Thiol-linked alkylation for the metabolic sequencing of RNA (SLAMseq)

Stefan Ameres (stefan.ameres@imba.oeaw.ac.at)
Institute of Molecular Biotechnology (IMBA), Vienna Biocenter Campus (VBC), 1030 Vienna, Austria
Veronika A. Herzog
Institute of Molecular Biotechnology (IMBA), Vienna Biocenter Campus (VBC), 1030 Vienna, Austria
Brian Reichholf
Institute of Molecular Biotechnology (IMBA), Vienna Biocenter Campus (VBC), 1030 Vienna, Austria

Method Article

Keywords: 4-thiouridine, RNA, RNA modifications, metabolic labeling, RNA kinetics, gene expression

Posted Date: September 25th, 2017

DOI: https://doi.org/10.1038/protex.2017.105

Abstract

Gene expression profiling by high-throughput sequencing reveals qualitative and quantitative changes in RNA species at steady-state but obscures the intracellular dynamics of RNA transcription, processing and decay. We developed thiol\(SH)-linked alkylation for the metabolic sequencing of RNA \(SLAMseq), an orthogonal chemistry-based epitranscriptomics-sequencing technology that uncovers 4-thiouridine \(s^4U)-incorporation in RNA species at single-nucleotide resolution. In combination with well-established metabolic RNA labeling protocols and coupled to standard, low-input, high-throughput RNA sequencing methods, SLAMseq enables rapid access to RNA polymerase II-dependent gene expression dynamics in the context of total RNA. SLAMseq facilitates the dissection of fundamental mechanisms that control gene expression in an accessible, cost-effective, and scalable manner. This protocol accompanies Herzog et al., Nature Methods, DOI: 10.1038/nmeth.4435.

Introduction

In thiol-linked alkylation for the metabolic sequencing of RNA \(SLAMseq), nascent RNA is labeled by the addition of 4-thiouridine \(s^4U) to cell culture growth medium. After RNA extraction, RNA is subjected to iodoacetamide treatment that covalently attaches a carboxyamidomethyl-group to s^4U by nucleophilic substitution. Reverse-transcription during library preparation specifically misincorporates G instead of A complementary to alkylated s^4U positions. Hence, s^4U incorporation events can be identified in sequencing data sets as T>C conversions. Working time for alkylation and QuantSeq library preparation are indicated. See Figure 1.

Reagents

Reagents - 2-Propanol \(Sigma-Aldrich, cat. no. 59300) - 4-Thiouridine \(Sigma-Aldrich, cat. no. T4509), protect from light - Chloroform:Isoamyl Alcohol 24:1 \(Applichem, cat. no. A1935) - DMSO \(Sigma-Aldrich, cat. no. 41640) - Ethanol absolute \(VWR Chemicals, cat. no 20821.321) - Glycerol \(Sigma-Aldrich, cat. no G5516) - Glycogen \(20 mg/ml) \(Roche, cat. no. 10901393001) \(optional) - Iodoacetamide \(Sigma-Aldrich, cat. no. I1149), protect from light - $MgCl_2$ \(Sigma-Aldrich, cat. no. M2670) - Sodium phosphate monobasic monohydrate \(NaH_2PO_4, monobasic; Sigma-Aldrich, cat. no. S9638) - Di-Sodium hydrogen phosphate dehydrate \(Na_2 HPO_4, dibasic; Merck, cat. no. 106580) - Nuclease-free water - OmniPur® DTT \(Merck, cat. no 3870) - Phosphate buffered saline \(PBS; 10×, commercially available) - QuantSeq 3´ mRNA-Seq Library Kit for Illumina \(Lexogen, cat. no. 015) - Sodium acetate \(NaOAc, Sigma Aldrich, cat. no. 71180) - Sodium chloride \(NaCl, Merck, cat. no. 106404) - Tris base \(Sigma-Aldrich, cat. no. T6066) - TRIzol™ Reagent \(Ambion, Life Technologies, cat. no. 15596018) - Optional: Bacterial Alkaline Phosphatase \(Invitrogen, cat. no. 18011015) - Optional: Cell viability assay, e.g. CellTiterGlo® Luminescent Cell Viability assay \(Promega, cat. no. G7570) - Optional: mESCs \(clone AN3-12), derived from C57BL/6x129 F1 females, obtained from IMBA Haplobank \(U. Elling _et al._, accepted for publication in Nature) - Optional: Phosphodiesterase I \(Worthington

