Analysis of pseudouridines and other RNA modifications using hydraPsiSeq protocol

Virginie Marchand, Valérie Bourguignon-Igel, Mark Helm, Yuri Motorin

METHODS special issue 2021 Revised R1

Analysis of pseudouridines and other RNA modifications using HydraPsiSeq protocol

Virginie Marchand[1], Valérie Bourguignon-Igel[1,2], Mark Helm[3], Yuri Motorin[1,2*]

[1] Université de Lorraine, CNRS, UMR7365 IMoPA, F-54000 Nancy, France.

[2] Université de Lorraine, CNRS, INSERM, UMS2008/US40 IBSLor, EpiRNA-Seq Core facility, F-54000 Nancy, France.

[3] Institute of Pharmacy of Pharmacy and Biochemistry, Johannes Gutenberg University Mainz, Mainz, Germany.

* corresponding author: Yuri MOTORIN

Email: Yuri.Motorin@univ-lorraine.fr

Keywords: deep sequencing, RNA modification, pseudouridylation, epitranscriptome, regulation

Abstract

Detection of RNA modified nucleotides using deep sequencing can be performed by several approaches, including antibody-driven enrichment and natural or chemically induced RT signatures. However, only very few RNA modified nucleotides generate natural RT signatures and antibody-driven enrichment heavily depends on the quality of antibodies used and may be highly biased. Thus, the use of chemically-induced RT signatures is now considered as the most trusted experimental approach. In addition, the use of chemical reagents allows inclusion of simple “mock-treated” controls, to exclude spontaneous RT arrests, SNPs and other misincorporation-prone sites. Hydrazine is a well-known RNA-specific reagent, already extensively used in the past for RNA sequencing and structural probing. Hydrazine is highly reactive to U and shows low reaction rates with ψ residues, allowing their distinction by deep sequencing-based protocols. However, other modified RNA residues also show particular behavior upon hydrazine treatment. Here we present methodological developments allowing to use HydraPsiSeq for precise quantification of RNA pseudouridylation and also detection and quantification of some other RNA modifications, in addition to ψ residues.

1 Introduction

Current methods of high-throughput RNA epitranscriptomic analysis (that is, detection/mapping and quantification of RNA modifications) are based on three main principles: antibody-driven enrichment of modified RNA fragments [1–3], analysis of natural RT signatures in RNA by second or third generation of cDNA/RNA sequencing [4–8] and the use of specific chemical reagents allowing the revelation of RT-silent modification by so-called "enhanced" RT signatures [9–14]. All these three analytical areas are dynamically developing at the moment, featuring better characterization of available RNA modification-specific antibodies [15–19], and introduction of new ones [20], with applications of ultra-deep sequencing to uncover even barely discernable RT signatures [16,21–24] or the use of direct RNA sequencing by nanopores to distinguish RNA modifications from their parental counterparts [25–28]. Major efforts in the field were also placed on the use of specific chemistry and specific reagent allowing to distinguish RNA modifications by their intrinsic reactivity [9–11,13,29–35]. Such methods are particularly successful, since they allow single nucleotide resolution, as well as quantification of RNA modification level using appropriate calibration, and also since they provide easy inclusion of internal controls (e.g. untreated sample), which, in turn, allow to drastically reduce false-positive hits [36,37]. Thus, such approaches were developed and applied for mapping of Nm residues (RiboMethSeq and Nm-Seq/RibOxi-Seq) [13,14,38–40], pseudouridines (CMCT and its "clickable" derivative) [9,10,41,42], ac^4C by reduction with $NaCNBH_3$ [35,43], m^7G by reduction with $NaBH_4$ followed or not by alkaline hydrolysis [11,12,34,44,45], m^3C by hydrazine treatment and subsequent cleavage [46], or NO^+-catalyzed deamination as an alternative for m^6A detection/quantification [33]. While not exempted from false positive/negative identifications [36,37], these methods generally show superior performance in terms of specificity and selectivity, in comparison with other approaches.

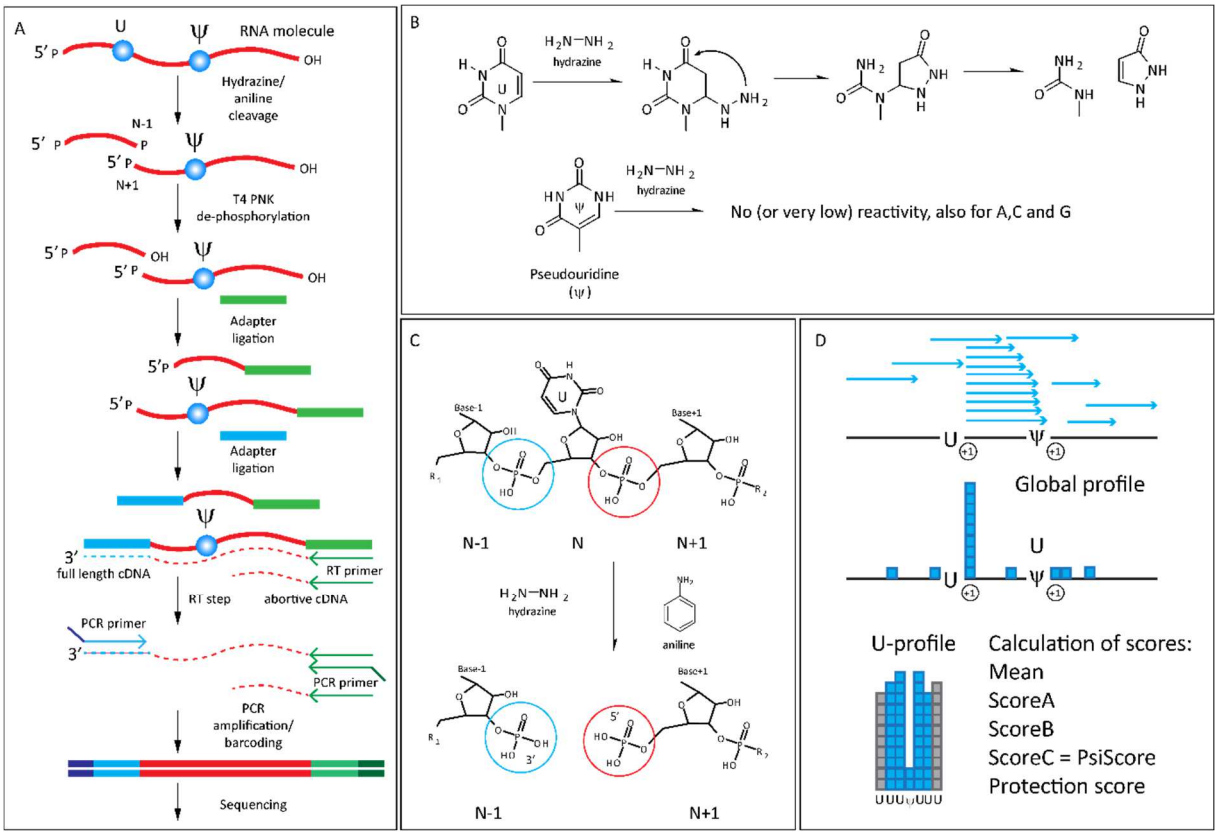
Hydrazine (NH_2-NH_2) is a well-known reagent in RNA chemistry. It readily reacts with U residues in RNA, and, under high salt conditions, with C, leading to the ablation of aromaticity and destruction of the pyrimidine ring [47,48]. Subsequent treatment with aniline at mild acidic pH leads to the cleavage of the glycosidic bond, resulting in an RNA abasic site. This, in turn promotes cleavage of the adjacent phosphodiester bond. These reactions have been routinely used in the past for direct chemical RNA sequencing, whereby analysis was performed using end-labeled RNA for detection of the cleaved fragment(s) [47]. Analysis of naturally modified RNAs by hydrazine cleavage pointed out the resistance of ψ residues to hydrazine cleavage, and this feature was even explored for ψ mapping (e.g. in rRNAs and UsnRNAs) [49–51], but such "negative" detection was supplanted by more specific ψ mapping by CMCT derivatization (reviewed in [30]). Combined with dimethylsulfate (DMS) treatment, creating chemically-induced m^1A, m^3C and m^7G modifications, hydrazine/aniline cleavage

at m^3C residues was also used for RNA structural probing [48]. These results clearly demonstrated that differential chemical reactivity (both increased or reduced) towards hydrazine can be used for mapping (and quantification) of RNA modifications.

