"Lost and Found": snoRNA Annotation in the *Xenopus* Genome and Implications for Evolutionary Studies

Svetlana Deryusheva,[†,1] Gaëlle J.S. Talhouarne,[†‡,1] and Joseph G. Gall[*,1]

[1]Department of Embryology, Carnegie Institution for Science, Baltimore, MD

[†]These authors contributed equally to this work.

[‡]Present address: Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT

***Corresponding author:** E-mail: gall@carnegiescience.edu.

Associate editor: Joanna Kelley

Abstract

Small nucleolar RNAs (snoRNAs) function primarily as guide RNAs for posttranscriptional modification of rRNAs and spliceosomal snRNAs, both of which are functionally important and evolutionarily conserved molecules. It is commonly believed that snoRNAs and the modifications they mediate are highly conserved across species. However, most relevant data on snoRNA annotation and RNA modification are limited to studies on human and yeast. Here, we used RNA-sequencing data from the giant oocyte nucleus of the frog *Xenopus tropicalis* to annotate a nearly complete set of snoRNAs. We compared the frog data with snoRNA sets from human and other vertebrate genomes, including mammals, birds, reptiles, and fish. We identified many *Xenopus*-specific (or nonhuman) snoRNAs and *Xenopus*-specific domains in snoRNAs from conserved RNA families. We predicted that some of these nonhuman snoRNAs and domains mediate modifications at unexpected positions in rRNAs and snRNAs. These modifications were mapped as predicted when RNA modification assays were applied to RNA from nine vertebrate species: frogs *X. tropicalis* and *X. laevis*, newt *Notophthalmus viridescens*, axolotl *Ambystoma mexicanum*, whiptail lizard *Aspidoscelis neomexicana*, zebrafish *Danio rerio*, chicken, mouse, and human. This analysis revealed that only a subset of RNA modifications is evolutionarily conserved and that modification patterns may vary even between closely related species. We speculate that each functional domain in snoRNAs (half of an snoRNA) may evolve independently and shuffle between different snoRNAs.

Key words: **snoRNA annotation, guide activity, RNA modification, *Xenopus*, vertebrates.**

Introduction

Small nucleolar RNAs (snoRNAs) are nuclear noncoding RNAs that primarily guide 2′-*O*-ribose methylation and post-transcriptional isomerization of uridine to pseudouridine in various RNA molecules, including very abundant and heavily modified ribosomal RNAs and spliceosomal U snRNAs. Therefore, snoRNAs are often referred to as modification guide RNAs. All pseudouridylation guide RNAs are characterized by short motifs called H (ANANNA) and ACA (ACA or AUA trinucleotide) boxes, and a distinctive "hairpin–hinge–hairpin–tail" secondary structure with the H and ACA boxes located in the hinge and tail regions, respectively. The hairpins form internal loops, or pseudouridylation pockets. The single-stranded sides of these loops serve as antisense elements that base-pair with substrate RNAs flanking the target uridine. 2′-*O*-methylation guide RNAs are distinguished by C/C′ (RUGAUGA) and D/D′ (CUGA) boxes and sequences complementary to substrate RNAs, or antisense elements, located upstream of D-box motifs (reviewed by Yu et al. 2005). Thus, sequence conservation of this class of RNAs is basically limited to antisense elements and "boxes."

Evolutionary studies of snoRNA conservation and diversity are based mostly on snoRNA gene annotations from publicly available genome assemblies (Hoeppner and Poole 2012; Kehr et al. 2014). A number of genomic snoRNA scanning tools (Lowe and Eddy 1999; Schattner et al. 2006; Yang et al. 2006; Freyhult et al. 2008; Hertel et al. 2008; Bartschat et al. 2014) and target prediction tools (Bazeley et al. 2008; Tafer et al. 2010; Kehr et al. 2011, 2014) are available and they are crucial for snoRNA annotation. However, their usage is restricted by our knowledge of the functional interaction between a guide RNA and its corresponding substrate RNA, such as the minimal length of antisense elements, the maximal number of mismatches and/or noncanonical pairs, and the stringency and flexibility of RNA secondary structure. Moreover, the motif-based search for snoRNA sequences can pick up pseudogenes or snoRNA-like remnants that never express, while leaving genuine snoRNAs with noncanonical structure undetected. Experimental RNA coimmunoprecipitation has been productive in the discovery of novel snoRNAs (Kiss et al. 2004; Torchet et al. 2005; Jády et al. 2012; Deryusheva and Gall 2013; Kishore et al. 2013; Machyna et al. 2014). However, deep sequencing of such RNA samples shows that they may be highly contaminated (Jády et al. 2012). Thus, despite many efforts, annotation of modification guide RNAs in higher eukaryotes is still far from complete. Even in the well-studied mouse and human a number of conserved snoRNAs remain "orphan," whereas certain

Article

well-characterized modifications in rRNAs and snRNAs have no assigned snoRNAs (https://www-snorna.biotoul.fr/index.php, http://www.bioinf.uni-leipzig.de/publications/supplements/15-065; last accessed August 16, 2019) (Lestrade and Weber 2006; Jorjani et al. 2016). Furthermore, a better understanding of evolution-driven snoRNA conservation and divergence requires nearly complete sets of snoRNAs, especially from nonmammalian species.

Studies from our laboratory (Gardner et al. 2012; Talhouarne and Gall 2014) demonstrated the advantages of the *Xenopus* oocyte in searching for missing snoRNAs. Amphibian oocytes are enormously large cells with a giant nucleus (also called germinal vesicle, or GV). GVs can be isolated manually from oocytes, providing 2 ng of pure nuclear RNA from a single cell. The box C/D and box H/ACA modification guide RNAs normally reside in the nucleus where rRNAs and snRNAs are modified (Maden 1990; Yu et al. 2001; Jády et al. 2003). Most snoRNAs are encoded within introns, and their proper processing relies on splicing, which is a nuclear event as well. Thus, after rRNA depletion, GV RNA is a good source of snoRNAs without any additional enrichment or size selection. We chose *Xenopus tropicalis* GV RNA for our snoRNA search because 1) the *X. tropicalis* genome is relatively small and has been sequenced, 2) pure GV RNA samples have already been obtained and the high-throughput sequencing data are available for further analysis, and 3) many highly conserved snoRNAs have already been annotated in *Xenopus*. In this study we scanned through the nuclear transcriptomes of *X. tropicalis* and successfully annotated 90% of human orthologous snoRNAs. In addition, we identified 30 new *Xenopus* snoRNAs. Having "complete" snoRNA sets from *Amphibia* and *Mammalia*, supplemented with subsets of orthologous snoRNAs annotated and newly identified in birds, reptiles, and fish, we performed interspecies comparative analysis of snoRNA modification target predictions and modification patterns of rRNAs and snRNAs. Our findings suggest that snoRNA–substrate assignments are not as conserved as generally believed and that a point mutation in a highly conserved snoRNA can result in loss or gain of function in targeting a substrate RNA for modification. Furthermore, we hypothesize that 5′- and 3′-terminal domains in snoRNAs evolve and migrate within the genome as independent units.

Results

Completion of *Xenopus* snoRNA Annotation

Although the *Xenopus tropicalis* genome is now relatively well assembled (http://www.xenbase.org; last accessed August 16, 2019) (Karpinka et al. 2015), snoRNA annotation is still incomplete—only about 100 snoRNAs have been annotated so far. To find new box C/D and box H/ACA guide RNAs in the *Xenopus* genome, we analyzed abundant transcripts in a nuclear RNA fraction isolated from *X. tropicalis* oocytes, or in GV RNA (SRA data: SRP066274, SRX4895193). After rRNA depletion this fraction consists predominantly of stable intronic RNAs, including snoRNAs, and a few types of abundant nuclear RNAs, such as U snRNAs, 7SK, and RNaseP. The latter sequences can be easily identified and excluded from the analysis. Details of this analysis are described in the Materials and Methods section. This approach allowed us to identify 422 snoRNAs in the *X. tropicalis* genome (supplementary table 1, Supplementary Material online). The list includes additional copies of previously annotated snoRNAs and newly identified *Xenopus* snoRNAs. Among these are *Xenopus* orthologs of known human snoRNAs and snoRNAs that do not have annotated homologs in any other species; the latter are highlighted in supplementary table 1, Supplementary Material online. For snoRNA annotation in the xentro9 genome assembly, we provide these data in a BED file format ("BED" tab in supplementary table 1, Supplementary Material online).

Overall, we have identified most *Xenopus* orthologs of annotated human snoRNAs. We expected some guide RNAs to be absent from the *Xenopus* genome because the corresponding modifications do not occur in *Xenopus* 18S and 28S rRNAs (Maden 1986, 1988; our data). However, a few guide RNAs for highly conserved modifications were not found, presumably because the reference genome is not complete. SNORD69 is an example of such an snoRNA (fig. 1). In humans, this box C/D snoRNA is annotated as a guide RNA for 2′-*O*-methylation of 28S rRNA at position G4464 and is encoded in an intron of the *gnl3* gene, which also hosts SNORD19 (fig. 1*A*). In the *Xenopus* genome assembly, *gnl3* contains several gaps, of which one is in the region where SNORD69 is expected (fig. 1*A* and *B*). We amplified and sequenced a genomic fragment from this region of the *Xenopus gnl3* gene. When this fragment was incorporated in the *gnl3* gene model and RNA-sequencing reads were mapped to the improved gene model, SNORD69 became evident (fig. 1*B*, blue). The expression of SNORD69 in *Xenopus* somatic cells and oocytes was verified by northern blot analysis (fig. 1*C*). Thus, missing snoRNAs will be found once a better genome assembly is available for *Xenopus*.

Target Predictions for Known and Newly Identified snoRNAs in *Xenopus* and Human

We analyzed all newly identified and previously annotated *Xenopus* and human snoRNAs for potential rRNA and snRNA targets. The results of this analysis are summarized in supplementary table 1 and supplementary figure 1, Supplementary Material online. A similar analysis was performed earlier by other investigators (Kehr et al. 2014; Jorjani et al. 2016). Many, but not all of our assignments coincide with previous predictions. All unexpected snoRNAs and predicted targets have been validated. To verify predicted modification guide RNA activities, we used an in vivo yeast cell system. *Saccharomyces cerevisiae* provides a favorable system because yeast cells support guide RNA-mediated posttranscriptional modification. At the same time, the number of modified positions is fewer than in higher eukaryotes and only five yeast snoRNAs are essential (https://people.biochem.umass.edu/sfournier/fournierlab/snornadb/mastertable.php; last accessed August 16, 2019, https://www.yeastgenome.org; last accessed August 16, 2019). In yeast, one can easily express exogenous guide RNAs from a plasmid, delete endogenous guide RNAs, and test predicted substrate RNAs for modifications. This system

