

Supporting information for
 Circular code motifs in the ribosome: a missing link in the evolution of translation?
 Gopal Dila¹, Raymond Ripp¹, Claudine Mayer^{1,2,3}, Olivier Poch¹, Christian J. Michel^{1,*} and Julie
 D. Thompson^{1,*}



Figure S1. Distribution of the number and total nucleotide lengths of *X* motifs in the SSU (16/18S) and LSU (23/28S) rRNA multiple alignments conserved in the 3 domains of life, in bacteria and archaea only, and 'universal' *X* motifs conserved in all 3 domains. *X* Motifs are considered to be conserved if they are present in >90% of sequences.

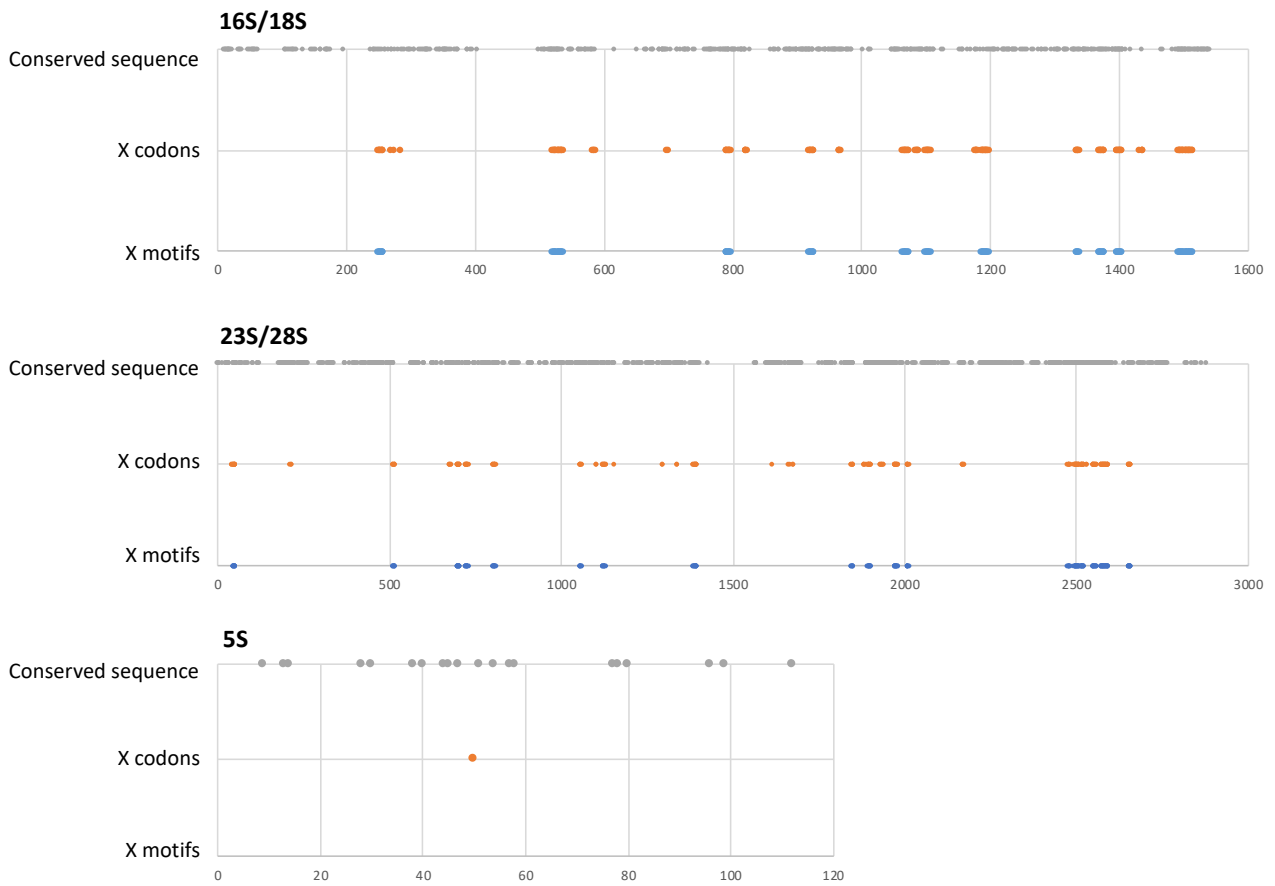


Figure S2. Comparison of sequence conservation (90% identity), universal *X* codons and the universal *X* motifs (*uX* motifs) in the rRNA multiple alignments containing 130 species. The x-axis shows the position in each of the 3 multiple alignments corresponding to the 16/18S rRNA, the 23/28S rRNA and the 5S rRNA. Positions with *X* codons in >90% of the sequences are identified as universal *X* codons. Universal *X* motifs indicate consecutive runs of universal *X* codons of length ≥ 6 nucleotides.