Modern deep sequencing-based applications of hydrazine treatment coupled to aniline cleavage of RNA abasic sites were recently proposed for mapping of rare m^3C residues in eukaryotic tRNAs (HAC-Seq, [46]) and for mapping and quantification of pseudouridine (ψ) residues by HydraPsiSeq [52]. Despite apparent similarity of these two protocols, HAC-Seq is designed for positive detection of the signal resulting from specific m^3C cleavage induced by hydrazine. In contrast, under HydraPsiSeq cleavage conditions, hydrazine introduces random cleavages at all accessible U residues, while ψ remains resistant (Figure 1). This implies totally different data extraction and analysis pipelines. In HydraPsiSeq, remnants of U residues react first with the aniline, followed by β- (and subsequent δ-elimination) reactions at the opened ribose cycle, producing 3'-phosphodiester bond cleavage in RNA. The resulting 5'-phosphate is thus used for specific ligation of the adapter, providing high specificity of detection. The resulting amplified library is sequenced in a single-end mode, allowing mapping of reads' 5'-extremities, mapped to N+1 nucleotide in RNA sequence.

In contract to Us, RNA residues C, A and G, as well as some modified Us (ψ and others, discussed below) are unreactive towards hydrazine and therefore do not lead to a signal. Certain modified uridines can hence be detected based on this protection. On the other hand, some modified residues derived from otherwise non-reactive C, A and G may also show increased hydrazine sensitivity, and thus their detection is based on unexpectedly high signals appearing at C/A or G.

In this work we describe applications of HydraPsiSeq protocol for mapping and quantification of ψ in rRNAs and tRNAs, their quantification, and also mapping of other hydrazine/aniline sensitive and insensitive RNA modified residues (namely m^3C, m^7G, k^2C, m^5U).



Marchand *et al* Figure 1

Figure 1 Overview of HydraPsiSeq protocol and essential steps of the hydrazine/aniline cleavage and library preparation by adapter ligation. A – Hydrazine reacts with unmodified U (and rare other modified residues, see text), creating RNA abasic sites. Those are converted to cleavage of the RNA phosphodiester bonds by aniline-driven elimination. Resulting 5'-P are used for adapter ligation. De-phosphorylation by PNK (no ATP conditions) allows selective removal of the 3'-P moieties. Conversion of RNA to cDNA is done after ligation of both adapters to RNA fragments. Amplification and simultaneous barcoding allow to obtain sequencing library. B – Chemical reaction between unmodified U and hydrazine leading to the opening of the pyrimidine cycle. Pseudouridine (ψ) is resistant to hydrazine. C – Positions of 3'-end (nucleotide N-1) and 5'-end (nucleotide N+1) phosphate residues after elimination of RNA abasic site by aniline (nucleotide N). D – Schematic representation of the alignment results and global and U-profiles used for calculations of HydraPsiSeq scores. Scores used in HydraPsiSeq are derived from the corresponding RiboMethSeq counterparts.

2 Materials and Methods

2.1 RNA extraction

Total RNA from bacteria or eukaryotic cells was isolated either using TRIzol Reagent (Invitrogen, ThermoFisher Scientific, ref #15596026) according to the manufacturer's protocol or using hot acid phenol extraction [53,54]. If a particular RNA of interest needs to be analyzed, an enrichment step may be necessary (e.g. ribodepletion of rRNAs, tRNA enrichment using columns [55], antibody enrichment, *etc*).

2.2 RNA quality control and quantification

Quantity and quality of the RNA preparation was assessed using Nanodrop One (or equivalent) and capillary electrophoresis using a PicoRNA chip on Bioanalyzer 2100 (Agilent Technologies, USA).

2.3 Overview and detailed description of HydraPsiSeq protocol

2.3.1 Hydrazine treatment

RNA solution (about 200 ng of RNA) was mixed with pure Hydrazine (Sigma Aldrich ref #207942-5G) at 50% final concentration and incubated for 30-60 min on ice. The reaction was stopped by ethanol precipitation using 0.3M NaOAc pH5.2 and glycoblue and incubated at -80°C for at least 30 min. After centrifugation, the RNA pellet was washed twice with 80% Ethanol and dried for 2 min at 37°C.

2.3.2 Aniline cleavage

The pellet was resuspended in 20 µL of 1 M aniline pH4.5 (Sigma Aldrich ref #242284-5 mL) and incubated in the dark for 15 min at 60°C. The reaction was stopped by ethanol precipitation using 0.3M NaOAc pH5.2 and glycoblue and incubated at -80°C for at least 2h. After centrifugation, the pellet was washed twice with 80% Ethanol and dried for 2 min at 37°C.

2.3.3 De-phosphorylation by T4 PNK

RNA fragments were dephosphorylated at the 3'-ends using 10 U of T4 PNK (New England Biolabs, USA, ref #M0201L) in 50 µL reaction buffer containing 100 mM Tris-HCl pH6.5, 100 mM MgOAc and 5 mM β-mercaptoethanol and incubated for 6h at 37°C. T4 PNK was inactivated by an incubation for 20 min at 65°C. RNA was recovered using a phenol:chloroform:isoamyl alcohol (25:24:1) extraction followed by Ethanol precipitation using 0.3M NaOAc pH5.2 and glycoblue and incubated overnight at -80°C.

2.3.4 Library preparation

After centrifugation, the RNA pellet was washed twice with 80% Ethanol, dried for 2 min at 37°C and resuspended in RNase free water and converted to library using NEBNext small RNA Library kit (New

England Biolabs, USA, ref #E7730L) following the manufacturer's instructions. After PCR, the reaction was cleaned up using silica-based membrane spin columns (GeneJET PCR purification kit, K0702, Thermofisher Scientific).

Experimental design of HydraPsiSeq analysis does not include removal of potential PCR duplicates neither at the primer design level (UMIs) nor during subsequent bioinformatic treatment (e.g. during trimming). First, the amount of input RNA used in the protocol is compatible with a PCR amplification step limited to 12-15 cycles only, thus reducing potential over-amplification of unique RT products and PCR duplicates. Secondly, considering the size of the reference sequence (rRNA or collection of tRNA sequences, <10 kb used in this study) and the number of sequencing reads used for analysis (>20 mln), it is fully anticipated that some of products/sequencing reads have exactly the same sequence, and thus should not be artificially removed.

NEBNext small RNA Library kit (as well as all other similarly designed protocols dedicated to small RNA/miRNA analysis) includes RT step after two subsequent adapter ligation reactions, and thus abortive cDNAs resulting from RT-arresting RNA modifications are not captured in the final library and not sequenced. This certainly creates a bias in regions where such RT-arresting nucleotides (e.g. m^1A, m^1G, $m^{22}G$, etc) are present. However, this bias is not visible here as a sharp signal, but rather as a progressive reduction of sequencing coverage upstream of the modification. This also explains rather variable coverage in different RNA domains (particularly visible for tRNAs).

Libraries were loaded on Agilent High Sensitivity DNA chip (for Agilent Bioanalyzer 2100) after appropriate dilution to concentration in the range of 5–500 pg/µL to check for the quality of the library and the eventual presence of adapter dimers. The mean size of each library was calculated using integrated software. Each library was carefully quantified by fluorometry (Qubit 3.0 fluorometer, Invitrogen, USA) using Qubit dsDNA HS assay kit (Thermo Fischer Scientific, ref #Q32851).

2.3.5 Sequencing

Libraries were multiplexed and subjected to sequencing using an Illumina HiSeq1000 or NextSeq2000 instruments with 50 bp single-read runs. Libraries were loaded at a final concentration of 10-12 pM and sequenced at a depth of 20 million reads per sample (for typical analysis of rRNA pseudouridylation).

2.3.6 Trimming

Adapter removal was performed using the Trimmomatic utility v39.0 [56], with a stringency parameter of 7, very short reads <8 nt were excluded.

2.3.7 Alignment

The alignment of raw reads was conducted by Bowtie2 [57] in end-to-end mode. Only uniquely

mapped reads were retained for further analysis and calculation of protection scores.