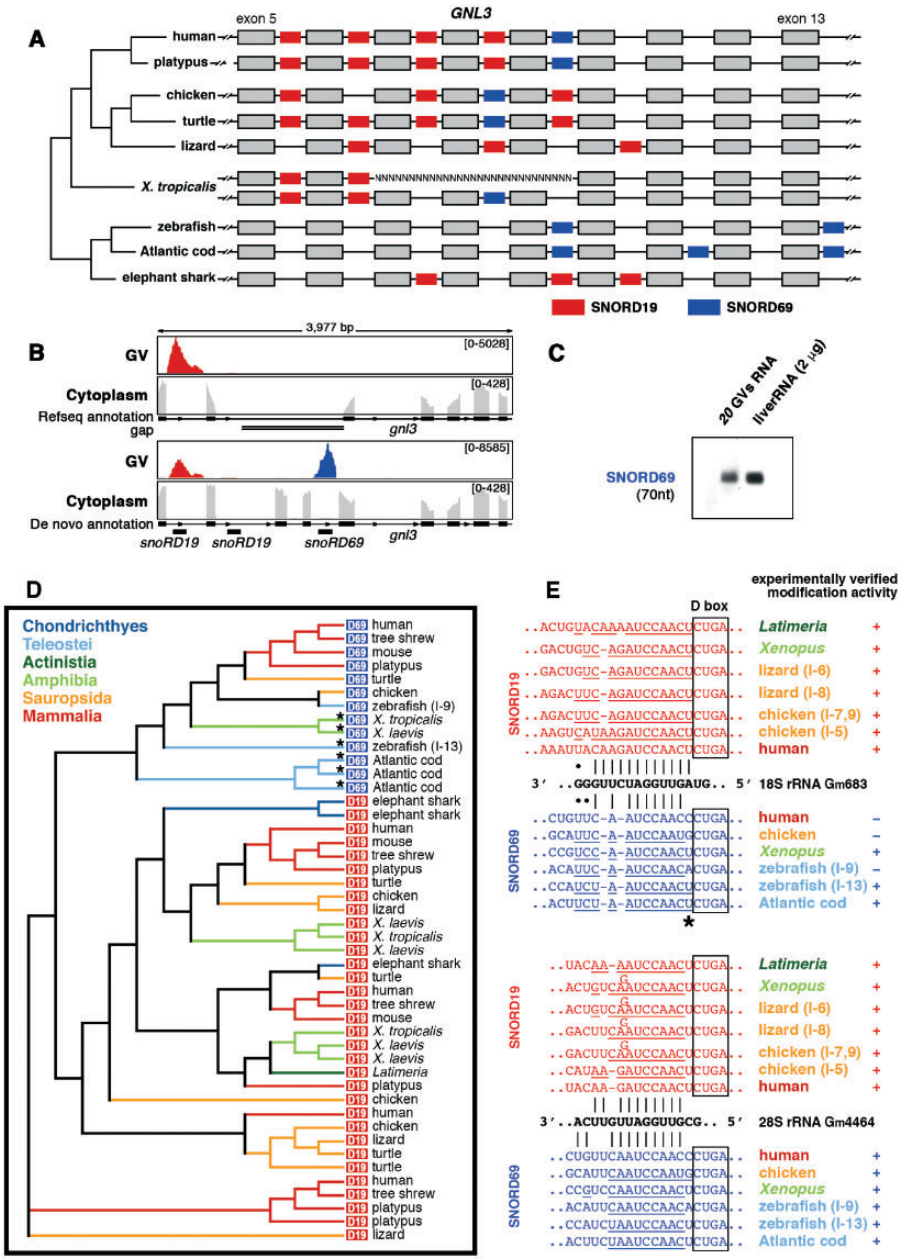

FIG. 1. SNORD19/SNORD69 snoRNA family. (*A*) Schematic gene models of the *GNL3* gene harboring SNORD19 (red) and SNORD69 (blue) snoRNAs in different vertebrate species. The current assembly of the *Xenopus tropicalis* genome contains gaps indicated with Ns (top of the two *X. tropicalis* schemes). Using PCR amplification and sequencing, we filled the gap in *X. tropicalis gnl3* and annotated SNORD69 in this region (bottom of the two *X. tropicalis* schemes). (*B*) IGV browser view of deep-sequencing reads generated for RNA from cytoplasm and nuclei (germinal vesicles, or GV) of *X. tropicalis* oocytes and aligned to the *X. tropicalis* 9.1 genome (top two panels) or to the improved gene model (two bottom panels). Expression of SNORD69 (blue) along with SNORD19 (red) became evident when the de novo gene annotation was used. (*C*) Northern blot analysis of newly identified SNORD69 in *X. tropicalis*. RNA was isolated from 20 oocyte nuclei (GVs) and liver. (*D*) Clustal analysis of SNORD19/SNORD69 snoRNA family sequences from different taxonomic groups. All SNORD69 (blue) sequences cluster together regardless of their modification activity on 18S rRNA. Stars indicate SNORD69 sequences that are functional on both 28S and 18S rRNAs. (*E*) Postulated base-pairing of SNORD19 and SNORD69 snoRNAs from different species with 18S and 28S rRNAs. Experimentally verified guide activity is indicated for each snoRNA with a plus or minus sign (functional or nonfunctional at the listed position). For more details on the guide-RNA modification activity assay, see supplementary figure 2*A* and *D*, Supplementary Material online. The positions of one-nucleotide variations critical for SNORD69 modification activity on 18S rRNA are marked with a star.

was previously used to test modification guide activity of *Drosophila* and vertebrate snoRNAs (Xiao et al. 2009; Qu et al. 2011; Deryusheva and Gall 2013, 2017, 2018).

A subset of the verified predictions is listed in table 1 and shown in supplementary figures 2 and 3, Supplementary Material online. Specifically, we identified the orphan snoRNAs SNORA11 and SNORD89 as guide RNAs for positioning Ψ58 and Gm12 in U2 snRNA, well-known highly conserved modifications that had no assigned guide RNAs (fig. 2). The predicted modification guide RNA activities were

Table 1. List of Experimentally Verified Novel Target Predictions for Evolutionarily Conserved Modification Guide RNAs.

Guide RNA	Target Predictions Made[a]			Verification
	Elsewhere[b]	This Study		
	Human	Human	*Xenopus* (and/or other species)	
SCARNA4	U2-41	U2-41	U2-41	See Deryusheva and Gall (2018) for details
	U2-39	—	—	
SCARNA15	—	—	U2-41 (all except placental)	See Deryusheva and Gall (2018) for details
	U2-37	U2-37/39	U2-37/39	
SNORD19	18S-G683	18S-G683/28S-G4464	18S-G683/28S-G4464 (all species)	Tested in yeast cell system
SNORD69	28S-G4464	28S-G4464	28S-G4464/18S-G683 (fish)	Tested in yeast cell system
SNORD89	orphan	U2-G12	U2-G12	Tested in yeast cell system
SNORA11	18S-1350	U2-58	U2-58	Tested in yeast cell system
SNORA14	18S-681	—	18S-892 (amphibians, fish)	Tested in yeast cell system, and in HeLa cells;
	18S-966	18S-966	18S-966	SNORA55 modifies 18S-681
SNORA22	28S-4966/4467	—	—	Tested in yeast, HeLa and XTC cells.
	28S-4975/18S-918	18S-918/28S-4975/ 4501 (no Ψ4501)	18S-918/28S-4501 (no Ψ4501)	SNORA22 can modify 4501 in artificial RNA substrate, yet actual modified position in 28S is Ψ4502, not 4501[c]
SNORA25	18S-801	18S-801	18S-801	Modification mapping in different species
	18S-814/28S-4597	—	28S-4558 (amphibians, reptiles, birds)	Tested in yeast cell system
SNORA29[d]	28S-4492[e]	28S-45 (no Ψ in human)	28S-45 (all, except human)	Modification mapping in different species, northern blot analysis of SNORA29 expression
	18S-220	18S-220 (no Ψ in human)	18S-220 (all, except human)	
SNORA35	18S-566	Orphan	18S-566 (amphibians, reptiles)	Modification mapping in different species
	U7-7		18S-350 (all, C in mammals)	
SNORA49	28S-2826	28S-2826	28S-2826	Modification mapping in different species
	28S-2830	28S-2830	— (28S-2830 in birds)	
SNORA57	18S-1004	18S-1004	18S-1004	Modification mapping in different species
	18S-1046	18S-1045/1046	18S-1045/1046	Tested in yeast cell system
SNORA61	28S-2495	28S-2495	28S-2495	Tested in yeast cell system
	18S-918	18S-918	18S-918	
SNORA63	—	28S-4491	28S-4491	*Xenopus* SNORA10 lacks antisense element for 28S-Ψ4491, but the modification is mediated by SNORA63
	28S-4390	28S-4390	28S-4390	

[a]Positions are numbered based on reference human sequences.
[b]Compiled from: https://www-snorna.biotoul.fr//index.php; last accessed August 16, 2019 (Lestrade and Weber 2006), http://www.bioinf.uni-leipzig.de/publications/supplements/12-022; last accessed August 16, 2019 (Kehr et al. 2014), http://snoatlas.bioinf.uni-leipzig.de; last accessed August 16, 2019 (Jorjani et al. 2016).
[c]Position 4501 in human 28S rRNA was mistakenly annotated as pseudouridine in the original article (Ofengand and Bakin 1997); the actual modified position was 4502; the corresponding pseudouridine was correctly numbered in mouse 28S rRNA.
[d]Mature SNORA29 is not expressed in human.
[e]This prediction resulted from a misprint in Kehr et al (2014). The correct prediction is 28S-45.

confirmed using a yeast cell system (supplementary fig. 2*B* and supplementary table 2, Supplementary Material online). Because these snoRNAs modify U2 snRNA and contain canonical Cajal body localization signals, they should be reannotated as scaRNAs. We also predicted that SNORA57 could mediate pseudouridylation of 18S rRNA at position 1045 (fig. 2) in addition to the previously assigned positions 1004 and 1046 (Jorjani et al. 2016); pseudouridines at positions 1045 and 1046 have been recently detected in human 18S rRNA (Carlile et al. 2014; Taoka et al. 2018), but 1045 has no guide RNA assigned so far. When experimentally tested in yeast cells, SNORA57 was able to modify all three predicted positions (supplementary fig. 2*C*, Supplementary Material online). In the *Xenopus* snoRNA set, we also identified an elusive guide RNA for positioning another recently identified pseudouridine in human 18S rRNA at position 1177. One of the two copies of *Xenopus* SNORA26 contains a pseudouridylation pocket specific for this position (fig. 2). Notably, none of the SNORA26 copies from other species has an antisense element specific to 18S-Ψ1177. Similarly, in *Xenopus* SNORA4 and SNORA55, we identified antisense elements for positioning a highly conserved pseudouridine in 28S rRNA that is equivalent to human 28S-Ψ4323 (fig. 2 and supplementary fig. 1, Supplementary Material online), yet orthologous snoRNAs in other species are missing this antisense element. In birds, mammals and reptiles SNORA55 targets position 36 in 18S rRNA instead.

Among *Xenopus* snoRNAs that showed no homology with any known snoRNAs, we found three box H/ACA RNAs for other highly conserved modifications in snRNAs and 18S rRNA; these modified positions have no guide RNAs assigned so far in any other species, including human (supplementary table 1, Supplementary Material online, highlighted in red). These new snoRNAs are pseudouridylation guide RNAs for

position 6 in U2 snRNA, position 12 in U4atac snRNA, and 2′-*O*-methylated U1288 in *Xenopus* 18S rRNA (equivalent to Ψm1326 in human 18S rRNA). The existence of these guide RNAs was verified by RACE (rapid amplification of cDNA ends) and northern blot analysis. We previously described a *Xenopus* guide RNA for positioning conserved Ψ43 in U2 snRNA (Deryusheva and Gall 2017). So far, we have not found corresponding guide RNAs in human and other genome assemblies based on sequence and structure similarity.

Intriguingly, five other *Xenopus*-specific snoRNAs show complementarity with 18S rRNA, at positions that have never been reported as modified in any vertebrate species; four additional *Xenopus*-specific snoRNAs were assigned to unexpected positions in 28S rRNA and U6 snRNA (supplementary table 1, highlighted in green; supplementary fig. 1, Supplementary Material online). Furthermore, in 21 evolutionarily conserved snoRNAs we predicted additional antisense elements for positioning modifications in 18S, 28S, and 5.8S rRNAs at unexpected positions (figs. 2–7, supplementary figs. 1 and 3 and supplementary table 1, Supplementary Material online). The majority of these antisense elements were found in *Xenopus* snoRNAs but not in human homologs. At the same time, at least ten antisense elements identified in human snoRNAs are missing from the homologous *Xenopus* snoRNAs (fig. 8 and supplementary table 1, Supplementary Material online). These findings made us curious about the existence of modifications at the predicted positions in rRNAs from different vertebrate species.

rRNA Modification in Vertebrate Species

rRNA modification mapping is often regarded as a completed story. In fact, rRNA modifications were mostly studied in budding yeast *S. cerevisiae* and human (Maden and Wakeman 1988; Maden 1990; Carlile et al. 2014; Krogh et al. 2016; Taoka et al. 2016, 2018; Sharma et al. 2017). Although 2′-*O*-methylated positions were meticulously mapped in both human and *X. laevis* 18S rRNA (Maden 1986), only eight pseudouridines have been mapped so far more or less precisely in *X. laevis* 18S rRNA (Salim and Maden 1980; Maden 1986; supplementary fig. 4*A*, highlighted in gray, Supplementary Material online). The total number of pseudouridines in *X. laevis* 18S rRNA was estimated to be between 45 and 49 ± 2 (Hughes and Maden 1978; McCallum and Maden 1985). *Xenopus tropicalis* rRNAs have never been examined for modifications. Therefore, we decided to map pseudouridines and 2′-*O*-methylated residues in both 18S rRNA and most of 28S rRNA from *X. tropicalis* and *X. laevis*; human rRNAs served as a control. In addition, we examined pseudouridylation patterns of 18S rRNA in six other species from different vertebrate classes: newt *Notophthalmus viridescens* and axolotl *Ambystoma mexicanum* (Amphibia), mouse *Mus musculus* (Mammalia), domestic chicken *Gallus gallus* (Aves), Mexican whiptail lizard *Aspidoscelis neomexicana* (Reptilia), and zebrafish *Danio rerio* (Teleostei). All modifications detected in 18S rRNA from nine vertebrate species are compiled in figure 9 and supplementary figure 4*A*, Supplementary Material online. We found five pseudouridines that were recently identified in human 18S rRNA to be highly conserved: Ψ256, Ψ1045, Ψ1046, Ψ1177, and Ψ1232 (Jorjani et al. 2016; Taoka et al. 2018). They have been detected unequivocally in all analyzed species (supplementary fig. 4*A*, highlighted in yellow, Supplementary Material online).

It is worth emphasizing that *Xenopus*-specific snoRNAs we identified and assigned to unexpected positions in 18S and 28S rRNAs coincide with corresponding modifications in *X. tropicalis* and *X. laevis* rRNAs, but not in human. Also, we found modifications at unexpected positions assigned to novel antisense elements in *Xenopus* snoRNAs from conserved snoRNA families. Again, these modifications exist in *Xenopus* but not in human (supplementary fig. 4, Supplementary Material online). At the same time, at least 14 snoRNAs identified in human are not expressed in *X. tropicalis*. Additionally, as mentioned above, ten human snoRNAs from highly conserved snoRNA families contain antisense elements missing from *Xenopus* orthologs (fig. 8). This resulted in the absence of 17 modifications in *Xenopus* rRNAs as compared with human. Modification patterns in other vertebrate species may vary; nonetheless, these patterns generally correlate with corresponding snoRNA modification activities (supplementary fig. 4, Supplementary Material online).