Biochemical, cat. no. LS003926) - Optional: Uridine \(Sigma-Aldrich, cat. no. U6381) **Buffers** _Bacterial Alkaline Phosphatase storage buffer:_ 10 mM Tris-HCl pH 8, 120 mM NaCl, 50% \(v/v) glycerol _Phosphodiesterase I storage buffer:_ 0.1 M NaCl, 15 mM $MgCl_2$ _Sodium phosphate buffer \($NaPO_4$, pH 8)_ Prepare 1 M stocks solutions of NaH_2PO_4 \(monobasic) and Na_2HPO_4 \(dibasic): Dissolve 138 g of NaH_2PO_4•H_2O \(monobasic) in H_2O to make a final volume of 1 L and dissolve 142 g of Na_2HPO_4 \(dibasic) in H_2O to make a final volume of 1 L. To prepare 200 ml of 0.5 M sodium phosphate buffer \(pH 8), mix 93.2 mL of 1 M Na_2HPO_4 and 6.8 ml of 1 M NaH_2PO_4 and add 100 ml of H_2O. Control pH after buffer preparation using a pH meter.

Equipment

- Standard molecular biology and cell culture lab equipment. - Illumina high-throughput sequencing machine \(such as Illumina HiSeq 2500). - Optional: HPLC with a Supelco Discovery C18 \(bonded phase silica 5 µM particle, 250 × 4.6 mm) reverse phase column \(Bellefonte, PA, USA)

Procedure

General considerations: The ability to determine de novo synthesized \(s^4U labeled) transcripts by SLAMseq will depend on \(1) the cellular s^4U uptake kinetics, labeling time and s^4U concentration, \(2) the overall transcriptional activity of the cell type and \(3) the library sequencing depth. These parameters need to be taken into account when designing a SLAMseq experiment, particularly when employing short s^4U pulse labeling, where sequencing depth demands adjustments to the given cellular parameters. We highly recommend testing optimal concentrations of s^4U for the cell line of interest by assessing cellular toxicity \(e.g. using CellTiterGlo® Luminescent Cell Viability assay, Promega after manufacturer's instructions) and s^4U incorporation rates \(e.g. by HPLC, see below) prior to the sequencing experiment in order to meet s^4U -labeling conditions that do not affect gene expression or cellular viability. **s^4U Labeling of the cells \(pulse/chase)** _Protect s^4U from light._ Pulse labeling: - Seed cells the day before the labelling experiment in a density that allows exponential growth for the duration of the experiment \(50-80% confluency). - Prepare cell culture growth medium supplemented with non-toxic concentrations of s^4U \(i.e. 100 µM for mESCs; Note, this concentration can vary for different cell types). - Remove media from the cells and replace with s^4U -containing media to start the pulse labeling. _Note, that regular exchange of fresh s^4U-containing media \(i.e. every three hours) can significantly enhance s4U incorporation._ - Take off media at the time-points of interest and lyse cells directly in TRIzol®. - Store samples at -80°C or directly proceed with RNA isolation. Pulse/chase labeling: - After pulse labeling \(see above), remove s^4U -containing media from the cells. - Wash cells twice with 1 X PBS. - Add cell culture growth medium containing 100 X excess of uridine \(compared to s^4U concentration in the pulse) to cells. - Take off media at the time-points of interest and lyse the cells directly in TRIzol®. - Store samples at -80°C or directly proceed with RNA isolation. **RNA isolation** _DTT is added during RNA isolation to keep the sample under reducing conditions._ - Thaw lysate and incubate 5 min at room temperature - Add