2.3.8 Score calculations

Mapped and sorted *.bam file was transformed to *.bed format. Locations of 5'-extremities of mapped unique reads were counted from *.bed file, giving raw cleavage profile. Reads' 5'-end counts were normalized to local background in rolling window of 10 nucleotides, values for U residues were excluded to calculate the median. If RNA reference contains multiple U-rich stretches >10 in length, normalization window can be extended to 16 nt or more. Locally normalized U profiles (NormUcount, full profiles shown in figures) were used as protection score values and further transformed to U only cleavage profiles, by dropping values for other nucleotides. Resulting U profiles were used for calculation of 'RiboMethSeq-like' scores (ScoreMEAN, A, B and PsiScore, equivalent to ScoreC in RiboMethSeq). Window of four neighboring nucleotides (±2 nt) was used for score calculation [58]. ScoreMEAN for each position is calculated in two steps, as follows: first, a ratio of number of cumulated 5'/3'-reads ends between preceding and following position is defined and, second, ScoreMEAN is calculated as a ratio of a drop for a given position compared to the average and variation for 4 neighboring positions (−2/+2).

Conversion of NormUcount profiles into actual modification ratio at a given site is done by measurement of the relative "gap" depth corresponding to modified uridine in the NormUcount profile. Since the reactivity of U residues is not uniform and the actual number of sequencing reads for the surrounding nucleotides varies due to the random nature of deep sequencing and multiple biases (see also below), the "gap" depth for a given nucleotide is compared with the variability observed for 4 neighboring nucleotides. PsiScore(ScoreC2 in RiboMethSeq), is thus calculated using the following formula: RiboMethScore = $1 - n_i/(0.5 \times (\mathrm{SUM}(n_j \times W_j)/\mathrm{SUM}(W_j) + \mathrm{SUM}(n_k \times W_k)/\mathrm{SUM}(W_k))$, where n_i is cumulated 5'/3'-end count for a given position, j varies from $i - 2$ to $i - 1$, k varies from $i + 1$ to $i + 2$, Weight parameters are defined as 1.0 for −1 /+1 and 0.9 for −2/+2 positions. ScoreA and ScoreB were previously reported [14].

2.3.9 Data Analysis

All further calculations were performed in R environment ver 4.0.3.

3 Results

3.1 Applications of HydraPsiSeq protocol

Pseudouridine residues are common in all the most-studied non-coding RNAs (e.g. rRNA, tRNA, snRNA) [59] and were also reported in eukaryotic mRNA [9,10]. Considering the requirement for high sequence coverage of the HydraPsiSeq protocol, this method is mostly appropriate for analysis of highly abundant RNA species, like rRNAs and tRNAs. Lower abundant RNAs can be analyzed, but

should first be specifically enriched either by group-affinity selection (e.g. polyA+-enrichment [52], cap-specific antibody) or by sequence -specific selection of a given RNA using e.g. a capture protocol with a complementary DNA oligonucleotide. A detailed description of these protocols is available elsewhere [60] and out of scope here.

The application of HydraPsiSeq to ψ quantification in rRNA analysis was already described in the past [52], so only tRNA ψ profiling will be discussed here. We illustrate these data with representative profiles obtained for *S. cerevisiae* and *B. subtilis* tRNA, but similar results have been obtained for other model organisms (e.g. *E. coli* and *S. aureus*)

3.2 Profiling of tRNA ψ content in WT *S. cerevisiae* strain and in ΔPUS mutants

Yeast *S. cerevisiae* tRNAs represent an ideal working model for tRNA ψ profiling, since all enzymes acting on cytoplasmic (as well as mitochondrial tRNAs, not considered here due to low content in tRNA fraction) are known and corresponding deleted strains are viable (even ΔPUS8/RIB2 is viable, but requires special growth conditions, [61]) and even commercially available. Here we present the profiling of tRNA ψ residues in a WT yeast strain and four deletion strains (ΔPUS1, ΔPUS3, ΔPUS4 and ΔPUS7), of enzymes that ensure the great majority of cytoplasmic tRNA ψ modifications. Previously reported specificity of Pus1, Pus3, Pus4 and Pus7 enzymes is illustrated in Figure 2A. Representative profiles for the WT strain (hydrazine-protected ψ residues) and deletion strains (hydrazine-cleaved U residue) are shown in the panels of Figure 2B-F. Visual inspection of the latter clearly shows, that the signal for unmodified U at a given typical ψ position is present only in tRNA extracted from a deletion strain lacking the corresponding enzyme. In some cases, a partial cleavage signal is already present in the WT, indicating a substoichiometric modification, which is logically further increased in the KO mutant. Other modified RNA nucleotides (e.g. Cm, m^1G, m^5C) present in $tRNA^{Gly}$(GCC) do not yield a HydraPsiSeq signal. Cleavage at D sites seems to be very variable and, taking into account that D tRNA modification was shown to be understoichiometric in many instances [11], likely depends on the modification content at a given site.

Similar data were also obtained for tRNAs from WT and ΔyjbO *B. subtilis* strains. The recently characterized *B. subtilis* tRNA:ψ-synthase YjbO is a functional homolog of RluA and catalyzes ψ31/ψ32 formation in several tRNAs [62]. Two representative profiles for WT (panels B,D) and ΔyjbO strains (panels C,E) are shown (Figure 3). The signal for unmodified U at position 31 appears only in the deleted strain.

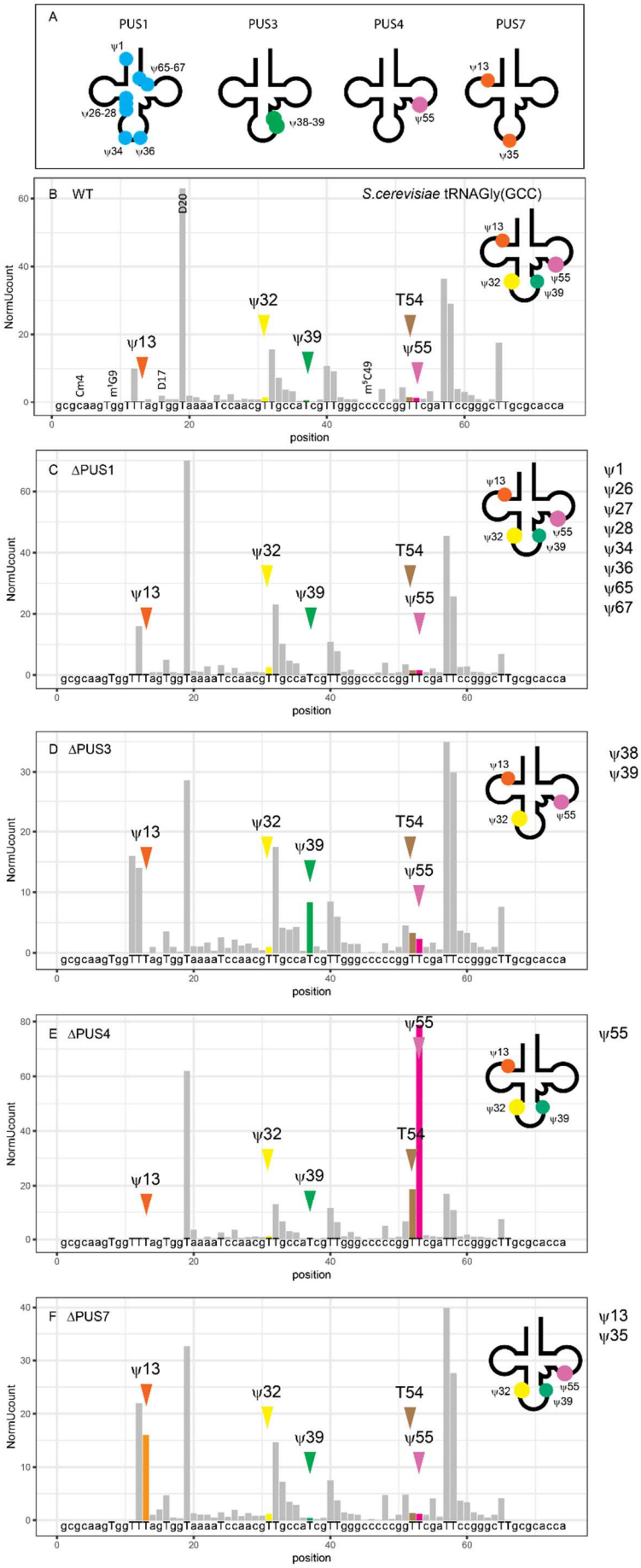
3.3 Detection of non-ψ residues in HydraPsiSeq pipeline

Initially, the HydraPsiSeq protocol was designed for mapping and quantification of pseudouridine residues in RNAs. However, the combination of hydrazine treatment and aniline cleavage for

revelation of the resulting abasic sites turned up signals for some additional modified RNA residues, apparently caused by their peculiar resistance to the hydrazine cleavage (for modified U residues) or as particular sensitivity to the treatment, for derivative of A, G or C nucleotides.

