RNA BIOLOGY
2021, VOL. 18, NO. 11, 1760–1777
https://doi.org/10.1080/15476286.2020.1869892

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Mapping rRNA 2′-O-methylations and identification of C/D snoRNAs in Arabidopsis thaliana plants

J. Azevedo-Favory[a,b,*], C. Gaspin[c,d,*], L. Ayadi[e,f], C. Montacié[a,b], V. Marchand [e], E. Jobet[a,b], M. Rompais[g], C. Carapito[g], Y. Motorin [e,f], and J. Sáez-Vásquez [a,b]

[a]CNRS, Laboratoire Génome et Développement des Plantes (LGDP), UMR 5096, 66860 Perpignan, France; [b]Univ. Perpignan Via Domitia, LGDP, UMR5096, 66860 Perpignan, France; [c]Université Fédérale de Toulouse, INRAE, MIAT, 31326, Castanet-Tolosan, France; [d]Université Fédérale de Toulouse, INRAE, BioinfOmics, Genotoul Bioinformatics facility, 31326; [e]Université de Lorraine, CNRS, INSERM, IBSLor, (UMS2008/US40), Epitranscriptomics and RNA Sequencing (EpiRNA-Seq) Core Facility, F-54000 Nancy, France; [f]Université de Lorraine, CNRS, IMoPA (UMR7365), F-54000 Nancy, France; [g]Laboratoire de Spectrométrie de Masse BioOrganique, Institut Pluridisciplinaire Hubert Curien, UMR7178 CNRS/Université de Strasbourg, Strasbourg, France

ABSTRACT
In all eukaryotic cells, the most abundant modification of ribosomal RNA (rRNA) is methylation at the ribose moiety (2′-*O*-methylation). Ribose methylation at specific rRNA sites is guided by small nucleolar RNAs (snoRNAs) of C/D-box type (C/D snoRNA) and achieved by the methyltransferase Fibrillarin (FIB). Here we used the Illumina-based RiboMethSeq approach for mapping rRNA 2′-*O*-methylation sites in *A. thaliana* Col-0 (WT) plants. This analysis detected novel C/D snoRNA-guided rRNA 2′-*O*-methylation positions and also some orphan sites without a matching C/D snoRNA. Furthermore, immunoprecipitation of Arabidopsis FIB2 identified and demonstrated expression of C/D snoRNAs corresponding to majority of mapped rRNA sites. On the other hand, we show that disruption of Arabidopsis Nucleolin 1 gene (NUC1), encoding a major nucleolar protein, decreases 2′-*O*-methylation at specific rRNA sites suggesting functional/structural interconnections of 2′-*O*-methylation with nucleolus organization and plant development. Finally, based on our findings and existent database sets, we introduce a new nomenclature system for C/D snoRNA in Arabidopsis plants.

ARTICLE HISTORY
Received 13 October 2020
Revised 10 December 2020
Accepted 23 December 2020

KEYWORDS
2′-*O*-Methylation; C/D snoRNA; rRNA; fibrillarin; nucleolin; arabidopsis

Introduction

In all eukaryotic cells, ribosomal RNAs (rRNAs) precursors are substrate of two major types of nucleotide modifications: methylation of sugars (2′-*O*-ribose methylation) and isomerization of uridine to pseudouridine. Purine and pyrimidine rings in rRNAs can be also be methylated (m1N, m6N and m7N), aminocarboxypropilated (acp3N) and/or acetylated (ac4N) [1–4].

2′-*O*-methylation of rRNA might stabilize rRNA-mRNA, rRNA-tRNA or rRNA-protein interactions [5,6]. Furthermore, the significance of 2′-*O*-methylation of rRNAs is highlighted through studies in animal cells showing that alterations of rRNA 2′-*O*-methylation can be associated to diseases, mainly cancer and autoimmune syndromes. Indeed, ribosomes with altered rRNA 2′-*O*-methylation profile translate mRNA with lower fidelity and increase internal ribosome entry site (IRES)-dependent translation initiation of key cancer genes [4–9].

2′-*O*-ribose methylations are guided by small nucleolar RNAs (snoRNAs), referred as C/D-box snoRNA (C/D snoRNA). The box C (5′PuUGAUGA3′) and D (5′CUGA3′) of C/D snoRNAs are short consensus sequences that localize a few nucleotides away from the 5′- and 3′-ends, respectively. In the central part, the C/D snoRNA might contain also less conserved C' and D' motifs. One or two of so-called antisense elements are located upstream of the D and/or D' box. These antisense sequences are about 10–21 nucleotides long and are complementary to the region surrounding the site of 2′-*O*-ribose methylation. The rRNA nucleotide to be methylated is located precisely at the fifth position upstream from the D or D' box. The C/D snoRNA associates to four nucleolar proteins called the C/D-box core proteins: Fibrillarin/Nop1p, Nop56p/NOP56, Nop58p/NOP58 and snu13/L7Ae. Fibrillarin contains a characteristic SAM-binding methyltransferase motif required for the 2′-*O*-ribose methylation activity Reviewed in [10–13].

In plants, 2′-*O*-methylation of rRNAs has been demonstrated at specific rRNA sites and/or predicted throughout *in silico* and functional characterization of C/D snoRNAs. On one hand, in *Arabidopsis thaliana* plants, over one hundred of C/D snoRNAs were first identified by computational screening of genomic sequences [14,15] and a plant snoRNA database was created [16]. Later, sequencing of plant small RNA reported the expression of 151 C/D snoRNAs [17] and sequencing of nucleolar RNA fraction identified 9 additional C/D snoRNA candidates [18]. These results expanded the number

CONTACT J. Sáez-Vásquez saez@univ-perp.fr CNRS, Laboratoire Génome Et Développement Des Plantes (LGDP), Perpignan, France.
*First co-authors
Supplemental data for this article can be accessed here.

of known C/D snoRNAs in Arabidopsis plants to over 200 (including variants). Similarly, hundreds of C/D snoRNAs predicted to guide 2′-*O*-methylation of rRNA, were reported in *O. sativa* [19,20] and also other plant species [21]. Notably, studies in Arabidopsis demonstrate that knockout of a single snoRNA (HID2) triggers strong developmental and growing defects [22], while gene disruption of the C/D-box snoRNP assembly factor NUFIP inhibits 2′-*O*-methylation at specific rRNA sites and leads to severe developmental phenotypes [23]. On the other hand, over 125 2′-*O*-methylated rRNA sites have been predicted in Arabidopsis [24]. However, only about half of these potential rRNA modification sites has been mapped and/or validated [14,15,22,23,25]. This is essentially due to the limits of current experimental approaches, like RT primer extension at low [dNTP] used to determine the methylation state of a single target site at a time.

Here, we used RiboMethSeq approach for global mapping all rRNA 2′-*O*-methylation sites in leaves from *Arabidopsis thaliana* Col-0 (WT) and nucleolin 1 (*nuc1*) mutant plants. RiboMethSeq is based on the resistance of RNA at 2′-*O*-methylated sites to alkaline fragmentation and employs Illumina sequencing of cloned RNA fragments for simultaneous mapping and quantification of hundreds ribose methylated sites in RNA [26]. Furthermore, to identify C/D snoRNAs linked to the mapped 2′-*O*-methylated rRNA candidate sites, we performed bioinformatics screening of the Arabidopsis genome as well as experimental identification of expressed snoRNAs in a Fibrillarin immunoprecipitated fraction by small RNA-seq.

Results

rRNA 2′-*O*-methylation in *A. thaliana* leaves

To map rRNA 2′-*O*-methylated sites in *A. thaliana*, we extracted total RNA from Arabidopsis leaves (3 biological replicates, Figure S1) and performed RiboMethSeq analysis as previously reported [10,26].

After trimming and alignment of reads to the reference *A. thaliana* rRNA sequence [25–29], we found substantial differences between the reference sequences for 18S (GenBank X16077.1) and 25S rRNA (GenBank X52320.1) and the observed rRNA reads. These rRNA sequences were thus corrected to fit to the observed sequencing data (Supplementary Information and Methods). Then, in order to map all possible candidate sites, we performed the same approach as reported in [30]. We used threshold values for ScoreMEAN and Score A2, respectively 0.93 and 0.5; although in some cases, less strict ScoreMEAN limit (0.92 or lower) was applied. Combination of these parameters was found to give the best results for human rRNA having now well-established 2′-*O*-methylation profile and thus to limit the number of false-positive/false-negative hits. We also compared the obtained RiboMethSeq hits with previously known [14,25] or tentatively assigned locations and rRNA 2′-*O*-methylation list [24]. See Supplemental Information and Methods and Table S1 for details.

RiboMethSeq mapped 117 potential ribose methylated sites: 38 in the 18S rRNA, 2 in the 5.8S rRNA and 77 in the 25S rRNA sequences (Table 1 and S1 for details): Among the 117 RiboMethSeq mapped sites, 52 were previously mapped [14,15,22,23,25] while 25 others were only predicted from the sequences of C/D snoRNA guides [24] but not experimentally validated in the previous studies. Notably, RiboMethSeq also mapped 40 potential rRNA methylated sites not reported in [14,13,25] or predicted previously (Table S2). Nine (18S Um123, Cm1219 and 25S Cm2198, Am2257, Am2362, Gm2396, Um2494, Um2922 and Gm2923) of these 40 mapped candidates have an equivalent position in human, and for 31 of newly mapped positions we assigned the corresponding snoRNA. However, 7 candidate sites (18S Am812, Am1188 and Um1554 and 25S Um378, Cm2294, Gm2410 and Am2561) showing high RiboMethSeq signal still have no assigned snoRNA guide (Table 1).

To validate 2′-*O*-methylated candidates mapped by RiboMethSeq we performed additional orthogonal mapping of *A. thaliana* rRNA residues by high-throughput version of primer extension at low [dNTP]/low [Mg2+] conditions. The protocol was derived from published 2OMe-Seq [31], with minor modifications at the adapter ligation step and uses RT enzymes AMV and MMLV (Supplemental Information and methods). Using this technique we validated 97 sites over the 117 mapped by RiboMethSeq, including 5 rRNA sites without a corresponding C/D snoRNA guide (Figure S2 and Table S2).

All sites mapped by RiboMethSeq are shown in the predicted 18S and 25S/5.8S RNA secondary structures (Figure 1). All domains in the 18S, 25S and 5.8 rRNA sequences are 2′-*O*-methylated at different extents. Noteworthy, the 5′ domain in the 18S rRNA and the domain V in the 25S rRNA contain the highest number of 2′-*O*-methylation sites under our plant growth conditions.

Identification and characterization of C/D snoRNAs associated to Arabidopsis FIB2

To identify C/D snoRNAs that might guide 2′-*O*-methylation of mapped rRNA sites, we characterized C/D snoRNAs that co-purify with Arabidopsis Fibrillarin 2 (FIB2). We performed immunoprecipitation (IP) experiments in triplicate using Arabidopsis plants expressing the 35S:FIB2-YFP (Yellow Fluorescent Protein) construct. The FIB2-YFP is a 61kDa protein and localizes in the nucleolus (Figure S3(A-B) and [32,33]).

Following IP with antibodies against GFP, we analysed by Western blot whole cell extract (WCE-input, lanes 1 and 4), unbound (lanes 2, 5, 7 and 9) and co-immunoprecipitated (CoIP, lanes 3, 6, 8 and 10) protein fractions. We observed that anti-GFP antibodies immunoprecipitate FIB2-YFP protein in all three FIB2-YFP samples (CoIP_1 to 3), but not from Col-0 (CoIP_1 mock) protein extract (Figure S3C).

Firstly, we performed LC-MS/MS and differential analysis on FIB2-YFP and Col-0 IP fractions to verify that C/D snoRNP proteins co-immunoprecipitate with FIB2-YFP. A total of 197 proteins were specifically identified in fractions FIB2-YFP compared to the Col-0 CoIP fractions (Table S3). Arabidopsis orthologues of Fibrillarin, NOP56, NOP58 and L7Ae proteins were identified and the first three were among the most enriched (top 10) proteins. The genome of Arabidopsis contains two Fibrillarin (FIB1 and FIB2), NOP56 and NOP58 and at least five L7Ae protein genes

Table 1. List of RiboMethSeq detected 2′-*O*-methylation sites in *A. thaliana* 5.8S, 18S and 25S rRNAs. Mapped rRNA Nm sites ('Mapped site') are compared with previously known locations ('Position 3D rRNAdb' and associated C/D snoRNA [22]). Conservation of rRNA modification sites with human ('Position in human') and yeast ('Position in yeast') profile was determined based on 2D rRNA structure and sequence [8]. Human rRNA sites located in the same structural context, but not strictly conserved, are shown in grey. Newly detected snoRNA guides corresponding to the modified positions are shown ('Assigned snoRNA'). The absence of identified snoRNA guide is shown in grey. Reduction of 2′-*O*-methylation in *nuc1-2* plants is shown in columns 'ScoreC nuc1' with corresponding *p-value*. Asterisks ****(≤0.0001), ***(≤0.001), **(≤ 0.01), *(≤0.05) represent significance level and ns, non-significant value ≥ 0.05.