Gain and Loss of snoRNA Modification Activities in Vertebrates

Although many eukaryotic snoRNAs show evolutionary conservation, a number of mammalian snoRNAs cannot be traced to ancestral snoRNA families. Similarly, several snoRNAs identified in *Xenopus* have no relationship to any known snoRNAs in other organisms. Furthermore, their expression might be limited to frogs, for example, pug*xt*18S-211 encoded in an intron of *Taf1d* harboring multiple conserved snoRNAs (fig. 3*A* and *B*) or only to the genus *Xenopus*, for example, pug*xt*18S-124 hosted by the *Rpl26* gene. Although *xt*pug18S-211 provides a clear example of gain of modification through acquisition of both novel intron and intronic snoRNA in a frog-restricted manner, pug*xt*18S-1279 found in an intron of the *Rpl35* gene represents a novel evolutionarily conserved snoRNA. The corresponding snoRNA sequences and 18S rRNA pseudouridilation at the equivalent position are evident in frogs, lizard, and zebrafish (supplementary figs. 1 and 4*A*, Supplementary Material online). The expression of pug18S-1279 homologs in these species also coincides with an additional intron in the *Rpl35* gene. This intron seems to be lost along with the snoRNA in mammals and birds. The loss of pug18S-1279 is accompanied by the U-to-C change in 18S rRNA (supplementary fig. 4*A*, Supplementary Material online).

Another mechanism for loss of snoRNA function is a transposable element insertion. That is how SNORA10 was destroyed in rodents and bats (fig. 4*A*). The loss of SNORA10 expression led to loss of pseudouridylation of 18S rRNA at position 210 (supplementary fig. 4*A*, Supplementary Material online), which coincided with the U-to-C transition at the corresponding position in some

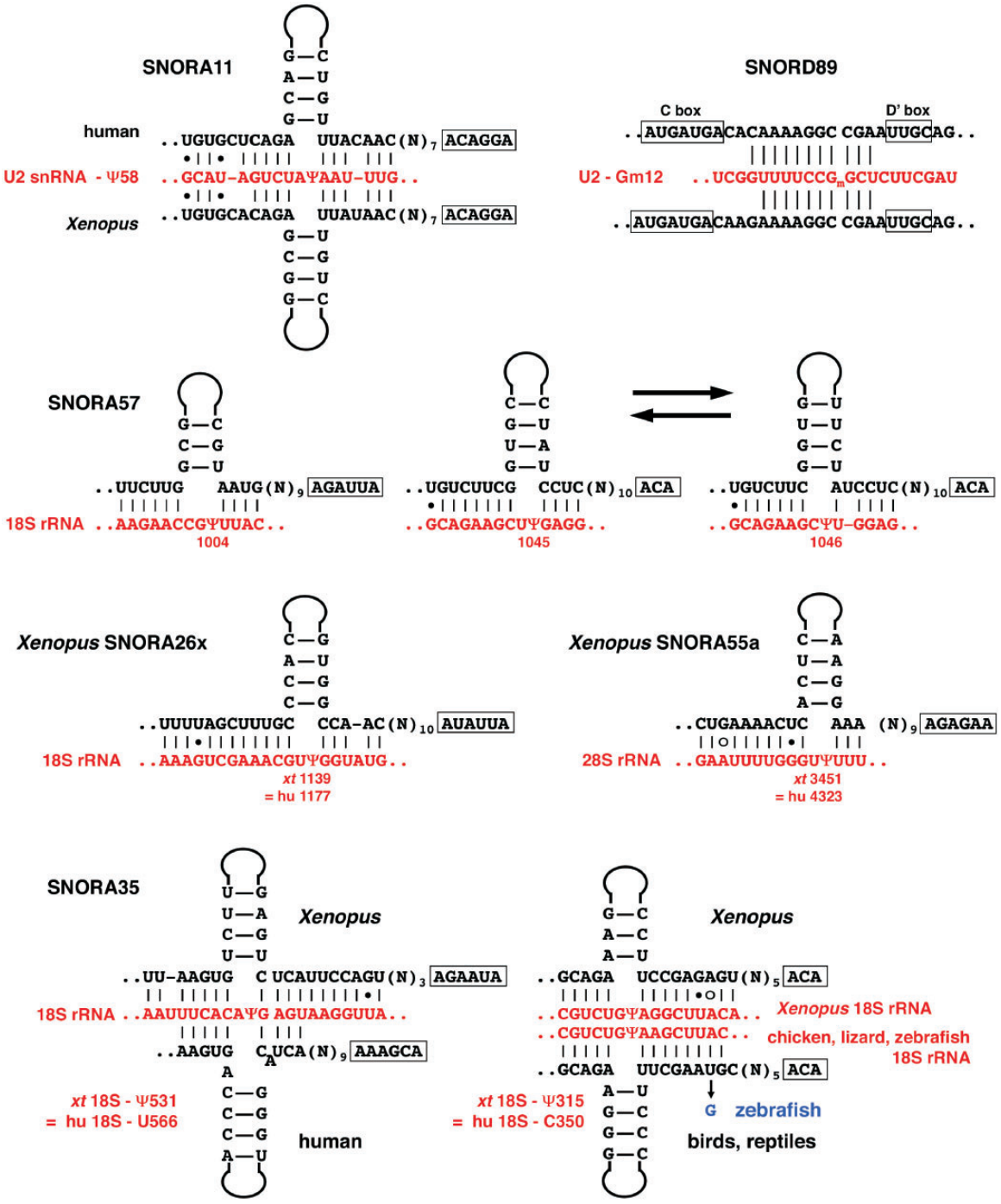

FIG. 2. Novel snoRNA assignments to previously known and recently identified orphan modifications in U2 snRNA, 18S and 28S rRNAs. For experimental verification of the predicted guide activities, see supplementary figures 2 and 4, Supplementary Material online.

species (fig. 4A). SNORA10 was identified also as a guide for 28S rRNA pseudouridylation at the position corresponding to human 4491 (fig. 4B). However, we did not observe the loss of this modification in the absence of SNORA10 expression. Furthermore, the *Xenopus* SNORA10 sequence is slightly diverged; that is, this snoRNA targets 5.8S rRNA instead of 28S rRNA (fig. 4B), yet modification of 28S rRNA is not affected in *Xenopus* either. In our target prediction analysis, we identified an additional antisense element for positioning 28S-Ψ4491 within an orphan pseudouridylation pocket of SNORA63 (fig. 4B); this antisense element is highly conserved evolutionarily, explaining the persisting pseudouridylation of 28S rRNA in the absence of SNORA10 activity.

Large genomic changes are not necessary for snoRNA distortion. Primary sequences homologous to SNORA75 are almost identical in two *Xenopus* species. The SNORA75 coding sequence in *X. laevis* has a short 10-nt deletion in the upper loop of the 5′-terminal domain compared with *X. tropicalis*; otherwise the sequences align with more than 82% identity, including conserved H/ACA boxes, antisense element sequence, and overall secondary structure.

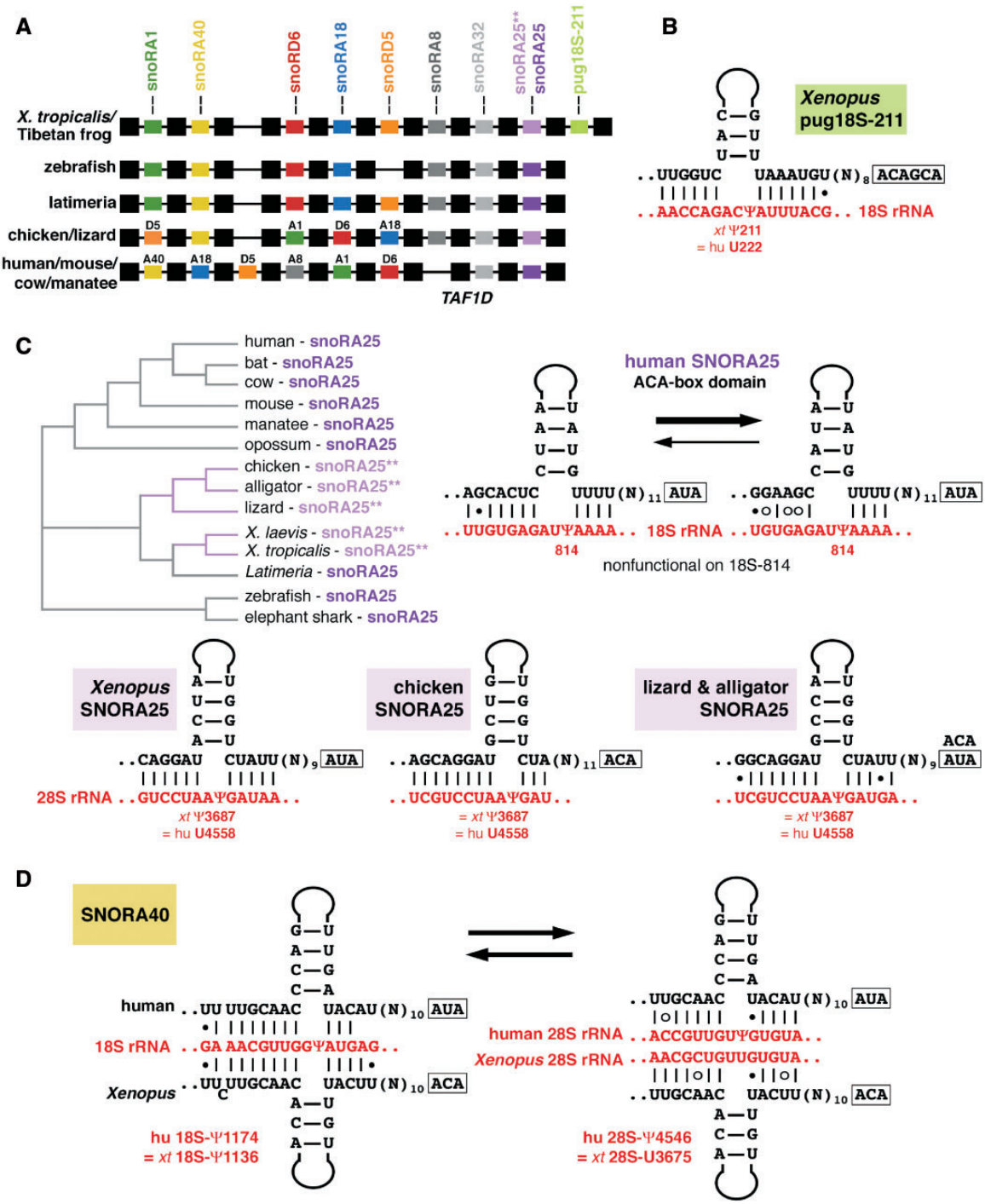

FIG. 3. An evolutionarily conserved snoRNA cluster hosted by the *TAF1D* locus. (*A*) Schematic drawing of *TAF1D* gene models in different vertebrate species. Homologous intron-encoded snoRNAs are color-coded. (*B*) Predicted base-pairing between 18S rRNA and *Xenopus*-specific snoRNA pug18S-211. (*C*) Clustal analysis of SNORA25 sequences from different species. Those that have two functional pseudouridylation pockets are highlighted with stars. Predicted interactions with rRNAs are shown for the 3′ terminal domains of SNORA25 from human, *Xenopus*, chicken, lizard and alligator. Strong base-pairing of human SNORA25 with 18S rRNA for positioning Ψ814 is possible only in a highly unstable configuration. For experimental verification of the predicted guide activities, see supplementary figures 3A and 4, Supplementary Material online. (*D*) Predicted base-pairing between SNORA40 and rRNAs: An alternative configuration for positioning of pseudouridine in 28S rRNA is functional in human but not in *Xenopus*.