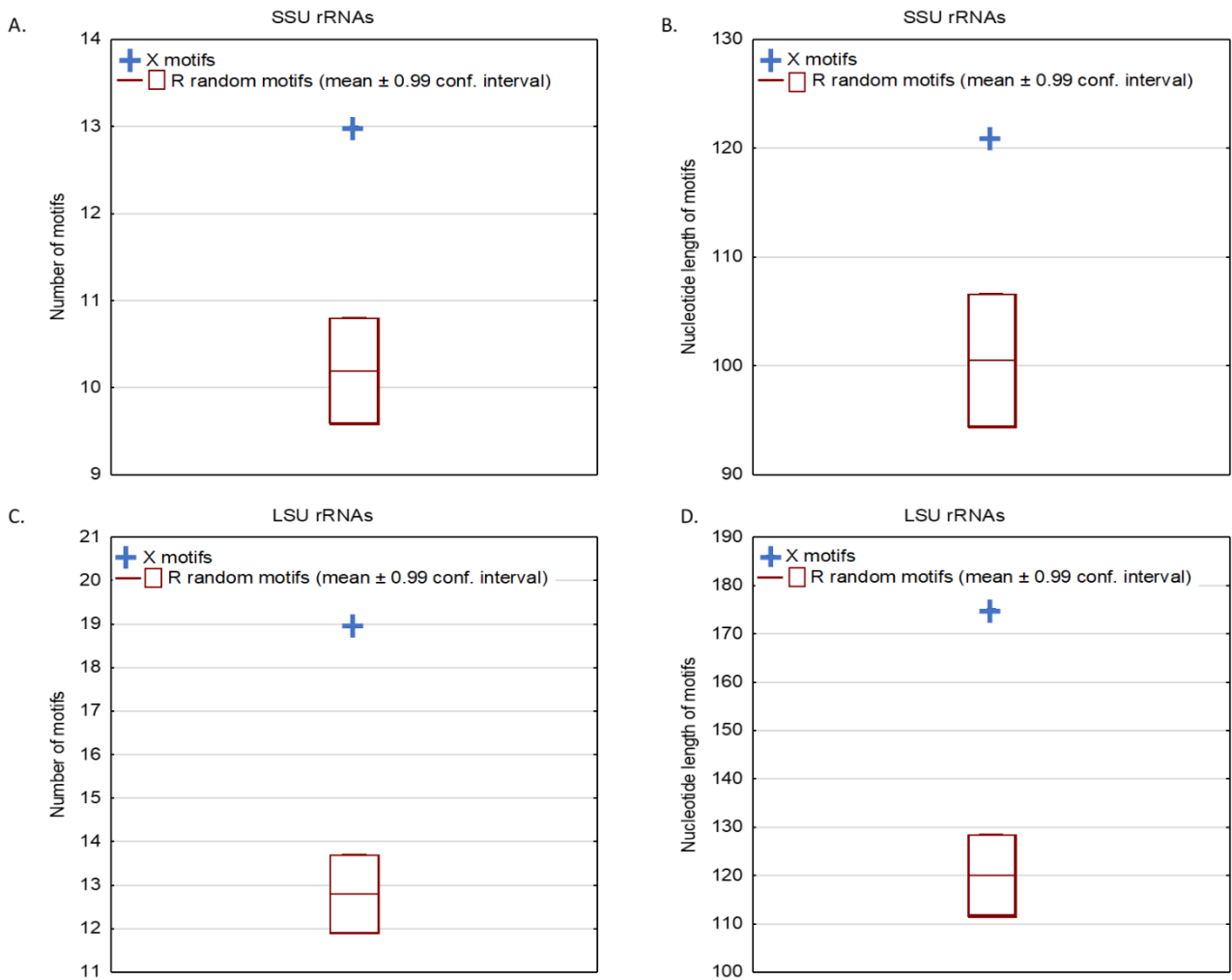


Figure S3. Comparison of the number of *uX* and *uR* random motifs and their nucleotide lengths in the SSU and LSU rRNA alignments. The number and total length of *uX* motifs are represented by a blue cross. The distribution of the *uR* random motifs obtained from 100 random codes *R* is indicated by boxplots representing the mean and ± 0.99 confidence interval. A. The number of *uX* motifs in the SSU (=13) is significantly higher than the mean number of *uR* motifs (=10.2) according to a one-sided Student's t-test ($p \approx 10^{-21}$). B. The nucleotide length of the *uX* motifs in the SSU (=121) is also significantly higher than the mean nucleotide length of *uR* motifs (=100.5) according to a one-sided Student's t-test ($p \approx 10^{-14}$). C. Similarly, the number of *uX* motifs in the LSU (=19) is significantly higher than the mean number of *uR* motifs (=12.8) according to a one-sided Student's t-test ($p \approx 10^{-33}$). The nucleotide length of the *uX* motifs in the LSU (=175) is also significantly higher than the mean nucleotide length of *uR* motifs (=120.1) according to a one-sided Student's t-test ($p \approx 10^{-32}$).

	Total sequence length (<i>E. coli</i>)	No. of conserved positions and percentage of total length	No. of <i>uX</i> motif positions and percentage of total length	No. (and %) of <i>uX</i> motif positions that are not conserved	No. (and %) of conserved positions that are not <i>uX</i> motifs
16/18S	1542	490 (31.8%)	121 (7.8%)	34 (28.1%)	403 (82.2%)
23/28S	2904	870 (30.0%)	175 (6.0%)	50 (28.6%)	745 (85.6%)
5S	120	20(16.7%)	0 (0%)	0 (0%)	20 (100%)

Table S1. Sequence conservation and *uX* motifs in the rRNA multiple alignments containing 130 species. In columns 2 and 3, the number of conserved positions and number of *uX* motif positions are shown as a percentage of the total sequence length. Column 4 shows the proportion of positions that are in *uX* motifs but are not conserved. Column 5 shows the proportion of positions that are conserved but are not in *uX* motifs.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>a</i>	<i>X</i> universality	97.7	97.7	97.7	97.7	97.7	97.7	97.7	97.7	97.7						
	Sequence conservation	97	100	100	100	95.6	100	100	100	100	98.5					
<i>b</i>	<i>X</i> universality	100	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3	99.3					
	Sequence conservation	100	100	97	100	100	98.5	98.5	98.5	100	100					
<i>c</i>	<i>X</i> universality	94.7	94.7	94.7	94.7	95.5	95.5	95.5	98.5	95.5	95.5	94	94			
	Sequence conservation	100	100	100	100	68.7	56.1	98.5	84.4	95.6	100	100	57.1			
<i>d</i>	<i>X</i> universality	98.5	98.5	98.5	98.5	98.5	98.5	98.5	98.5	98.5						
	Sequence conservation	100	100	100	97	100	59.5	98.5	100	100						
<i>e</i>	<i>X</i> universality	100	100	99.3	99.3	99.3	99.3	99.3	100	100						
	Sequence conservation	72.2	100	100	100	94.1	100	100	98.5	63.9						
<i>f</i>	<i>X</i> universality	97.7	98.5	98.5	98.5	99.3	99.3	99.3	99.3	99.3	93.2					
	Sequence conservation	70.9	92.7	92.6	61.6	89.9	100	100	66.3	97	83.1					
<i>g</i>	<i>X</i> universality	98.5	98.5	98.5	98.5	98.5	98.5									
	Sequence conservation	100	100	100	61.7	97	100									
<i>h</i>	<i>X</i> universality	97.7	97.7	97	97	97	97.7	97.7	97.7	97.7	97.7	91	91.7	93.2	93.2	
	Sequence conservation	100	100	100	100	98.5	97	100	75.9	64.8	98.5	46.8	33.4	48.9	48.2	
<i>i</i>	<i>X</i> universality	91	90.2	92.5	92.5	92.5	92.5	92.5	94	94	94	100	94.7	90.2	90.2	90.2
	Sequence conservation	80.9	79.7	48.3	50.3	44.5	100	64.8	64.8	37.3	74.3	97	49.7	79.8	100	97
<i>j</i>	<i>X</i> universality	94	94	94	94	94	94									
	Sequence conservation	98.5	54.6	33.4	33.8	100	100									
<i>k</i>	<i>X</i> universality	92.5	93.2	97.7	97.7	97.7	97.7	97.7	98.5	98.5						
	Sequence conservation	92.7	45.3	100	57.8	72	100	87.3	67.7	55.6						
<i>l</i>	<i>X</i> universality	97.7	97.7	97.7	97.7	97	97	97	96.2	95.5	95.5					
	Sequence conservation	63.6	82.1	64.8	95.6	100	57	95.6	67.1	100	100					
<i>m</i>	<i>X</i> universality	99.3	99.3	99.3	99.3	99.3	99.3	99.3	100	100						
	Sequence conservation	58.1	55	100	100	98.5	97	94.1	65.6	100						