3.3.1 Modified U residues showing high resistance to cleavage

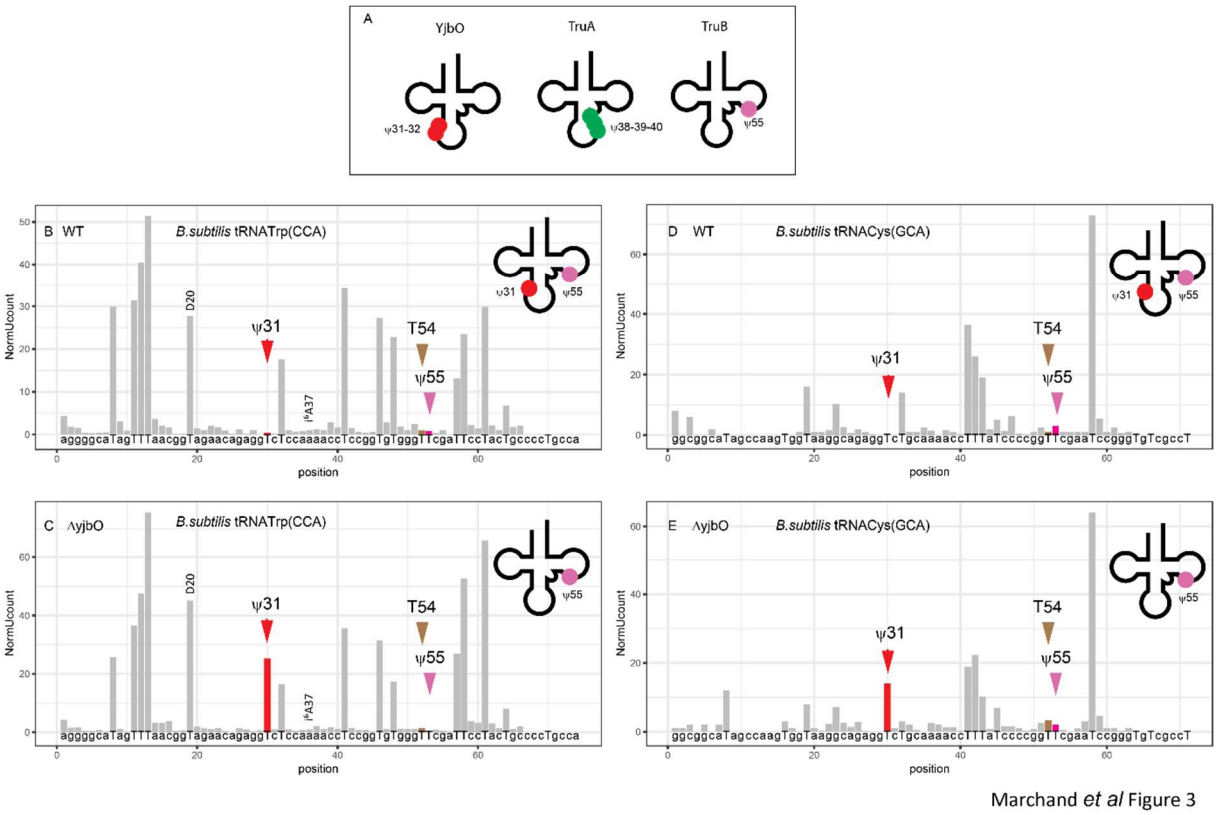
The chemistry of RNA cleavage by hydrazine at different conditions was previously explored in a number of studies [49–51]. In addition to specific cleavages at U residues, hydrazine treatment can be specifically re-directed to C by modulation of the buffer ionic strength (3M NaCl) [46,47]. However, the nucleophilic addition to the C5=C6 double bond of the pyrimidine ring may be dampened or impeded by many of the known substituents occurring at position 5, thus many modified derivatives of the U residues may be also resistant to cleavage. Inspection of the available tRNA HydraPsiSeq profiles for *S. cerevisiae* and *B. subtilis* showed that this was the case for the highly conserved ribothymidine (rT or m^5U) at position 54 in tRNAs (Figure 2 and Figure 3).



Marchand *et al* Figure 2

Figure 2 Analysis of tRNA pseudouridylation profiles in WT and four *S. cerevisiae* deletion strains. A – Known specificity of four major cytoplasmic RNA:ψ-synthases encoded by genes PUS1, PUS3, PUSS4 and PUS7 in *S. cerevisiae*. B – Global cleavage profile of $tRNA^{Gly}$(GCC) in total RNA fraction extracted from WT yeast strain. Positions of signals corresponding to 4 ψ residues (ψ13, ψ32, ψ39, ψ55) as well as protected T54 are shown by color arrows. Positions of other HydraPsiSeq-silent RNA modified nucleotides (Cm, m^1G, D, m^5C) are shown in the panel B. Panels C – F show HydraPsiSeq cleavage profiles for the same tRNA from ΔPUS1, ΔPUS3, ΔPUS4 and ΔPUS7 strains missing the corresponding enzymatic activity. Positions, modified by these enzymes are given at the right. Notice that the loss of ψ55 upon PUS4 deletion (E) likely results in partial loss of T54 methylation.

The HydraPsiSeq signal observed for T54 is at the background level for the majority of explored tRNAs, and only in rare cases cleavage was observed at T54, most probably reflecting incomplete methylation (see Figure 2E and Figure 4D shown in brown). Along the same line, modified U*34 variants also show very low cleavage, allowing detection and possibly even quantification of those residues, similar to pseudouridines (typical profiles are shown in Figure 4AC, shown in blue).



Marchand *et al* Figure 3

Figure 3 Analysis of *B. subtilis* tRNA pseudouridylation in the WT and ΔyjbO strains. Specificity of *B. subtilis* RNA:ψ-synthases is shown in (A). Comparative global HydraPsiSeq profiles for two tRNA substrates of YjbO, top panels for the WT strain, bottom for the KO mutant. Positions of ψ residues as well as other RNA modifications are shown. The signal for ψ31 is shown in red.

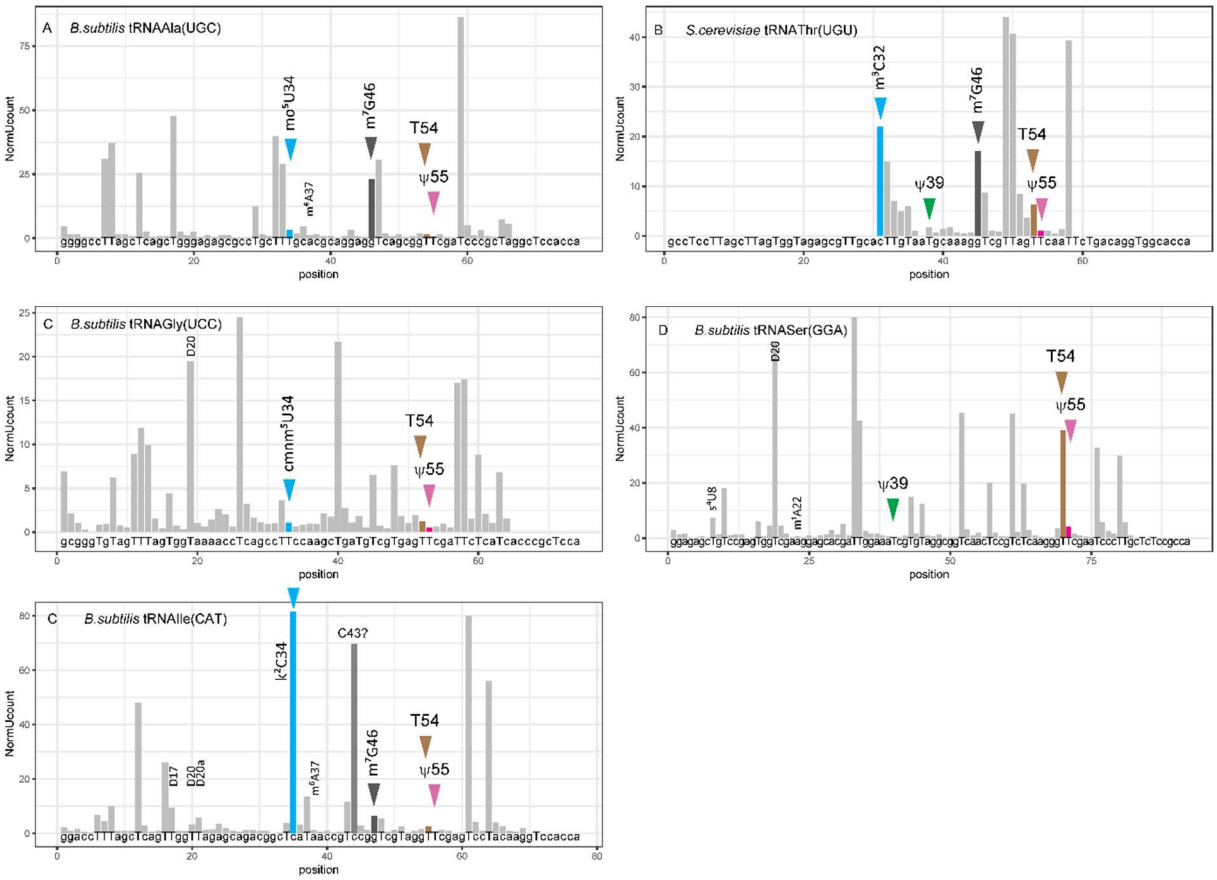
3.3.2 Modified G, C and A residues showing high sensitivity to hydrazine