Nevertheless, SNORA75 is not expressed in *X. laevis*, which results in the absence of 18S-Ψ93 (supplementary fig. 4*A*, Supplementary Material online). SNORA29 expression in human is even more compelling. In this snoRNA, we predicted antisense elements for position 45 in 28S rRNA and position 220 in 18S rRNA; the latter position was also predicted by Jorjani et al. (2016). The sequence of this snoRNA is highly conserved cross-species (fig. 5*A*); however, the

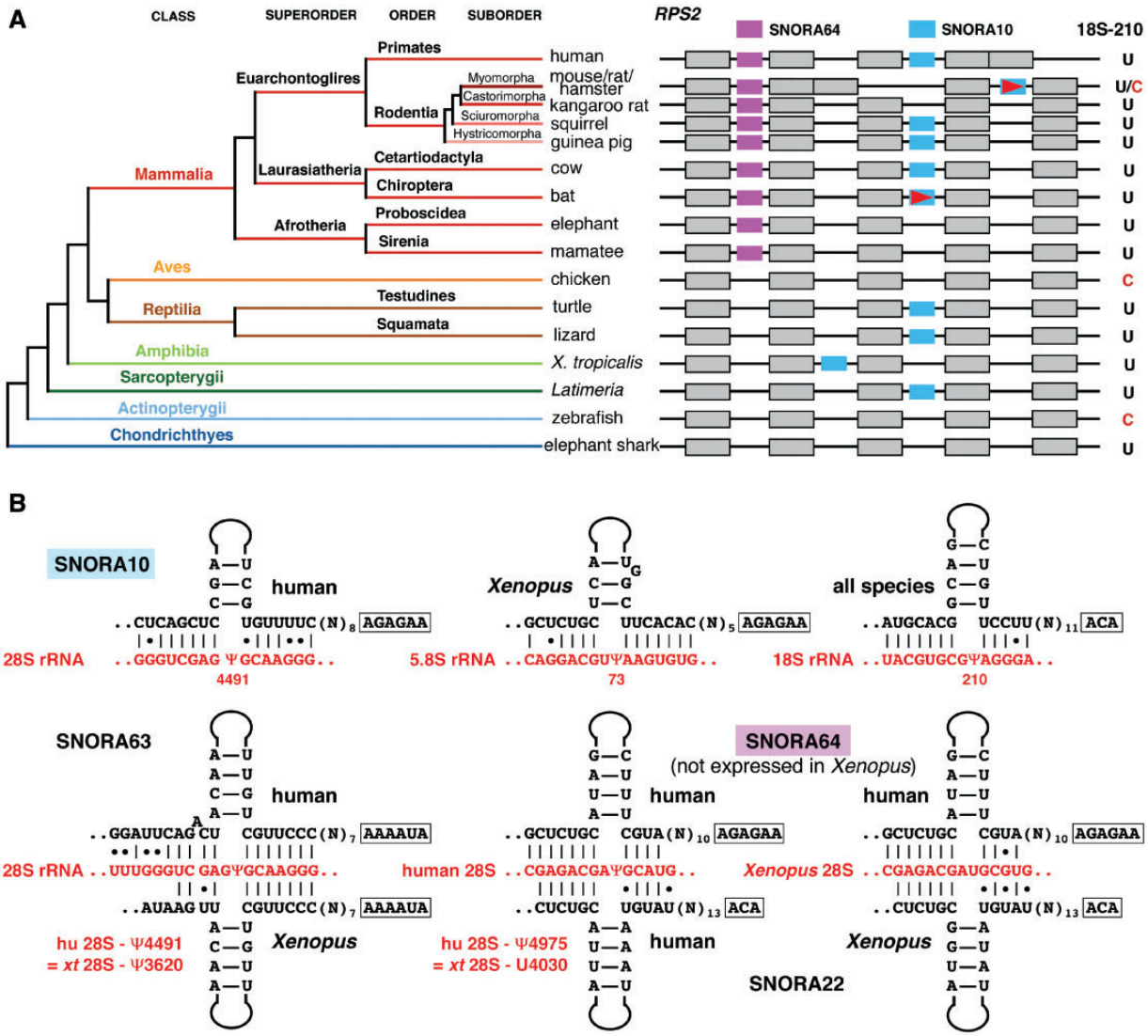

FIG. 4. (*A*) Expression of the SNORA10/SNORA64 snoRNA family. Insertions of transposable elements (indicated with red triangles) destroyed SNORA10 in rodents and bats. No SNORA10 sequences were found in fish (except *Latimeria*) and chicken. In mammalian species SNORA10 gave rise to SNORA64, targeting 28S rRNA at position 4975. (*B*) Postulated base-pairing between SNORA10 and rRNAs: human SNORA10 targets 18S-210 and 28S-4491, whereas *Xenopus* SNORA10 targets 5.8S rRNA instead of 28S rRNA. SNORA63 is a newly identified conserved guide RNA for positioning 28S-Ψ4491. SNORA64 and SNORA22 can modify human 28S rRNA at position 4975. The equivalent position in *Xenopus* is not modified, because SNORA22 interaction with *Xenopus* 28S rRNA is weak, and SNORA64 is not expressed in nonmammalian species.

predicted modifications are evident in *Xenopus* and other amphibians, lizard, chicken, and mouse, but they have never been detected in human, both in our experiments (supplementary fig. 4*A* and *B*, Supplementary Material online) and other studies (Carlile et al. 2014; Stanley et al. 2016; Taoka et al. 2018). Despite the fact that SNORA29 was originally identified in HeLa cells (Kiss et al. 2004), we could not detect processed SNORA29 in human fibroblasts and HeLa cell lines by 5′- and 3′-RACE or by northern blot analysis (fig. 5*B*). The only structural difference that discriminates the human SNORA29 coding sequence from other primate species and various vertebrates is a point mutation at the box ACA (fig. 5*A*). Thus, this mutation presumably causes misprocessing of human SNORA29 and loss of evolutionarily highly conserved modifications in one particular species.

Orphan Domains in snoRNAs Find Their Targets

As we mentioned above, several evolutionarily conserved snoRNAs in *Xenopus* contain antisense elements specific to positions in rRNAs that are not modified in mammals. We will describe three notable examples of gain and loss of guide activity in functional domains of snoRNAs during evolution.

SNORA14 provides an example of a conserved snoRNA that has lost its modification activity on rRNA at a highly conserved region. In all vertebrate species, the 3′-terminal domain of this snoRNA targets position 966 in 18S rRNA for pseudouridylation. At the same time, only in amphibians and zebrafish does the 5′-terminal domain form a pseudouridylation pocket that specifically modifies 18S rRNA at the position equivalent to human U892 (supplementary fig. 3*B*, Supplementary Material online). Furthermore, in *Xenopus* the 5′-terminal pseudouridylation pocket is invariably present in

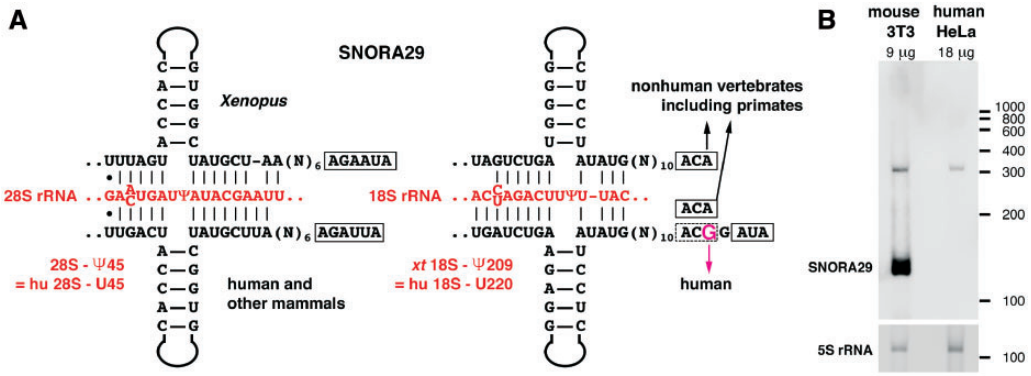

FIG. 5. (*A*) Postulated base-pairing between SNORA29 and 28S rRNA (H-box domain) and 18S rRNA (ACA-box domain). The canonical ACA box is characteristic of SNORA29 in all nonhuman vertebrates that have been analyzed. A point mutation in the human SNORA29 sequence is shown in magenta. We found the same A-to-G substitution in the Neandertal and Denisovan genomes (https://genome.ucsc.edu/Neandertal/; last accessed August 16, 2019). Pseudouridylation of 28S rRNA at position 45 and of 18S rRNA at position 220 was not detected in human RNA samples (see supplementary fig. 4, Supplementary Material online). (*B*) Northern blot analysis of total RNA isolated from mouse 3T3 and human HeLa cells with a probe specific for the human SNORA29 sequence. Mature SNORA29 was detected in mouse RNA samples, but not in human. Amount of RNA loaded on the gel is indicated at the top of each lane. 5S rRNA served as a loading control.

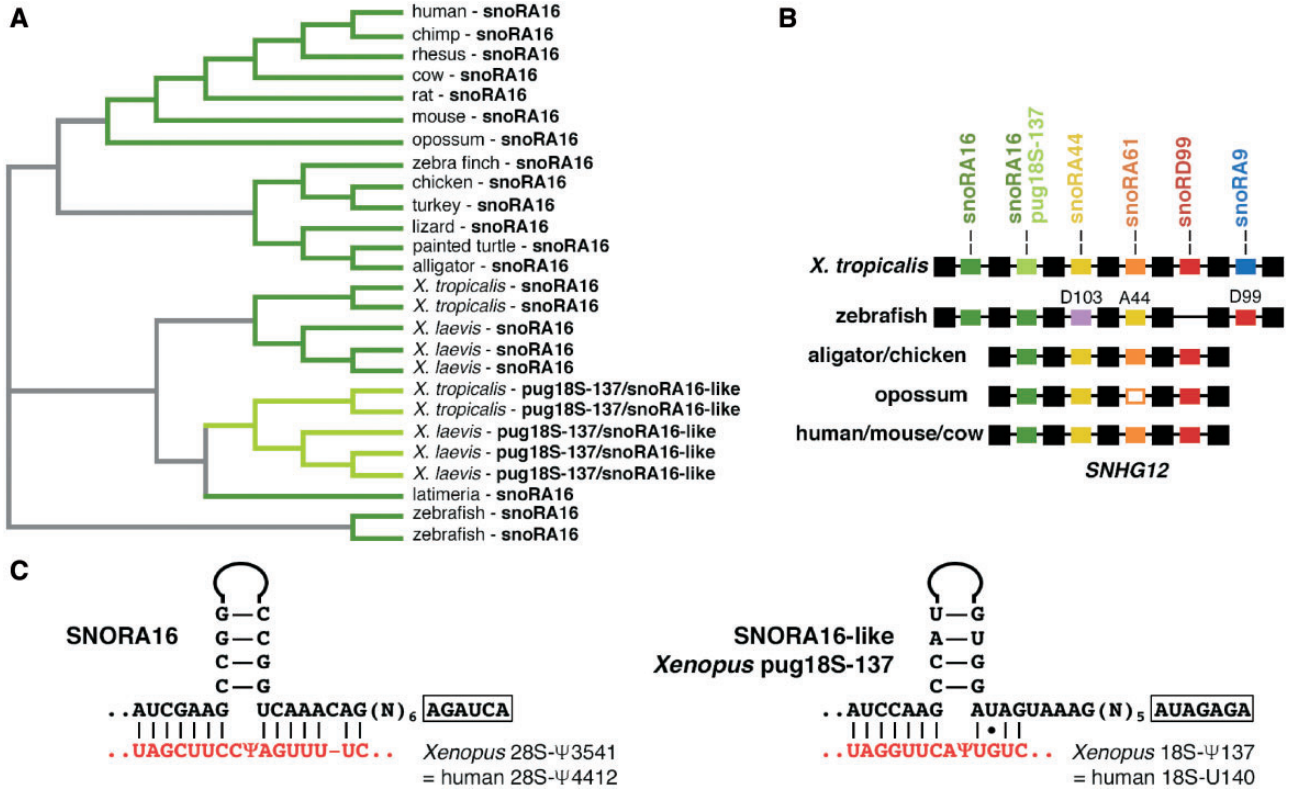

FIG. 6. (*A*) The SNORA16 snoRNA family includes *Xenopus*-specific SNORA16-like guide RNA pug18S-137 clustered together with SNORA16 from diverged vertebrate species. (*B*) *Xenopus* SNORA16-like pug18S-137 arose from an additional copy of SNORA16 encoded within the nonprotein-coding *SNHG12*, which harbors a conserved snoRNA cluster; conserved snoRNAs in this cluster are indicated and color-coded. (*C*) Postulated base-pairing between evolutionarily conserved SNORA16 and 28S rRNA and between *Xenopus*-specific SNORA16-like and 18S rRNA.

all five copies of this snoRNA, even though one of these copies does not contain the conserved antisense element in the 3′-terminal domain. Yet in reptiles, birds, and mammals, the 3′-terminal antisense element, but not the 5′-terminal antisense element, remains preserved. Importantly, *Xenopus* SNORA14, when expressed exogenously from a plasmid, was able to modify the corresponding positions in yeast (supplementary fig. 3*A*, Supplementary Material online) and human 18S rRNA (Deryusheva and Gall 2019). That is, the corresponding position of 18S rRNA is promiscuous for pseudouridylation, yet the ability to modify this position has been lost in reptiles, birds, and mammals.