	rRNA	Mapped RiboMethSeq	Mapped published	Position, 3D rRNA db	snoRNA, 3D rRNA db	Position in human	Position in yeast	ScoreC nuc1 reduction	p-value	Assigned snoRNA			
1	5.8S	Am47	y	Am47	snoR9	-		0.000609127	***	At1gCDbox6.1	At5gCDbox6.2	-	-
2	5.8S	Gm79	y	Gm79	snoR39BY	Gm75		6.47E-06	****	At2gCDbox68.1	At2gCDbox68.2		
1	18S	Am28	y	Am28	AtU27	Am27	Am28	0.261456285	ns	At3gCDbox92.1	At3gCDbox92.2	At4gCDbox92.3	
2	18S	Cm38	n	Cm398				0.015982453	*	At3gCDbox87.1			
3	18S	Um123	n			Um121		0.000345542	***	At1gCDbox51.1			
4	18S	Am162	n	Am162	AtsnoR18	Am166		2.80281E-05	****	At4gCDbox105.1	At4gCDbox105.2		
5	18S	Um213	n					7.72863E-05	****	At1gCDbox33.1			
6	18S	Gm246	n					0.000240838	***	At5gCDbox141.1			
7	18S	Gm392	y	Gm390	AtsnoR30	Gm436		0.02012768	*	At5gCDbox138.1			
8	18S	Cm418	n	Cm416	AtU14	Cm462	Cm414	0.247486665	ns	At4gCDbox120.1	At4gCDbox120.2	At4gCDbox120.3	At4gCDbox120.4
9	18S	Am440	y	Am438	AtU16/AtsnoR15	Am484		0.056664405	ns	At1gCDbox32.1	At2gCDbox64.1		
10	18S	Am468	y	Am466	AtsnoR17	Am512		0.035529289	*	At1gCDbox46.1			
11	18S	Cm473	y	Cm471	AtU56	Cm517		0.409665359	ns	At3gCDbox102.1	At5gCDbox102.2		
12	18S	Am545	y	Am543	AtsnoR41Y/ AtsnoR43	Am590	Am541	0.008517727	**	At1gCDbox7.1*	At1gCDbox8.1*	At4gCDbox107.1	
13	18S	Um582	n	Um580	AtsnoR77Y	Um627	Um578	0.059605143	ns	At2gCDbox53.1	At2gCDbox53.2		
14	18S	Gm599	n	Gm597	AtU54	Gm644		0.002222822	**	At5gCDbox128.1			
15	18S	Um604	n					0.099999745	ns	At1gCDbox17.1	At1gCDbox17.2		
16	18S	Um615	n					0.181336315	ns	At3gCDbox78.1	At5gCDbox78.2		
17	18S	Am623	n	Am621	AtU36	Am668	Am619	0.295248869	ns	At5gCDbox127.1			
18	18S	Am780	n					0.068617102	ns	At5gCDbox122.1			
19	18S	Am796	n					0.00010828	***	At1gCDbox37.1			
20	18S	Am801	y	Am799	AtsnoR53Y		Am796	0.017023118	*	At1gCDbox36.1			
21	18S	Am812	n						ns	snoRNA not found			
22	18S	Am978	y	Am975	AtU59	Am1031	Am974	0.013456274	*	At1gCDbox22.1	At1gCDbox22.2		
23	18S	Um1013	y	Um1010	AtsnoR20.1			0.00012,7129	***	At3gCDbox99.1			
24	18S	Am1188	n					0.441879451	ns	snoRNA not found			
25	18S	Cm1219	n			Cm1272		0.000222391	***	At1gCDbox47.1			
26	18S	Um1235	n	Um1232	AtsnoR14	Um1288		2.81944E-06	****	At1gCDbox31.1	At1gCDbox31.2		
27	18S	Um1264	n					0.156858336	ns	At3gCDbox95.1	At3gCDbox95.2	At5gCDbox144.1	
28	18S	Um1266	y	Um1263	AtsnoR32			0.385921318	ns	At3gCDbox93.1	At4gCDbox93.2		
29	18S	Um1273	y	Um1270	AtU33/AtsnoR34	Um or Psi? 1326	Um1269	0.038906432	*	At1gCDbox27.1	At1gCDbox27.2		
30	18S	Gm1275	y	Gm1272	AtsnoR21	Gm1328	Gm1271	0.1109212	ns	At2gCDbox67.1	At2gCDbox67.2		
31	18S	Am1330	n	Am1327	AtsnoR32	Am1383		0.874209847	ns	At3gCDbox93.1	At4gCDbox93.2		
32	18S	Um1384	n	Um1381	AtU61	Um1442		0.005144689	**	At2gCDbox66.1			
33	18S	Gm1434	n	Gm1431	AtsnoR19	Gm1490	Gm1428	0.154113634	ns	At3gCDbox97.1	At5gCDbox97.2		
34	18S	Um1448	y	Um1445	AtsnoR19			0.000392463	***	At3gCDbox97.1	At5gCDbox97.2		
35	18S	Um1554	n					0.008120092	**	snoRNA not found			
36	18S	Am1579	n					0.246143245	ns	At3gCDbox95.1	At3gCDbox95.2		
37	18S	Cm1645	n	Cm1641	AtU43	Cm1703	Cm1639	0.06961813	ns	At1gCDbox19.1	At1gCDbox19.2		
38	18S	Am1758	y	Am1754	AtsnoR23/AtsnoR70/ AtsnoR12.1			0.009606764	**	At3gCDbox90.1	At4gCDbox111.1	At4gCDbox111.2	
1	25S	Um44	n					0.002444844	**	At5gCDbox125.1			
2	25S	Um48	n	Um48				0.00024019	***	At1gCDbox18.1	At1gCDbox18.2		
3	25S	Um144	n					0.001361127	**	At4gCDbox108.1			
4	25S	Um378	n					0.565722843	ns	snoRNA not found			
5	25S	Gm399	n	Gm398		Am398, Am400		0.000874608	***	At1gCDbox33.1			
6	25S	Am661	n	Am660	AtU18	Am1326	Am649	0.002801163	**	At3gCDbox77.1	At5gCDbox77.2		
7	25S	Cm675	n	Cm674	AtsnoR58Y	Cm1340	Cm663	0.001863058	**	At3gCDbox76.1	At5gCDbox76.2		

(*Continued*)

Table 1. (Continued).

	rRNA	Mapped RiboMethSeq	Mapped published	Position, 3D rRNA db	snoRNA, 3D rRNA db	Position in human	Position in yeast	ScoreC nuc1 reduction p-value		Assigned snoRNA			
8	25S	Um803	n					0.000249067	***	At1gCDbox31.1	At1gCDbox31.2		
9	25S	Gm814	y	Gm812	AtsnoR39Y	Gm1522	Gm805	0.000473233	***	At2gCDbox68.1	At2gCDbox68.2		
10	25S	Am816	y	Am814	AtU51	Am1524	Am807	0.001726194	**	At1gCDbox28.1	At1gCDbox28.2		
11	25S	Am826	n	Am824	AtsnoR77	Am1534	Am817	0.013932955	*	At3gCDbox91.1	At4gCDbox91.2		
12	25S	Am885	n	Am883	AtsnoR72Y		Am876	0.010638536	*	At2gCDbox73.1	At2gCDbox74.1	At2gCDbox74.2	At2gCDbox74.3
13	25S	Gm917	n	Gm915	AtU80	Gm1625	Gm908	0.031093264	*	At3gCDbox91.1	At4gCDbox91.2		
14	25S	Am945	n					0.000320149	***	At4gCDbox114.1	At4gCDbox114.2		
15	25S	Um1067	n					0.045524699	*	At4gCDbox107.1			
16	25S	Am1143	n	Am1140	AtU38	Am1871	Am1133	0.000579152	***	At1gCDbox2.1	At1gCDbox2.2	At4gCDbox2.3	
17	25S	Am1263	y	Am1260	AtsnoR22			8.26471E-05	****	At3gCDbox89.1	At4gCDbox89.2	At4gCDbox89.3	At4gCDbox89.4
18	25S	Um1278	y	Um1275	AtsnoR22			0.001101115	**	At3gCDbox89.1	At4gCDbox89.3	At4gCDbox89.3	At4gCDbox89.4
19	25S	Am1377	n					0.00599488	**	At3gCDbox102.1	At5gCDbox102.2		
20	25S	Cm1447	y	Cm1439	AtU24	Cm2351	Cm1437	0.003846846	**	At4gCDbox113.1	At5gCDbox113.2		
21	25S	Am1459	y	Am1451	AtU24	Am2363	Am1449	0.002105908	**	At4gCDbox113.1	At5gCDbox113.2		
22	25S	Gm1460	y	Gm1452	AtU24	Gm2364	Gm1450	0.145653216	ns	snoRNA not found			
23	25S	Cm1479	n					0.000130194	***	At1gCDbox23.1			
24	25S	Cm1518	y	Cm1510	AtU49	Cm2422		0.001840695	**	At3gCDbox85.1			
25	25S	Cm1847	n					7.91855E-05	****	At5gCDbox123.1			
26	25S	Cm1850	y	Cm1840	Z42			5.71437E-06	****	At2gCDbox69.1			
27	25S	Gm1855	y	Gm1845	AtU59			0.005039323	**	At1gCDbox22.1	At1gCDbox22.2		
28	25S	Cm1860	y	Cm1850	AtU55/AtsnoR15	Cm2804		1.50041E-05	****	At1gCDbox32.1	At2gCDbox65.1		
29	25S	Am1871	y	Am1861	AtsnoR33	Am2815		0.365959473	ns	At2gCDbox55.1			
30	25S	Um1892	y	Um1882	AtU34	Um2837	Um1888	0.00106097	**	At2gCDbox57.1	At2gCDbox58.1	At2gCDbox59.1	At5gCDbox124.1
31	25S	Um2114	n					0.002058348	**	At1gCDbox52.1			
32	25S	Gm2125	y	Gm2114	AtsnoR60	Gm3627		0.152002072	ns	At4gCDbox112.1	At5gCDbox112.2		
33	25S	Am2127	y	Am2116	AtsnoR12			1.49825E-05	****	At4gCDbox114.1	At4gCDbox114.2	At5gCDbox114.3	
34	25S	Cm2198	n			Cm3701		0.000237866	***	At3gCDbox86.1			
35	25S	Am2215	y	Am2204	AtU37	Am3718		0.949154016	ns	At1gCDbox12.1			
36	25S	Am2221	y	Am2210	AtU36a	Am3724	Am2220	0.000775621	***	At1gCDbox3.1	At1gCDbox3.2	At4gCDbox3.3	
37	25S	Gm2237	n					0.012020963	*	At1gCDbox3.1	At1gCDbox3.2		
38	25S	Am2257	n		AtU40	Am3760		0.833140865	ns	At4gCDbox104.1	At5gCDbox121.1		
39	25S	Am2282	y	Am2271	AtU15	Am3785	Am2281	0.957190634	ns	At3gCDbox103.1	At5gCDbox103.2	At5gCDbox103.3	
40	25S	Gm2289	y	Gm2278	AtU15	Gm3792	Gm2288	0.108252147	ns	At3gCDbox103.1	At5gCDbox103.2	At5gCDbox103.3	
41	25S	Cm2294	n					0.002019999	**	snoRNA not found			
42	25S	Am2322	y	Am2311	AtU30	Am3825		0.02232348	*	At3gCDbox83.1			
43	25S	Am2327	y	Am2316	AtsnoR44	Am3830		0.012534812	*	At5gCDbox16.4			
44	25S	Cm2338	y	Cm2327	AtsnoR44	Cm3841	Cm2337	0.017495006	*	At1gCDbox16.1	At1gCDbox16.2	At1gCDbox16.3	At5gCDbox16.4
45	25S	Am2362	n			Am3867		4.70613E-07	****	At2gCDbox63.1*			
46	25S	Cm2366	y	Cm2355	AtU53/AtsnoR37	Cm3869		0.000616952	***	At3gCDbox84.1	At3gCDbox88.1	At4gCDbox88.2	At4gCDbox88.3
47	25S	Gm2392	n	Gm2381				2.27238E-05	****	At5gCDbox137.1	At5gCDbox142.1		
48	25S	Gm2396	n			Gm3899		0.102961299	ns	snoRNA not found			
49	25S	Gm2410	n					3.17807E-06	****	snoRNA not found			
50	25S	Um2411	y	Um2400	AtsnoR53			1.91313E-05	****	At5gCDbox137.1	At5gCDbox142.1		
51	25S	Um2422	n	Um2411	AtsnoR37	Um3925	Um2421	5.03695E-06	****	At3gCDbox88.1	At4gCDbox88.2	At4gCDbox88.3	
52	25S	Um2456	n					0.000153676	***	At1gCDbox18.1			
53	25S	Um2494	n			Cm4054		1.55423E-05	****	At4gCDbox117.1	At4gCDbox118.1		
54	25S	Am2561	n						ns	snoRNA not found			
55	25S	Gm2620	y	Gm2610	AtU31/AtsnoR35	Gm4196	Gm2619	0.071590958	ns	At1gCDbox25.1	At1gCDbox25.2		
56	25S	Am2641	y	Am2631	AtsnoR68Y		Am2640	7.3765E-06	****	At2gCDbox71.1			
57	25S	Um2651	y	Um2641	AtsnoR10	Um4227		0.058787757	ns	At1gCDbox5.1	At5gCDbox5.2		
58	25S	Gm2652	y	Gm2642	?	Gm4228		8.40177E-05	****	snoRNA not found			
59	25S	Cm2683	n					2.0631E-05	****	At5gCDbox136.1			

(*Continued*)

Table 1. (Continued).