Table S2. Comparison of *X* motif universality (number of species having an *X* motif) and sequence conservation (percent sequence identity) for each position in the 13 universal *X* motifs (*uX* motifs) in the SSU rRNA alignments (prokaryotic 16S and eukaryotic 18S). Each row represents one *uX* motif and each column represents one position within the *uX* motif. There is no correlation between *X* motif universality and sequence conservation: Pearson correlation coefficient $r = 0.37$ ($p < 10^{-4}$), Spearman correlation coefficient $\rho = 0.25$ ($p = 0.006$) and Kendall coefficient $\tau = 0.19$ ($p = 0.007$) are weakly positively related. Furthermore, a two-tailed matched sample signed ranks Wilcoxon test shows that the 2 distributions are significantly different ($p < 10^{-3}$).

Fully conserved nucleotides in the SSU rRNA alignment are provided in a separate file: [supp table S2b.pdf](#).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>A</i> <i>X</i> universality	97	97.7	97.7	97.7	97.7	97.7									
Sequence conservation	63.2	98.5	100	95.6	98.5	57									
<i>B</i> <i>X</i> universality	99.3	99.3	99.3	100	99.3	99.3	100	98.5	98.5	98.5	98.5	98.5	98.5	98.5	98.5
Sequence conservation	63.6	100	55.3	95.6	94.1	98.5	97	95.6	100	100	98.5	100	88.4	74.3	100
<i>C</i> <i>X</i> universality	100	100	94	94	94	97	97.7	97.7	97.7	93.2					
Sequence conservation	34	62.4	100	81.7	100	65.3	100	62	86	40.7					
<i>D</i> <i>X</i> universality	94	94	94	94	94	94	94	94	97	97	97	97	96.2		
Sequence conservation	95.5	60.2	100	100	100	88.4	100	100	100	98.5	100	100	54.4		
<i>E</i> <i>X</i> universality	100	100	100	100	100	100	100	100	100	100					
Sequence conservation	100	83.2	95.6	49.6	100	98.5	98.5	100	100	84.7					
<i>F</i> <i>X</i> universality	91.7	91.7	93.2	93.2	93.2	93.2	93.2	93.2	93.2	94	94	91.7			
Sequence conservation	64.8	50.7	100	100	86	87.3	100	62.4	100	100	39.1	60.1			
<i>G</i> <i>X</i> universality	90.2	92.5	93.2	93.2	93.2	93.2									
Sequence conservation	89.8	49.7	100	100	98.5	51									
<i>H</i> <i>X</i> universality	94.7	94.7	94	94	93.2	94	94.7								
Sequence conservation	100	98.5	94.1	69.2	77.4	62	37.3								
<i>I</i> <i>X</i> universality	91	90.2	90.2	90.2	91	96.2	95.5	96.2	96.2						
Sequence conservation	97	100	98.5	100	87.3	100	29.5	79.8	88.6						
<i>J</i> <i>X</i> universality	97.7	98.5	98.5	98.5	97	97	97.7	97	97	96.2					
Sequence conservation	97	88.6	94.1	31.2	75.1	95.5	100	100	92.7	40.3					
<i>K</i> <i>X</i> universality	94.7	94.7	100	100	100	100	100								
Sequence conservation	100	39.5	46.5	54.9	100	100	67.4								
<i>L</i> <i>X</i> universality	97.7	92.5	92.5	94	94.7	94	94	94	94						
Sequence conservation	94.1	95.6	100	97	83	100	100	94.1	98.5						
<i>M</i> <i>X</i> universality	90.2	91	94	94	91	91	91	91	91	90.2					
Sequence conservation	42.6	84.6	97	61.8	100	100	83.2	94.1	86	100					
<i>N</i> <i>X</i> universality	94.7	94.7	96.2	96.2	96.2	96.2									
Sequence conservation	89.8	39.2	71.5	100	39	100									
<i>O</i> <i>X</i> universality	99.3	98.5	98.5	98.5	98.5	98.5	99.3	99.3	99.3						
Sequence conservation	97	100	98.5	100	98.5	100	100	97	100						
<i>P</i> <i>X</i> universality	96.2	95.5	95.5	95.5	95.5	94.7	97	95.5							
Sequence conservation	37.9	82.9	95.5	79.2	57.8	100	97	67.1							
<i>Q</i> <i>X</i> universality	94	94	94	94	93.2	93.2									
Sequence conservation	62.4	58.4	98.5	100	46.4	70									
<i>R</i> <i>X</i> universality	97.7	97.7	97.7	97	97	97	97	97	94						
Sequence conservation	83.3	92.7	95.5	49.1	100	100	95.6	94.1	66.5						
<i>S</i> <i>X</i> universality	93.2	93.2	93.2	93.2	94	94	93.2	94	94	94	93.2				
Sequence conservation	87	95.5	98.5	75.5	56.7	94.1	64.2	98.5	94.1	86	94.1				

Table S3. Comparison of *X* motif universality (number of species having an *X* motif) and sequence conservation (percent sequence identity) for each position in the 19 universal *X*

motifs (*uX* motifs) in the LSU rRNA alignments (prokaryotic 23S and eukaryotic 25/28S). Each row represents one *uX* motif and each column represents one position within the *uX* motif. There is no correlation between *X* motif universality and sequence conservation: Pearson correlation coefficient $r = 0.04$, Spearman correlation coefficient $\rho = 0.07$ and Kendall coefficient $\tau = 0.05$. Furthermore, a two-tailed matched sample signed ranks Wilcoxon test shows that the 2 distributions are significantly different ($p < 10^{-4}$).

Fully conserved nucleotides in the LSU rRNA alignment are provided in a separate file: supp table S2c.pdf.