200 µl chloroform per 1 ml of TRIzol® - Shake tube vigorously for 15 sec - Incubate at room temperature for 2-3 min - Spin down at 16,000 x g for 15 min at 4°C - Transfer aqueous phase to new tube - Add 1/100 volume of 10 mM DTT \(0.1 mM final concentration), 1 volume 2-propanol and optionally 1 µl glycogen \(20 mg/ml) - Vortex well - Incubate 10 min at room temperature - Spin down at 16,000 x g for 20 min at 4°C - Discard supernatant - Add 500 µl 75% EtOH and 5 µl of 10 mM DTT, vortex well - Spin down at 7,500 x g for 5 min at room temperature - Remove supernatant, let the pellet dry for ~5-10 min and resuspend in appropriate amount of water supplemented with 1 mM DTT \(final concentration) - Incubate for 10 min at 55°C - Measure concentration by Nanodrop and freeze RNA at -80°C or proceed to the next step Proceed with _"Digestion to single Nucleosides"_ in order to prepare the samples for HPLC analysis and control s^4U incorporation rates or to _"Thiol modification"_ in order to alkylate s^4U for subsequent library preparation and sequencing. **Optional: Digestion to single nucleosides and HPLC analysis** _Note: Single nucleoside digestion and HPLC analysis is optional but highly recommended to confirm adequate RNA labeling conditions, a necessary requirement for a successful SLAMseq experiment. Because HPLC-analysis is less sensitive compared to SLAMseq, s^4U -incorporation is only detectable at later labeling timepoints \(i.e. ≥3 h after labeling with 100 µM s^4U in mESC)._ - Prepare the reaction to digest and dephosphorylate total RNA to single nucleosides as described in Table 1. - Incubate overnight \(≥16h) at 37°C - Add 6 µl 3 M NaOAc \(pH 5.2), 150 µl ice-cold 100% EtOH and 30 µl 10 mM DTT \(1 mM final concentration), vortex - Incubate 10 min at -80°C - Spin down at 12,500 x g for 5 min at room temperature - Transfer supernatant to a new tube - Add 30 µl 10 mM DTT and 270 µl ice-cold 100% EtOH, incubate for 10 min at -80°C - Spin down at 12,500 x g for 5 min at room temperature - Transfer supernatant to a new tube - Evaporate supernatant in speed-vac to complete dryness - Dissolve sample in 50 µl H_2O. Store at -20°C if necessary. - Take 25 µl of the digested RNA sample and add 75 µl H_2O - For the uridine and s^4U standards, prepare the stock solutions from 2 mg/ml uridine and 0.1 mg/ml s^4U stock solutions as described in Table 2. Prepare the standard dilutions as described in Table 3. - Load the samples and the standards on HPLC. HPLC is performed on a Supelco Discovery C18 \(bonded phase silica 5 µM particle, 250 x 4.6 mm) reverse phase column. For running conditions refer to \(Spitzer et al., 2014) and \(Andrus and Kuimelis, 2001). **Thiol modification \(Iodoacetamide treatment)** - Prepare the reaction mix for each Iodoacetamide \(IAA) treatment reaction as described in Table 4. - Incubate reaction at 50°C for 15 min - Stop reaction by quenching the reaction with 1 µl 1M DTT - Add 1 µl glycogen \(20 mg/ml), 5 µl NaOAc \(3M, pH 5.2), 125 µl EtOH 100%, vortex and precipitate for 30 min at -80°C - Spin down at 16,000 x g for 30 min - Wash with 1 ml 75% EtOH, vortex - Spin down at 16,000 x g for 10 min - Remove supernatant and let the pellet dry for 5-10 min - Resuspend in appropriate volume \(5-10 µl) H_2O - Proceed with RNA quality control and library preparation **Library preparation for High-Throughput Sequencing** For mRNA SLAMseq, we recommend 3´ end sequencing using the QuantSeq 3' mRNA-Seq Library Prep Kit for Illumina \(Lexogen) according to the manufacturer's instructions \(QuantSeq FWD kit). _Considerations for sample sequencing: As standard approach, we performed SLAMseq on an Illumina HiSeq 2500 machine. Sequencing depth needs to be adjusted to s^4U incorporation, which is dependent on the experimental s^4U labeling time, the cellular s^4U uptake kinetics and the overall transcriptional activity of the cell type. As a rough estimate \(based on experiments using mouse embryonic stem cells), we

recommend for longer labeling times \(i.e. >3h) 10-20 million and for short pulse labelling \(i.e. <3h) 30-50 million reads per library for efficient quantification of s^4U-labeled transcripts. We recommend sequencing reactions in single read 100 \(SR100) mode, which recovers ~70% of all labeled transcripts in mESCs under steady-state labeling conditions \(i.e. 100 µM s^4U-labeling for 24 h)._ **Data analysis** In principle, any self-assembled bioinformatics analysis pipeline can be employed to recover and quantify T>C-conversion containing reads. Bioinformatics analysis in the publication associated with this protocol has been performed using Digital Unmasking of Nucleotide conversion-containing K-mers \(DUNK) \(see full description and availability from http://t-neumann.github.io/slamdunk). SLAM-DUNK \(Neumann et al., manuscript in preparation) is a fully automated, modular and T>C-conversion-aware alignment software package streamlining SLAMseq data analysis. Briefly, SLAM-DUNK comes with several statistics and diagnostic plotting functions as well as a MultiQC \(http://multiqc.info) plugin to make SLAMseq data analysis and integration available to bench-scientists. SLAM-DUNK executes the following four main steps: \(1) SLAM-DUNK uses NextGenMap \(Sedlazeck et al., 2013) to map the SLAMseq reads to a reference genome. \(2) Reads with a high number of mismatches and reads that map to more than one annotated 3' UTR are discarded. \(3) SLAM-DUNK uses VarScan2 \(Koboldt et al., 2012) to identify genuine SNPs between the reference and the sequenced genomes. \(4) SLAM-DUNK counts the number of T>C containing reads and estimates the fraction of labeled transcripts for each gene individually. SNP positions and read positions with a low base quality value are automatically excluded. The final output of SLAM-DUNK reports for each transcript the genomic coordinates and the UTR region, the T content of the UTR region, the number of mapped reads \(raw counts and norm. counts), the number of T>C conversion-containing reads, the coverage over T positions, the T>C conversion rate, and the number of multimapping reads.