SNORA25 shows gain and loss of modification activity of one pseudouridylation pocket, whereas the activity of the other pocket remains preserved during evolution (fig. 3C). When this snoRNA was first identified in human RNA, it

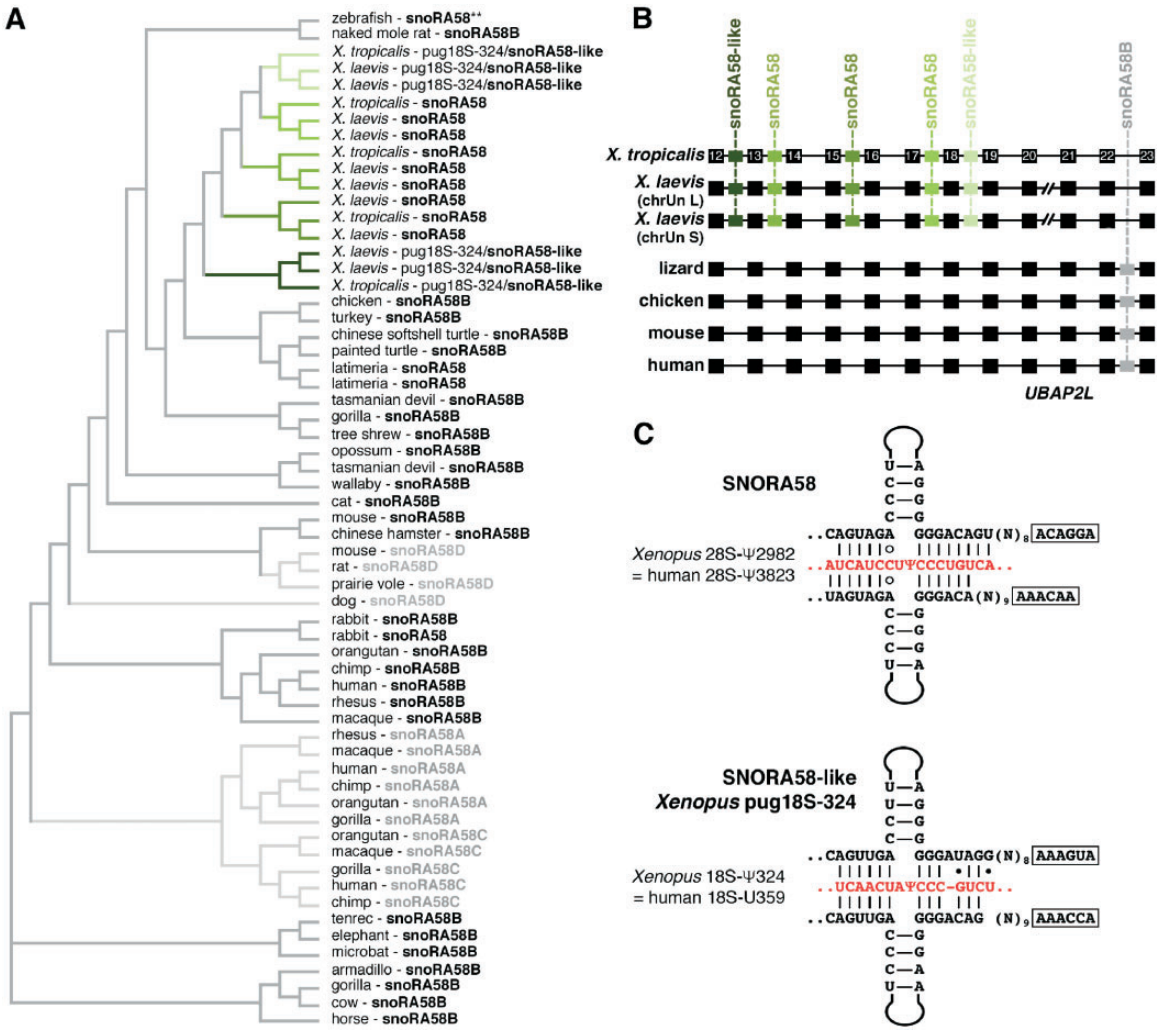

FIG. 7. (*A*) The SNORA58 snoRNA family includes *Xenopus*-specific SNORA58-like guide RNA pug18S-324 clustered together with SNORA58 from diverged vertebrate species. Zebrafish snoRNA (top line marked with two stars), although similar in sequence to other SNORA58s, does not have the conserved pseudouridylation pocket for positioning pseudouridine in 28S rRNA. (*B*) *Xenopus* SNORA58 and SNORA58-like RNAs are encoded in the evolutionarily conserved SNORA58 host gene *UBAP2L*. (*C*) Postulated base-pairing between evolutionarily conserved SNORA58 and 28S rRNA and between *Xenopus*-specific SNORA58-like and 18S rRNA.

was annotated as a guide RNA for positions 801 and 814 in 18S rRNA (Kiss et al 2004). Based on our analysis, the 3′-terminal domain of human SNORA25 does not form a structurally stable pseudouridylation pocket for positioning of 18S-Ψ814 (fig. 3C) and, accordingly, this snoRNA could not modify the corresponding position in yeast (supplementary fig. 3*A*, Supplementary Material online). Thus, we exclude 18S-814 from potential SNORA25 targets, which leaves no target assigned to the 3′-terminal domain of human SNORA25. However, in *Xenopus*, reptile, and chicken SNORA25 orthologs, the 3′-terminal domain contains an antisense element that targets a position equivalent to 4558 in human 28S rRNA (fig. 3C). We found pseudouridine at this position in four amphibian species, lizard and chicken as expected, but not in mammals and fish (supplementary fig. 4*B*, Supplementary Material online).

SNORA35 is an even more intriguing example of loss of modification activity. This highly conserved orphan box H/ACA snoRNA has a predicted antisense element for position 566 in human 18S rRNA (Jorjani et al. 2016), yet no pseudouridine has been detected in human and mouse (Carlile et al. 2014; Stanley et al. 2016; Taoka et al. 2018; our data). The equivalent position, however, is pseudouridylated in the four amphibian species and the lizard we tested (supplementary fig. 4*A*, Supplementary Material online). Furthermore, in all species except mammals, SNORA35 also targets position 350 for pseudouridylation (fig. 2); mammalian 18S rRNA has a C at this position (supplementary fig. 4*A*, Supplementary Material online). Thus, in nonmammal species, SNORA35 is a guide for 18S rRNA pseudouridylation, but in mammals it is well-preserved and still remains "orphan," despite the loss of function on 18S rRNA.

Target Switch by Point Mutations

The comparative analysis of human and *Xenopus* snoRNA functions revealed a substantial number of differences in snoRNAs and corresponding modifications between these two species (fig. 8 and supplementary table 1,

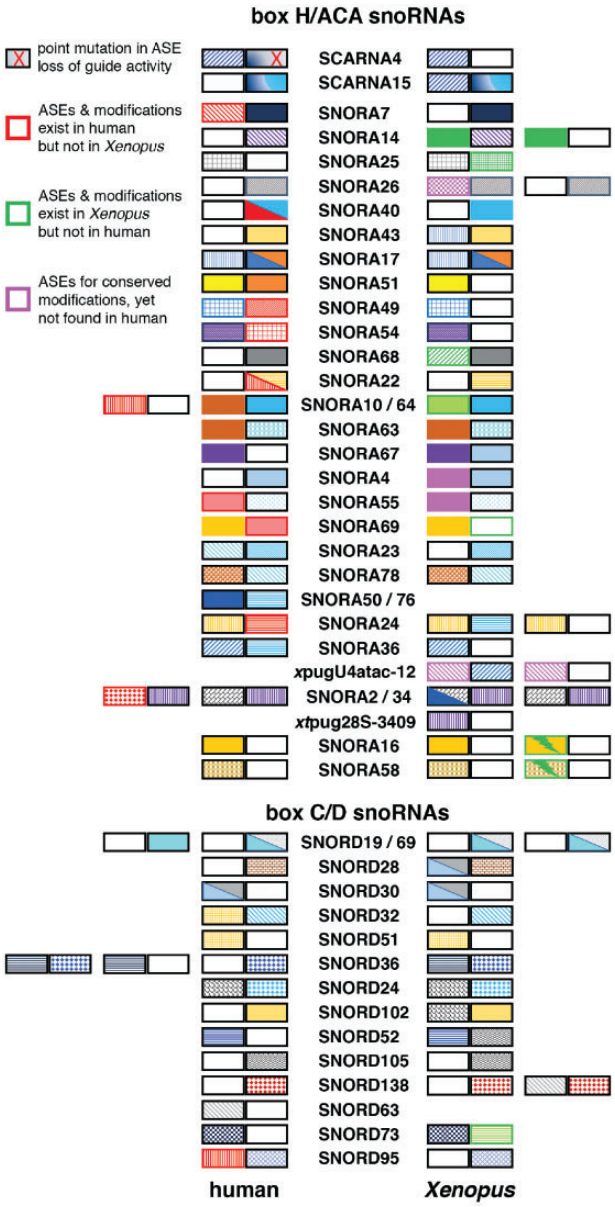

FIG. 8. Antisense element shuffling between conserved snoRNA families in *Xenopus* and human. When a family is represented by multiple copies of snoRNAs with diverged antisense elements, all such variants are depicted. Identical and/or diverged antisense element targets are color-coded. The actual target positions and snoRNA copy numbers are listed in supplementary table 1, Supplementary Material online.

Supplementary Material online). Typically, novel modifications are associated with newly acquired snoRNAs and significantly diverged copies of preexisting or recently duplicated snoRNA sequences. These diverged snoRNAs with different targets show some similarities and are usually grouped in families, such as the SNORA10/SNORA64 or SNORA2/SNORA34 families. Thus, mammalian SNORA64 must be derived from a recently duplicated ancestral SNORA10, whereas a preexisting copy of SNORA2 gave rise to SNORA34. Notably, human SNORA34 shows more similarity to both copies of *Xenopus* SNORA2 (54% and 55% identity) than to a human copy of SNORA2 (46% identity). Although this moderate or low level of similarity is quite common for snoRNA families with diverged substrate specificities, we argue here that such a significant sequence divergence is not required for a distinctive change in substrate recognition. Even a single nucleotide difference can convert a fully functional snoRNA (or a particular antisense element) into a nonfunctional copy or can completely change its modification guide specificity.

When we identified the *Xenopus* SNORD69 sequence and performed a comparative analysis of target predictions for it and SNORD19 from different vertebrate species, we found dual-guide modification activities for all SNORD19s and for SNORD69 from *Xenopus* and fish: namely, 2′-*O*-methylation of 28S rRNA at the position corresponding to human G4464 and of 18S rRNA at the position corresponding to human G683 (fig. 1*E*). Our analysis allowed us to group SNORD19 and SNORD69 in an snoRNA family, both having originated from the same ancestral snoRNA by duplication and divergence. Essentially, *Xenopus* SNORD69 clusters with SNORD69 sequences from other species (fig. 1*D*), yet it differs functionally. That is, a single nucleotide difference between SNORD69 antisense elements from different species led to the loss of modification guide activity on 18S rRNA in reptiles, birds, and mammals (fig. 1*E*, star). Two other examples of highly conserved snoRNAs having their target specificities changed by point mutations are *Xenopus* SNORA16 and SNORA58, which gave rise to pug18S-137 and pug18S-324, respectively (figs. 6 and 7). These novel modification guide RNAs cluster indistinguishably with other members of the families from which they originated. Thus, these data clearly demonstrate that snoRNA sequence conservation is not necessarily indicative of functional conservation.

Another intriguing case is SNORA28. This highly conserved snoRNA is a well-characterized guide RNA for two pseudouridines in 18S rRNA, 18S-Ψ815 and 18S-Ψ866 (Kiss et al. 2004; Xiao et al. 2009). Although 18S-Ψ815 was detected in all vertebrate species we analyzed, pseudouridylation at the position equivalent to 866 varied between species. Even the closely related frogs, *X. laevis* and *X. tropicalis*, showed differential modification at this position (supplementary fig. 4*A*, Supplementary Material online). Remarkably, both copies of *X. laevis* SNORA28 have point mutations in the antisense element for positioning 18S-Ψ866 (supplementary fig. 3*B*, Supplementary Material online). One of these copies has only a single nucleotide mismatch and exhibits modification guide activity on the appropriate artificial substrate RNA (supplementary fig. 3*C*, Supplementary Material online), but cannot induce pseudouridylation of endogenous 18S rRNA when tested in a yeast cell system (supplementary fig. 3*A*, Supplementary Material online). It is worth noting that both yeast and *X. laevis* 18S rRNAs are promiscuous for pseudouridylation at the position equivalent to human 18S-Ψ866. When *X. tropicalis* SNORA28 was expressed in yeast (supplementary fig. 3*A*, Supplementary Material online) or in the *X. laevis* XTC cell line (Deryusheva and Gall 2019), 18S rRNA was efficiently modified at the expected position. Why some modifications are stabilized or lost only in particular species within a taxon is an open question.