Nucleotide frequencies (%)	A	C	G	T
rRNA sequences	24.6	23.7	31.1	20.5
<i>X</i> circular code	25.0	25.0	25.0	25.0

Trinucleotide frequencies (%)					
AAA	1.9	CTA	1.1	AAC	1.7
AAG	2.4	CTT	1.1	AAT	1.4
ACA	1.1	GCA	1.5	ACC	1.6
ACG	1.4	GCG	2.0	ATC	1.1
ACT	1.2	GCT	1.5	ATT	1.0
AGA	1.6	GGA	2.3	CAG	1.5
AGC	1.9	GGG	3.3	CTC	1.2
AGG	2.4	GTG	2.1	CTG	1.7
AGT	1.6	TAA	1.7	GAA	2.6
ATA	1.0	TAG	1.4	GAC	1.4
ATG	1.3	TAT	0.7	GAG	2.2
CAA	1.3	TCA	1.0	GAT	1.5
CAC	1.0	TCC	1.4	GCC	2.0
CAT	0.9	TCG	1.3	GGC	2.2
CCA	1.2	TCT	1.0	GGT	2.2
CCC	1.9	TGA	1.9	GTA	1.6
CCG	2.3	TGC	1.3	GTC	1.4
CCT	1.5	TGG	2.1	GTT	1.4
CGA	1.8	TGT	1.2	TAC	1.0
CGC	1.5	TTA	1.1	TTC	1.0
CGG	2.3	TTG	1.4		
CGT	1.4	TTT	1.0		

Table S4. Nucleotide and trinucleotide composition of sequences in the SSU and LSU rRNA alignments. Trinucleotides belonging to the *X* circular code are highlighted in red. The mean frequency of *X* trinucleotides (1.58) is not significantly greater than the mean frequency of non-*X* trinucleotides (1.55), according to a Mann-Whitney U test $z = -0.51419$ ($p = 0.61$).

	1	2	3	4	5	6	7	8	9	10	11	12	13	
a	A	131(98)	0(0)	133(100)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)			ACACCGCCC
	C	0(0)	133(100)	0(0)	133(100)	130(98)	0(0)	133(100)	133(100)	132(99)				
	G	0(0)	0(0)	0(0)	0(0)	0(0)	133(100)	0(0)	0(0)	0(0)				
	T	2(2)	0(0)	0(0)	0(0)	3(2)	0(0)	0(0)	0(0)	1(1)				
b	A	133(100)	133(100)	0(0)	0(0)	0(0)	1(1)	1(1)	132(99)	133(100)	0(0)			AAGTCGTAAC
	C	0(0)	0(0)	1(1)	0(0)	133(100)	0(0)	0(0)	1(1)	0(0)	133(100)			
	G	0(0)	0(0)	131(98)	0(0)	0(0)	132(99)	0(0)	0(0)	0(0)	0(0)			
	T	0(0)	0(0)	1(1)	133(100)	0(0)	0(0)	132(99)	0(0)	0(0)	0(0)			
c	A	133(100)	0(0)	0(0)	0(0)	108(81)	8(6)	0(0)	1(1)	0(0)	0(0)	133(100)	0(0)	AGGTAGCCGTAG
	C	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	132(99)	122(92)	3(2)	0(0)	0(0)	25(19)	
	G	0(0)	133(100)	133(100)	0(0)	2(2)	95(71)	0(0)	2(2)	130(98)	0(0)	0(0)	97(73)	
	T	0(0)	0(0)	0(0)	133(100)	23(17)	30(23)	1(1)	8(6)	0(0)	133(100)	0(0)	11(8)	
d	A		133(100)	133(100)	0(0)	2(2)	0(0)	96(72)	0(0)	0(0)	0(0)			AATTGACGG
	C		0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	132(99)	0(0)	133(100)			
	G		0(0)	0(0)	0(0)	0(0)	133(100)	37(28)	0(0)	133(100)	0(0)			
	T		0(0)	0(0)	133(100)	131(98)	0(0)	0(0)	1(1)	0(0)	0(0)			
e	A	0(0)	133(100)	0(0)	133(100)	0(0)	133(100)	0(0)	0(0)	3(2)				TAGATACCc
	C	22(17)	0(0)	0(0)	0(0)	0(0)	0(0)	133(100)	132(99)	103(77)				
	G	0(0)	0(0)	133(100)	0(0)	4(3)	0(0)	0(0)	0(0)	27(20)				
	T	111(83)	0(0)	0(0)	0(0)	129(97)	0(0)	0(0)	1(1)	0(0)				
f	A	22(17)	0(0)	3(2)	2(2)	0(0)	0(0)	133(100)	106(80)	0(0)	121(92)			GCGGTGAATA
	C	1(1)	128(96)	2(2)	7(5)	6(5)	0(0)	0(0)	3(2)	1(1)	0(0)			
	G	110(83)	0(0)	128(96)	102(77)	0(0)	133(100)	0(0)	0(0)	0(0)	8(6)			
	T	0(0)	5(4)	0(0)	22(17)	126(95)	0(0)	0(0)	23(17)	131(99)	3(2)			
g	A	133(100)	0(0)	0(0)	99(74)	0(0)	0(0)							AGCAGC
	C	0(0)	0(0)	133(100)	34(26)	2(2)	133(100)							
	G	0(0)	133(100)	0(0)	0(0)	131(98)	0(0)							
	T	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)							
h	A	0(0)	0(0)	0(0)	0(0)	0(0)	131(98)	133(100)	0(0)	103(77)	0(0)			GCGGTAAATAC
	C	0(0)	133(100)	0(0)	0(0)	0(0)	0(0)	0(0)	15(11)	0(0)	132(99)			
	G	133(100)	0(0)	133(100)	133(100)	1(1)	2(2)	0(0)	3(2)	0(0)	1(1)			
	T	0(0.0)	0(0)	0(0)	0(0)	132(99)	0(0)	0(0)	115(86)	30(23)	0(0)			
i	A		13(10)	14(11)	69(52)	0(0)	42(32)	133(100)	0(0)	30(23)	30(23)	0(0)	131(98) 64(48)	GGACACGCAAG
	C		1(1)	0(0)	0(0)	74(56)	9(7)	0(0)	103(77)	0(0)	0(0)	113(85)	0(0) 0(0)	
	G		119(89)	118(89)	62(47)	0(0)	0(0)	78(59)	0(0)	0(0)	103(77)	36(27)	0(0) 69(52)	
	T		0(0)	0(0)	2(2)	59(44)	4(3)	0(0)	30(23)	0(0)	67(50)	20(15)	2(2) 0(0)	
j	A		132(99)	0(0)	0(0)	34(26)	0(0)	0(0)						AGGG
	C		1(1)	40(30)	39(29)	47(35)	0(0)	0(0)						
	G		0(0)	90(68)	52(39)	0(0)	133(100)	133(100)						
	T		0(0)	3(2)	42(32)	52(39)	0(0)	0(0)						
k	A	0(0)	77(59)	0(0)	38(29)	12(9)	0(0)	0(0)	1(1)	40(30)				TAGTTGGTG
	C	5(4)	2(2)	0(0)	1(1)	9(7)	0(0)	0(0)	18(14)	2(2)				
	G	0(0)	6(5)	133(100)	0(0)	0(0)	133(100)	124(93)	6(5)	91(68)				
	T	128(96)	46(35)	0(0)	94(71)	112(84)	0(0)	9(7)	108(81)	0(0)				
l	A	0(0)	0(0)	0(0)	130(98)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)			GTcAGcTcGT
	C	26(20)	12(9)	103(78)	0(0)	0(0)	92(69)	3(2)	106(80)	0(0)	0(0)			
	G	103(78)	1(1)	0(0)	3(2)	133(100)	0(0)	0(0)	26(20)	133(100)	0(0)			
	T	3(2)	120(90)	30(23)	0(0)	0(0)	41(31)	130(98)	1(1)	0(0)	133(100)			
m	A	25(19)	1(1)	133(100)	133(100)	0(0)	1(1)	129(97)	29(22)	0(0)				GC AACGAGC
	C	0(0)	90(68)	0(0)	0(0)	132(99)	1(1)	0(0)	0(0)	133(100)				
	G	98(74)	1(1)	0(0)	0(0)	0(0)	131(98)	4(3)	104(78)	0(0)				
	T	10(8)	41(31)	0(0)	0(0)	1(1)	0(0)	0(0)	0(0)	0(0)				