While unmodified G, C and A residues are fairly resistant to the hydrazine treatment, some of their modified counterparts show high sensitivity towards the HydraPsiSeq protocol. Notably, this was observed for methylated m^7G and m^3C, as well as for lysidine (L or k^2C) nucleotide, derived from parental C.

Methylated 7-methylguanosine (m^7G) is relatively common in rRNA (both 16S and 23S rRNA in bacteria and 18S rRNA in eukaryotes), and also at the position 46 in tRNAs [59]. Specific signals, were readily detected at all positions, where this RNA modification was previously reported (representative profiles are shown in Figure 4AB, shown in dark grey). Plausibly, the N-glycosidic bond within m^7G is unstable under the conditions of the hydrazine reaction, releasing the methylated base, which is an excellent leaving group. This would be followed by formation of an abasic site, which is known to decompose under the reaction conditions of aniline cleavage. Signals observed for m^7G are comparable to the signals of unmodified Us, and thus can be potentially used for analysis of m^7G-modification stoichiometry.

The second prominent signals in HydraPsiSeq come from modified m^3C (see Figure 4B), this modification is frequently present at position 32 in eukaryotic tRNAs (mostly $tRNA^{Ser}$ and $tRNA^{Thr}$). Those signals are not surprising, since hydrazine is known to specifically cleave m^3C [47,48]. This feature of m^3C was also recently explored for specific mapping of m^3C in RNAs by the protocol called HAC-Seq [46].

The last group of well-detectable HydraPsiSeq signals stem from modified k^2C (lysidine) residues found in bacterial $tRNA^{Ile}$(CAT). Indeed, for correct decoding of the rare AUA Ile codon, anticodon of the $tRNA^{Ile}$(CAT) primary transcript is converted to k^2CAU, after this modification $tRNA^{Ile}$(k^2CAU) ensures specific basepairing with AUA [63–65]. We observed rather prominent cleavage signals (see Figure 4E, shown in blue) at the pos 34 in $tRNA^{Ile}$(CAT), showing that k^2C is also sensitive under HydraPsiSeq treatment conditions. Hydrazine pre-treatment is essential, since k^2C does not show up in AlkAnilineSeq protocol [11], which uses the same aniline cleavage as at the second step. One can hypothesize that the nucleobase in k^2C is protonated and because of the positive charge more electrophilic. It would hence be more susceptible to an attack of the supernucleophile hydrazine.



Marchand *et al* Figure 4

Figure 4 Detection of non-ψ RNA modifications by protection against hydrazine treatment (T, U*) or by particular sensitivity (m^7G, m^3C, k^2C). Profiles obtained for *B. subtilis* tRNAs are shown as examples. Positions of anomalous cleavage ($cmnm^5U$, k^2C, mo^5U, m^7G) are shown in light blue or dark grey. Notice a low protection of T54 in (D), indication low methylation level under the experimental conditions used.

These data show that some rare non-ψ residues can also be reliably mapped by HydraPsiSeq protocol, but additional validation is required to ascertain the exact chemical nature of such nucleotides. Except m^7G, which is rather wide widespread, m^3C is found only in eukaryotic tRNAs [46] and modified 5-substituted U*, as well as, k^2C are restricted to specific tRNAs at the wobble position. This also opens a possibility for quantification of such residues in a high-throughput manner, in parallel with the pseudouridines, provided that calibration curve between HydraPsiSeq signal and modification content can be obtained. Alternatively, this observed sensitivity may be further explored for more specific protocol dedicated for such tRNA(rRNA) modifications.

3.4 False-positive and false-negative hits in HydraPsiSeq

False-positive/negative hits observed in HydraPsiSeq concern erroneous identification of “protected” U residues as ψ (false-positive hits), or non-detection of partially modified ψ site (false-negative). Both types of errors are common, since, despite almost a complete denaturation of RNA tertiary and

even secondary structure upon hydrazine reaction, some RNA regions may be still low accessible for the reagent or less reactive due to the local environment. This is also visible by very low spontaneous cleavages at G/C/A in these locations. Other factors leading to a non-uniform U-cleavage profile are the random nature of the RNA chemical fragmentation and deep sequencing pipeline, as well as the known ligation and amplification biases during the library preparation step. Moreover, variable sequencing coverage at different RNA regions may stem from the presence of RT-arresting RNA modifications (this is in particular the case for tRNAs, which are rich in such residues). However, since only double-tagged cDNAs (and not the abortive cDNA molecules) are amplified and converted to library using the library kit protocol, these RT-arresting residues do not generate discrete signals, but rather lead to a general decrease of coverage upstream of such RNA modifications. Altogether, these lead to an unequal coverage and the absolute values of NormUcount may vary substantially along the RNA molecule, even for unmodified U residues.

Other types of false-positive signals result from spontaneous loss of G or A bases (de-purination) in RNA. Those abasic sites are subsequently cleaved by aniline in the second step of the protocol and can be also found in mock-treated samples (no hydrazine). Such false-positive signals not corresponding to any RNA modification were found in some tRNA anticodon loops (frequently positions 37-38). It is noteworthy that some tRNA regions may have rather low coverage in HydraPsiSeq and no analysis can be performed in a reliable way for these positions, unless much higher sequencing depth is obtained.

4 Discussion

4.1 Resistance of selected modified U residues to hydrazine cleavage

The HydraPsiSeq protocol is based on specific U cleavage by hydrazine, thus it is not surprising that certain modified U residues show reduced reactivity similar to the one of ψ. Our data show that this is clearly the case for m^5U=rT and some of the 5-substituted U* that typically occur at position 34 in tRNAs. Low reactivity of m^5U was noticed before [49,50], but reduced reactivity of other similarly substituted nucleotides was not reported. This reduced reactivity towards hydrazine is readily rationalized by the increased electronic density of the C5=C6 double bond in 5-substituted pyrimidines, which repels the nucleophilic attack of hydrazine. This is indeed similar to the situation in m^5C, where the increased electron density repels an analogous nucleophilic attack to the C5=C6 double by bisulfite. In contrast, neither thiolated U (s^4U) nor other methylated U (m^3U) are not giving noticeable protection signals in HydraPsiSeq. The protection status of the dihydrouridine (D) residues remains unclear, this modification is frequently sub-stoichiometric in tRNA, and thus residual signal from unmodified U can be observed.