	rRNA	Mapped RiboMethSeq	Mapped published	Position, 3D rRNA db	snoRNA, 3D rRNA db	Position in human	Position in yeast	ScoreC nuc1 reduction p-value			Assigned	snoRNA	
60	25S	Um2736	n					0.002043458	**	At2gCDbox75.1			
61	25S	Gm2792	y	Gm2781	AtsnoR1		Gm2791	0.200622753	ns	At3gCDbox94.1			
62	25S	Gm2794	y	Gm2783	?	Gm4370	Gm2793	0.205258169	ns	At3gCDbox101.1			
63	25S	Gm2816	n	Gm2805	AtsnoR38Y	Gm4392	Gm2815	4.17794E-05	****	At3gCDbox100.1	At5gCDbox100.2		
64	25S	Cm2837	y	Cm2826	AtsnoR24			0.005432031	**	At4gCDbox115.1	At4gCDbox115.2	At4gCDbox115.3	At4gCDbox115.4
65	25S	Cm2880	y	Cm2869	AtU49/ZmU49	Cm4456		0.009929072	**	At2gCDbox54.1	At2gCDbox54.2	At4gCDbox54.3	
66	25S	Um2884	y	Um2873	AtsnoR64			0.000863839	***	At3gCDbox98.1			
67	25S	Am2912	n					0.614273202	ns	At5gCDbox139.1			
68	25S	Gm2918	n	Gm2907	AtsnoR34	Gm4494		0.178388459	ns	At1gCDbox13.1			
69	25S	Um2922	n			Um4498		0.305290635	ns	snoRNA not found			
70	25S	Gm2923	n			Gm4499		0.948521456	ns	snoRNA not found			
71	25S	Am2935	n					0.000782662	***	At4gCDbox105.1	At4gCDbox105.2		
72	25S	Am2947	n	Am2936	AtU29	Am4523	Am2946	0.000255634	***	At1gCDbox38.1			
73	25S	Cm2949	n	Cm2938	snoR69Y		Cm2948	0.002038487	**	At1gCDbox39.1			
74	25S	Um2954	y	Um2943	?			0.033711211	*	At2gCDbox56.1			
75	25S	Cm2960	y	Cm2949	AtU35	Cm4536	Cm2959	0.00734276	**	At1gCDbox20.1			
76	25S	Gm3292	n					0.001637291	**	At1gCDbox27.1	At1gCDbox27.2		
77	25S	Um3301	n					0.930282471	ns	At3gCDbox78.1	At5gCDbox78.2		

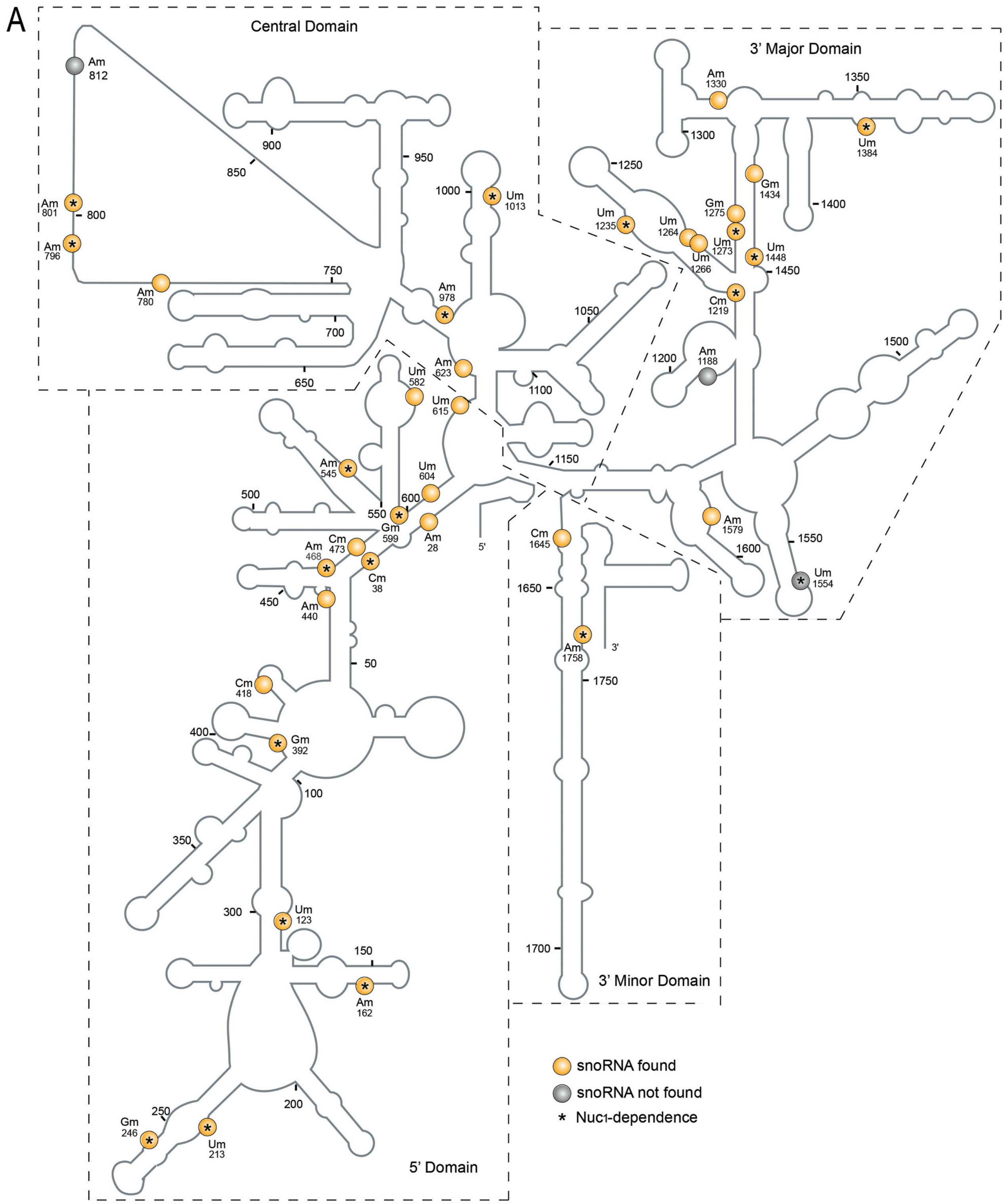

Figure 1. Predicted 18S, 5.8S and 25S rRNA structures with mapped 2′-*O*-methylations sites **(A)** 18S and **(B)** 5.8S/25S rRNA structures were generated on the basis of 2D structural predictions [24]. 2′-*O*-methylated sites in yellow circles or grey circles show rRNA sites with or without matching C/D snoRNAs (respectively). Star inside the circles indicates observed dependency of rRNA 2′-*O*-methylation on NUC1 gene expression. In B, the 5.8S rRNA structure is coloured in green and two 25S fragments (5′-fragment and 3′-fragment, a and b, respectively) are split in the main structure for clarity. The 18S (5′, central, 3′Major and 3′Minor) and 25S (I–VI) rRNA domains are designated according to [1]. P and A loop are indicated by arrows.

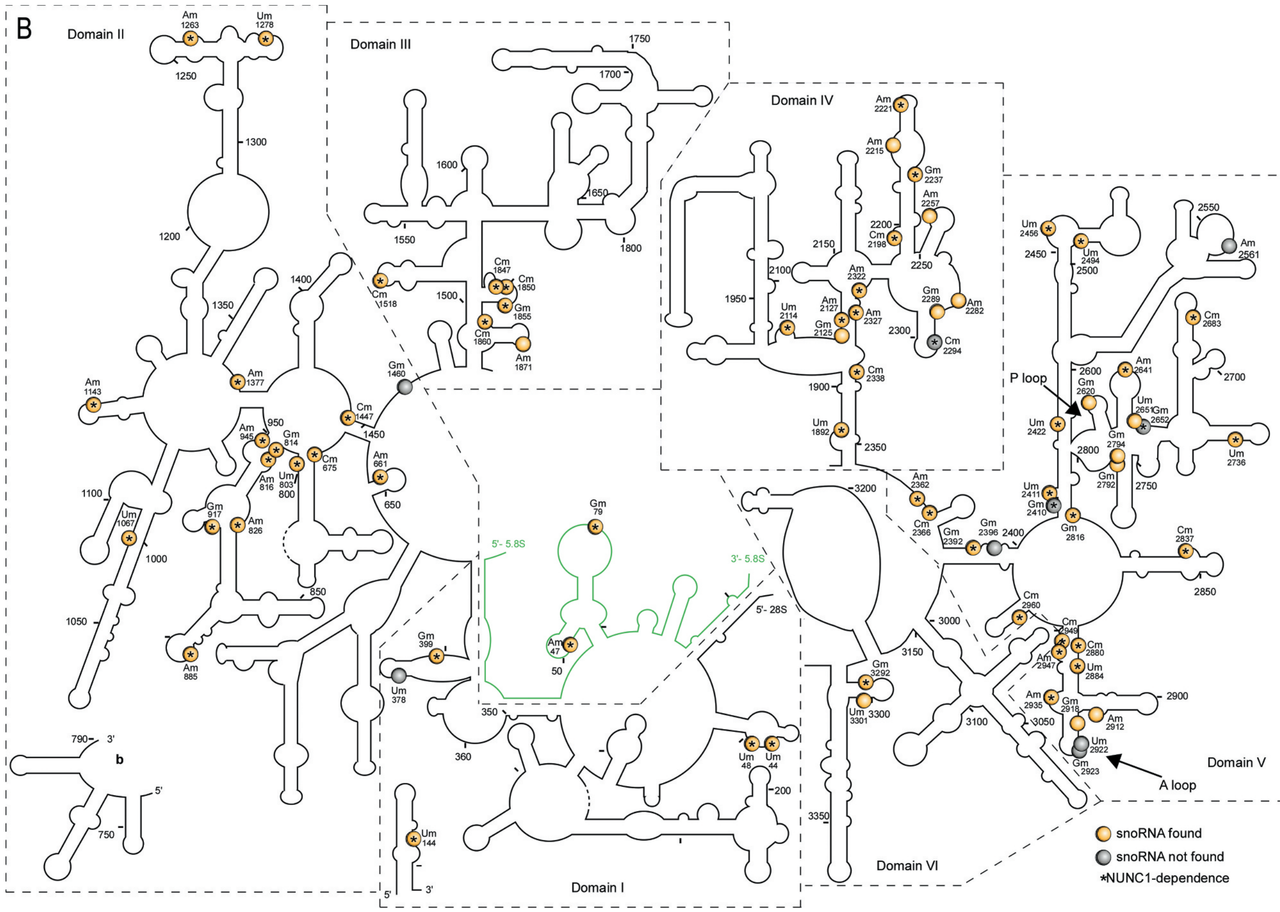

Figure 1. (Continued).

A

S. cerevisiae	*H. sapiens*	*A. thaliana*		Peptides / Coverage
Nop1	FBL/Fibrillarin	FIB1	At5g52470	22 / 62%
		FIB2	At4g25630	25 / 67%
Nop56/Sik1	NOP56	NOP56	At1g56110	23 / 48%
			At3g12860	15 / 26%
Nop58/5	NOP58	NOP58	At5g27120	19 / 42%
			At3g05060	28 / 53%
Snu13	15,5K/NHP2-like 1	L7Ae/L30e	At4g12600	1 / 12%
			At4g22380	ND
			At5g20160	ND
			At2g47610	4 / 18%
			At3g62870	N/D

141 C/D snoRNA	25S	18S	5.8S	Total
Predicted target (5')	66	41	0	107
Predicted target (3')	54	23	4	81
Unique target sites	**66**	**37**	**2**	**105**

B

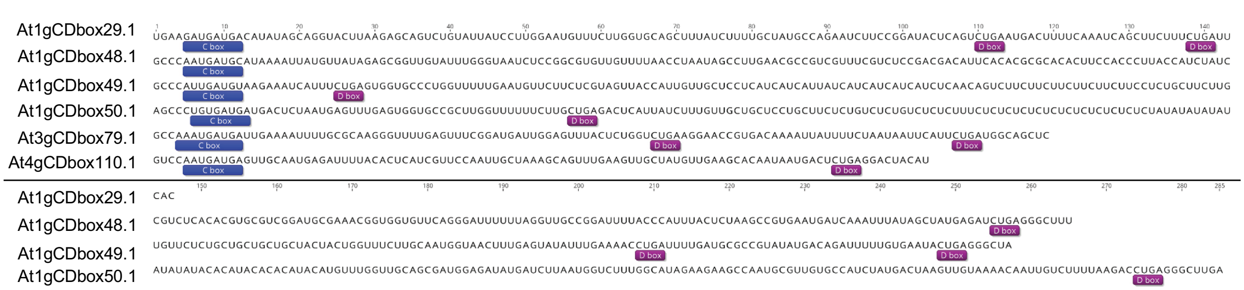

Figure 2. Characterization of FIB2-YFP snoRNP **(A)** Arabidopsis C/D-box protein orthologues (left) and C/D snoRNA (right) detected in the FIB2-YFP IP fractions using LC-MS/MS and RNA-seq approaches. The number of peptides and % of protein coverage for each protein is shown and the number of predicted 25S, 18S and 5.8S rRNA targets identified in the 5′ and 3′ antisense sequences of 141 C/D snoRNA. **(B)** Representation of 6 novel C/D snoRNA detected in the FIB2-IP fraction. C and D boxes are boxed. These C/D snoRNA do not contain a matching rRNA target sequence.

Table 2. List of C/D box snoRNAs identified from RNA-seq data. The number of C/D snoRNA reads are provided for each Co_IP fraction. Are indicated, rRNA positions mapped as 2′-*O*-methylated in this study, when predicted to be guided by a snoRNA and the names of known C/D snoRNA in databases Are highlighted in red snoRNAs with predicted rRNA methylation sites but not detected by RiboMethSeq. ID with an asterisk identifies a snoRNA which is dicistronic with a tRNA. ID with two asterisks represents a snoRNA which is polycistronic with a miRNA.