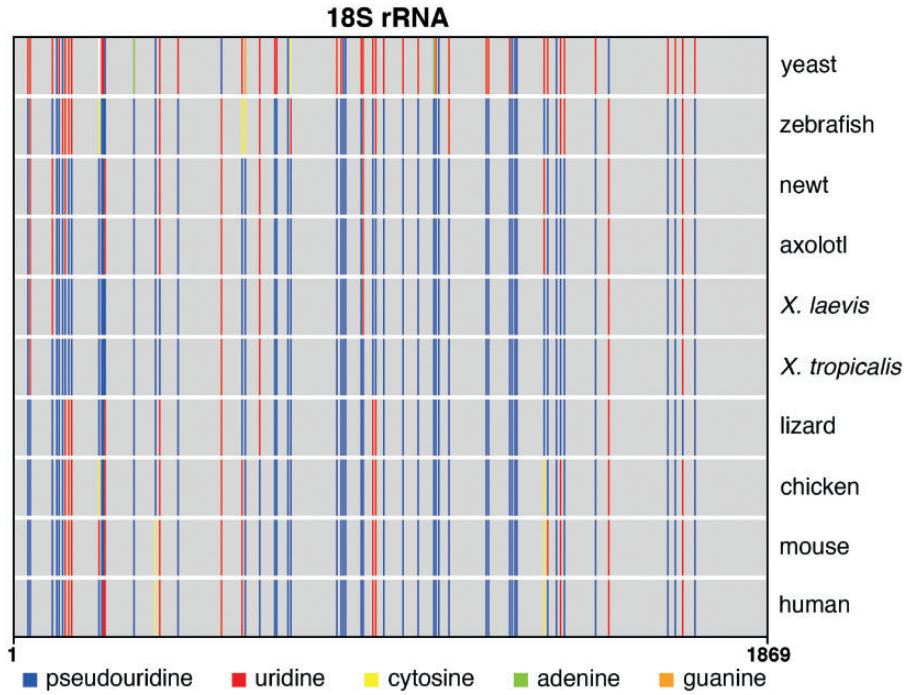

FIG. 9. Graphical representation of comparative pseudouridylation analysis of 18S rRNA from nine vertebrate species and yeast *Saccharomyces cerevisiae*. All positions that contain pseudouridine in at least one species are highlighted.

Point mutations in substrate RNAs also contribute to the generation of functional antisense elements within a conserved snoRNA to target a novel position for modification. In fact, a single U-to-A mutation in human SNORA40 accompanied by a C-to-U substitution in 28S rRNA resulted in the formation of 28S-Ψ4546 in human. However, the equivalent position in *Xenopus* is unmodified (supplementary fig. 4*B*, Supplementary Material online). Instead, as we mentioned earlier, the position equivalent to human 4558 is modified by SNORA25 in amphibians, reptiles, and birds, but not in mammals. Notably, the mutation in human SNORA40 that introduces a new 28S rRNA substrate specificity does not affect modification activity of the same pseudouridylation pocket at the highly conserved position in 18S rRNA (fig. 3*D* and supplementary fig. 4*A*, Supplementary Material online).

Human SNORA22 illustrates many of the same issues. It was originally annotated as a guide RNA for 28S rRNA: Its 5′- and 3′-terminal domains were predicted to guide 28S-Ψ4966 (Kiss et al. 2004) and 28S-Ψ4975 (Schattner et al. 2006), respectively. We find that the 5′-terminal antisense element is weak for positioning of 28S-Ψ4966. When tested in yeast, SNORA22 did not mediate pseudouridylation of an artificial substrate RNA corresponding to 28S-Ψ4966 (supplementary table 2, Supplementary Material online); that is, SNORA33, but not SNORA22 is the genuine guide RNA for this modification. As for the 3′-terminal domain, we determined that SNORA22 could mediate pseudouridylation at position 4975 of human 28S RNA, but not at the equivalent position of *Xenopus* 28S rRNA. This activity occurred despite the unusually long distance from the ACA box to the target uridine (18 nt). The corresponding sequence in *Xenopus* 28S rRNA generates an additional noncanonical U–G base pair between SNORA22 and 28S rRNA (fig. 4*B*). In our guide activity assay using cultured vertebrate cells, the human-specific substrate RNA became modified when expressed in both human (HeLa) and *Xenopus* (XTC) cell lines. At the same time, the *Xenopus*-specific substrate RNA became modified only in HeLa cells (supplementary table 2, Supplementary Material online). Importantly, endogenous *Xenopus* 28S rRNA is not pseudouridylated at the position equivalent to human 4975. How can we interpret these results? In mammals, SNORA64 is an additional guide RNA for this position (fig. 4*B*). This is a diverged copy of SNORA10 that is missing in nonmammalian species. Nonetheless, both human and *Xenopus*-specific sequences of 28S rRNA are good substrates for human SNORA64 (fig. 4*B*). Thus, a single G-to-A substitution in 28S rRNA in mammals generated a good substrate for a preexisting pseudouridylation pocket in SNORA22. Somehow this novel pseudouridine became beneficial, and mammals acquired an additional better guide RNA for the same position, SNORA64. It is worth noting that the position in *Xenopus* equivalent to human 28S-5011 is modified instead of 4975; the corresponding guide RNA pug*xt*28S-4066 has been identified in our snoRNA search (supplementary table 1 and supplementary fig. 1, Supplementary Material online). The recently annotated SNORA100 was predicted to modify 28S-5011 in human and mouse (Machyna et al. 2014). However, we could not detect expression of this snoRNA in either species. Moreover, modification itself has not been detected in these species (Stanley et al. 2016; Taoka et al. 2018; our data). These observations make us propose that the so-called SNORA100 is an snoRNA-like sequence representing a remnant of the actual guide RNA identified in *Xenopus*.

Recently, the 3′-terminal domain of SNORA22 and the 5′-terminal domain of SNORA61 were predicted to mediate pseudouridylation of position 918 in human 18S rRNA and the equivalent position in other species (Kehr et al. 2014; Jorjani et al. 2016; our predictions). For a long time, this highly conserved modification had no assigned guide RNA. Because the binding affinity of snoRNAs with 18S rRNA is not very strong (supplementary fig. 1, Supplementary Material online), we urge that predicted modification guide activities be tested experimentally. We expressed human and *Xenopus* SNORA22 and SNORA61 in yeast cells and observed the predicted pseudouridylation in an artificial substrate RNA that contained the corresponding fragment of human 18S rRNA (supplementary table 2, Supplementary Material online). These tests verified that SNORA22 and SNORA61 are genuine pseudouridylation guide RNAs for 18S rRNA. Remarkably, despite the relatively weak binding affinity of these snoRNAs with 18S rRNA, this guide activity is evolutionarily conserved in vertebrate SNORA22.

Thus, our data clearly demonstrate that miniscule changes in an antisense element may result in loss or gain of snoRNA modification guide activity, even when overall sequence conservation is preserved. Furthermore, 5′- and 3′-terminal domains in snoRNAs may evolve independently. Thus, one domain may remain intact during evolution whereas the other gains a novel modification activity or loses a conserved function. These events can even occur in a single species. We stress that careful target predictions and experimental verifications are essential steps in studies of snoRNA evolution.

Discussion

Previous evolutionary studies of snoRNAs focused mostly on the conservation and ancestral tracing of these RNAs (Hoeppner and Poole 2012; Kehr et al. 2014). The analysis of modification guide RNA specificities was carried out primarily on human snoRNAs (Lestrade and Weber 2006; Jorjani et al. 2016). In this study, we performed modification target prediction analysis independently on snoRNA sets from two diverged species: human and the frog *X. tropicalis*. Then, we compared our functional predictions across species. This approach allowed us to identify novel antisense elements in *Xenopus* snoRNAs from highly conserved snoRNA families; these antisense elements were predicted to target unexpected positions in rRNAs. Furthermore, we found several novel *Xenopus* snoRNAs that also target unusual positions in rRNAs and U6 snRNA (supplementary table 1 and supplementary fig. 1, Supplementary Material online). Three such nonmammalian box C/D snoRNAs had been identified earlier, but the assigned novel modifications had never been verified (Makarova and Kramerov 2009). We performed RNA modification mapping and found the predicted modifications in *Xenopus*, but not in human. Similarly, at least 14 snoRNAs are expressed in human but are absent in *X. tropicalis*, including nine unmodified positions in *Xenopus* rRNAs. In addition, ten highly conserved human snoRNAs contain antisense elements missing from *Xenopus* (fig. 8), resulting in nine additional modifications detected in human but not in *Xenopus*. Our interspecies analysis thus questions the evolutionary conservation of snoRNAs and the posttranscriptional modifications they mediate. An earlier bioinformatic analysis of fungal snoRNAs also revealed extensive lineage-specific gain and loss of entire snoRNA families and specific guide functions during fungal evolution (Canzler et al. 2018).

The evolutionarily stable association of snoRNAs with their host genes has been emphasized in mammalian genomic studies (Hoeppner et al. 2009). However, snoRNA distribution does not show perfect synteny between more diverged organisms. For instance, only one-third of orthologous chicken and human host genes have orthologous snoRNA sequences in the same introns and in the same order (Shao et al. 2009). Extending this analysis to *Xenopus* and other vertebrates, we always find differences between taxa in snoRNA order, position in a particular intron, and specific snoRNA family members within orthologous host genes (figs. 1*A*, 3*A*, 4*A*, 6*B*, and 7*B*). In addition, there are more snoRNA copies in *Xenopus* than in human (supplementary table 1, Supplementary Material online). A smaller number of gene copies was previously reported for box C/D snoRNAs in placental mammals (Makarova and Kramerov 2009), and we find the same trend for H/ACA snoRNAs. Thus, over evolutionary time snoRNA sequences move from intron to intron within the same gene and between different genes. The most plausible mechanisms for these intra- and intergenic snoRNA movements are recombination, segmental duplication, and retrotransposition. All of these mechanisms seem to be exploited (Weber 2006; Shao et al. 2009; Hoeppner and Poole 2012).

It is usually postulated that modifications in snRNAs and rRNAs are highly conserved. It is true that the majority of pseudouridylated and 2′-*O*-methylated positions concentrate in highly conserved and functionally important regions (Karijolich and Yu 2010; Taoka et al. 2018), but the specific modified positions within these regions may vary in different species. When we analyzed the pseudouridylation patterns of 18S rRNA from nine vertebrate species, we found that only half of the detected pseudouridines were present in all species (30 out of 57 positions; fig. 9 and supplementary fig. 4*A*, Supplementary Material online). Undoubtedly, such variation is common and more examples will be found when additional species are analyzed. We also noticed that loss of pseudouridylation activity was sometimes accompanied by an U-to-C transition at the corresponding position (fig. 9 and supplementary fig. 4*A* and *B*, Supplementary Material online). Although the ability of pseudouridine to base-pair with other bases is similar to that of uridine (Spenkuch et al. 2014), when modified RNA passes through the nanopore of an Oxford Nanopore sequencing device, pseudouridine induces the same or very similar current fluctuation as cytosine (supplementary fig. 4*C*, Supplementary Material online). This finding suggests that the U-to-C transition plays a compensatory role in the absence of pseudouridylation, and that pseudouridine has an additional hidden feature of structural importance. In our comparative analysis of RNA modifications, we were struck by unpredicted differences between closely related species. How can we explain the loss of SNORA29 and two

highly conserved pseudouridines, 18S-Ψ220 and 28S-Ψ45 only in human? Why do two pseudouridines positioned by SNORA28 and SNORA75 exist in *X. tropicalis* 18S rRNA but not in *X. laevis*? One possibility is that those pseudouridines play a role in temperature adaptation. Heat shock-inducible pseudouridines have been detected in yeast and human RNAs, U2 snRNA (Wu et al. 2011), and numerous mRNAs (Carlile et al. 2014; Lovejoy et al. 2014; Schwartz et al. 2014; Karijolich et al. 2015). Accordingly, the *X. tropicalis* habitat (23–28 °C) correlates with a higher level of pseudouridylation of rRNA (fig. 9) compared with that of *X. laevis* (16–22 °C), axolotl (12–20 °C), and newt (21 °C). In support of this hypothesis, the recent study of pseudouridines in *Trypanosoma brucei* found higher level of pseudouridylation in the bloodstream form of the parasite, which develops at elevated temperatures, compared with the procyclic form (Rajan et al. 2019). Thus, the adaptive role of RNA modification at certain positions could explain some unexpected differences in modification patterns between evolutionarily close but environmentally diverged species.

Another common notion is that conserved snoRNAs target orthologous sites for modification (Hoeppner and Poole 2012; Patra Bhattacharya et al. 2016), especially in vertebrates (Kehr et al. 2014). Indeed, about 73% of orthologous box C/D snoRNAs contain antisense elements that are well preserved from *Xenopus* to human. However, only 39% of box H/ACA snoRNAs are functionally conserved in the sense that homologous snoRNAs from different species possess antisense elements for modification of orthologous positions (supplementary table 1, Supplementary Material online). Furthermore, these numbers could be lower if we consider orphan domains in snoRNAs. Their unassigned potential antisense element sequences often vary significantly in different species. The usual assumption is that these highly diverged sequences do not experience any selective pressure because they are nonfunctional. Alternatively, they may mediate modification of nonorthologous positions in diverged RNA molecules. In support of this postulate, we identified genuine modification guide activities in conserved snoRNAs, which were previously thought to contain orphan domains. In many cases, these guide activities are species or taxon-specific. In addition, certain modification guide activities at orthologous positions may shuffle between nonorthologous snoRNAs. These are not rare or exceptional cases (fig. 8). We postulate that each domain within a particular snoRNA can evolve independently as a functional unit. It has been shown experimentally that an individual H-box or ACA-box stem upstream of a box C/D RNA sequence produces a stable RNA molecule (Bortolin et al. 1999). Furthermore, stable single-domain box H/ACA RNAs, namely AluACA RNAs, have been identified in human (Jády et al. 2012). Single hairpin pseudouridylation guide RNAs exist in Archaea (Dennis and Omer 2005) and are typical for trypanosomes (Rajan et al. 2019). The same is true for one half of a box C/D snoRNA (Kiss-László et al. 1998; Tycowski et al. 2004; Qu et al. 2011; our unpublished data).

