0.03	0.96 ±0.01	0.92 ±0.06	0.81 ±0.02
	C63	0.93 ±0.04	0.96 ±0.02	0.83 ±0.04	0.80 ±0.05	0.94 ±0.01	0.94 ±0.02	0.75 ±0.09
	A70	0.99 ±0.01	0.99 ±0.00	0.91 ±0.01	0.95 ±0.02	0.97 ±0.01	0.98 ±0.01	0.89 ±0.01
	C77	0.98 ±0.02	0.95 ±0.01	0.86 ±0.09	0.92 ±0.02	0.97 ±0.01	0.98 ±0.01	0.78 ±0.08

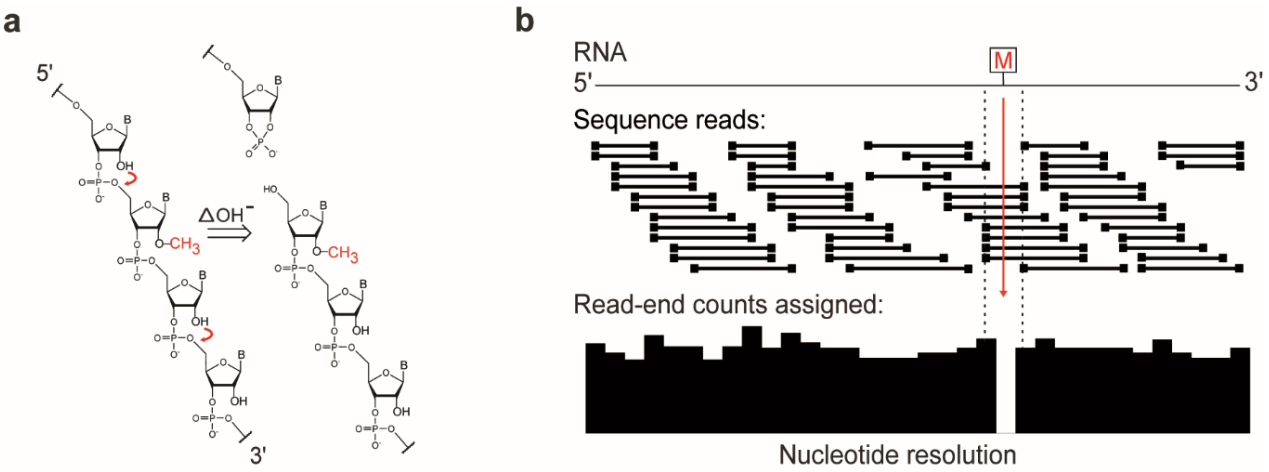

Figure S1. Detection of ribose methylations in RNA using RiboMeth-seq. (a) The principle of RiboMeth-seq is that ribose methylation renders the neighboring phosphodiester bond resistant to alkaline degradation. (b) 20–40 nt fragments from partial alkaline hydrolysis are gel purified, ligated to adaptors, reverse transcribed, and sequenced. The sequence is mapped to a reference sequence and the first and last nucleotide of the library fragment recorded. Read-ends at methylated residues will be underrepresented compared to flanking positions. The stoichiometry of ribose methyls is calculated from a comparison of the read-ends at the queried site and six flanking sites at either side.[1]

