Supplementary Table 1: Detailed results of the native screen and the random control screen.

Native screen							
Set		Clusters	Size (MB)	% of input	% of genome	cluster length average	cluster length maximum
A: Set 1	$P > 0.5$	91,676	12.47	15.09	0.44	136	1320
B: Set 1	$P > 0.9$	35,985	5.48	6.62	0.19	152	1320
C: Set 2	$P > 0.5$	20,391	2.80	11.52	0.10	137	665
D: Set 2	$P > 0.9$	8,802	1.34	5.50	0.05	152	665
E: Set 3	$P > 0.5$	2,916	0.38	5.57	0.01	131	488
F: Set 3	$P > 0.9$	996	0.14	2.03	0.00	139	488

Randomized screen								
Set		Clusters	Overlap native A	Size (MB)	% of input	% of genome	Cluster length average	Cluster length maximum
Set 1	$P > 0.5$	26,508	9039	3.20	3.87	0.11	121	496
Set 1	$P > 0.9$	6,898	2555	0.89	1.08	0.03	130	496
Set 2	$P > 0.5$	6,551	2158	0.81	3.35	0.03	124	394
Set 2	$P > 0.9$	2,281	881	0.31	1.26	0.01	134	394
Set 3	$P > 0.5$	795	179	0.096	1.40	0.00	121	338
Set 3	$P > 0.9$	208	63	0.026	0.38	0.00	127	279

Set 1: human/mouse/rat/dog, Set 2 = Set 1 + chicken, Set 3 = Set 2 + fugu or zebrafish
"cluster" refers to clustered regions of overlapping `RNAz` hits as described in "Methods"

Supplementary Table 2: MicroRNAs missing from our input set

Name	Conservation	Repeat	Other
hsa-let-7g	rat missing		
hsa-let-7i	gap in dog		
hsa-mir-9-1		simple Repeat	
hsa-mir-15a	rat missing		
hsa-mir-16-1	rat missing		
hsa-mir-22			overlap with coding region
hsa-mir-23a			`PhastCons` artifact[1]
hsa-mir-28		LINE	
hsa-mir-95		LINE	
hsa-mir-130b		SINE	
hsa-mir-133a-2			overlap with coding region
hsa-mir-135a-1	part of rat sequence missing		
hsa-mir-138-1	mouse missing		
hsa-mir-147	`PhastCons` region <50		
hsa-mir-148a	rat missing		
hsa-mir-149	rat missing		
hsa-mir-150			overlap with coding region
hsa-mir-151		LINE	
hsa-mir-155	rat missing		
hsa-mir-182	part of rat sequence missing		
hsa-mir-197	long gap in mouse		
hsa-mir-198	rat missing		
hsa-mir-199b	rat missing		
hsa-mir-203			`PhastCons` artifact[1]
hsa-mir-205			overlap with coding region
hsa-mir-212		low complexity	
hsa-mir-302a	rat missing		
hsa-mir-302b	rat missing		
hsa-mir-302c	rat missing		
hsa-mir-302d	rat missing		
hsa-mir-321		tRNA	
hsa-mir-325		LINE	
hsa-mir-326		Arthur 1	
hsa-mir-328	`PhastCons` region <50		
hsa-mir-330		SINE	
hsa-mir-335	rat missing	SINE	
hsa-mir-337	dog missing		
hsa-mir-340	rat missing	MARNA	
hsa-mir-345		SINE	
hsa-mir-367	rat missing		
hsa-mir-370		SINE	
hsa-mir-371	`PhastCons` region <50		
hsa-mir-372	`PhastCons` region <50		
hsa-mir-373	rat and mouse missing		
hsa-mir-374	gaps in mouse and rat	LINE	

[1] `PhastCons` region extends into the very gap-rich surrounding of the miRNA. Alignment discarded because it contains too many gaps.

Supplementary Table 3: H/ACA snoRNAs missing from our input set

Name	Conservation	Repeat	Other
ACA2A	gap in mouse and rat		
ACA5	`PhastCons` region <50		
ACA5b	`PhastCons` region <50		
ACA10	`PhastCons` region <50		
ACA11	gap in mouse		
ACA29			alignment artifact[1]
ACA33	`PhastCons` region <50		
ACA39	`PhastCons` region <50		
ACA42	not detected by `PhastCons`		
ACA48	not detected by `PhastCons`		
ACA56	rat missing		
ACA59 (Chr. 1)		SINE	
ACA59 (Chr. 17)		SINE	
ACA67	`PhastCons` region <50		
U17a		other	
U17b		other	
U64			alignment artifact[1]
U66	`PhastCons` region <50		
U71a	`PhastCons` region <50		
U71b	rat missing		
U98b	`PhastCons` region <50		

[1] The sequence in chicken is much longer and opens up long gaps in the other sequences, which are thus discarded.

Supplementary Table 4: Selected ncRNAs from literature with conserved RNA secondary structures detected in our screen.