Additionally, functional conservation of homologous snoRNAs is somewhat overestimated due to lack of experimental verification of predicted guide RNA modification activities. In most eukaryotic species, except for yeast, snoRNA substrate assignment is based on computational analysis and sequence complementarity predictions. Only a few of the predicted interactions have been experimentally tested so far (Tycowski et al. 1996, 1998; Jády and Kiss 2001; Zhao et al. 2002; Deryusheva and Gall 2009, 2013, 2017, 2018; Xiao et al. 2009; Qu et al. 2011). Our data clearly demonstrate that even a single nucleotide change may result in the generation of a novel antisense element in a highly conserved snoRNA and consequently induce a novel modification. Likewise, a one-nucleotide mismatch within a conserved antisense element may convert a functional domain into a nonfunctional one (Deryusheva and Gall 2018, 2019; this study). At the same time, some snoRNAs with weak antisense elements were found fully functional when they were tested experimentally (Deryusheva and Gall 2013, 2017, 2018; this study). That is, almost identical paralogous and orthologous snoRNAs, even when they cluster together and can be traced to a common ancestor, can function completely differently. Thus, the functional identity of snoRNAs from different species cannot be deduced from sequence conservation alone, but must be demonstrated directly by functional assays.

Our data discussed so far demonstrate that RNA modification patterns and snoRNA guide activities are more adaptable than is often assumed. Nevertheless, many snoRNAs show exceptional conservation across different taxa in their modification guide activities and/or in their genomic positions, being encoded in orthologous host genes and grouped there with the same sets of other snoRNAs. Examples include the well-characterized guide RNAs for multiple positions in the heavily modified branch point recognition region in spliceosomal U2 snRNA or for two adjacent positions in the 5′-terminal loop of U5 snRNA: SCARNA8, SCARNA15, SCARNA10. The conservation of these RNAs is evident in both vertebrate and invertebrate species. These combinations of multiple activities in one guide RNA prevent interference between different modified positions during the modification process (Deryusheva and Gall 2018). Similarly, evolutionarily well-preserved snoRNA clusters might be essential to regulate coordinated expression of functionally interconnected snoRNAs. Notably, when mammalian SNORA25 lost its 28S rRNA specificity, a "compensatory" guide activity on the nearby position was acquired by SNORA40, an snoRNA encoded by the same host gene (fig. 3). Thus, we propose that the positioning of cross-talking modifications is a major force for evolutionary conservation of snoRNAs. Functionally important but independent modifications can easily switch specific antisense elements between nonhomologous modification guide RNAs. Furthermore, minimal variations in snoRNA sequence may result in rapid change of guide specificity. Such flexibility of snoRNAs provides an adaptive advantage by introducing novel posttranscriptional modifications or erasing those that became detrimental. These changes of RNA modification patterns—either adaptive or neutral—can occur in one particular taxon or even in a single species.

Closing Remarks

It is commonly assumed that rRNA posttranscriptional modifications and the snoRNAs that mediate them are evolutionarily highly conserved. In fact, this is true for only a subset of modified positions and snoRNAs. The overall rate of gain and loss of function for guide RNA-mediated posttranscriptional modifications is profoundly underestimated. Although most evolutionary studies of snoRNAs focus on the comparative analysis of overall RNA sequence conservation, our data clearly demonstrate the urgent need for further studies on evolutionary conservation and variability of snoRNA guide specificities. Furthermore, experimental verification of predicted modifications and guide activities is an essential step in these studies.

Materials and Methods

RNA Preparation

RNA samples were obtained from the following species: human (HeLa cells from Yixian Zheng, Carnegie Institution; fibroblast cell line GM3814), mouse (3T3 cells from Yixian Zheng, Carnegie Institution), chicken (DF1 cells from Karen Beemon, Johns Hopkins University, and 5-day embryos from fertilized eggs purchased from Amazon), whiptail lizard *As. neomexicana* (total RNA and frozen embryos from Peter Baumann, Stowers Institute for Medical Research), frog *X. tropicalis* (liver and oocytes from females purchased from Xenopus 1), frog *X. laevis* (XTC cell line from Hélène Cousin, Boston University, liver and oocytes from females purchased from Xenopus 1), axolotl *A. mexicanum* (oocytes from females purchased from the Ambystoma Genetic Stock Center), newt *N. viridescens* (oocytes), zebrafish *D. rerio* (eggs, whole adult males and females of wild-type AB line from the Carnegie Institution fish facility). Trizol reagent was used to extract RNA. RNA purification was performed using the Direct-zol RNA MiniPrep kit (Zymo Research). RNA isolation from nuclei of *X. tropicalis* oocytes or GVs was described previously (Talhouarne and Gall 2014).

After ribosomal RNA depletion with the Ribozero-Gold kit (Epicenter), strand-specific libraries were generated according to the TruSeq total RNA sample preparation protocol (Illumina) using extracted human, mouse, chicken, zebrafish, and *Xenopus* RNAs. Sequencing was performed on an Illumina HiSeq 2000 sequencer as described (Talhouarne and Gall 2014, 2018). Raw RNA sequencing data were deposited at Sequence Read Archive (PRJNA302326, PRJNA479418).

Identification of snoRNA Sequences

A previously generated data set of *X. tropicalis* oocyte nucleus RNA-seq (SRX4895193) was mapped with TopHat (v2.0.7) to the xenTro9 genome assembly, and intronic regions were quantified using BEDtools (v2.15.0). To detect potential snoRNA sequences, we inspected three types of intronic sequences in the Integrative Genomics Viewer (Broad Institute, v2.4.5) and looked for snoRNA motifs: 1) the top 500 most abundant introns, 2) the most abundant regions of large introns (>1,000 bp) that were binned into 200-bp intronic fragments prior to quantification, and 3) all introns of conserved snoRNA host genes. For box C/D snoRNAs, we searched for box C (RUGAUGA) and box D (CUGA) motifs flanked by short inverted sequences. For box H/ACA snoRNAs, we screened for a "hairpin–H box (ANANNA) hinge–hairpin–ACA box tail" structure. Additionally, we performed BLAST searches for similarity with snoRNAs annotated in human and other vertebrate species.

Standard RACE-PCR (polymerase chain reaction) was used to verify 5′- and 3′-ends of newly identified *Xenopus* snoRNAs. The expression of these snoRNAs was confirmed by northern blot analysis. Total RNA was separated on 8% polyacrylamide-8M urea gels and transferred onto a nylon membrane (Zeta Probe, Bio-Rad). Hybridization and detection of digoxigenin-labeled probes were performed according to standard protocols. The presence of snoRNAs in other species was verified by examination of the recently published multispecies RNA-seq data sets (SRX4479788, SRX4479793, SRX4479799, SRX4479800, SRX4481492). The 100- or 150-bp single-end reads were mapped with TopHat (v2.0.7) to the following genomes: human (v19), mouse (v10), chicken (v5), zebrafish (v10), and *X. laevis* (v9.0). Expression of SNORA29 and SNORA100 in human and mouse was tested by northern blots.

To characterize snoRNA families, we generated cladograms based on sequence alignment using Clustal Omega with default parameters (Chojnacki et al. 2017). We chose Clustal Omega because it is fast, efficient, suitable for this type of analysis and has been widely employed. Orthologous sequences from different taxa were selected from publicly available databases containing snoRNA annotation, such as Rfam and Ensembl. To identify unannotated orthologous snoRNAs, we used the short sequence BLAST search with query sequences corresponding to conserved antisense elements and structural motifs. Hits were screened for typical snoRNA secondary structures.

RNA Modification Analysis

For RNA modification mapping, we used primer extension-based methods as previously described (Deryusheva and Gall 2009, 2017). In brief, CMC treatment of test RNA was used to map pseudouridines, whereas a low concentration of dNTP in the reaction mix was used for 2′-*O*-methylation mapping. 6-FAM-labeled oligos were used for primer extension; some oligos were previously described (Deryusheva et al. 2012; Deryusheva and Gall 2013, 2019), and the newly designed ones are depicted in supplementary figure 4, Supplementary Material online. The primer extension reactions were performed using AMV (New England Biolabs) or EpiScript (Epicentre) reverse transcriptases. Single-stranded DNA fragments were precipitated, dissolved in formamide, and separated on an ABI3730xl capillary electrophoresis instrument (Applied Biosystems). The GeneScan-500 LIZ Size Standard was added to each sample. The data were analyzed using GeneMapper software (Applied Biosystems). Also, we used Oxford Nanopore technology for native RNA sequencing. Poly(A) tails were added to total RNA using poly(U) polymerase (New England Biolabs), and the Direct RNA

sequencing kit (SQK-RNA001) was used to add sequencing adapters and tethers. MinKNOW software was used to run samples on MinION and to call bases. Sequences were aligned to *Xenopus* 18S, 5.8S, and 28S rRNAs and visualized using the IGV browser.

Guide RNA Activity Assays

We searched for potential antisense elements in all previously known and newly identified snoRNAs and screened their sequences for complementarity against U snRNAs and rRNAs. Each snoRNA in *Xenopus* and human snoRNA sets was analyzed individually. Multiple alternative configurations of each pseudouridylation pocket and alternative D/D′-box motifs with adjacent sequences were analyzed. Based on results of our previous functional tests of snoRNA modification guide activities (Deryusheva and Gall 2013, 2017, 2018), the minimal length of an antisense element was considered eight nucleotides for a potential pseudouridylation guide, and seven nucleotides for a 2′-*O*-methylation guide. Unpaired target uridines were placed within a pseudouridylation pocket in three conformations: U alone, UN, and UNN.

The predicted guide activities were verified using an in vivo yeast cell system. To express vertebrate snoRNAs in yeast cells, we amplified the coding sequences from genomic DNA and cloned them into YEplac181 [LEU2 2 μ] or YEplac195 [URA3 2 μ] vectors containing a GPD promoter, an RNT1 cleavage site, and an snR13 terminator (Huang et al. 2011). To express artificial substrate RNAs that contain fragments of human and *Xenopus* rRNAs or U2 snRNA, U87 scaRNA-based constructs were generated as previously described (Deryusheva and Gall 2013). The list of tested snoRNAs and their substrates is provided in supplementary table 2, Supplementary Material online. The expression constructs were introduced into yeast *S. cerevisiae* by the standard lithium acetate method. Wild type (BY4741) and several mutant strains (*pus1*Δ, *pus7*Δ, *pus4*Δ) were used. When guide RNA activity was tested on endogenous yeast RNAs, nontransformed cells served as a control. When guide RNA activity was tested on an artificial substrate RNA, yeast cells were transformed either with a substrate RNA construct alone as a control or with substrate RNA and test guide RNA constructs together. RNA was extracted from transformed and control yeast cells using a hot acid phenol method. Expression of exogenous RNAs was verified by northern blot analysis, and RNA modification assays were performed in several replicates. In addition, modification guide activities of some *Xenopus*-specific and human-specific snoRNAs were verified using vertebrate cell lines as previously described (Deryusheva and Gall 2019).

As modification mapping based on reverse transcription is only semiquantitative, we performed all reactions on control and experimental RNAs simultaneously. A modification activity was scored as positive if GeneMapper software called the peak corresponding to the predicted modification and if the peak height was at least 2- to 3-fold higher than the control base line trace (no guide RNA).

Supplementary Material

Supplementary material are available at *Molecular Biology and Evolution* online.

Acknowledgments

We are grateful to Allison Pinder for expert technical help with RNA modification mapping and deep sequencing. We thank Peter Baumann for RNA samples and frozen embryos of whiptail lizards from the Reptile Facility of the Stowers Institute for Medical Research. This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (Grant Number R01 GM33397 to J.G.G.). J.G.G. is an American Cancer Society Professor of Developmental Genetics.

References

Bartschat S, Kehr S, Tafer H, Stadler PF, Hertel J. 2014. snoStrip: a snoRNA annotation pipeline. *Bioinformatics* 30(1):115–116.

Bazeley PS, Shepelev V, Talebizadeh Z, Butler MG, Fedorova L, Filatov V, Fedorov A. 2008. snoTARGET shows that human orphan snoRNA targets locate close to alternative splice junctions. *Gene* 408(1–2):172–179.

Bortolin ML, Ganot P, Kiss T. 1999. Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. *EMBO J.* 18(2):457–469.

Canzler S, Stadler PF, Schor J. 2018. The fungal snoRNAome. *RNA* 24(3):342–360.

Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. 2014. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* 515(7525):143–146.

Chojnacki S, Cowley A, Lee J, Foix A, Lopez R. 2017. Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. *Nucleic Acids Res.* 45(W1):W550–W553.