Table S5. Nucleotide composition of the 13 universal *X* motifs (*uX* motifs: number of species having an *X* motif) in the SSU rRNA alignments (prokaryotic 16S and eukaryotic 18S). Each row represents one *uX* motif and each column represents one position within the *uX* motif. Each cell contains the number (percentage) of species with a given nucleotide at each position in the *uX* motifs. The final column show the WebLogos for each of the *uX* motifs, constructed on the web server: <https://weblogo.berkeley.edu/logo.cgi>.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	0(0)	0(0)	0(0)	130(98)	0(0)	3(2)									
	0(0)	132(99)	0(0)	0(0)	132(99)	36(27)									
	32(24)	1(1)	133(100)	3(2)	0(0)	94(71)									
	101(76)	0(0)	0(0)	0(0)	1(1)	0(0)									
B	103(77)	0(0)	6(5)	0(0)	3(2)	0(0)	131(98)	0(0)	0(0)	0(0)	0(0)	0(0)	3(2)	0(0)	0(0)
	26(20)	133(100)	93(70)	3(2)	129(97)	0(0)	2(2)	3(2)	0(0)	0(0)	132(99)	0(0)	0(0)	113(85)	0(0)
	0(0)	0(0)	0(0)	0(0)	0(0)	132(99)	0(0)	0(0)	133(100)	0(0)	0(0)	133(100)	125(94)	0(0)	0(0)
	4(3)	0(0)	34(26)	130(98)	1(1)	1(1)	0(0)	130(98)	0(0)	133(100)	1(1)	0(0)	5(4)	20(15)	133(100)
C	27(20)	0(0)	133(100)	7(5)	0(0)	28(21)	0(0)	6(5)	0(0)	4(3)					
	65(49)	100(75)	0(0)	4(3)	133(100)	104(78)	0(0)	1(1)	10(8)	33(25)					
	33(25)	0(0)	0(0)	2(2)	0(0)	0(0)	0(0)	102(77)	123(92)	76(57)					
	8(6)	33(25)	0(0)	120(90)	0(0)	1(1)	133(100)	24(18)	0(0)	20(15)					
D	1(1)	0(0)	0(0)	133(100)	0(0)	5(4)	0(0)	0(0)	0(0)	1(1)	0(0)	0(0)	0(0)	0(0)	
	2(2)	36(27)	0(0)	0(0)	0(0)	125(94)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	46(35)	
	130(98)	0(0)	133(100)	0(0)	133(100)	0(0)	0(0)	133(100)	133(100)	132(99)	0(0)	0(0)	0(0)	0(0)	
	0(0)	97(73)	0(0)	0(0)	0(0)	3(2)	133(100)	0(0)	0(0)	0(0)	133(100)	133(100)	87(65)		
E	133(100)	10(8)	130(98)	67(50)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)					
	0(0)	2(2)	0(0)	66(50)	133(100)	0(0)	1(1)	133(100)	0(0)	11(8)					
	0(0)	121(91)	0(0)	0(0)	0(0)	132(99)	0(0)	0(0)	133(100)	0(0)					
	0(0)	0(0)	3(2)	0(0)	0(0)	1(1)	132(99)	0(0)	0(0)	122(92)					
F	30(23)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	33(25)	0(0)	0(0)	32(24)	100(75)		
	0(0)	82(62)	0(0)	0(0)	0(0)	10(8)	9(7)	133(100)	0(0)	133(100)	133(100)	75(56)	0(0)		
	103(77)	3(2)	0(0)	133(100)	0(0)	0(0)	0(0)	0(0)	100(75)	0(0)	0(0)	10(8)	8(6)		
	0(0)	48(36)	133(100)	0(0)	0(0)	123(92)	124(93)	0(0)	0(0)	0(0)	0(0)	16(12)	25(19)		
G	1(1)	0(0)	0(0)	133(100)	132(99)	90(68)									
	2(2)	64(48)	0(0)	0(0)	1(1)	28(21)									
	126(95)	0(0)	133(100)	0(0)	0(0)	14(11)									
	4(3)	69(52)	0(0)	0(0)	0(0)	1(1)									
H	133(100)	132(99)	129(97)	0(0)	0(0)	24(18)	8(6)								
	0(0)	1(1)	2(2)	108(81)	116(87)	1(1)	15(11)								
	0(0)	0(0)	1(1)	0(0)	1(1)	102(77)	42(32)								
	0(0)	0(0)	1(1)	25(19)	16(12)	6(5)	68(51)								
I	0(0)	0(0)	1(1)	133(100)	9(7)	0(0)	46(35)	0(0)	0(0)						
	2(2)	0(0)	0(0)	0(0)	0(0)	0(0)	43(33)	118(89)	125(94)						
	0(0)	133(100)	132(99)	0(0)	124(93)	133(100)	7(5)	0(0)	0(0)						
	131(98)	0(0)	0(0)	0(0)	0(0)	0(0)	36(27)	15(11)	8(6)						
J	131(98)	8(6)	4(3)	27(20)	0(0)	2(2)	133(100)	133(100)	5(4)	1(1)					
	0(0)	0(0)	0(0)	8(6)	18(14)	0(0)	0(0)	0(0)	0(0)	73(55)					
	2(2)	125(94)	129(97)	52(39)	0(0)	130(98)	0(0)	0(0)	128(96)	39(30)					
	0(0)	0(0)	0(0)	46(35)	114(86)	0(0)	0(0)	0(0)	0(0)	19(14)					
K	5(4)	1(1)	133(100)	133(100)	0(0)	3(2)	130(98)	2(2)	50(38)						
	0(0)	24(18)	0(0)	0(0)	106(80)	1(1)	1(1)	25(19)	18(14)						
	70(53)	13(10)	0(0)	0(0)	0(0)	129(97)	0(0)	64(48)	33(25)						
	58(44)	95(71)	0(0)	0(0)	27(20)	0(0)	2(2)	42(32)	32(24)						
L	129(97)	0(0)	0(0)	0(0)	5(4)	0(0)	0(0)	4(3)	1(1)						
	4(3)	0(0)	133(100)	0(0)	6(5)	0(0)	0(0)	0(0)	132(99)						
	0(0)	130(98)	0(0)	2(2)	121(91)	133(100)	0(0)	0(0)	0(0)						
	0(0)	3(2)	0(0)	131(98)	1(1)	0(0)	133(100)	129(97)	0(0)						
M	39(29)	1(1)	1(1)	0(0)	102(77)	133(100)	133(100)	0(0)	4(3)	0(0)	0(0)				
	11(8)	10(8)	2(2)	6(5)	6(5)	0(0)	0(0)	121(91)	0(0)	10(8)	133(100)				
	77(58)	122(92)	0(0)	23(17)	23(17)	0(0)	0(0)	10(8)	129(97)	123(92)	0(0)				
	6(5)	0(0)	131(98)	2(2)	2(2)	0(0)	0(0)	2(2)	0(0)	0(0)	0(0)				