Tables

Due to technical limitations, tables 1-4 have been placed in the supplementary file section for access.

Table 1 Reaction setup to digest and dephosphorylate total RNA to single nucleosides.

Table 2 Stock solutions of standards used in HPLC analysis.

Table 3 Dilution of standards used in HPLC analysis.

Table 4 Reaction setup for iodoacetamide-treatment of s^4^U-labeled total RNA.

References

Andrus, A., and Kuimelis, R.G. \(2001). Analysis and purification of synthetic nucleic acids using HPLC. Curr Protoc Nucleic Acid Chem Chapter 10, Unit 10.5–10.5.13. Koboldt, D. C. et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22, 568–576 \(2012). Sedlazeck, F. J., Rescheneder, P. & Haeseler, von, A. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. Bioinformatics 29, 2790–2791 \(2013). Spitzer, J., Hafner, M.,

Landthaler, M., Ascano, M., Farazi, T., Wardle, G., Nusbaum, J., Khorshid, M., Burger, L., Zavolan, M., et al. \(2014). PAR-CLIP \(Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. Meth. Enzymol. 539, 113–161.

Acknowledgements

We thank G. Krssakova and K. Mechtler \(IMP) for assistance with high-performance liquid chromatography analysis and all laboratory members for support and discussions. Highthroughput sequencing was performed at the VBCF NGS Unit \(www.vbcf.ac.at). This work was supported by grants from the European Research Council \(ERC-StG-338252) and the Austrian Science Fund \(Y-733-B22 START, W-1207-B09 and SFB F43-22) to S.L.A.

Figures

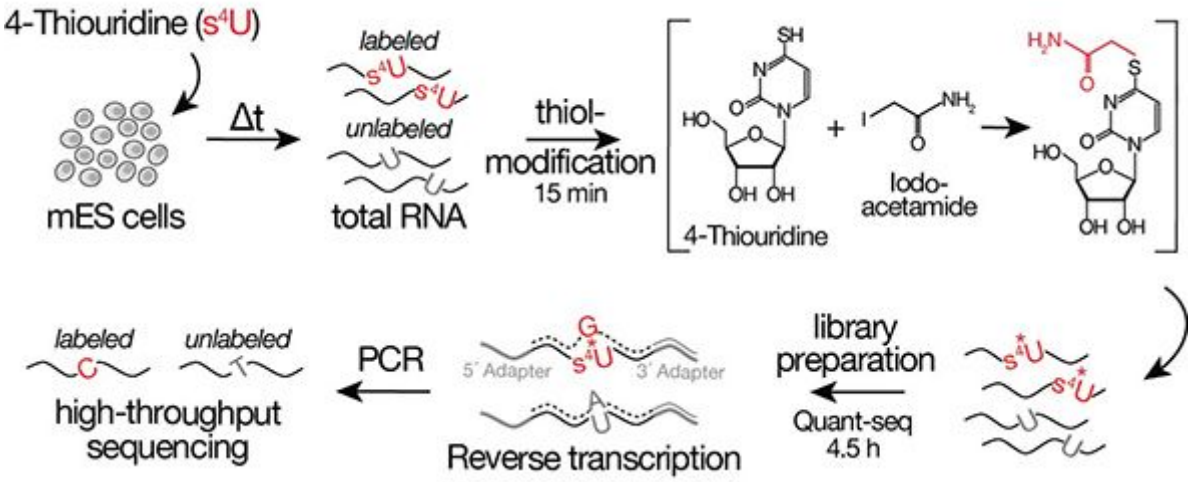

Figure 1

SLAMseq Workflow associated with thiol-linked alkylation for the metabolic sequencing of RNA (SLAMseq).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplement2.jpg
- supplement3.jpg
- supplement4.jpg
- supplement5.jpg