Name	Type	Max. P	Hits	Comment
U11	snRNA	0.98	1	
U12	snRNA	0.94	2	
U4atac	snRNA	0.71	3	
U6atac	snRNA	0.98	12	
RNAseP	Ribozyme	0.57	1	
UM 9(5)	Transcript of unknown function	1.0	8	Transcript was found to be differentially expressed in the brain, 7 of the 8 hits match the same region of this long (1241nt) transcript
HUC-1	Other functional transcript	0.95	1	Tissue specific transcript that enhances H19 transcription (an antisense transcript for imprinting)
MALAT-1	transcript of unknown function	1.0	3	three independent hits along this 8kb transcript, which was identified in lung cancer cells as ncRNA
NCRMS	Other functional transcript	0.90	3	three independent hits in this 1.8 kb transcript; identified in rhabdomyosarcoma (RMS); host gene of mir-135a-2
BCMS	Other functional transcript	0.71	1	B-cell neoplasia associated transcript
aHIF	antisense transcript	0.98	1	aHIF is complementary to the 3' untranslated region of HIF1alpha mRNA, which encodes a protein known to stabilize p53 protein during hypoxia and to act as a transcription factor for hypoxia inducible genes
Air	Antisense transcript	0.96	8	Classical mouse model for imprinted antisense transcription.
CNS1	Other functional transcript	0.83	1	Expression of CNS1 accompanies the induction of the hyperacetylation of histone H3 on nucleosomes associated with the interleukin (IL)-4, IL-13 and IL-5 genes in developing Th2 cells
HOXA11 AS	Antisense transcript	0.53	1	
GA3824	Transcript of unknown function	0.74	1	Homo sapiens noncoding RNA GA3824 implicated in autism
XIST	Other functional transcript	1.0	3	Three independent hits in the long transcript responsible for X-inactivation in mammals
TTTY11	Transcript of unknown function	0.98	12	Identified in testis
TTTY3	Transcript of unknown function	0.86	1	Identified in testis
TTTY23	Transcript of unknown function	0.54	1	Identified in testis
His-1	Transcript of unknown function	1.0	2	Two independent hits on the same transcript; activation of this transcript leads to carcinogenesis

Supplementary Table 5: Selected RNAz hits in intergenic regions overlapping with 'transfrags'

Cluster	Chr.	Structure	From	To	Strand	N	Length	ID	SCI	*z*	*P*
61977	chr20	116862	21,196,730	21,196,846	+	5	120	65.64	0.45	-2.4	0.989825
		116867	21,196,806	21,196,922	+	5	120	70.33	0.36	-1.91	0.831492
109551	chrX	196844	45,492,123	45,492,242	+	6	120	75.75	0.51	-1.9	0.775894
		196847	45,492,123	45,492,242	-	6	120	75.75	0.72	-2.49	0.989492
89549	chr6	161585	85,026,793	85,026,908	-	5	120	69.8	0.68	-2.07	0.994139
		161589	85,026,832	85,026,922	-	5	108	64.42	0.56	-2.01	0.976256
90357	chr6	163072	100,904,835	100,904,946	+	5	120	79.91	0.68	-2.25	0.958033
		163067	100,904,795	100,904,915	+	5	120	88.49	0.83	-2.38	0.990587
29420	chr13	57463	104,046,105	104,046,204	+	5	99	75.66	0.43	-2.48	0.990074
		57467	104,046,105	104,046,204	-	5	99	75.66	0.28	-1.95	0.852306
63887	chr21	488	25,637,739	25,637,825	+	5	92	75	0.64	-1.8	0.903541
		489	25,637,739	25,637,825	-	5	92	75	0.67	-2.03	0.976266
62565	chr20	118029	37,552,028	37,552,147	+	4	120	70.15	0.58	-2.1	0.988436
		118033	37,552,068	37,552,163	-	4	96	73.61	0.65	-2.16	0.991237
		118032	37,552,068	37,552,163	+	4	96	73.61	0.69	-1.5	0.742403
89785	chr6	162009	91,142,625	91,142,738	-	4	120	72.99	0.45	-2.09	0.555579
		162008	91,142,585	91,142,698	-	4	120	78.25	0.64	-2.5	0.989123
		162007	91,142,585	91,142,698	+	4	120	78.25	0.54	-2.44	0.907347
62382	chr20	117675	34,330,779	34,330,896	-	4	120	69.94	0.56	-2.55	0.99889
91828	chr6	165586	148,404,320	148,404,440	+	5	120	81.75	0.74	-2.21	0.944193
		165590	148,404,320	148,404,440	-	5	120	81.75	0.61	-1.78	0.537876
		165591	148,404,360	148,404,476	+	5	120	83.33	0.58	-3.24	0.973039
		165593	148,404,360	148,404,476	-	5	120	83.33	0.58	-2.98	0.95201
27930	chr13	54890	64,513,126	64,513,242	-	4	120	81.21	0.87	-1.7	0.854404
		54893	64,513,162	64,513,257	-	4	95	89.07	0.95	-2.71	0.996902
92734	chr7	167098	12,673,969	12,674,073	+	5	120	74	0.66	-2.4	0.99671
		167099	12,673,969	12,674,073	-	5	120	74	0.58	-1.95	0.932432
96359	chr7	173573	115,002,715	115,002,822	+	5	120	78.7	0.73	-1.79	0.876986