Box C/D sRNA	Known Id	CoIP_1	CoIP_2	CoIP_3	Predicted target (5′)	Predicted target (3′)
At1gCDbox1.1	AtsnoR6-1	29,792	22,096	21,345		
At1gCDbox3.1	AtU36a-1	3110	3338	3171	Gm2237_25S	Am2221_25S
At1gCDbox5.1	AtsnoR10-1	176,840	127,190	126,633	Um2651_25S	
At1gCDbox6.1	AtsnoR9-1	6842	5905	5520		Am47_5.8S
At1gCDbox3.2	AtU36a-2	2534	2350	2117	Gm2237_25S	Am2221_25S
At1gCDbox1.2	AtsnoR6-2	30,690	23,145	22,939		
At1gCDbox7.1*	SnoR43.2	5422	4728	4674	Am545_18S	
At1gCDbox8.1*	SnoR43.1	197,481	168,698	168,566	Am545_18S	
At1gCDbox11.1	U3	945	857	999		
At1gCDbox12.1	AtU37	8944	7347	7129	Am2215_25S	
At1gCDbox13.1	AtsnoR34	191,705	149,781	141,542	Gm2918_25S	
At1gCDbox14.1	AtsnoRNA_R38	291	254	240		
At1gCDbox15.1	snoR122	117	136	137		
At1gCDbox16.1	AtU79-1b ou AtsnoR44-1b	31,030	26,453	26,368	Cm2338_25S	Gm2327_25S
At1gCDbox16.2	AtU79-1a ou AtsnoR44-1a	53,154	47,539	46,891	Cm2338_25S	Gm2327_25S
At1gCDbox17.1	Z267	6378	5913	6178	Um604_18S	
At1gCDbox18.1	AtsnoR16-1	23,730	20,825	19,757	Um2456_25S	Um48_25S
At1gCDbox19.1	AtU43-1	86,681	72,783	73,809	Cm1645_18S	
At1gCDbox20.1	AtU35	30,832	23,099	22,599	Cm2960_25S	
At1gCDbox21.1	AtsnoR101	9314	6949	6911		
At1gCDbox22.1	AtU59b	9988	8160	8183	Am978_18S	Gm1855_25S
At1gCDbox22.2	AtU59a	125,292	106,451	112,345	Am978_18S	Gm1855_25S
At1gCDbox23.1	AtsnoTAIRsnoRNA6	78,189	68,892	64,353		Cm1479_25S
At1gCDbox24.1	AtsnoR105	17,441	14,677	14,231		
At1gCDbox25.1	AtU31a/HID2	43,635	42,762	39,045	Gm2620_25S	
At1gCDbox26.1	AtsnoR4a	167,211	140,369	132,738		
At1gCDbox27.1	AtU33a	90,503	71,938	72,290	Gm3292_25S	Um1273_18S
At1gCDbox28.1	AtU51a	114	92	73		Am816_25S
At1gCDbox25.2	AtU31b	9205	8916	8560	Gm2620_25S	
At1gCDbox26.2	AtsnoR4b	17,348	14,959	14,460		
At1gCDbox27.2	AtU33b	18,424	13,612	13,136	Gm3292_25S	Um1273_18S
At1gCDbox28.2	AtU51b	84	57	56		Am816_25S
At1gCDbox29.1		376	287	302		
At1gCDbox31.1	AtsnoR14-2	633	459	461	Um1235_18S	Um803_25S
At1gCDbox32.1	AtsnoR15	11,650	10,327	10,208	Am440_18S	Cm1860_25S
At1gCDbox9.2*	SnoR43.10	2331	1963	2095		
At1gCDbox9.3*	SnoR43.9	58	78	74		
At1gCDbox33.1	AtsnoR65	40,967	35,353	32,666	Um213_18S	Gm399_25S
At1gCDbox34.1	snoR113	9635	8275	8385		
At1gCDbox35.1	snoR114	7223	4748	5284		
At1gCDbox36.1	AtsnoR53Y	7874	7704	7403	Am801_18S	
At1gCDbox37.1	AtsnoR25	4936	4092	3856	Am796_18S or Cm797_18S	
At1gCDbox38.1	AtU29	22,365	21,744	21,262	Am2947_25S	
At1gCDbox39.1	AtsnoR69Y	19,662	17,455	16,644	Cm2949_25S	
At1gCDbox40.1*	SnoR43.7	58,869	50,112	49,216		
At1gCDbox41.1*	SnoR43.8	178,707	139,589	137,628		
At1gCDbox43.1*	SnoR43.6	1257	1332	1410		
At1gCDbox24.2	Z279 ou snoR105	2491	2134	2056		
At1gCDbox44.1	AtsnoR102	1798	1743	1714		
At1gCDbox45.1	snoR17	2226	1733	1681		
At1gCDbox46.1	AtsnoR17	74,719	57,795	57,805		Am468_18S
At1gCDbox16.3	AtU79-2 ou AtsnoR44-2	197,875	174,821	167,417	Cm2338_25S	Gm2327_25S
At1gCDbox17.2	Z267	52,829	42,299	43,150	Um604_18S	
At1gCDbox18.2	AtsnoR16-2	12,596	10,424	10,348		Um48_25S
At1gCDbox19.2	AtU43-2	91,375	73,859	73,358	Cm1645_18S	
At1gCDbox47.1	At_snoTAIRsnoRNA16	37,870	30,906	29,897	Cm1219_18S	
At1gCDbox48.1		13,102	10,348	10,696		
At1gCDbox49.1		284	277	263		
At1gCDbox50.1		198	134	148		
At1gCDbox51.1	AtsnoTAIRsnoRNA16	193,801	143,224	141,301	Um123_18S	
At1gCDbox52.1	AtsnoTAIRsnoRNA17	34,154	24,685	24,197	Um2114_25S	
At2gCDbox53.1	At77Y-2	216,644	171,303	167,532	Um582_18S	
At2gCDbox54.1	AtU49-2	96,660	80,342	79,563	Cm2880_25S	
At2gCDbox54.2	AtU49-1	171	148	141	Cm2880_25S	
At2gCDbox53.2	At77Y-1	1593	1398	1616	Um582_18S	
At2gCDbox56.1	SnoRNA R4	729	598	636	Um2954_25S	
At2gCDbox57.1	AtU34c	49,130	40,248	38,388	Um1892_25S	
At2gCDbox58.1	AtU34b	3809	3712	3848	Um1892_25S	
At2gCDbox59.1	AtU34a	98,590	90,168	83,378	Um1892_25S	
At2gCDbox60.1**	AtsnoTAIRsnoRNA19	137,218	125,343	125,025		
At2gCDbox61.1**	AtsnoTAIRsnoRNA20	7490	6932	7099		
At2gCDbox62.1*	AtsnoTAIRsnoRNA21	6346	4876	5005		
At2gCDbox63.1*	AtsnoTAIRsnoRNA22	19,528	18,462	18,567	Am2362_25S	
At2gCDbox62.2*	AtsnoTAIRsnoRNA23	221,115	149,030	154,503		
At2gCDbox64.1	AtU16	62,676	57,905	55,828	Am440_18S	
At2gCDbox65.1	AtU55	60,488	54,352	52,021		Cm1860_25S

(*Continued*)

Table 2. (Continued).

Box C/D sRNA	Known Id	CoIP_1	CoIP_2	CoIP_3	Predicted target (5′)	Predicted target (3′)
At2gCDbox31.2	AtsnoR14-1	118,361	101,090	100,066	Um1235_18S	Um803_25S
At2gCDbox66.1	AtU61	13,714	9708	10,316		Um1384_18S
At2gCDbox67.1	AtsnoR21b	21,923	17,104	16,779	Gm1275_18S	
At2gCDbox68.1	At39BYb	2212	1817	1773	Gm814_25S	Gm79_5.8S
At2gCDbox69.1	AtsnoTAIRsnoRNA24	301,655	212,851	212,975	Cm1850_25S	
At2gCDbox67.2	AtsnoR21a	208,675	175,574	165,999	Gm1275_18S	
At2gCDbox68.2	At39BYa	216,657	191,596	191,892	Gm814_25S	Gm79_5.8S
At2gCDbox70.1	AtsnoR68Y	9865	10,468	10,162		
At2gCDbox71.1	AtsnoR27	6920	6760	6127	Am2641_25S	
At2gCDbox72.1	AtsnoR26	83,519	60,542	61,582		
At2gCDbox73.1	At72Ya	5141	4911	4638		Am885_25S
At2gCDbox74.1	At72Yb	6337	6026	5710		Am885_25S
At2gCDbox74.2	At72Yc	10,327	8427	8361		Am885_25S
At2gCDbox74.3	At72Yd	11,629	10,050	9554		Am885_25S
At2gCDbox75.1	AtsnoR68	37,403	36,963	35,138		Um2736_25S
At3gCDbox76.1	At58Y-2	104,547	99,295	91,003	Cm675_25S	
At3gCDbox77.1	AtU18-2	157,637	123,863	127,897	Am661_25S	
At3gCDbox78.1	AtsnoR13-2	88,121	75,126	68,762	Um615_18S	Um3301_25S
At3gCDbox79.1		25,488	24,764	24,198		
At3gCDbox80.1	AtsnoTAIRsnoRNA25	13,804	11,156	10,920		
At3gCDbox81.1	AtsnoTAIRsnoRNA26	6621	5725	5723		
At3gCDbox82.1	AtsnoTAIRsnoRNA27	67,825	51,683	50,940		
At3gCDbox83.1	AtU30	17,636	16,160	15,596	Am2322_25S	
At3gCDbox85.1	AtsnoTAIRsnoRNA28	104,257	76,646	79,098		Cm1518_25S
At3gCDbox86.1	AtsnoTAIRsnoRNA29	40,571	34,299	34,742	Cm2198_25S	
At3gCDbox87.1	AtsnoR66	40,877	34,380	32,516	Cm38_18S	
At3gCDbox88.1	AtsnoR37-1	24,265	18,831	18,854	Um2422_25S	Cm2366_25S
At3gCDbox89.1	AtsnoR22-1	2580	2509	2564	Am1263_25S	Um1278_25S
At3gCDbox90.1	AtsnoR23-1	12,037	9239	9142		Am1758_18S
At3gCDbox91.1	AtU80-1	7413	5920	5890	Am826_25S	Gm917_25S
At3gCDbox92.1	AtU27-2	23,212	18,833	19,003	Am28_18S	
At3gCDbox92.2	AtU27-1	88,895	73,932	71,541	Am28_18S	
At3gCDbox93.1	AtsnoR32-1	23,849	21,273	20,109	Am1330_18S	Um1266_18S
At3gCDbox94.1	AtsnoR1a	30,339	24,173	23,591	Gm2792_25S	
At3gCDbox95.1	AtsnoR8a	114,537	96,738	95,050	Am1579_18S	Um1264_18S
At3gCDbox94.2	AtsnoR1b	6269	5347	5288	Gm2792_25S	
At3gCDbox95.2	AtsnoR8b	114,537	96,738	95,050	Am1579_18S	Um1264_18S
At3gCDbox96.3	AtsnoR72c	36	41	40		
At3gCDbox97.1	AtsnoR19-2	101	98	86	Gm1434_18S	Um1448_18S
At3gCDbox98.1	AtsnoR64	4272	3732	3905		Um2884_25S
At3gCDbox99.1	AtsnoR20-1	300,894	227,990	221,098		Um1013_18S
At3gCDbox100.1	AtsnoR38Y-1	102,352	72,888	66,762	Gm2816_25S	
At3gCDbox102.1	AtsnoR7-2	34,720	29,003	27,824	Cm473_18S	Am1377_25S
At3gCDbox103.1	AtU15-2	71,018	54,889	52,322	Gm2289_25S	Am2282_25S
At3gCDbox104.1	AtsnoTAIRsnoRNA30	874	755	712		Am2257_25S
At3gCDbox105.1	AtsnoR18a	32,569	25,911	25,280	Am2935_25S	Am162_18S
At3gCDbox105.2	AtsnoR18b	119,756	101,487	93,449	Am2935_25S	Am162_18S
At4gCDbox11.3	U3	16	17	14		
At4gCDbox107.1	At41Y	75,750	56,450	56,188	Um1067_25S	Am545_18S
At4gCDbox108.1	AtsnoR36	94,730	75,162	70,439		Um144_25S
At4gCDbox1.3	AtsnoR6-3	102,050	85,752	80,379		
At4gCDbox2.3	AtU38-3	117,493	97,592	94,243	Am1143_25S	
At4gCDbox3.3	AtU36a-4	12,645	12,449	12,223		Am2221_25S
At4gCDbox54.3	AtU49-3	32	20	18	Cm2880_25S	
At4gCDbox110.1		18,344	15,554	16,083		
At4gCDbox93.2	AtsnoR32-2	1727	1580	1531	Am1330_18S	Um1266_18S
At4gCDbox92.3	AtsnoTAIRsnoRNA32	3145	2915	3005	Am28_18S	
At4gCDbox91.2	AtU80-2	1892	1340	1414	Am826_25S	Gm917_25S
At4gCDbox89.2	AtsnoR22-2	249	245	283	Am1263_25S	Um1278_25S
At4gCDbox111.1	AtsnoR23-2	3739	3730	3522		Am1758_18S
At4gCDbox111.2	AtsnoR23-3	19,137	17,750	17,045		Am1758_18S
At4gCDbox89.3	AtsnoR22-3b	174	194	224	Am1263_25S	Um1278_25S
At4gCDbox89.4	AtsnoR22-3a	814	722	770	Am1263_25S	Um1278_25S
At4gCDbox112.1	AtU60.2 F	5027	4993	4957	Gm2125_25S	
At4gCDbox113.1	AtU24-1	17,050	12712	12,613	Am1459_25S	Cm1447_25S
At4gCDbox114.1	AtsnoR12-1b	80,896	52,758	53,486	Am2127_25S	Am945_25S
At4gCDbox114.2	AtsnoR12-1a	39,043	26,538	26,191	Am2127_25S	Am945_25S
At4gCDbox115.1	AtsnoR24d	10,657	8288	8309		Cm2837_25S
At4gCDbox115.2	AtsnoR24c	6282	4477	4456		Cm2837_25S
At4gCDbox115.3	AtsnoR24b	18,253	13,263	13,022		Cm2837_25S
At4gCDbox115.4	AtsnoR24a	12,920	9877	9467		Cm2837_25S
At4gCDbox44.2	AtsnoR102-2	420	355	379		
At4gCDbox117.1	AtsnoTAIRsnoRNA33	70,721	61,804	60,659	Um2494_25S	
At4gCDbox118.1	AtsnoTAIRsnoRNA34	80,587	65,550	65,323	Um2494_25S	
At4gCDbox120.1	AtU14a	2246	1831	1760		Cm418_18S
At4gCDbox120.2	AtU14b	852	693	758		Cm418_18S
At4gCDbox120.3	AtU14c	2363	1828	1793		Cm418_18S
At4gCDbox120.4	AtU14d	8375	6556	6362		Cm418_18S

(*Continued*)

Table 2. (Continued).