4.2 Cleavage at Lysidine (k^2C), m^3C and m^7G

Non-U RNA nucleotides (A, G and C) are supposed to be resistant to hydrazine under conditions of the protocol. However, some modified counterparts turned out to be sensitive to such treatment. None of the examined *N6*-A modifications (i^6A, ms^2i^6A, m^6A, etc) were detected, but prominent signals were found for two C derivatives (namely m^3C and k^2C) as well as for m^7G. Other C or G-derivatives present in bacterial and yeast tRNAs were not reactive (e.g. m^1G, m^2G, $m^{22}G$, m^5C, etc). Reactivity of m^3C to hydrazine was fully anticipated, given that this modified C was shown to undergo ring opening under standard U-cleavage conditions [47], and this was already reported as a basis for its detection by deep sequencing. Given, that several other methods have already been proposed for m^3C detection, HydraPsiSeq may be either used as opportunity for simultaneous detection of m^3C and pseudouridines, or for relative quantification of the modification level, since other described methods (HAC-Seq and AlkAnilineSeq) are less quantitative [11,12,46]. Similar conclusion can be applied to the quantification of m^7G by HydraPsiSeq, since this RNA modification too, can be detected by a number of approaches. However, mapping and quantification of lysidine (k^2C) is of great interest since there is no other straightforward and high-throughput approach for its mapping and quantification. Mapping of k^2C is likely not the most relevant, since this nucleotide is only found in one tRNA at positions 34, but its quantification opens perspectives for fine analysis of lysidine dynamics under stress conditions, for example.

4.3 Non-reactive modifications

As stated above, all other RNA base and ribose modifications known in bacterial and yeast tRNAs and rRNAs are not hydrazine sensitive and behave as normal A, C, G and U residues. The non-exhaustive list if those is given in Table 1.

Table 1 HydraPsiSeq detectable and silent RNA modifications

Parental nucleotide	HydraPsiSeq detectable (RNA)	HydraPsiSeq silent (RNA)
U	Ψ (tRNA, rRNA) m^5U (tRNA, rRNA) U* (tRNA)	D (tRNA, rRNA) m^3U (rRNA) s^4U (tRNA) Um (tRNA, rRNA)
C	m^3C (tRNA) k^2C (tRNA)	m^5C (tRNA, rRNA) Cm (rRNA) ac^4C (tRNA, rRNA)
G	m^7G (tRNA, rRNA)	m^1G (tRNA, rRNA) m^2G (tRNA, rRNA)

		$m^{22}G$ (tRNA) Gm (tRNA, rRNA) yW (tRNA)
A	none	Am (rRNA) I (tRNA) m^1A (tRNA, rRNA) t^6A (tRNA) i^6A (tRNA)

5 Acknowledgements

This work was supported by ANR PRCI grant “DERASE” (ANR-20-CE92-0030-02) to VM and MH, and Epi-ARN FRCR2018 grant from Grand Est Region, France to YM and DFG grants HE3397/21-1, TRR 319 TP C01, and HE 3397/13-2 in the frame of the SPP1784 to MH. This work was performed in the frame of COST EPITRAN action CA16120.

6 References

[1] D. Dominissini, S. Moshitch-Moshkovitz, M. Salmon-Divon, N. Amariglio, G. Rechavi, Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing, Nat Protoc. 8 (2013) 176–189. https://doi.org/10.1038/nprot.2012.148.

[2] B. Linder, A.V. Grozhik, A.O. Olarerin-George, C. Meydan, C.E. Mason, S.R. Jaffrey, Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome, Nat. Methods. 12 (2015) 767–772. https://doi.org/10.1038/nmeth.3453.

[3] A.V. Grozhik, B. Linder, A.O. Olarerin-George, S.R. Jaffrey, Mapping m6A at Individual-Nucleotide Resolution Using Crosslinking and Immunoprecipitation (miCLIP), Methods Mol. Biol. 1562 (2017) 55–78. https://doi.org/10.1007/978-1-4939-6807-7_5.

[4] R. Hauenschild, L. Tserovski, K. Schmid, K. Thüring, M.-L. Winz, S. Sharma, K.-D. Entian, L. Wacheul, D.L.J. Lafontaine, J. Anderson, J. Alfonzo, A. Hildebrandt, A. Jäschke, Y. Motorin, M. Helm, The reverse transcription signature of N-1-methyladenosine in RNA-Seq is sequence dependent, Nucleic Acids Res. 43 (2015) 9950–9964. https://doi.org/10.1093/nar/gkv895.

[5] S. Werner, L. Schmidt, V. Marchand, T. Kemmer, C. Falschlunger, M.V. Sednev, G. Bec, E. Ennifar, C. Höbartner, R. Micura, Y. Motorin, A. Hildebrandt, M. Helm, Machine learning of reverse transcription signatures of variegated polymerases allows mapping and discrimination of methylated purines in limited transcriptomes, Nucleic Acids Res. 48 (2020) 3734–3746. https://doi.org/10.1093/nar/gkaa113.

[6] Y. Wang, Y. Xiao, S. Dong, Q. Yu, G. Jia, Antibody-free enzyme-assisted chemical approach for detection of N6-methyladenosine, Nat Chem Biol. 16 (2020) 896–903. https://doi.org/10.1038/s41589-020-0525-x.

[7] E.M. Novoa, C.E. Mason, J.S. Mattick, Charting the unknown epitranscriptome, Nat. Rev. Mol. Cell Biol. 18 (2017) 339–340. https://doi.org/10.1038/nrm.2017.49.

[8] L. Xu, M. Seki, Recent advances in the detection of base modifications using the Nanopore sequencer, Journal of Human Genetics. 65 (2020) 25–33. https://doi.org/10.1038/s10038-019-0679-0.
[9] T.M. Carlile, M.F. Rojas-Duran, B. Zinshteyn, H. Shin, K.M. Bartoli, W.V. Gilbert, Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells, Nature. 515 (2014) 143–146. https://doi.org/10.1038/nature13802.
[10] S. Schwartz, D.A. Bernstein, M.R. Mumbach, M. Jovanovic, R.H. Herbst, B.X. León-Ricardo, J.M. Engreitz, M. Guttman, R. Satija, E.S. Lander, G. Fink, A. Regev, Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA, Cell. 159 (2014) 148–162. https://doi.org/10.1016/j.cell.2014.08.028.
[11] V. Marchand, L. Ayadi, F.G.M. Ernst, J. Hertler, V. Bourguignon-Igel, A. Galvanin, A. Kotter, M. Helm, D.L.J. Lafontaine, Y. Motorin, AlkAniline-Seq: Profiling of m7G and m3C RNA Modifications at Single Nucleotide Resolution, Angew. Chem. Int. Ed. Engl. 57 (2018) 16785–16790. https://doi.org/10.1002/anie.201810946.
[12] S. Lin, Q. Liu, Y.-Z. Jiang, R.I. Gregory, Nucleotide resolution profiling of m7G tRNA modification by TRAC-Seq, Nature Protocols. 14 (2019) 3220–3242. https://doi.org/10.1038/s41596-019-0226-7.
[13] V. Marchand, F. Blanloeil-Oillo, M. Helm, Y. Motorin, Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA, Nucleic Acids Res. 44 (2016) e135. https://doi.org/10.1093/nar/gkw547.
[14] U. Birkedal, M. Christensen-Dalsgaard, N. Krogh, R. Sabarinathan, J. Gorodkin, H. Nielsen, Profiling of ribose methylations in RNA by high-throughput sequencing, Angew. Chem. Int. Ed. Engl. 54 (2015) 451–455. https://doi.org/10.1002/anie.201408362.
[15] E. Mishima, D. Jinno, Y. Akiyama, K. Itoh, S. Nankumo, H. Shima, K. Kikuchi, Y. Takeuchi, A. Elkordy, T. Suzuki, K. Niizuma, S. Ito, Y. Tomioka, T. Abe, Immuno-Northern Blotting: Detection of RNA Modifications by Using Antibodies against Modified Nucleosides, PLoS ONE. 10 (2015) e0143756. https://doi.org/10.1371/journal.pone.0143756.
[16] A.V. Grozhik, A.O. Olarerin-George, M. Sindelar, X. Li, S.S. Gross, S.R. Jaffrey, Antibody cross-reactivity accounts for widespread appearance of m1A in 5'UTRs, Nature Communications. 10 (2019) 5126. https://doi.org/10.1038/s41467-019-13146-w.
[17] K. Slama, A. Galliot, F. Weichmann, J. Hertler, R. Feederle, G. Meister, M. Helm, Determination of enrichment factors for modified RNA in MeRIP experiments, Methods (San Diego, Calif.). 156 (2019) 102–109. https://doi.org/10.1016/j.ymeth.2018.10.020.
[18] M. Helm, F. Lyko, Y. Motorin, Limited antibody specificity compromises epitranscriptomic analyses, Nat Commun. 10 (2019) 5669. https://doi.org/10.1038/s41467-019-13684-3.
[19] A.B.R. McIntyre, N.S. Gokhale, L. Cerchietti, S.R. Jaffrey, S.M. Horner, C.E. Mason, Limits in the detection of m6A changes using MeRIP/m6A-seq, Scientific Reports. 10 (2020) 6590. https://doi.org/10.1038/s41598-020-63355-3.
[20] R. Feederle, A. Schepers, Antibodies specific for nucleic acid modifications, RNA Biology. 14 (2017) 1089–1098. https://doi.org/10.1080/15476286.2017.1295905.
[21] A.V. Grozhik, S.R. Jaffrey, Epitranscriptomics: Shrinking maps of RNA modifications, Nature. 551 (2017) 174–176. https://doi.org/10.1038/nature24156.
[22] B. Linder, S.R. Jaffrey, Discovering and Mapping the Modified Nucleotides That Comprise the Epitranscriptome of mRNA, Cold Spring Harb Perspect Biol. 11 (2019). https://doi.org/10.1101/cshperspect.a032201.
[23] M. Safra, A. Sas-Chen, R. Nir, R. Winkler, A. Nachshon, D. Bar-Yaacov, M. Erlacher, W. Rossmanith, N. Stern-Ginossar, S. Schwartz, The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution, Nature. 551 (2017) 251–255. https://doi.org/10.1038/nature24456.
[24] S. Schwartz, m1A within cytoplasmic mRNAs at single nucleotide resolution: a reconciled transcriptome-wide map, RNA. 24 (2018) 1427–1436. https://doi.org/10.1261/rna.067348.118.
[25] A.M. Smith, M. Jain, L. Mulroney, D.R. Garalde, M. Akeson, Reading canonical and modified

nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing, PloS One. 14 (2019) e0216709. https://doi.org/10.1371/journal.pone.0216709.
[26] L. Zhao, H. Zhang, M.V. Kohnen, K.V.S.K. Prasad, L. Gu, A.S.N. Reddy, Analysis of Transcriptome and Epitranscriptome in Plants Using PacBio Iso-Seq and Nanopore-Based Direct RNA Sequencing, Front Genet. 10 (2019) 253. https://doi.org/10.3389/fgene.2019.00253.
[27] P. Jenjaroenpun, T. Wongsurawat, T.D. Wadley, T.M. Wassenaar, J. Liu, Q. Dai, V. Wanchai, N.S. Akel, A. Jamshidi-Parsian, A.T. Franco, G. Boysen, M.L. Jennings, D.W. Ussery, C. He, I. Nookaew, Decoding the epitranscriptional landscape from native RNA sequences, Nucleic Acids Research. (2020). https://doi.org/10.1093/nar/gkaa620.
[28] M.T. Parker, K. Knop, A.V. Sherwood, N.J. Schurch, K. Mackinnon, P.D. Gould, A.J. Hall, G.J. Barton, G.G. Simpson, Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m6A modification, Elife. 9 (2020). https://doi.org/10.7554/eLife.49658.
[29] S. Schwartz, Y. Motorin, Next-generation sequencing technologies for detection of modified nucleotides in RNAs, RNA Biol. 14 (2017) 1124–1137. https://doi.org/10.1080/15476286.2016.1251543.
[30] I. Behm-Ansmant, M. Helm, Y. Motorin, Use of specific chemical reagents for detection of modified nucleotides in RNA, J Nucleic Acids. 2011 (2011) 408053. https://doi.org/10.4061/2011/408053.
[31] Y. Motorin, V. Marchand, Detection and Analysis of RNA Ribose 2'-O-Methylations: Challenges and Solutions, Genes (Basel). 9 (2018). https://doi.org/10.3390/genes9120642.
[32] Y. Motorin, M. Helm, Methods for RNA Modification Mapping Using Deep Sequencing: Established and New Emerging Technologies, Genes (Basel). 10 (2019). https://doi.org/10.3390/genes10010035.
[33] S. Werner, A. Galliot, F. Pichot, T. Kemmer, V. Marchand, M.V. Sednev, T. Lence, J.-Y. Roignant, J. König, C. Höbartner, Y. Motorin, A. Hildebrandt, M. Helm, NOseq: amplicon sequencing evaluation method for RNA m6A sites after chemical deamination, Nucleic Acids Res. (2020). https://doi.org/10.1093/nar/gkaa1173.
[34] S. Lin, Q. Liu, V.S. Lelyveld, J. Choe, J.W. Szostak, R.I. Gregory, Mettl1/Wdr4-Mediated m7G tRNA Methylome Is Required for Normal mRNA Translation and Embryonic Stem Cell Self-Renewal and Differentiation, Mol. Cell. 71 (2018) 244-255.e5. https://doi.org/10.1016/j.molcel.2018.06.001.
[35] A. Sas-Chen, J.M. Thomas, D. Matzov, M. Taoka, K.D. Nance, R. Nir, K.M. Bryson, R. Shachar, G.L.S. Liman, B.W. Burkhart, S.T. Gamage, Y. Nobe, C.A. Briney, M.J. Levy, R.T. Fuchs, G.B. Robb, J. Hartmann, S. Sharma, Q. Lin, L. Florens, M.P. Washburn, T. Isobe, T.J. Santangelo, M. Shalev-Benami, J.L. Meier, S. Schwartz, Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping, Nature. 583 (2020) 638–643. https://doi.org/10.1038/s41586-020-2418-2.
[36] D. Wiener, S. Schwartz, The epitranscriptome beyond m6A, Nat Rev Genet. (2020). https://doi.org/10.1038/s41576-020-00295-8.
[37] A.V. Grozhik, S.R. Jaffrey, Distinguishing RNA modifications from noise in epitranscriptome maps, Nat. Chem. Biol. 14 (2018) 215–225. https://doi.org/10.1038/nchembio.2546.
[38] Q. Dai, S. Moshitch-Moshkovitz, D. Han, N. Kol, N. Amariglio, G. Rechavi, D. Dominissini, C. He, Nm-seq maps 2'-O-methylation sites in human mRNA with base precision, Nat. Methods. 14 (2017) 695–698. https://doi.org/10.1038/nmeth.4294.
[39] Y. Zhu, S.P. Pirnie, G.G. Carmichael, High-throughput and site-specific identification of 2'-O-methylation sites using ribose oxidation sequencing (RibOxi-seq), RNA. 23 (2017) 1303–1314. https://doi.org/10.1261/rna.061549.117.
[40] K.S. Rajan, Y. Zhu, K. Adler, T. Doniger, S. Cohen-Chalamish, A. Srivastava, M. Shalev-Benami, D. Matzov, R. Unger, C. Tschudi, A. Günzl, G.G. Carmichael, S. Michaeli, The large repertoire of 2'-O-methylation guided by C/D snoRNAs on Trypanosoma brucei rRNA, RNA Biology. 17 (2020) 1018–1039. https://doi.org/10.1080/15476286.2020.1750842.
[41] A.F. Lovejoy, D.P. Riordan, P.O. Brown, Transcriptome-wide mapping of pseudouridines:

pseudouridine synthases modify specific mRNAs in S. cerevisiae, PLoS ONE. 9 (2014) e110799. https://doi.org/10.1371/journal.pone.0110799.
[42] X. Li, P. Zhu, S. Ma, J. Song, J. Bai, F. Sun, C. Yi, Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome, Nat. Chem. Biol. 11 (2015) 592–597. https://doi.org/10.1038/nchembio.1836.
[43] J.M. Thomas, C.A. Briney, K.D. Nance, J.E. Lopez, A.L. Thorpe, S.D. Fox, M.-L. Bortolin-Cavaille, A. Sas-Chen, D. Arango, S. Oberdoerffer, J. Cavaille, T. Andresson, J.L. Meier, A Chemical Signature for Cytidine Acetylation in RNA, J. Am. Chem. Soc. 140 (2018) 12667–12670. https://doi.org/10.1021/jacs.8b06636.
[44] C. Enroth, L.D. Poulsen, S. Iversen, F. Kirpekar, A. Albrechtsen, J. Vinther, Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing, Nucleic Acids Research. 47 (2019) e126. https://doi.org/10.1093/nar/gkz736.
[45] L. Pandolfini, I. Barbieri, A.J. Bannister, A. Hendrick, B. Andrews, N. Webster, P. Murat, P. Mach, R. Brandi, S.C. Robson, V. Migliori, A. Alendar, M. d'Onofrio, S. Balasubramanian, T. Kouzarides, METTL1 Promotes let-7 MicroRNA Processing via m7G Methylation, Molecular Cell. 74 (2019) 1278-1290.e9. https://doi.org/10.1016/j.molcel.2019.03.040.
[46] J. Cui, Q. Liu, E. Sendinc, Y. Shi, R.I. Gregory, Nucleotide resolution profiling of m3C RNA modification by HAC-seq, Nucleic Acids Res. (2020). https://doi.org/10.1093/nar/gkaa1186.
[47] D.A. Peattie, Direct chemical method for sequencing RNA, Proc. Natl. Acad. Sci. U.S.A. 76 (1979) 1760–1764. https://doi.org/10.1073/pnas.76.4.1760.
[48] D.A. Peattie, W. Gilbert, Chemical probes for higher-order structure in RNA, Proc Natl Acad Sci U S A. 77 (1980) 4679–4682. https://doi.org/10.1073/pnas.77.8.4679.
[49] A. Bakin, J. Ofengand, Four newly located pseudouridylate residues in Escherichia coli 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique, Biochemistry. 32 (1993) 9754–9762.
[50] A. Bakin, B.G. Lane, J. Ofengand, Clustering of pseudouridine residues around the peptidyltransferase center of yeast cytoplasmic and mitochondrial ribosomes, Biochemistry. 33 (1994) 13475–13483.
[51] S. Massenet, Y. Motorin, D.L. Lafontaine, E.C. Hurt, H. Grosjean, C. Branlant, Pseudouridine mapping in the Saccharomyces cerevisiae spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA, Mol. Cell. Biol. 19 (1999) 2142–2154.
[52] V. Marchand, F. Pichot, P. Neybecker, L. Ayadi, V. Bourguignon-Igel, L. Wacheul, D.L.J. Lafontaine, A. Pinzano, M. Helm, Y. Motorin, HydraPsiSeq: a method for systematic and quantitative mapping of pseudouridines in RNA, Nucleic Acids Research. (2020). https://doi.org/10.1093/nar/gkaa769.
[53] M.E. Schmitt, T.A. Brown, B.L. Trumpower, A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae., Nucleic Acids Res. 18 (1990) 3091–3092.
[54] M.A. Collart, S. Oliviero, Preparation of yeast RNA, Curr Protoc Mol Biol. Chapter 13 (2001) Unit13.12. https://doi.org/10.1002/0471142727.mb1312s23.
[55] Genomic DNA and RNA purification. NucleoBond® RNA / DNA, Macherey-Nagel User manual, rev 09 (2016).
[56] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics. 30 (2014) 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
[57] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods. 9 (2012) 357–359. https://doi.org/10.1038/nmeth.1923.
[58] F. Pichot, V. Marchand, L. Ayadi, V. Bourguignon-Igel, M. Helm, Y. Motorin, Holistic Optimization of Bioinformatic Analysis Pipeline for Detection and Quantification of 2'-O-Methylations in RNA by RiboMethSeq, Front Genet. 11 (2020) 38. https://doi.org/10.3389/fgene.2020.00038.
[59] P. Boccaletto, M.A. Machnicka, E. Purta, P. Piatkowski, B. Baginski, T.K. Wirecki, V. de Crécy-Lagard, R. Ross, P.A. Limbach, A. Kotter, M. Helm, J.M. Bujnicki, MODOMICS: a database of RNA modification pathways. 2017 update, Nucleic Acids Res. 46 (2018) D303–D307.

https://doi.org/10.1093/nar/gkx1030.
[60] F. Kanwal, C. Lu, A review on native and denaturing purification methods for non-coding RNA (ncRNA), J Chromatogr B Analyt Technol Biomed Life Sci. 1120 (2019) 71–79. https://doi.org/10.1016/j.jchromb.2019.04.034.
[61] I. Behm-Ansmant, H. Grosjean, S. Massenet, Y. Motorin, C. Branlant, Pseudouridylation at position 32 of mitochondrial and cytoplasmic tRNAs requires two distinct enzymes in Saccharomyces cerevisiae, J. Biol. Chem. 279 (2004) 52998–53006. https://doi.org/10.1074/jbc.M409581200.
[62] V. de Crécy-Lagard, R.L. Ross, M. Jaroch, V. Marchand, C. Eisenhart, D. Brégeon, Y. Motorin, P.A. Limbach, Survey and Validation of tRNA Modifications and Their Corresponding Genes in Bacillus subtilis sp Subtilis Strain 168, Biomolecules. 10 (2020) E977. https://doi.org/10.3390/biom10070977.
[63] T. Suzuki, K. Miyauchi, Discovery and characterization of tRNAIle lysidine synthetase (TilS), FEBS Lett. 584 (2010) 272–277. https://doi.org/10.1016/j.febslet.2009.11.085.
[64] H. Grosjean, G.R. Björk, Enzymatic conversion of cytidine to lysidine in anticodon of bacterial isoleucyl-tRNA--an alternative way of RNA editing, Trends Biochem Sci. 29 (2004) 165–168. https://doi.org/10.1016/j.tibs.2004.02.009.
[65] T. Numata, Mechanisms of the tRNA wobble cytidine modification essential for AUA codon decoding in prokaryotes, Biosci Biotechnol Biochem. 79 (2015) 347–353. https://doi.org/10.1080/09168451.2014.975185.

Figure legends

Figure 1 Overview of HydraPsiSeq protocol and essential steps of the hydrazine/aniline cleavage and library preparation by adapter ligation. A – Hydrazine reacts with unmodified U (and rare other modified residues, see text), creating RNA abasic sites. Those are converted to cleavage of the RNA phosphodiester bonds by aniline-driven elimination. Resulting 5'-P are used for adapter ligation. De-phosphorylation by PNK (no ATP conditions) allows selective removal of the 3'-P moieties. B – Chemical reaction between unmodified U and hydrazine leading to the opening of the pyrimidine cycle. Pseudouridine (ψ) is resistant to hydrazine. C – Positions of 3'-end (nucleotide N-1) and 5'-end (nucleotide N+1) phosphate residues after elimination of RNA abasic site by aniline (nucleotide N). D – Schematic representation of the alignment results and global and U-profiles used for calculations of HydraPsiSeq scores. Scores used in HydraPsiSeq are derived from the corresponding RiboMethSeq counterparts.

Figure 2 Analysis of tRNA pseudouridylation profiles in WT and four *S. cerevisiae* deletion strains. A – Known specificity of four major cytoplasmic RNA:ψ-synthases encoded by genes PUS1, PUS3, PUSS4 and PUS7 in *S. cerevisiae*. B – Global cleavage profile of $tRNA^{Gly}$(GCC) in total RNA fraction extracted from WT yeast strain. Positions of signals corresponding to 4 ψ residues (ψ13, ψ32, ψ39, ψ55) as well as protected T54 are shown by color arrows. Positions of other HydraPsiSeq-silent RNA modified nucleotides (Cm, m^1G, D, m^5C) are shown in the panel B. Panels C – F show HydraPsiSeq cleavage profiles for the same tRNA from ΔPUS1, ΔPUS3, ΔPUS4 and ΔPUS7 strains missing the corresponding enzymatic activity. Positions, modified by these enzymes are given at the right. Notice that the loss of ψ55 upon PUS4 deletion (E) likely results in partial loss of T54 methylation.

Figure 3 Analysis of *B. subtilis* tRNA pseudouridylation in the WT and ΔyjbO strains. Specificity of *B. subtilis* RNA:ψ-synthases is shown in (A). Comparative global HydraPsiSeq profiles for two tRNA substrates of YjbO, top panels for the WT strain, bottom for the KO mutant. Positions of ψ residues as well as other RNA modifications are shown. The signal for ψ31 is shown in red.

Figure 4 Detection of non-ψ RNA modifications by protection against hydrazine treatment (T, U*) or by particular sensitivity (m^7G, m^3C, k^2C). Profiles obtained for *B. subtilis* tRNAs are shown as examples. Positions of anomalous cleavage ($cmnm^5U$, k^2C, mo^5U, m^7G) are shown in light blue or dark grey. Notice a low protection of T54 in (D), indication low methylation level under the experimental conditions used.