N	A	126(95)	0(0)	3(2)	0(0)	0(0)	133(100)						
	C	5(4)	22(17)	19(14)	0(0)	28(21)	0(0)						
	G	0(0)	69(52)	111(83)	0(0)	34(26)	0(0)						
	T	2(2)	42(32)	0(0)	133(100)	71(53)	0(0)						
O	A	131(98)	0(0)	0(0)	133(100)	0(0)	133(100)	2(2)	133(100)				
	C	0(0)	0(0)	1(1)	0(0)	132(99)	0(0)	0(0)	0(0)				
	G	1(1)	133(100)	0(0)	0(0)	0(0)	133(100)	0(0)	131(98)				
	T	1(1)	0(0)	132(99)	0(0)	1(1)	0(0)	0(0)	0(0)				
P	A	4(3)	4(3)	130(98)	118(89)	94(71)	133(100)	0(0)	10(8)				
	C	71(53)	5(4)	2(2)	6(5)	0(0)	0(0)	2(2)	1(1)				
	G	24(18)	121(91)	0(0)	0(0)	38(29)	0(0)	0(0)	108(82)				
	T	34(26)	3(2)	1(1)	9(7)	0(0)	0(0)	131(98)	12(9)				
Q	A	100(75)	1(1)	0(0)	0(0)	7(5)	17(13)						
	C	33(25)	4(3)	1(1)	0(0)	0(0)	0(0)						
	G	0(0)	97(73)	132(99)	0(0)	51(38)	110(83)						
	T	0(0)	31(23)	0(0)	133(100)	75(56)	6(5)						
R	A	0(0)	0(0)	130(98)	25(19)	0(0)	133(100)	130(98)	4(3)	28(21)			
	C	121(92)	0(0)	0(0)	13(10)	0(0)	0(0)	3(2)	0(0)	0(0)			
	G	0(0)	128(97)	0(0)	6(5)	133(100)	0(0)	0(0)	129(97)	105(79)			
	T	11(8)	4(3)	2(2)	89(67)	0(0)	0(0)	0(0)	0(0)	0(0)			
S	A	3(2)	2(2)	1(1)	3(2)	98(74)	129(97)	5(4)	132(99)	3(2)	0(0)	1(1)	
	C	5(4)	0(0)	0(0)	12(9)	11(8)	2(2)	5(4)	1(1)	0(0.0)	10(8)	129(97)	
	G	124(93)	130(98)	0(0)	3(2)	18(14)	0(0)	18(14)	0(0)	1(1)	0(0)	0(0)	
	T	1(1)	1(1)	132(99)	115(86)	6(5)	2(2)	105(79)	0(0)	129(97)	123(92)	3(2)	

Table S6. Nucleotide composition of the 19 universal uX motifs (uX motifs: number of species having an uX motif) in the LSU rRNA alignments (prokaryotic 23S and eukaryotic 25/28S). Each row represents one uX motif and each column represents one position within the uX motif. Each cell contains the number (percentage) of species with a given nucleotide at each position in the uX motifs. The final column show the WebLogos for each of the uX motifs, constructed on the web server: <https://weblogo.berkeley.edu/logo.cgi>.

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>	<i>m</i>
AAC	0(0)	1(0)	8(2)	0(0)	0(0)	1(0)	0(0)	15(3)	46(9)	32(11)	0(0)	0(0)	155(40)
AAT	0(0)	0(0)	0(0)	0(0)	0(0)	3(1)	0(0)	113(23)	24(5)	2(1)	0(0)	0(0)	1(0)
ACC	130(50)	0(0)	7(2)	0(0)	131(34)	0(0)	0(0)	35(7)	0(0)	0(0)	0(0)	0(0)	1(0)
ATC	0(0)	0(0)	6(1)	0(0)	1(0)	0(0)	0(0)	0(0)	3(1)	0(0)	0(0)	1(0)	0(0)
ATT	0(0)	0(0)	1(0)	131(49)	2(1)	3(1)	0(0)	30(6)	11(2)	0(0)	1(0)	0(0)	0(0)
CAG	0(0)	0(0)	0(0)	0(0)	22(6)	1(0)	194(74)	38(8)	50(10)	0(0)	0(0)	101(21)	4(1)
CTC	2(1)	1(0)	0(0)	0(0)	1(0)	0(0)	0(0)	0(0)	0(0)	0(0)	1(0)	90(18)	0(0)
CTG	0(0)	0(0)	2(0)	0(0)	55(14)	5(1)	1(0)	0(0)	1(0)	31(11)	5(1)	0(0)	0(0)
GAA	0(0)	83(24)	0(0)	8(3)	0(0)	105(28)	0(0)	9(2)	0(0)	114(40)	0(0)	0(0)	26(7)
GAC	0(0)	1(0)	0(0)	95(35)	0(0)	4(1)	0(0)	1(0)	84(16)	1(0)	8(2)	0(0)	3(1)
GAG	0(0)	1(0)	1(0)	0(0)	4(1)	0(0)	0(0)	0(0)	16(3)	34(12)	18(5)	0(0)	131(34)
GAT	0(0)	0(0)	0(0)	0(0)	149(38)	22(6)	0(0)	2(0)	65(13)	9(3)	30(8)	0(0)	23(6)
GCC	130(50)	0(0)	88(20)	0(0)	4(1)	3(1)	68(26)	0(0)	0(0)	8(3)	0(0)	25(5)	0(0)
GGC	0(0)	0(0)	2(0)	35(13)	0(0)	4(1)	0(0)	0(0)	45(9)	8(3)	17(5)	1(0)	29(7)
GGT	0(0)	0(0)	72(17)	1(0)	0(0)	97(26)	0(0)	129(26)	36(7)	2(1)	107(30)	26(5)	5(1)
GTA	0(0)	131(38)	217(50)	0(0)	0(0)	0(0)	0(0)	55(11)	21(4)	0(0)	56(16)	25(5)	3(1)
GTC	0(0)	130(37)	0(0)	0(0)	22(6)	0(0)	0(0)	0(0)	114(22)	29(10)	7(2)	182(37)	2(1)
GTT	0(0)	0(0)	0(0)	0(0)	0(0)	20(5)	0(0)	0(0)	0(0)	12(4)	92(26)	25(5)	7(2)
TAC	0(0)	0(0)	2(0)	0(0)	0(0)	109(29)	0(0)	41(8)	3(1)	0(0)	0(0)	0(0)	0(0)
TTC	0(0)	1(0)	24(6)	0(0)	0(0)	3(1)	0(0)	30(6)	0(0)	0(0)	17(5)	12(2)	0(0)