Box C/D sRNA	Known Id	CoIP_1	CoIP_2	CoIP_3	Predicted target (5′)	Predicted target (3′)
At5gCDbox121.1	AtU40-2	856	845	765		Am2257_25S
At5gCDbox122.1	AtsnoTAIRsnoRNA35	40,647	35,080	36,274	Am780_18S	
At5gCDbox125.1	AtsnoTAIRsnoRNA36	9618	9481	8923		Um44_25S
At5gCDbox126.1	AtsnoTAIRsnoRNA37	106,792	92,982	92,873		
At5gCDbox127.1	AtU36	78,308	60,602	60,263		Am623_18S
At5gCDbox78.2	AtsnoR13-1	12,081	10,646	9726	Um615_18S	Um3301_25S
At5gCDbox77.2	AtU18-1	179,675	139,901	136,562	Am661_25S	
At5gCDbox128.1	AtU54	31,942	29,406	27,817	Gm599_18S	
At5gCDbox76.2	At58Y-1	82,390	80,850	77,589	Cm675_25S	
At5gCDbox130.1	snoR106	63,042	59,937	59,720		
At5gCDbox131.1	atsnoR106b	44,824	40,029	40,207		
At5gCDbox132.1	AtsnoR28-1 c	5120	4176	4128		
At5gCDbox132.2	AtsnoR28-1b	11,497	10,145	9652		
At5gCDbox133.1	AtsnoTAIRsnoRNA28	68,747	61,581	57,194		
At5gCDbox134.1	AtsnoTAIRsnoRNA39	869	778	645		
At5gCDbox132.3	SnoR28-2b	20	12	2		
At5gCDbox24.3	AtsnoR108	1692	1400	1241		
At5gCDbox136.1	AtsnoTAIRsnoRNA40	463	397	355		Cm2683_25S
At5gCDbox6.2	AtsnoR9-2	13,136	13,580	12,599		Am47_5.8S
At5gCDbox5.2	AtsnoR10-2	58,815	50,880	49,942	Um2651_25S	
At5gCDbox137.1	AtsnoR29-1	291,496	222,950	215,519	Um2411_25S	Gm2392_25S
At5gCDbox138.1	AtsnoR30	156,956	127,555	122,733	Gm392_18S	
At5gCDbox139.1	AtsnoR31	71,940	54,894	48,793	Am2912_25S	
At5gCDbox112.2	AtU60.1 F	22,146	19,138	19,022	Gm2125_25S	
At5gCDbox11.4	U3	658	619	635		
At5gCDbox11.5	U3	175	154	154		
At5gCDbox11.8	U3	1119	1093	1152		
At5gCDbox114.3	AtsnoR12-2	331	357	332	Am2127_25S	
At5gCDbox100.2	SAtsnoR38Y-2	7651	8224	7944	Gm2816_25S	
At5gCDbox140.1	AtsnoR20-2	438	289	325		
At5gCDbox97.2	AtsnoR19-1	23	24	28	Gm1434_18S	Um1448_18S
At5gCDbox141.1	AtsnoTAIRsnoRNA42	8447	8529	8296	Gm246_18S	
At5gCDbox16.4	AtsnoR79-3	17,608	15,165	14,993	Cm2338_25S	Am2327_25S
At5gCDbox142.1	AtsnoR29-2	539	402	378	Um2411_25S	Gm2392_25S
At5gCDbox143.1	AtsnoTAIRsnoRNA43	19,162	15,772	15,927		
At5gCDbox144.1	AtsnoR67	93,237	79,011	80,230		Um1264_18S
At5gCDbox103.2	AtU15-1a	12,819	11,273	11,365	Gm2289_25S	Am2282_25S
At5gCDbox103.3	AtU15-1b	6327	6159	5853	Gm2289_25S	Am2282_25S
At5gCDbox102.2	AtsnoR7-1	5143	3915	3799	Cm473_18S	Am1377_25S

(Figure 2). LC-MS/MS data did not allow precise discrimination of orthologue(s) co-immuno-precipitating with FIB2-YFP, but clearly detected at least one orthologue of each protein factor. Altogether, these data indicate that FIB2-YFP interacts with C/D-box snoRNP core proteins and likely forms a functional C/D-box snoRNP complex.

Next, to identify C/D snoRNAs in the FIB2-YFP fractions (CoIP_1 to 3), we performed RNA-seq analysis. Total RNAs were extracted from FIB2-YFP CoIP fractions, converted to library and sequenced. Between 7–8.7 × 10^6 C/D snoRNA reads are detected in each RNA samples (Figure S3D) and count for a total of 193 C/D snoRNAs (with at least 10 reads in one of the CoIP fractions; Table 2). Among these, 187 were known C/D snoRNAs while the other 6 were novel species. Noticeable, 141 C/D snoRNAs target respectively 66, 37 and 2 sites in the 25S, 18S and 5.8S rRNA sequences while the other 52 C/D snoRNAs do not target rRNAs sites identified by RiboMethSeq (Fig. 2B and Table 2 and S4).

Up to four C/D snoRNA might target each of the 105 ribose methylated sites detected by RiboMetSeq (Tables 1 and 2). However, based on the number of mapped reads in the same replicate, some C/D snoRNAs targeting the same rRNA methylation site seem to be differentially expressed. For instance ~80 thousands and ~3 thousands reads are respectively counted for At3gCDbox92.2 and At4gCDbox92.3 targeting 18S-Am28 while ~14 thousands reads are counted for At4gCDbox113.1 and none for At5gCDbox113.2 (not detected), both targeting 25S-Am1459. In contrast, for 14 rRNA 2′-*O*-methylated sites (all with at least one annotated C/D snoRNA) we have not detected a corresponding C/D snoRNA in the FIB2-IP fraction, including 25S-Gm2794 and 25S-Cm1847 sites for which C/D snoRNAs were identified *in silico* (Table 1 and S4).

In conclusion, this analysis demonstrated expression for most of C/D snoRNA driving 2′-*O*-methylation of rRNA sites detected by RiboMethSeq and also identified novel FIB2 interacting C/D snoRNAs without rRNA target site(s).

C/D snoRNA bioinformatics search

To identify RNA guides of ribose methylated rRNA sites (Table 1) without known and/or immunoprecipitated C/D box snoRNAs, we performed a bioinformatics analysis of the Arabidopsis genome (Table 2).

A first catalogue of 217 C/D box snoRNA sequences was established by collecting data from existing resources including SNOOPY [34], ARAPORT [35] and TAIR [36]. This first catalogue was enriched by 3 novel C/D snoRNAs (At3gCDbox101.1, At5gCDbox123.1, At5gCDbox129.1) identified by searching the genome sequence for C/D snoRNA targeting orphan rRNA 2′-*O*-methylated sites

(Fig. 3 and Table 2 and S4) . FIB2-YFP CoIP sequencing gave experimental evidence for 6 novel species mentioned before (Fig. 2B) resulting in an extended catalogue of 226 C/D snoRNAs. Altogether, 193 of the 226 C/D snoRNAs show expression evidence and represent 144 families. Most of them are organized into clusters of C/D snoRNAs. Notably, as previously described, 12 C/D box snoRNAs share a dicistronic organization with tRNAs [14] and 2 snoRNA clusters share a polycistronic organization with miRNAs [37].

All C/D box snoRNA genes were (re)named as At*chr*gCDbox*number.isoform* and consecutively numbered. For example, At4gCDbox94.3 is the name of the C/D box (CDbox) snoRNA which is located on chromosome 4 of the genome (At4g) and is the third isoform of a previously annotated C/D box snoRNA which is numbered 94 (94.3). This nomenclature simplifies current annotation and provide useful information concerning the genomic organization of C/D snoRNA variants.

All identified C/D box snoRNAs were searched for their ability to guide known predicted and/or mapped rRNA 2′-*O*-methylations (Table S4). We found 181 C/D box snoRNAs that contain predicted guide sequences upstream of their D' or D box. Among them, 111 have predicted guide sequences upstream of each box, 155 have a predicted guide sequence upstream the D' box and 136 have a predicted guide sequence upstream the D box (predicted target 5′ and 3′ respectively in Table S4). Several box C/D snoRNAs may have close homologs that are able to guide methylation at similar rRNA sites. Members of different families may also guide a modification at the same rRNA site. For instance, 3 different C/D snoRNAs potentially guide methylation of 18S-Am545 and 4 others guide methylation of 25S-Cm2338 (Tables 1, 2 and S4). Altogether, 4 snoRNAs are able to guide 2′-*O*-methylation at the 2 positions of 5.8S rRNA, 101 at the 65 mapped/predicted positions of 18S rRNA and 183 at the 100 mapped/predicted positions of 25S rRNA. Exploring the updated catalogue of *Arabidopsis thaliana* C/D box snoRNAs with pairing constraints as defined in Materials and Methods, we could not find C/D box snoRNA guides for 12 of the mapped rRNA modifications (Table 1).

Differential 2′-*O*-Methylation in *nuc1*-2 mutant plants

To determine if variations of rRNA 2′-*O*-methylation might occur in Arabidopsis plants, we performed RiboMethSeq analysis of Nucleolin 1 (*nuc1-2*) mutants. Nucleolin is a phylogenetically conserved multifunctional protein required for transcription and processing of 45S pre-rRNA and assembly of ribosomes [38,39]. In Arabidopsis, NUC1 gene disruption affects rRNA expression and functional nucleolar structure and provokes major growth and development defects [40–42]. RiboMethSeq analysis was performed in 3 biological replicates from 21-days-old Arabidopsis *nuc1-2* mutant plants and compared with results from Col-0 (WT) control plants.

Quantitative 2′-*O*-methylation score was calculated for each mapped position in 18S, 5.8S and 25S rRNAs both in Col-0 and *nuc1*-2 plants (Table S1) and compiled in a heat map (Fig. 4). All methylation scores were clustered with the software 'R' (hclust). This analysis revealed 65 sites that were significantly (*p-value* <0.01) hypo-methylated in the *nuc1-2* compared to WT plants (Table 1). Remarkably, the RiboMethSeq analysis has not detected any hypermethylation of mapped sites in *nuc1*-2 mutant compared to Col-0 plants.

65 rRNA 2′-*O*-methylation sites were differentially down regulated in the *nuc1-2* plants. 61 are guided by C/D snoRNA, while the other 4 (18S Um1554 and 25S Cm2294, Gm2410, Gm2652) have no matching C/D snoRNA (Table 1 and yellow circles and grey circles with star in Fig. 2). Noteworthy, the rRNA site 25S-Gm2652, is located in the domain V and close to the P loop (Fig. 1B). We also noticed that out of 65 sites undermethylated in *nuc1-2* plants, 30 and 16 are conserved in human and yeast rRNA respectively, including the 25S-Gm2652, equivalent to the 28S-Gm4228 in human (Table 1).

Altogether, this analysis supports the idea that 2′-*O*-methylation of conserved and non-conserved rRNA sites could be modulated in Arabidopsis plants.

Discussion

In this work we used RiboMethSeq and mapped 117 rRNA 2′-O-methylation sites in leaves from 21-days-old *Arabidopsis thaliana* plants; among these sites, 65 are mapped for the first time. Notably, 38 of the 65 sites mapped in leaves were recently reported as well in plant stems using a nuclease-base detection protocol [43], ratifying that they are truthfully rRNA methylation targets in *Arabidopsis thaliana* rRNA (Table S2). Only 6 rRNA sites mapped by RiboMethSeq were not supported by at least two independent protocols or reported elsewhere (Table S2). Sites without such orthogonal validation were thus considered as potential false-positive hits. Thus, together with 66 previously [14,15,22,23,25] and the 20 newly mapped sites in stems reported in [41] a total of 151 ribose methylated sites have been now mapped in Arabidopsis (Fig. 5).

We noted that the majority of the mapped rRNA sites are guided by ~150 distinct C/D snoRNAs (Table 1). As expected, a number of these rRNA sites is targeted by different C/D snoRNA isoforms resulting from gene duplications of the Arabidopsis genome [14,15,25].

Interestingly, 49 rRNA 2′-*O*-methylation sites previously mapped or predicted were not detected by RiboMethSeq (Table S2). Among these, 4 sites were mapped in young seedlings [15,25], indicating that 2′-*O*-methylation of these sites might occur specifically only at early plant growth and/or developmental stages. Then, it can be expected, that a fraction of rRNA methylated sites detected by RiboMetSeq could be specific to later stage of plant organs, including leaves.

The total number of rRNA mapped sites (117, Table 1) in our plant growth conditions is relatively similar to human with 110 positions guided by 118 C/D snoRNA [10,44] but much higher compared to yeast with 55 positions guided by 43 C/D snoRNA [45]. A subset of 63 and 36 rRNA 2′-*O*-methylated sites in Arabidopsis are conserved in human and/or yeast, respectively (Table 1), indicating that these rRNA sites may be important for ribosome assembly and/or function in eukaryotic cells. On the other hand, most of the Arabidopsis specific sites are located in the 25S rRNA and more precisely in the domain V (Fig. 1), responsible for peptidyl-transferase activity and tRNA binding [4,46]. This might suggest that rRNA methylation at these specific sites in domain V could be required for tuning translation in plants, and likely linked to the usage of synonymous codons and cognate tRNAs to optimize protein translation at particular conditions, including tissue-specific translation in Arabidopsis [47].