Dennis PP, Omer A. 2005. Small non-coding RNAs in Archaea. *Curr Opin Microbiol.* 8(6):685–694.

Deryusheva S, Choleza M, Barbarossa A, Gall JG, Bordonne R. 2012. Post-transcriptional modification of spliceosomal RNAs is normal in SMN-deficient cells. *RNA* 18(1):31–36.

Deryusheva S, Gall JG. 2009. Small Cajal body-specific RNAs of *Drosophila* function in the absence of Cajal bodies. *Mol Biol Cell.* 20(24):5250–5259.

Deryusheva S, Gall JG. 2013. Novel small Cajal-body-specific RNAs identified in *Drosophila*: probing guide RNA function. *RNA* 19(12):1802–1814.

Deryusheva S, Gall JG. 2017. Dual nature of pseudouridylation in U2 snRNA: Pus1p-dependent and Pus1p-independent activities in yeasts and higher eukaryotes. *RNA* 23(7):1060–1067.

Deryusheva S, Gall JG. 2018. Orchestrated positioning of post-transcriptional modifications at the branch point recognition region of U2 snRNA. *RNA* 24(1):30–42.

Deryusheva S, Gall JG. 2019. scaRNAs and snoRNAs: are they limited to specific classes of substrate RNAs? *RNA* 25(1):17–22.

Freyhult E, Edvardsson S, Tamas I, Moulton V, Poole AM. 2008. Fisher: a program for the detection of H/ACA snoRNAs using MFE secondary structure prediction and comparative genomics—assessment and update. *BMC Res Notes.* 1:49.

Gardner EJ, Nizami ZF, Talbot CC, Gall JG. 2012. Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of *Xenopus tropicalis*. *Genes Dev.* 26(22):2550–2559.

Hertel J, Hofacker IL, Stadler PF. 2008. SnoReport: computational identification of snoRNAs with unknown targets. *Bioinformatics* 24(2):158–164.

Hoeppner MP, Poole AM. 2012. Comparative genomics of eukaryotic small nucleolar RNAs reveals deep evolutionary ancestry amidst ongoing intragenomic mobility. *BMC Evol Biol.* 12:183.

Hoeppner MP, White S, Jeffares DC, Poole AM. 2009. Evolutionarily stable association of intronic snoRNAs and microRNAs with their host genes. *Genome Biol Evol*. 1:420–428.

Huang C, Karijolich J, Yu YT. 2011. Post-transcriptional modification of RNAs by artificial box H/ACA and box C/D RNPs. *Methods Mol Biol*. 718:227–244.

Hughes DG, Maden B. 1978. The pseudouridine contents of the ribosomal ribonucleic acids of three vertebrate species. Numerical correspondence between pseudouridine residues and 2′-*O*-methyl groups is not always conserved. *Biochem J*. 171(3):781–786.

Jády BE, Darzacq X, Tucker KE, Gregory Matera A, Bertrand E, Kiss T. 2003. Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO J*. 22(8):1878–1888.

Jády BE, Ketele A, Kiss T. 2012. Human intron-encoded Alu RNAs are processed and packaged into Wdr79-associated nucleoplasmic box H/ACA RNPs. *Genes Dev*. 26(17):1897–1910.

Jády BE, Kiss T. 2001. A small nucleolar guide RNA functions both in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J*. 20(3):541–551.

Jorjani H, Kehr S, Jedlinski DJ, Gumienny R, Hertel J, Stadler PF, Zavolan M, Gruber AR. 2016. An updated human snoRNAome. *Nucleic Acids Res*. 44(11):5068–5082.

Karijolich J, Yi C, Yu YT. 2015. Transcriptome-wide dynamics of RNA pseudouridylation. *Nat Rev Mol Cell Biol*. 16(10):581–585.

Karijolich J, Yu YT. 2010. Spliceosomal snRNA modifications and their function. *RNA Biol*. 7(2):192–204.

Karpinka JB, Fortriede JD, Burns KA, James-Zorn C, Ponferrada VG, Lee J, Karimi K, Zorn AM, Vize PD. 2015. Xenbase, the *Xenopus* model organism database; new virtualized system, data types and genomes. *Nucleic Acids Res*. 43(D1):D756–D763.

Kehr S, Bartschat S, Stadler PF, Tafer H. 2011. PLEXY: efficient target prediction for box C/D snoRNAs. *Bioinformatics* 27(2):279–280.

Kehr S, Bartschat S, Tafer H, Stadler PF, Hertel J. 2014. Matching of soulmates: coevolution of snoRNAs and their targets. *Mol Biol Evol*. 31(2):455–467.

Kishore S, Gruber AR, Jedlinski DJ, Syed AP, Jorjani H, Zavolan M. 2013. Insights into snoRNA biogenesis and processing from PAR-CLIP of snoRNA core proteins and small RNA sequencing. *Genome Biol*. 14(5):R45.

Kiss-László Z, Henry Y, Kiss T. 1998. Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. *EMBO J*. 17(3):797–807.

Kiss AM, Jady BE, Bertrand E, Kiss T. 2004. Human box H/ACA pseudouridylation guide RNA machinery. *Mol Cell Biol*. 24(13):5797–5807.

Krogh N, Jansson MD, Häfner SJ, Tehler D, Birkedal U, Christensen-Dalsgaard M, Lund AH, Nielsen H. 2016. Profiling of 2′-*O*-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res*. 44(16):7884–7895.

Lestrade L, Weber MJ. 2006. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res*. 34(Database issue):D158–D162.

Lovejoy AF, Riordan DP, Brown PO. 2014. Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in *S. cerevisiae*. *PLoS One* 9(10):e110799.

Lowe TM, Eddy SR. 1999. A computational screen for methylation guide snoRNAs in yeast. *Science* 283(5405):1168–1171.

Machyna M, Kehr S, Straube K, Kappei D, Buchholz F, Butter F, Ule J, Hertel J, Stadler PF, Neugebauer KM. 2014. The coilin interactome identifies hundreds of small noncoding RNAs that traffic through Cajal bodies. *Mol Cell*. 56(3):389–399.

Maden B. 1986. Identification of the locations of the methyl groups in 18S ribosomal RNA from *Xenopus laevis* and man. *J Mol Biol*. 189(4):681–699.

Maden B. 1988. Locations of methyl groups in 28S rRNA of *Xenopus laevis* and man. Clustering in the conserved core of molecule. *J Mol Biol*. 201(2):289–314.

Maden BEH. 1990. The numerous modified nucleotides in eukaryotic ribosomal RNA. In: Cohn WE, Moldave K, editors. Progress in nucleic acid research and molecular biology. Vol. 39. San Diego: Academic Press. p. 241–303.

Maden BEH, Wakeman JA. 1988. Pseudouridine distribution in mammalian 18S ribosomal RNA. A major cluster in the central region of the molecule. *Biochem J*. 249(2):459–464.

Makarova JA, Kramerov DA. 2009. Analysis of C/D box snoRNA genes in vertebrates: the number of copies decreases in placental mammals. *Genomics* 94(1):11–19.

McCallum FS, Maden B. 1985. Human 18S ribosomal RNA sequence inferred from DNA sequence. Variations in 18S sequences and secondary modification patterns between vertebrates. *Biochem J*. 232(3):725–733.

Ofengand J, Bakin A. 1997. Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts. *J Mol Biol*. 266(2):246–268.

Patra Bhattacharya D, Canzler S, Kehr S, Hertel J, Grosse I, Stadler PF. 2016. Phylogenetic distribution of plant snoRNA families. *BMC Genomics* 17(1):969.

Qu G, van Nues RW, Watkins NJ, Maxwell ES. 2011. The spatial-functional coupling of box C/D and C′/D′ RNPs is an evolutionarily conserved feature of the eukaryotic box C/D snoRNP nucleotide modification complex. *Mol Cell Biol*. 31(2):365–374.

Rajan KS, Doniger T, Cohen-Chalamish S, Chen D, Semo O, Aryal S, Glick Saar E, Chikne V, Gerber D, Unger R. 2019. Pseudouridines on *Trypanosoma brucei* spliceosomal small nuclear RNAs and their implication for RNA and protein interactions. *Nucleic Acids Res*. 47(14):7633–7647.

Salim M, Maden B. 1980. Nucleotide sequence encoding the 5′ end of *Xenopus laevis* 18S rRNA. *Nucleic Acids Res*. 8(13):2871–2884.

Schattner P, Barberan-Soler S, Lowe TM. 2006. A computational screen for mammalian pseudouridylation guide H/ACA RNAs. *RNA* 12(1):15–25.

Schwartz S, Bernstein DA, Mumbach MR, Jovanovic M, Herbst RH, León-Ricardo BX, Engreitz JM, Guttman M, Satija R, Lander ES, et al. 2014. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell* 159(1):148–162.

Shao P, Yang J-H, Zhou H, Guan D-G, Qu L-H. 2009. Genome-wide analysis of chicken snoRNAs provides unique implications for the evolution of vertebrate snoRNAs. *BMC Genomics* 10:86.

Sharma S, Marchand V, Motorin Y, Lafontaine D. 2017. Identification of sites of 2′-*O*-methylation vulnerability in human ribosomal RNAs by systematic mapping. *Sci Rep*. 7(1):11490.

Spenkuch F, Motorin Y, Helm M. 2014. Pseudouridine: still mysterious, but never a fake (uridine)! *RNA Biol*. 11(12):1540–1554.

Stanley SE, Gable DL, Wagner CL, Carlile TM, Hanumanthu VS, Podlevsky JD, Khalil SE, DeZern AE, Rojas-Duran MF, Applegate CD, et al. 2016. Loss-of-function mutations in the RNA biogenesis factor *NAF1* predispose to pulmonary fibrosis–emphysema. *Sci Transl Med*. 8(351):351ra107.

Tafer H, Kehr S, Hertel J, Hofacker IL, Stadler PF. 2010. RNAsnoop: efficient target prediction for H/ACA snoRNAs. *Bioinformatics* 26(5):610–616.

Talhouarne GJS, Gall JG. 2014. Lariat intronic RNAs in the cytoplasm of *Xenopus tropicalis* oocytes. *RNA* 20(9):1476–1487.

Talhouarne GJS, Gall JG. 2018. Lariat intronic RNAs in the cytoplasm of vertebrate cells. *Proc Natl Acad Sci U S A*. 115(34):E7970–E7977.

Taoka M, Nobe Y, Yamaki Y, Sato K, Ishikawa H, Izumikawa K, Yamauchi Y, Hirota K, Nakayama H, Takahashi N, et al. 2018. Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res*. 46(18):9289–9298.

Taoka M, Nobe Y, Yamaki Y, Yamauchi Y, Ishikawa H, Takahashi N, Nakayama H, Isobe T. 2016. The complete chemical structure of *Saccharomyces cerevisiae* rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. *Nucleic Acids Res*. 44(18):8951–8961.

Torchet C, Badis G, Devaux F, Costanzo G, Werner M, Jacquier A. 2005. The complete set of H/ACA snoRNAs that guide rRNA pseudouridylations in *Saccharomyces cerevisiae*. *RNA* 11(6):928–938.

Tycowski KT, Aab A, Steitz JA. 2004. Guide RNAs with 5′ caps and novel box C/D snoRNA-like domains for modification of snRNAs in Metazoa. *Curr Biol*. 14(22):1985–1995.

Tycowski KT, Smith CM, Shu M-D, Steitz JA. 1996. A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in *Xenopus*. *Proc Natl Acad Sci U S A*. 93(25):14480–14485.

Tycowski KT, You ZH, Graham PJ, Steitz JA. 1998. Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol Cell*. 2(5):629–638.

Weber MJ. 2006. Mammalian small nucleolar RNAs are mobile genetic elements. *PLoS Genet*. 2(12):e205.

Wu G, Xiao M, Yang C, Yu YT. 2011. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J*. 30(1):79–89.

Xiao M, Yang C, Schattner P, Yu YT. 2009. Functionality and substrate specificity of human box H/ACA guide RNAs. *RNA* 15(1):176–186.

Yang J-H, Zhang X-C, Huang Z-P, Zhou H, Huang M-B, Zhang S, Chen Y-Q, Qu L-H. 2006. snoSeeker: an advanced computational package for screening of guide and orphan snoRNA genes in the human genome. *Nucleic Acids Res*. 34(18):5112–5123.

Yu Y-T, Terns RM, Terns MP. 2005. Mechanisms and functions of RNA-guided RNA modification. In: Grosjean H, editor. Fine-tuning of RNA functions by modification and editing. Topics in Current Genetics. Vol. 12. Berlin/Heidelberg: Springer. p. 223–262.

Yu YT, Shu M, Di Narayanan A, Terns RM, Terns MP, Steitz JA. 2001. Internal modification of U2 small nuclear (sn)RNA occurs in nucleoli of *Xenopus* oocytes. *J Cell Biol*. 152(6):1279–1288.

Zhao X, Li Z-H, Terns RM, Terns MP, Yu Y-T. 2002. An H/ACA guide RNA directs U2 pseudouridylation at two different sites in the branchpoint recognition region in *Xenopus* oocytes. *RNA* 8(12):1515–1525.