Table S7. Trinucleotide composition of the 13 universal *X* motifs (*uX* motifs) in the SSU rRNA alignments (prokaryotic 16S and eukaryotic 18S). Each row represents an *X* trinucleotide and each column represents a *uX* motif. Each cell contains the number (percentage) of trinucleotides observed in each *uX* motif.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
AAC	0(0)	0(0)	5(2)	0(0)	64(21)	0(0)	6(2)	14(4)	0(0)	0(0)	107(19)	0(0)	121(25)	16(5)	0(0)	1(0)	0(0)	2(0)	4(1)
AAT	0(0)	0(0)	0(0)	0(0)	0(0)	8(2)	0(0)	7(2)	0(0)	2(0)	29(5)	0(0)	2(0)	47(15)	0(0)	92(25)	1(0)	0(0)	90(20)
ACC	6(2)	71(10)	8(3)	0(0)	66(21)	33(7)	28(10)	99(30)	35(9)	1(0)	2(0)	0(0)	3(1)	9(3)	0(0)	3(1)	1(0)	3(1)	5(1)
ATC	101(34)	6(1)	116(40)	2(0)	0(0)	0(0)	2(1)	14(4)	2(1)	1(0)	20(4)	3(1)	0(0)	7(2)	0(0)	1(0)	0(0)	11(3)	10(2)
ATT	3(1)	0(0)	24(8)	0(0)	0(0)	54(11)	53(19)	12(4)	0(0)	2(0)	52(9)	1(0)	0(0)	0(0)	0(0)	9(2)	0(0)	0(0)	109(24)
CAG	0(0)	3(0)	0(0)	43(9)	43(14)	0(0)	1(0)	12(4)	0(0)	101(22)	2(0)	14(4)	2(0)	0(0)	19(7)	1(0)	0(0)	0(0)	61(13)
CTC	0(0)	90(13)	8(3)	0(0)	0(0)	1(0)	0(0)	1(0)	8(2)	0(0)	2(0)	79(22)	4(1)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)
CTG	2(1)	0(0)	100(35)	127(27)	0(0)	73(15)	6(2)	8(2)	40(10)	0(0)	11(2)	123(34)	0(0)	19(6)	0(0)	1(0)	0(0)	0(0)	13(3)
GAA	0(0)	0(0)	4(1)	0(0)	0(0)	0(0)	120(42)	125(37)	51(13)	131(29)	15(3)	0(0)	20(4)	64(21)	2(1)	113(30)	12(3)	174(42)	3(1)
GAC	128(43)	8(1)	0(0)	0(0)	0(0)	0(0)	0(0)	1(0)	44(11)	1(0)	24(4)	0(0)	0(0)	0(0)	0(0)	10(3)	30(9)	19(5)	1(0)
GAG	4(1)	1(0)	3(1)	125(27)	0(0)	16(3)	0(0)	1(0)	116(29)	1(0)	59(11)	0(0)	0(0)	0(0)	130(45)	0(0)	53(15)	7(2)	4(1)
GAT	2(1)	127(18)	1(0)	1(0)	0(0)	8(2)	0(0)	0(0)	2(1)	25(6)	53(9)	3(1)	0(0)	3(1)	0(0)	49(13)	30(9)	84(20)	14(3)
GCC	0(0)	0(0)	9(3)	0(0)	1(0)	91(19)	3(1)	2(1)	41(10)	30(7)	0(0)	0(0)	6(1)	0(0)	0(0)	34(9)	1(0)	40(10)	6(1)
GGC	27(9)	171(24)	4(1)	0(0)	0(0)	0(0)	18(6)	2(1)	10(3)	17(4)	73(13)	0(0)	189(39)	0(0)	0(0)	3(1)	92(26)	51(12)	1(0)
GGT	2(1)	23(3)	5(2)	124(27)	0(0)	5(1)	34(12)	2(1)	0(0)	74(16)	10(2)	0(0)	23(5)	80(26)	0(0)	6(2)	54(15)	2(0)	17(4)
GTA	1(0)	0(0)	0(0)	0(0)	0(0)	0(0)	0(0)	11(3)	9(2)	4(1)	66(12)	4(1)	101(21)	1(0)	6(2)	47(13)	8(2)	0(0)	4(1)
GTC	13(4)	137(19)	0(0)	0(0)	131(42)	0(0)	0(0)	2(1)	33(8)	13(3)	12(2)	0(0)	6(1)	27(9)	0(0)	2(1)	0(0)	11(3)	11(2)
GTT	11(4)	2(0)	0(0)	43(9)	0(0)	69(14)	13(5)	21(6)	3(1)	47(10)	1(0)	124(34)	2(0)	22(7)	0(0)	0(0)	70(20)	3(1)	102(22)
TAC	0(0)	32(5)	0(0)	0(0)	3(1)	0(0)	1(0)	1(0)	0(0)	0(0)	0(0)	0(0)	0(0)	10(3)	132(46)	1(0)	0(0)	0(0)	2(0)
TTC	0(0)	37(5)	0(0)	0(0)	1(0)	123(26)	0(0)	0(0)	1(0)	0(0)	20(4)	11(3)	0(0)	0(0)	0(0)	0(0)	0(0)	3(1)	2(0)

Table S8. Trinucleotide composition of the 19 universal *X* motifs (*uX* motifs) in the LSU rRNA alignments (prokaryotic 23S and eukaryotic 25/28S). Each row represents an *X* trinucleotide and each column represents a *uX* motif. Each cell contains the number (percentage) of trinucleotides observed in each *uX* motif.