Plants are constantly subjected to different cellular and environmental stress conditions and might require additional RNA modifications to protect ribosome integrity or activity. We do not know yet if the Arabidopsis-specific rRNA sites are conserved in other plant species. However, specific rRNA methylation sites can be anticipated in other plant species, as reported in rice [19].

Among the 12 rRNA sites without predictable guiding C/D snoRNAs, 6 were also mapped in Arabidopsis stems [42] including 18S Um1554 and 25S Gm1460, Cm2294, Gm2396, Gm2652 and Gm2923 (Table S2). However, it is possible that some snoRNAs might guide methylation of the rRNA site adjacent to the targeted site. Indeed as shown for U24 in yeast [48] and hypothesized in snoRNAdb for U24 in human [49], the snoRNA At4gCDbox113.1, proposed to guide the modification of 25S-Am1459, might also guide the conserved 2′-*O*-methylation of 25S-Gm1460. Similarly, the snoRNA At1gCDbox5.1 proposed to guide methylation at 25S-Um2651, might also guide the conserved 2′-*O*-methylation of 25S-Gm2652. Interestingly, we have not found any C/D snoRNA guide able to guide modification at positions Gm2396, Um2922 and Gm2923 in 25S rRNA, as it is the case for conserved 2′-*O*-methylated positions in human.

While most of the eukaryotic rRNA ribose modifications are guided by snoRNA, four different modifying enzymes direct ribose methylation in *E. coli*: rsmH/rsmL methylates 16S-Cm1402 while rlmB, rlmM and rlmE/rrmJ methylate Gm2251, Cm2498 and Um2552 in the 23S, respectively [50]. For Arabidopsis 18S-Cm1645 and 25S-Gm2620, equivalent to *E. coli* 16S-Cm1402 and 23S-Gm2251, respectively, we have identified at least 2 C/D snoRNAs for each. In contrast, for 25S-Um2922, the equivalent of *E. coli* 23S-Um2552, we have not found either a specific and/or adjacent snoRNA that might guide methylation. This is an significant observation since methylation of *E. coli* 23S-U2552 influences the interaction of aminoacyl-tRNA with the ribosomal A-site and lack of methylation affect accuracy of translation [50]. An Arabidopsis protein gene phylogenetically related to yeast genes encoding close homologues of *E. coli* rlmE/rrmJ and able to 2′-*O*-methylate tRNA was reported [51]. However, potential 2′-*O*-metyltransferase activity of this Arabidopsis protein has not been demonstrated and requires further investigation.

For most of the rRNA methylated sites we detected, guiding C/D snoRNAs co-precipitate with FIB2 (Table 2). However, C/D snoRNAs At3gCDbox101.1 and At5gCDbox123.1 directing ribose methylation of Gm2794 and Cm1847 in the 25S, were not detected in any of the FIB2-IP fractions, suggesting that they are probably low expressed in our conditions. On the other hand, in the FIB2-IP fraction we found C/D snoRNAs At1gCDbox5.1 and At4gCDbox113.1, which, as mentioned before, might methylate adjacent rRNA sites without matching C/D snoRNA (Table 2). Interestingly, 6 novel C/D snoRNA that co-precipitated with FIB2, do not have recognized rRNA targets (Fig. 2) and then they might guide fibrillarin methylation activity to other coding and/or non-coding RNA. Indeed, in eukaryotes, ribose methylation has been found in snRNA, tRNA, snoRNA and also mRNA [16,52,53]. Alternatively, specialized C/D snoRNPs might also guide RNA base acetylation [54].

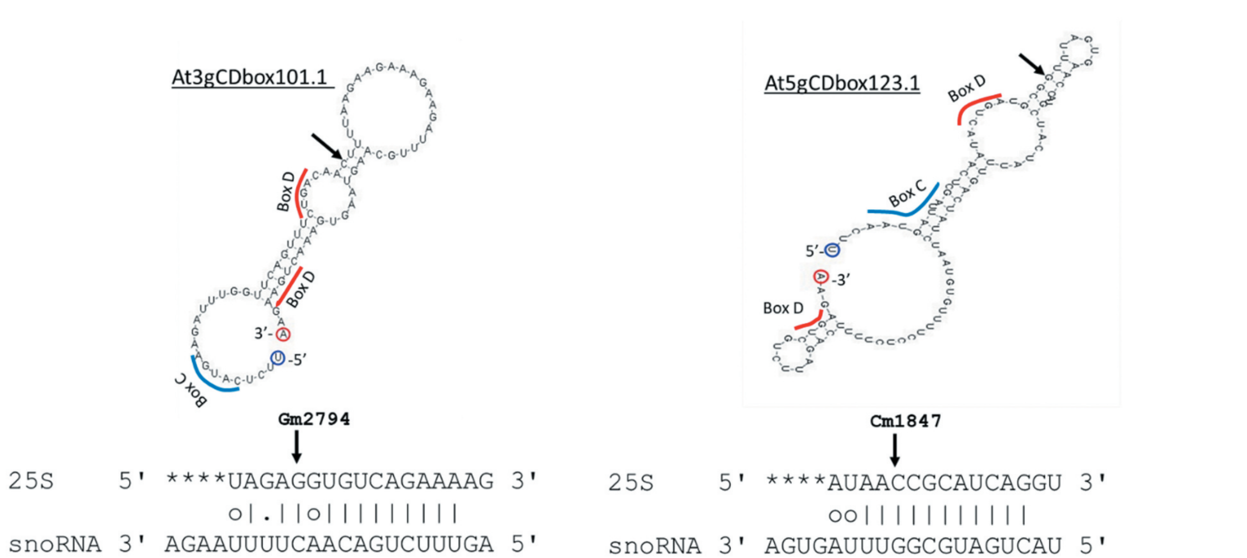

Figure 3. Mfold structure of At3gCDbox1011 and At5gDbox123.1 identified *in silico* and alignment with 25S rRNA sequences. Arrows show mapped rRNA sites Gm2794 and Cm1847 in the 25S and their counterpart in the snoRNA. Blue and red bars underline potential C and D boxes sequences.

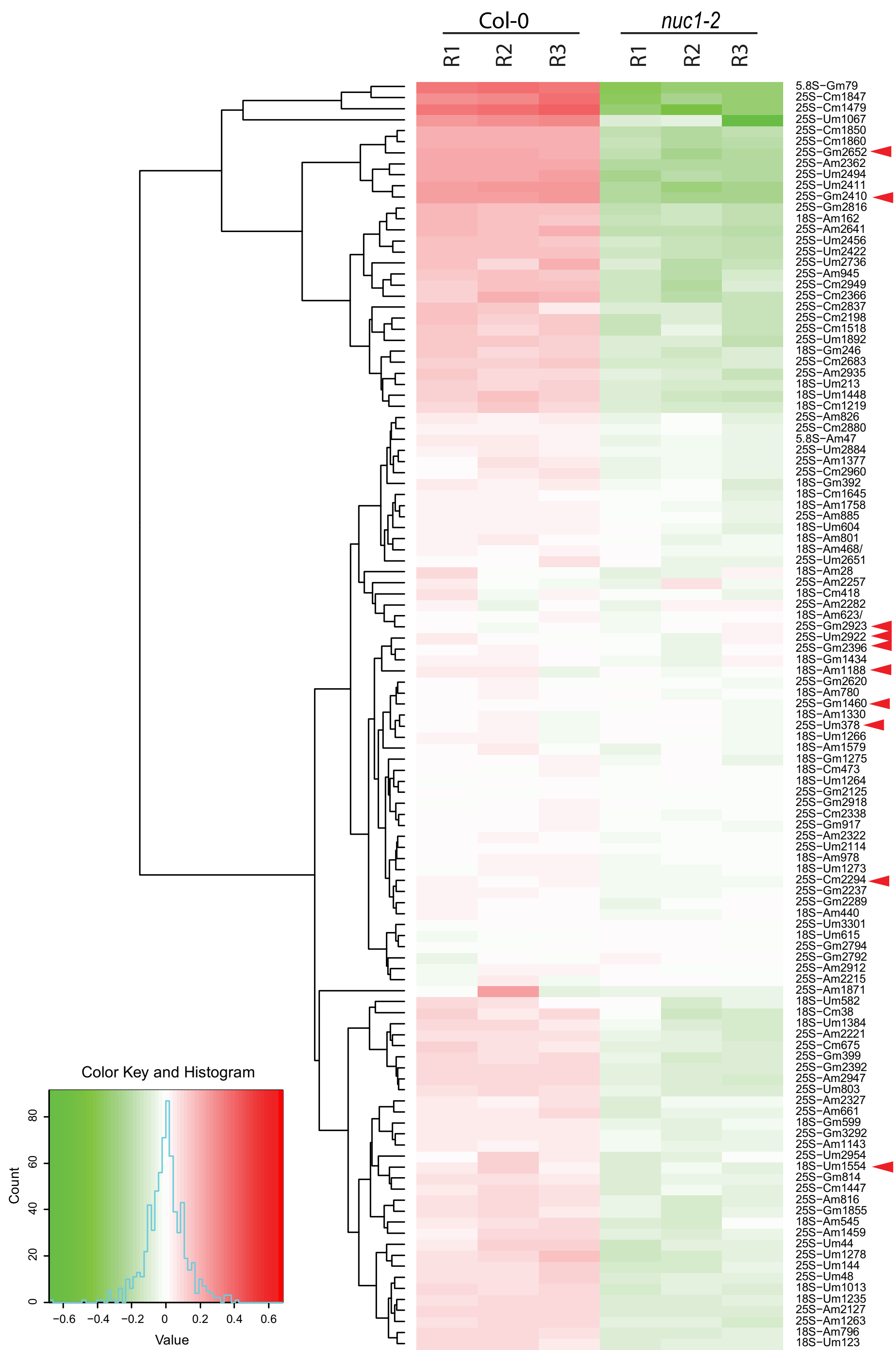

Figure 4. Heat map representation of rRNA 2′-*O*-methylation in Arabidopsis. Differential methscore levels for 115 rRNA 2′-*O*-methylated sites observed in three Col-0 and three *nuc1-2* biological replicates (R1-R3) are shown. The rRNA sites are clustered according to hclust/ward.D2 method. Arrow heads show rRNA sites without associated guiding snoRNA. The colour key, histogram and values are shown.

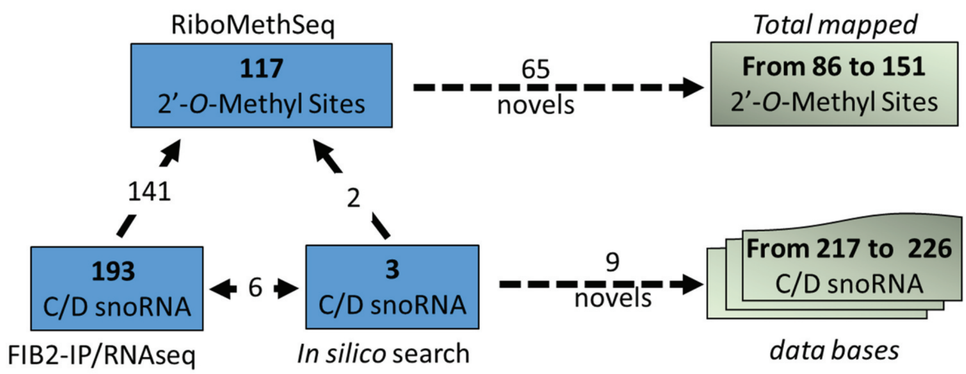

Figure 5. Illustration shows the number of 2′-*O*-methylated sites mapped by RiboMethSeq (117) and C/D snoRNA detected in the FIB2 immunoprecipitated fraction (193) and/or identified by bioinformatics search (9). From these analyses we account 65 first time mapped 2′-*O*-methylated sites and 9 novel C/D snoRNAs, increasing the number of 2′-*O*-methylated mapped sites and C/D snoRNA identified and listed in Arabidopsis databases to 151 and 226 respectively.

Further bioinformatics analysis revealed that 18 C/D snoRNAs detected in the FIB2-IP fraction target 2′-*O*-methylation of major U1, U2, U4, U6, and minor U12 and U6atac snRNAs (Figure S4). Among these 18 C/D snoRNA we have found that 5 are 'orphans' since do not have predicted rRNA target sites (Table S4). Furthermore, 5 other C/D box snoRNAs found in the FIB2-IP fraction and having no rRNA target are isoforms of U3 snoRNAs (Table 2). Unlike other C/D box snoRNAs, U3 is not directing 2′-*O* methylation of other RNAs. Rather, U3 snoRNA is required to guide site-specific cleavage of rRNA during pre-rRNA processing and contains a longer 5′ region downstream the C box that pairs with pre-rRNA rRNA [55–57]. All orphan snoRNAs (Table S4) were submitted to RNA central [58] and show high conservation of all sequences and C/D box motifs in *Arabidopsis lyrata.*

138 C/D snoRNA in the FIB2- IP fraction were also detected among 154 C/D snoRNA recently reported in a nuclear fraction from Arabidopsis cell suspension culture [59]. However, 16 others C/D snoRNAs detected in the nuclear fraction were not found to be associated to FIB2 in our conditions (Table S2). Altogether, the data and observations indicate specific expression of snoRNA in 21 days-old plant compared to Arabidopsis cell suspension culture. Indeed differential expression of C/D snoRNA has been reported in plants [22,60,61] and in animal cells [62,63]. However, we do not rule out the possibility that some C/D snoRNA not detected in the FIB2-IP fraction are in fact expressed, but not assembled into C/D snoRNP in our plant conditions. Indeed, proper assembly of C/D snoRNA with core proteins, including Fibrillarin, is essential for activity of C/D snoRNP [64]. Finally, the comparative analysis of FIB2-IP and *in silico* search analysis demonstrated expression of nearly two hundred C/D snoRNA species and identified 9 novel ones from which at least two direct rRNA 2′-*O*-methylation (Fig. 3), expanding the number of known C/D snoRNA to 226 (Fig. 5).