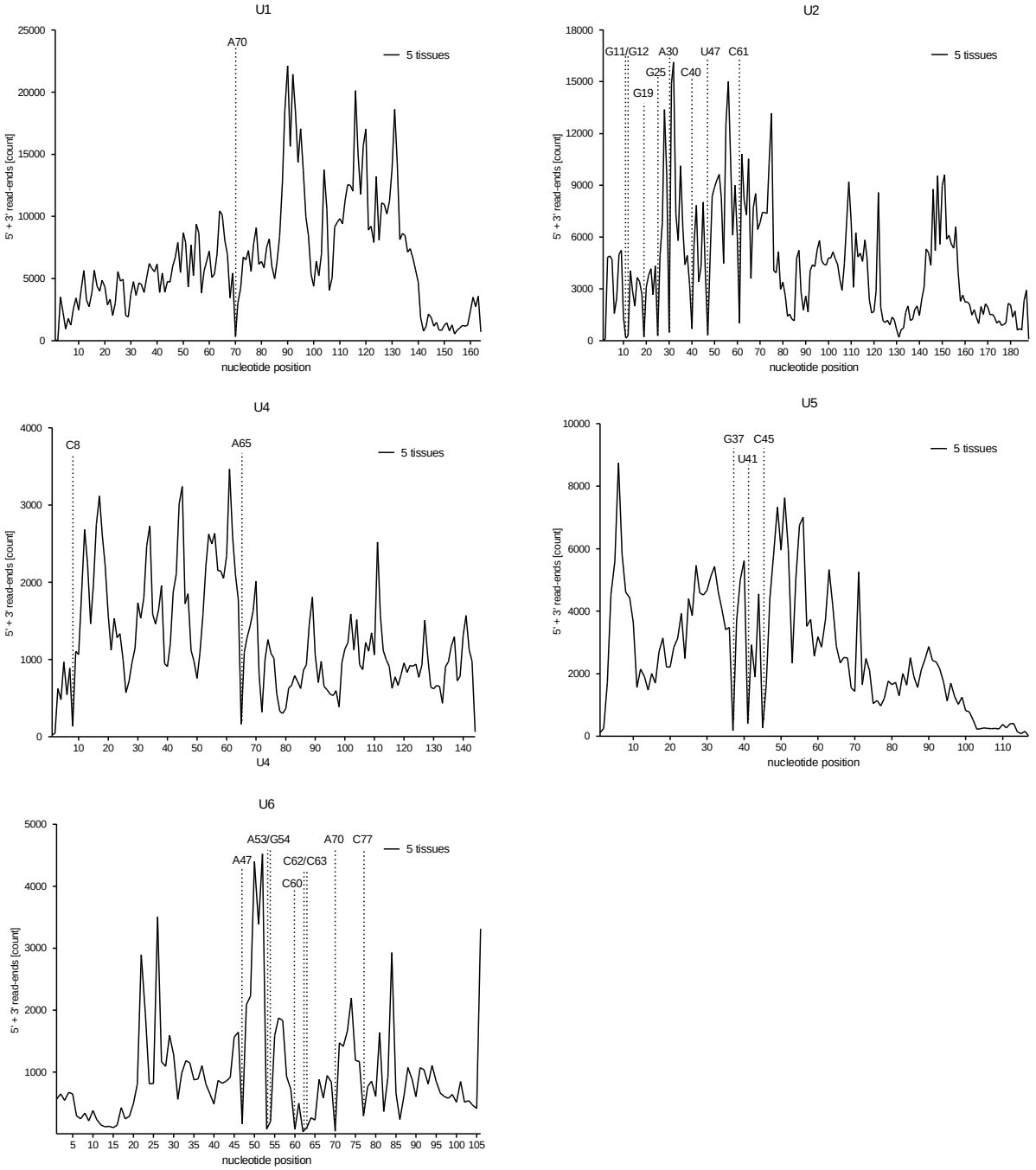

Fig. S2. Total read-end coverage at all nucleotide positions in snRNA from solid tissue datasets. 2'-O-Me sites are apparent from low coverage at the methylated nucleotides (internal 2'-O-Me sites are indicated by dashed lines).

Supplemental References

1. U. Birkedal, M. Christensen-Dalsgaard, N. Krogh, R. Sabarinathan, J. Gorodkin and H. Nielsen, *Angewandte Chemie*, 2015, **54**, 451-455.
2. N. Krogh, U. Birkedal and H. Nielsen, *Methods in molecular biology*, 2017, **1562**, 189-209.
3. M. Kongsbak, M. R. von Essen, L. Boding, T. B. Levring, P. Schjerling, J. P. Lauritsen, A. Woetmann, N. Odum, C. M. Bonefeld and C. Geisler, *PloS one*, 2014, **9**, e96695.
4. L. Lestrade and M. J. Weber, *Nucleic acids research*, 2006, **34**, D158-162.
5. J. Karijolich and Y. T. Yu, *RNA biology*, 2010, **7**, 192-204.
6. M. Machyna, S. Kehr, K. Straube, D. Kappei, F. Buchholz, F. Butter, J. Ule, J. Hertel, P. F. Stadler and K. M. Neugebauer, *Molecular cell*, 2014, **56**, 389-399.