	Mainly globular	Globular domain with a long unstructured extension	Hairpin extension	Helical and hairpin extension
SSU	S2, S3, S4, S14, S15	S7, S9, S11, S12, S13, S19	S5, S8, S10	S17
LSU	L1, L6, L7, L11, L12, L23, L29, L30	L2, L15, L16, L18, L24	L3, L4, L5, L13, L14, L22	

Table S9. Universal ribosomal proteins, classified according to their known 3D structure as defined in Smith et al. 2008. Extensions refer to protein segments that extend away from the more compact or globular part of the protein for a significant distance. Proteins in contact with a *uX* motif are shown in bold.

Bacteria	Archaea	Eukaryota
<i>Acinetobacter sp.</i>	<i>Aeropyrum pernix</i>	<i>Adineta vaga</i>
<i>Agrobacterium tumefaciens</i>	<i>Archaeoglobus fulgidus</i>	<i>Aedes albopictus</i>
<i>Anabaena variabilis</i>	<i>Caldivirga maquilingensis</i>	<i>Anolis carolinensis</i>
<i>Azoarcus sp.</i>	<i>Haloarcula marismortui</i>	<i>Arabidopsis thaliana</i>
<i>Bacillus anthracis</i>	<i>Halobacterium sp.</i>	<i>Caenorhabditis briggsae</i>
<i>Bacteroides thetaiotaomicron</i>	<i>Haloferax volcanii</i>	<i>Caenorhabditis elegans</i>
<i>Bartonella henselae</i>	<i>Haloquadratum walsbyi</i>	<i>Cryptosporidium hominis</i>
<i>Bifidobacterium longum</i>	<i>Halorubrum lacusprofundi</i>	<i>Cyanidioschyzon merolae</i>
<i>Blochmannia floridanus</i>	<i>Hyperthermus butylicus</i>	<i>Danio rerio</i>
<i>Bradyrhizobium japonicum</i>	<i>Ignicoccus hospitalis</i>	<i>Dictyostelium discoideum</i>
<i>Buchnera aphidicola</i>	<i>Metallosphaera sedula</i>	<i>Drosophila melanogaster</i>
<i>Burkholderia sp.</i>	<i>Methanocaldococcus jannaschii</i>	<i>Eremothecium gossypii</i>
<i>Caulobacter crescentus</i>	<i>Methanococcoides burtonii</i>	<i>Gallus gallus</i>
<i>Chlamydomonas reinhardtii</i>	<i>Methanococcus aeolicus</i>	<i>Guillardia theta</i>
<i>Chlorobium tepidum</i>	<i>Methanocorpusculum labreanum</i>	<i>Homo sapiens</i>
<i>Coxiella burnetii</i>	<i>Methanoculleus marisnigri</i>	<i>Latimeria chalumnae</i>
<i>Crocospaera watsonii</i>	<i>Methanopyrus kandleri</i>	<i>Leishmania major</i>
<i>Cytophaga hutchinsonii</i>	<i>Methanoregula boonei</i>	<i>Monodelphis domestica</i>
<i>Dechloromonas aromatica</i>	<i>Methanosaeta thermophila</i>	<i>Mus musculus</i>
<i>Dehalococcoides ethenogenes</i>	<i>Methanosarcina acetivorans</i>	<i>Oryza sativa</i>
<i>Deinococcus radiodurans</i>	<i>Methanosarcina barkeri</i>	<i>Pan troglodytes</i>
<i>Escherichia coli</i>	<i>Methanosarcina mazei</i>	<i>Plasmodium falciparum</i>
<i>Fusobacterium nucleatum</i>	<i>Methanosphaera stadtmanae</i>	<i>Rattus norvegicus</i>
<i>Geobacillus kaustophilus</i>	<i>Methanospirillum hungatei</i>	<i>Saccharomyces cerevisiae</i>
<i>Geobacter sulfurreducens</i>	<i>Methanothermobacter thermautotrophicus</i>	<i>Schizosaccharomyces pombe</i>
<i>Gloeobacter violaceus</i>	<i>Nanoarchaeum equitans</i>	<i>Tetrahymena thermophila</i>
<i>Gluconobacter oxydans</i>	<i>Natronomonas pharaonis</i>	<i>Thalassiosira pseudonana</i>
<i>Haemophilus influenzae</i>	<i>Picrophilus torridus</i>	<i>Trypanosoma brucei</i>
<i>Helicobacter hepaticus</i>	<i>Pyrobaculum calidifontis</i>	<i>Yarrowia lipolytica</i>
<i>Legionella pneumophila</i>	<i>Pyrococcus furiosus</i>	<i>Xenopus laevis</i>
<i>Leifsonia xyli</i>	<i>Staphylothermus marinus</i>	
<i>Listeria monocytogenes</i>	<i>Sulfolobus acidocaldarius</i>	
<i>Magnetococcus sp.</i>	<i>Sulfolobus tokodaii</i>	
<i>Magnetospirillum magnetotacticum</i>	<i>Thermococcus kodakarensis</i>	

<i>Mesoplasma florum</i>	<i>Thermofilum pendens</i>	
<i>Mycobacterium leprae</i>	<i>Thermoplasma volcanium</i>	
<i>Neisseria gonorrhoeae</i>		
<i>Nitrosomonas europaea</i>		
<i>Novosphingobium aromaticivorans</i>		
<i>Oceanobacillus iheyensis</i>		
<i>Photorhabdus luminescens</i>		
<i>Polaromonas sp.</i>		
<i>Porphyromonas gingivalis</i>		
<i>Propionibacterium acnes</i>		
<i>Pseudomonas aeruginosa</i>		
<i>Ralstonia eutropha</i>		
<i>Rhodobacter sphaeroides</i>		
<i>Rhodopirellula baltica</i>		
<i>Shewanella oneidensis</i>		
<i>Sinorhizobium meliloti</i>		
<i>Staphylococcus aureus</i>		
<i>Streptococcus pneumoniae</i>		
<i>Streptomyces coelicolor</i>		
<i>Symbiobacterium thermophilum</i>		
<i>Synechococcus sp.</i>		
<i>Synechocystis sp.</i>		
<i>Thermoanaerobacter tengcongensis</i>		
<i>Thermosynechococcus elongatus</i>		
<i>Thermotoga maritima</i>		
<i>Thermus thermophilus</i>		
<i>Treponema pallidum</i>		
<i>Tropheryma whipplei</i>		
<i>Wolbachia_endosymbiont</i>		
<i>Xanthomonas axonopodis</i>		
<i>Yersinia pestis</i>		
<i>Zymomonas mobilis</i>		

Table S10. List of the 133 organisms included in the multiple sequence alignments of LSU rRNAs (23S/28S and 5S) and SSU rRNAs (16S/18S).