Differentially 2′-*O*-ribose methylated ribosomes and their impact in cellular function have been reported in animal and yeast cells (reviewed in [4,9,10,65]). Our results revealed rRNA 2′-*O*-methylation alterations in the *nuc1-2* mutant plants (Fig. 4), which display strong developmental phenotypes [40–42]. Surprisingly, in the *nuc1-2* plants we have not detected hypomethylation of 25S-Gm2620. Indeed, lack of methylation of 25S-Gm2620, by inhibiting expression of the guiding C/D snoRNA HID2, provokes developmental defects, which are reminiscent of plant mutants for specific ribosomal proteins and ribosome biogenesis factors, including Nucleolin [22]. In contrast, the Um2422 and Am2641, located structurally closed to 25S-Gm2620 are strongly hypomethylated (Fig. 1B), proposing a functional and/or structural connection between these three rRNA sites.

rRNA ribose hypomethylation of specific sites and major growth and developmental phenotypes were also observed in *nufip* mutant plants [23]. NUFIP (Rsa1p in yeast) is a conserved and a central protein factor directing assembly of C/D snoRNPs [66,67]. Though, a functional link between Nucleolin and C/D snoRNP biogenesis/assembly has not been demonstrated yet, the mouse Nucleolin directly interacts with the MBII-52 snoRNA that assembles into a non-canonical snoRNPs and might function as a chaperone or assist shuttling of MBII-52 RNPs between nucleoplasm and nucleoli [68]. Assembly and maturation of C/D-box snoRNP occurs in the nucleoplasm in human cells while in yeast maturation initiates in the nucleoplasm and terminates in the nucleolus [69]. To our knowledge, it is not known where precisely assembly of snoRNP occurs in plants. The nucleolus is disorganized in *nuc1-2* mutant plants [42] and it is tempting to speculate its direct impact on C/D snoRNP biogenesis in plants and/or more directly on 2′-*O*-methylation activity. If the assembly and/or maturation of pre-mature snoRNPs in functional C/D snoRNP is affected or not in *nuc1* mutant plants remains an open question. Finally, how precisely rRNA ribose modification is impacted at different developmental stages and in response to environmental conditions, is also a next challenging question to be addressed.

Material and methods

Plant materials and growth conditions

All lines were derived from *Arabidopsis thaliana* Columbia (Col-0) ecotype. Plants mutant *nuc1-2* and plants expressing FIB2-YFP nucleolar marker constructs were previously described in [32,42,70]. Seeds were sown on soil (1/5 vermiculite and 4/5 soil) and left for 2 days at 4°C to synchronize. Plants were then grown in controlled growth chambers under a 16 h light/8 h dark cycle at 20°C. Light 100 µE. $m^{-2}.s^{-1}$ and Relative Humidity 60% (CLF Plant Climatics GmbH, Wertingen Germany). Aerial parts from three-week-old plant seedlings were collected, shock-frozen in liquid nitrogen and grinded in fine powder and store at −80°C.

RNA isolation

For RiboMethSeq, about 800 µL of frozen powder were supplemented with 5 mL of TRI Reagent® (Molecular Research Center, Inc), then 1 mL of cold chloroform was added and incubated for 3 minutes. The mix was centrifuged at 8,000x*g* for 15 min at 4°C. The aqueous phase was recovered and precipitated with 3 mL of cold isopropanol. After 30 min incubation, the mix was centrifuged at 8,000x*g* for 30 min at 4°C. Isopropanol was removed and 1.25 mL of 75% ethanol was added to the pellet and incubated over-night at −20°C. After centrifugation, the ethanol was removed and the pellet air dried. The RNA was suspended in RNAse free water and purity verified using Agilent RNA 6000 Pico Kit, analysed in an Agilent 2100 Bioanalyzer, according to the manufacturer's protocol (Figure S1).

2′-*O*-methylation sites cartography by RiboMethSeq

RiboMethSeq analysis was performed as previously described in [24]. Briefly, 100 ng of total RNA from WT plants was subjected to alkaline hydrolysis for 12 min at 96°C. RNA was precipitated and end-repaired before being converted to library using NEBNext©Small RNA Library kit (NEB, USA). Library quality and quantity were assessed using a High Sensitivity DNA chip on a Bioanalyzer 2100 and using Qubit 2.0 fluorometer, respectively. Libraries were multiplexed and subjected for high-throughput sequencing using an Illumina HiSeq 1000 instrument with a 50 bp single-end read mode.

Heat map

Heatmap to compare RiboMethSeq methylation score for Col-0 and *nuc1*-2 mutants was constructed using position-normalized variations of MethScore relative to average value observed for a given position. In this presentation overmethylated sites are in red, while undermethylation compared to average is represented by green colour. Highly constitutive and invariable sites are in white. Clustering was performed using heatmap.2/hclust R functions using ward.D2 method.

Immunoprecipitation: Protein and RNA extraction

Leaves from 3 weeks-old Col-0 (WT) and p35S:FIB2-YFP seedlings were grinded into fine powder in liquid nitrogen. The whole cell extracts (2.4 g of p35S-FIB-YFP and 0.8 g for Col-0) were prepared in 3 volumes of buffer EB150 (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, and EDTA free proteases inhibitor cocktail from Roche) supplemented with 0.1% NP-40. p35S:FIB2-YFP cell extracts (input) were divided into three reactions, and all samples were incubated with 20 µL of GFP-Trap_ MA beads (Chromotek) for 2 h at 4°C with gentle rotation. The unbound fractions were collected and the beads were then washed three times with EB150.

RNA and proteins in IP fraction were extracted using TRI-Reagent (MRC research). RNA separated in aqueous phase were precipitated using glycogen as a carrier (SIGMA, 20 mg) following supplier instructions. Proteins were precipitated from phenolic phase by adding three volumes of cold acetone. After washing, pellets were resuspended in 4X SDS-Laemmli buffer, denatured to perform western blot analysis.

Alternatively, proteins bound to the beads were eluted and recovered in 4X SDS-Laemmli buffer (250 mM Tris-HCl pH 6.7; 8% SDS; 40% glycerol; 0.2% bromophenol bleu, 0.4 M DDT) in the experiments dedicated to Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) analysis.

SDS-PAGE and Western blot analysis

For SDS-PAGE and Western blot, proteins extracts were diluted in SDS-Laemmli buffer, supplemented with β-Mercaptoethanol, heated at 95°C for 10 min and subjected to 10% SDS-polyacrylamide gel electrophoresis. After electrophoresis proteins were either visualized by coomassie-blue staining or transferred to PVDF (Millipore) or nitrocellulose membranes (Bio-Rad) according to manufacturer's instructions. The membranes were then blotted with α-GFP 1:5,000 (Tebu-Bio) and goat anti-rabbit coupled HRP antibodies (Bio-rad). Immunoreactive proteins were detected using Immobilon western chemiluminescent substrates (Millipore) and the acquisition of images was performed using Fusion Solo S camera (Vilber Lourmat). Once immunodetections performed, the membranes used to control IP experiments intended to RNAseq analysis were stained with colloidal blue coomassie solution.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) analyses

Immuno-precipitated protein fractions were diluted to 1X SDS-Laemmli buffer and supplemented with 10 mM DTT before being loaded on an in-house poured 4% acrylamide stacking gel. Gel was stained with Coomassie Blue and the stacking bands were manually excised. Proteins were then reduced, alkylated and digested overnight at 37°C with modified trypsin in a 1:100 enzyme:protein ratio (Promega, Madison, USA). Peptides were extracted during 1 hour with 80 µL of 80% acetonitrile, 0.1% formic acid, before being dried and suspended in water acidified with 0.1% formic acid prior to nanoLC-MS/MS analysis.

LC-MS/MS analyses were performed on a NanoAcquity LC-system (Waters, Milford, MA, USA) coupled to a Q-Exactive plus Orbitrap (Thermo Fisher Scientific, Waltham, MA, USA) mass spectrometer operated in Data-Dependent Acquisition mode as previously described [32]. Peptides/proteins were identified using the Mascot search engine (version 2.5.1, MatrixScience, London, UK) against an *Arabidopsis thaliana* protein sequences database downloaded from The Arabidopsis Information Resource TAIR site (TAIR10 version gene model) to which common contaminants and decoy sequences were added (total of 2 × 27 534 protein entries). Identifications were validated and label-free extracted ion chromatogram-based quantification was performed using the Proline software suite. False Discovery Rate was optimized to be below 1% at PSM level using Mascot Adjusted E-value and below 1% at Protein Level using Mascot Mudpit score. Statistical analysis was performed using the Prostar software suite (version 1.12.11). Pairwised Limma t-tests were performed. P-values calibration was corrected using adapted Benjamini-Hochsberg method, and FDR was set to <1-2%. (For more details see Supplementary Information)

RNAseq of RNA from FIB2:YFP fractions

RNA samples from CoIP_1 (150 ng), CoIP_2 (300 ng) and CoIP_3 (225 ng) samples were subjected to alkaline hydrolysis for 5 min at 96°C. RNA was precipitated and end-repaired before being converted to library using NEBNext©Small RNA Library kit (NEB, USA). Library quality and quantity were assessed using a High Sensitivity DNA chip on a Bioanalyzer 2100 and using Qubit 2.0 fluorimeter, respectively. Libraries were multiplexed and subjected for high-throughput sequencing using an Illumina HiSeq 1000 instrument with a 50 bp single-end read mode.

C/D snoRNA bioinformatic search and analysis

Data. The genomic sequence of *Arabidopsis thaliana* used in these analyses is available at ftp.ensemblgenomes.org/pub/plants/current/fasta/arabidopsis_thaliana/dna/Arabidopsis_thaliana.TAIR10.dna.toplevel.fa.gz. Annotations are from TAIR. Fasta sequences of used ribosomal RNAs and snRNA are given in Supplemental Information.

C/D box snoRNA identification. The catalogue of *Arabidopsis thaliana* C/D box snoRNA genes was established using three approaches. The initial dataset of C/D box snoRNA genes was built from sequences available in the snoRNA Orthological Gene Database (snOPY) [34] and in *Arabidopsis thaliana* repositories such as ARAPORT portal [35] and TAIR resource [36]. SnoRNA genes from ARAPORT and TAIR were manually curated to distinguish between H/ACA box and C/D box snoRNA sequences. Only C/D box snoRNA sequences that mapped on the genomic sequence with 100% of identity were kept for further analysis. The initial dataset of C/D box snoRNA sequences was enriched by using PatScan [71] to search the genomic sequence for patterns encoding new C/D box snoRNAs. Such patterns were defined to contain motifs corresponding to the C box (RUGAUGA allowing one mismatch), the C/D box snoRNA region of the snoRNA:rRNA interaction (defined from each rRNA position orphan of a C/D box snoRNA guide at the region around a mapped methylation site and with at most three mismatches in the first eleven base pairs) and the D'/D box motifs (CUGA allowing one mismatch). This updated catalogue was enriched with regions expressed in the sequenced FIB2-YFP coIP fractions and matching the C/D box snoRNA gene pattern not considering the snoRNA:rRNA interaction constraint. Reads from sequenced FIB2-YFP coIP fractions were trimmed for the Illumina 3′ adaptor sequence using Cutadapt [72] and aligned using the bowtie2 aligner [73]. Highly expressed intronic and intergenic regions without annotation were inspected in IGV [74], assembled in transcripts and searched for the C/D box snoRNA pattern. This dataset was also used to modify boundaries of previously identified C/D box snoRNA genes. All C/D box snoRNA genes were named as At*chr*gCDbox*number.isoform* and consecutively numbered with *chr* giving the chromosome number, *number* increasing for each new C/D box snoRNA seen for the first time (paralogs have the same *number*) and *isoform* giving the number+1 of times an isoform was already found in preceding chromosomes or before this C/D box snoRNA in the same chromosome.

RNA folding structure prediction of C/D snoRNAs was performed using the mfold Web server http://unafold.rna.albany.edu/?q=mfold

Acknowledgments

We thank our master student, Clarisse Mariez, who contributed to IP-LC MS/MS and Anne de Bures for technical assistance.

Disclosure and availability Statement

The authors report no conflict of interest. Raw data were generated at the Epitranscriptomics and RNA Sequencing (EpiRNA-Seq) Core Facility, Nancy, France. Derived data supporting the findings of this study are available from the corresponding author [JSV] on request.

Disclosure statement

No potential conflict of interest was reported by the authors.

Author contributions

J. A-F designed and performed protein related and IP experiments for sRNA analysis; and supervised master degree student training for IP LC-MS/MS experiments. C.M. and E.J performed RNA related experiments. Y.M., L. A. and V.M. performed RiboMethSeq analysis. M.R. and C.C performed LC-MS/MS and analyzed the data. C.G performed bioinformatics analysis. J. S-V, C.G and Y.M supervised and analyzed the data. J.S-V wrote the manuscript with the assistance of J. A-F., C.C., C.G and Y.M.

Funding

This work was supported by the CNRS, INRAE and by the ANR (Agence Nationale de la Recherche) under Grant RiboStress 17-CE12-0026-01 and MetRibo and a BQR (Bonus Qualité Recherche) from the UPVD to JSV. YM was supported by EpiARN FRCR project from Grand Est Region (France). This study is set within the framework of the "Laboratoires d'Excellence (LABEX) TULIP (ANR-10-LABX-41). Mass spectrometry experiments were supported by the French Proteomic Infrastructure (ProFI; ANR-10-INBS-08-03);ANR [17-CE12-0026-01]; ANR [MetRibo]; ANR [10-INBS-08-03]; Grand Est Region [EpiARN FRCR].

ORCID

V. Marchand http://orcid.org/0000-0002-8537-1139
Y. Motorin http://orcid.org/0000-0002-8018-334X
J. Sáez-Vásquez http://orcid.org/0000-0002-2717-7995

References

[1] Decatur WA, Fournier MJ. rRNA modifications and ribosome function. Trends Biochem Sci. 2002;27(7):344–351.

[2] Polikanov YS, Melnikov SV, Soll D, et al. Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. Nat Struct Mol Biol. 2015;22(4):342–344.

[3] Sharma S, Lafontaine DL. 'View from a bridge': a new perspective on eukaryotic rrna base modification. Trends Biochem Sci. 2015;40(10):560–575.

[4] Sloan KE, Warda AS, Sharma S, et al. Tuning the ribosome: the influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol. 2017;14(9):1138–1152.

[5] Blanchard SC, Puglisi JD. Solution structure of the A loop of 23S ribosomal RNA. Proc Natl Acad Sci USA. 2001;98:3720–3725.

[6] Liang XH, Liu Q, Fournier MJ. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. RNA. 2009;15:1716–1728.

[7] Erales J, Marchand V, Panthu B, et al. Evidence for rRNA 2′-O-methylation plasticity: control of intrinsic translational capabilities of human ribosomes. Proc Natl Acad Sci U S A. 2017;114(49):12934–12939.

[8] Marcel V, Ghayad SE, Belin S, et al. p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. Cancer Cell. 2013;24(3):318–330. .

[9] Monaco PL, Marcel V, Diaz JJ, et al. 2′-O-Methylation of Ribosomal RNA: towards an Epitranscriptomic control of translation? Biomolecules. 2018;8(4):1–13.

[10] Ayadi L, Galvanin A, Pichot F, et al. RNA ribose methylation (2′-O-methylation): occurrence, biosynthesis and biological functions. Biochim Biophys Acta Gene Regul Mech. 2019;1862(3):253–269.

[11] Massenet S, Bertrand E, Verheggen C. Assembly and trafficking of box C/D and H/ACA snoRNPs. RNA Biol. 2017;14(6):680–692.

[12] Watkins NJ, Bohnsack MT. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley Interdiscip Rev RNA. 2012;3(3):397–414.

[13] Yu G, Zhao Y, Li H. The multistructural forms of box C/D ribonucleoprotein particles. RNA. 2018;24(12):1625–1633.

[14] Brown JW, Clark GP, Leader DJ, et al. Multiple snoRNA gene clusters from Arabidopsis. RNA. 2001;7(12):1817–1832.

[15] Qu LH, Meng Q, Zhou H, et al. Identification of 10 novel snoRNA gene clusters from Arabidopsis thaliana. Nucleic Acids Res. 2001;29(7):1623–1630.

[16] Brown JW, Echeverria M, Qu LH. Plant snoRNAs: functional evolution and new modes of gene expression. Trends Plant Sci. 2003;8(1):42–49.

[17] Chen HM, Wu SH. Mining small RNA sequencing data: a new approach to identify small nucleolar RNAs in Arabidopsis. Nucleic Acids Res. 2009;37(9):e69.

[18] Kim SH, Spensley M, Choi SK, et al. Plant U13 orthologues and orphan snoRNAs identified by RNomics of RNA from Arabidopsis nucleoli. Nucleic Acids Res. 2010;38(9):3054–3067.

[19] Chen CL, Liang D, Zhou H, et al. The high diversity of snoRNAs in plants: identification and comparative study of 120 snoRNA genes from Oryza sativa. Nucleic Acids Res. 2003;31(10):2601–2613.

[20] Liu TT, Zhu D, Chen W, et al. A global identification and analysis of small nucleolar RNAs and possible intermediate-sized non-coding RNAs in Oryza sativa. Mol Plant. 2013;6(3):830–846.

[21] Patra Bhattacharya D, Canzler S, Kehr S, et al. Phylogenetic distribution of plant snoRNA families. BMC Genomics. 2016;17(1):1–12.

[22] Zhu P, Wang Y, Qin N, et al. Arabidopsis small nucleolar RNA monitors the efficient pre-rRNA processing during ribosome biogenesi. Proc Natl Acad Sci U S A. 2016;113(42):11967–11972.

[23] Rodor J, Jobet E, Bizarro J, et al. AtNUFIP, an essential protein for plant development, reveals the impact of snoRNA gene organisation on the assembly of snoRNPs and rRNA methylation in Arabidopsis thaliana. Plant J. 2011;65(5):807–819.

[24] Piekna-Przybylska D, Decatur WA, Fournier MJ. The 3D rRNA modification maps database: with interactive tools for ribosome analysis. Nucleic Acids Res. 2008;36: D178–D183.

[25] Barneche F, Gaspin C, Guyot R, et al. Identification of 66 box C/D snoRNAs in Arabidopsis thaliana: extensive gene duplications generated multiple isoforms predicting new ribosomal RNA 2′-O-methylation sites. J Mol Biol. 2001;311(1):57–73.

[26] Marchand V, Blanloeil-Oillo F, Helm M, et al. Illumina-based RiboMethSeq approach for mapping of 2′-O-Me residues in RNA. Nucleic Acids Res. 2016;44(16):e135.

[27] Gruendler P, Unfried I, Pointner R, et al. Nucleotide sequence of the 25S-18S ribosomal gene spacer from Arabidopsis thaliana. Nucleic Acids Res. 1989;17(15):6395–6396.

[28] Unfried I, Gruendler P. Nucleotide sequence of the 5.8S and 25S rRNA genes and of the internal transcribed spacers from Arabidopsis thaliana. Nucleic Acids Res. 1990;18(13):4011.

[29] Unfried I, Stocker U, Gruendler P. Nucleotide sequence of the 18S rRNA gene from Arabidopsis thaliana Co10. Nucleic Acids Res. 1989;17(18):7513.

[30] Pichot F, Marchand V, Ayadi L, et al. Holistic optimization of bioinformatic analysis pipeline for detection and quantification of 2′-O-Methylations in RNA by RiboMethSeq. Front Genet. 2020;11:38.

[31] Incarnato D, Anselmi F, Morandi E, et al. High-throughput single-base resolution mapping of RNA 2′-O-methylated residues. Nucleic Acids Res. 2017;45(3):1433–1441.

[32] Montacie C, Durut N, Opsomer A, et al. Nucleolar proteome analysis and proteasomal activity assays reveal a link between nucleolus and 26S Proteasome in A. thaliana. Front Plant Sci. 2017;8:1–13.

[33] Pontvianne F, Carpentier MC, Durut N, et al. Identification of nucleolus-associated chromatin domains reveals a role for the nucleolus in 3d organization of the a. thaliana genome. Cell Rep. 2016;16(6):1574–1587. .

[34] Yoshihama M, Nakao A, Kenmochi N. snOPY: a small nucleolar RNA orthological gene database. BMC Res Notes. 2013;6:6.

[35] Krishnakumar V, Hanlon MR, Contrino S, et al. Araport: the Arabidopsis information portal. Nucleic Acids Res. 2015;43(D1):D1003–1009.

[36] Berardini TZ, Reiser L, Li D, et al. The Arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. Genesis. 2015;53:474–485.

[37] Qu G, Kruszka K, Plewka P, et al. Promoter-based identification of novel non-coding RNAs reveals the presence of dicistronic snoRNA-miRNA genes in Arabidopsis thaliana. BMC Genomics. 2015;16(1):1–15.

[38] Tajrishi MM, Tuteja R, Tuteja N. Nucleolin: the most abundant multifunctional phosphoprotein of nucleolus. Commun Integr Biol. 2011;4(3):267–275.

[39] Ugrinova I, Petrova M, Chalabi-Dchar M, et al. Multifaceted nucleolin protein and its molecular partners in oncogenesis. Adv Protein Chem Struct Biol. 2018;111:133–164.

[40] Kojima H, Suzuki T, Kato T, et al. Sugar-inducible expression of the nucleolin-1 gene of Arabidopsis thaliana and its role in ribosome synthesis, growth and development. Plant J. 2007;49(6):1053–1063. .

[41] Petricka JJ, Nelson TM. Arabidopsis nucleolin affects plant development and patterning. Plant Physiol. 2007;144(1):173–186.

[42] Pontvianne F, Matia I, Douet J, et al. Characterization of AtNUC l1 reveals a central role of nucleolin in nucleolus organization and silencing of AtNUC L2 gene in Arabidopsis. Mol Biol Cell. 2007;18(2):369–379.

[43] Tang Y, Wu Y, Xu R, et al. Identification and exploration of 2′-O-methylation sites in rRNA and mRNA with a novel RNase based platform. bioRxiv. 2020;2020(2003):2027.011759.
[44] Dupuis-Sandoval F, Poirier M, Scott M. The emerging landscape of small nucleolar RNAs in cell biology. WIREs RNA. 2015;6(4):381–397.
[45] Decatur WA, Liang X-H, Piekna-Przybylska D, et al. Identifying effects of snorna-guided modifications on the synthesis and function of the yeast Ribosome. Methods Enzymol. 2007;425:283–316.
[46] Wu CC, Zinshteyn B, Wehner KA, et al. High-resolution ribosome profiling defines discrete ribosome elongation states and translational regulation during cellular stress. Mol Cell. 2019;73(959–970):e955.
[47] Camiolo S, Farina L, Porceddu A. The Relation of Codon Bias to Tissue-Specific Gene Expression in Arabidopsis thaliana. Genetics. 2012;192(2):641–649.
[48] Kiss-Laszlo Z, Henry Y, Bachellerie JP, et al. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell. 1996;85(7):1077–1088.
[49] Lestrade L, Weber MJ. snoRNA-LBME-db, a comprensive database of human H/ACA and C/D box snoRNA. Nucleic Acid Research. 2006;34(90001):158–162.
[50] Sergeeva OV, Bogdanov AA, Sergiev PV. What do we know about ribosomal RNA methylation in Escherichia coli. Biochimie. 2015;117:110–118.
[51] Pintard L, Lecointe F, Bijnicki JM, et al. Trm7p catalyses the formation of two 2′-O-methylriboses in yeast tRNA anticodin loop. Embo J. 2002;21(7):1811–1820.
[52] Bachellerie JP, Cavaille J, Huttenhofer A. The expanding snoRNA world. Biochimie. 2002;84(8):775–790.
[53] Decatur WA, Fournier MJ. RNA-guided nucleotide modification of ribosomal and other rNAs. J Biol Chem. 2003;278(2):695–698.
[54] Sharma S, Yang J, van Nues R, et al. Specialized box C/D snoRNPs act as antisense guides to target RNA base acetylation. PLoS Genet. 2017;13(5):e1006804.
[55] Beltrame M, Tollervey D. Base pairing between U3 and the pre-ribosomal RNA is required for 18S rRNA synthesis. Embo J. 1995;14(17):4350–4356.
[56] Borovjagin AV, Gerbi SA. U3 small nucleolar RNA is essential for cleavage at sites 1, 2 and 3 in pre-rRNA and determines which rRNA processing pathway is taken in Xenopus oocytes. J Mol Biol. 1999;286(5):1347–1363.
[57] Saez-Vasquez J, Caparros-Ruiz D, Barneche F, et al. A plant snoRNP complex containing snoRNAs, fibrillarin, and nucleolin-like proteins is competent for both rRNA gene binding and pre-rRNA processing in vitro. Mol Cell Biol. 2004;24(16):7284–7297.
[58] Consortium TR. RNAcentral: a hub of information for non-coding RNA sequences. NAR. 2019;47(D1):D221–D229.
[59] Streit D, Shanmugam T, Garbelyanski A, et al. The existence and localization of nuclear snoRNAs in Arabidopsis thaliana revisited. Plants (Basel). 2020;9(8):9.
[60] Wang Y, Li H, Sun Q, et al. Characterization of small rnas derived from trnas, rrnas and snornas and their response to heat stress in wheat seedlings. PLoS One. 2016;11(3):e0150933.
[61] Zheng J, Zeng E, Du Y, et al. Temporal small rna expression profiling under drought reveals a potential regulatory role of small nucleolar rnas in the drought responses of maize. Plant Genome. 2019;12(1):1–15. .
[62] Falaleeva M, Welden JR, Duncan MJ, et al. C/D-box snoRNAs form methylating and non-methylating ribonucleoprotein complexes: old dogs show new tricks. Bioessays. 2017;39(6):1–10.
[63] Stepanov GA, Filippova JA, Komissarov AB, et al. Regulatory role of small nucleolar RNAs in human diseases. Biomed Res Int. 2015;2015:206849.
[64] McKeegan KS, Debieux CM, Boulon S, et al. A dynamic scaffold of pre-snoRNP factors facilitates human box C/D snoRNP assembly. Mol Cell Biol. 2007;27(19):6782–6793.
[65] Dimitrova DG, Teysset L, Carre C. RNA 2′-O-Methylation (nm) modification in human diseases. Genes (Basel). 2019;10(2):1–23.
[66] Bizarro J, Charron C, Boulon S, et al. Proteomic and 3D structure analyses highlight the C/D box snoRNP assembly mechanism and its control. J Cell Biol. 2014;7:1–18.
[67] Rothe B, Manival X, Rolland N, et al. Implication of the box C/D snoRNP assembly factor Rsa1p in U3 snoRNP assembly. Nucleic Acids Res. 2017;45(12):7455–7473.
[68] Soeno Y, Taya Y, Stasyk T, et al. Identification of novel ribonucleo-protein complexes from the brain-specific snoRNA MBII-52. RNA. 2010;16(7):1293–1300.
[69] Verheggen C, Lafontaine DL, Samarsky D, et al. Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. Embo J. 2002;21(11):2736–2745.
[70] Pontvianne F, Abou-Ellail M, Douet J, et al. Nucleolin is required for DNA methylation state and the expression of rRNA gene variants in Arabidopsis thaliana. PLoS Genet. 2010;6(11):1–13. .
[71] Dsouza M, Larsen N, Overbeek R. Searching for patterns in genomic data. Trends Genet. 1997;13(12):497–498.
[72] Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17(1):10–12.
[73] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–359.
[74] Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol 2001;29:24–26.