Cluster	Chr.	Structure	From	To	Strand	N	Length	ID	SCI	*z*	*P*
		173578	115,002,715	115,002,822	-	5	120	78.7	0.69	-2.35	0.982449
94898	chr7	171020	72,754,119	72,754,238	-	4	120	74.09	0.74	-2	0.988917
		171039	72,754,198	72,754,318	-	4	120	68.61	0.49	-1.79	0.757058
		171011	72,754,079	72,754,198	-	4	120	77.44	0.75	-1.98	0.958662
		170968	72,754,002	72,754,119	+	4	120	80.97	0.79	-1.81	0.841935
63001	chr20	118831	44,330,281	44,330,386	+	4	120	72.03	0.62	-2.23	0.995101
		118835	44,330,320	44,330,425	+	4	120	74.75	0.61	-1.77	0.774982
110221	chrX	198313	68,360,418	68,360,538	-	4	120	80.08	0.65	-2.23	0.937577
		198338	68,360,458	68,360,578	-	4	120	83.15	0.8	-2.68	0.990857
96693	chr7	174122	126,489,350	126,489,467	+	4	120	79.8	0.69	-1.71	0.703828
		174115	126,489,310	126,489,428	+	4	120	75.85	0.65	-2.26	0.987136
61291	chr20	115658	6,772,144	6,772,247	+	4	104	77.53	0.77	-2.76	0.999149
		115652	6,772,104	6,772,223	+	4	120	78.44	0.88	-2.24	0.992374
110219	chrX	198299	68,326,811	68,326,931	+	4	120	84.03	0.77	-2.37	0.96475
		198309	68,326,851	68,326,957	+	4	120	78.17	0.7	-1.81	0.840742
29369	chr13	57378	102,471,382	102,471,502	+	5	120	81.14	0.62	-3.61	0.998111
		57380	102,471,382	102,471,502	-	5	120	81.14	0.73	-3.32	0.998255
28163	chr13	55318	70,554,914	70,555,077	-	5	163	82.09	0.7	-2.89	0.990487
		55315	70,554,914	70,555,077	+	5	163	82.09	0.73	-2.9	0.992806
29090	chr13	56926	94,774,638	94,774,757	-	4	120	81.4	0.72	-2.32	0.965953
		56925	94,774,638	94,774,757	+	4	120	81.4	0.62	-2.39	0.922944
109341	chrX	196480	40,105,701	40,105,816	+	4	120	77.59	0.68	-2.56	0.996204
91671	chr6	165322	143,190,370	143,190,461	-	4	101	75.13	0.63	-3.25	0.999892
33379	chr14	64344	97,086,718	97,086,838	-	4	120	75	0.62	-3.14	0.999842
		64346	97,086,758	97,086,878	+	4	120	73.89	0.59	-3.33	0.999913
		64342	97,086,718	97,086,838	+	4	120	75	0.72	-4.25	0.999986
96254	chr7	173394	114,006,710	114,006,826	+	5	120	78.88	0.59	-2.98	0.997166
91809	chr6	165563	148,188,892	148,188,980	+	5	110	77.04	0.61	-2.58	0.994454
		165560	148,188,813	148,188,930	-	5	120	83.76	0.73	-4.34	0.998788

Cluster	Chr.	Structure	From	To	Strand	N	Length	ID	SCI	*z*	*P*
31534	chr14	61162	56,593,996	56,594,068	+	4	72	81.48	0.77	-2.31	0.973024
		61163	56,593,996	56,594,068	-	4	72	81.48	0.88	-1.73	0.877635
89808	chr6	162046	91,496,514	91,496,632	-	5	120	84.68	0.85	-3.54	0.999456
		162040	91,496,514	91,496,632	+	5	120	84.68	0.7	-2.96	0.980763
26889	chr13	53069	30,551,559	30,551,737	-	4	183	80.79	0.8	-3.87	0.999706
		53068	30,551,559	30,551,737	+	4	183	80.79	0.69	-2.36	0.971574
27543	chr13	54180	53,677,616	53,677,733	-	4	120	87.22	0.67	-3.2	0.92269
		54185	53,677,655	53,677,773	-	4	120	87.82	0.83	-2.46	0.973109
		54191	53,677,693	53,677,812	-	4	120	88.22	0.75	-2	0.818394
61043	chr20	115082	496,890	497,004	-	4	120	80.31	0.71	-2.62	0.992929
		115090	496,929	497,044	-	4	120	83.78	0.72	-1.59	0.608496
109240	chrX	196268	39,527,756	39,527,861	-	4	107	80.66	0.76	-2.22	0.967603
112633	chrX	202757	128,565,991	128,566,111	+	4	120	82.06	0.84	-1.84	0.87596
		202764	128,566,031	128,566,151	+	4	120	82.87	0.88	-3.06	0.998176
		202770	128,566,031	128,566,151	-	4	120	82.87	0.83	-2.38	0.980258
27659	chr13	54415	57,553,726	57,553,846	+	4	120	81.09	0.73	-2.69	0.993413
		54421	57,553,766	57,553,886	-	4	120	82.63	0.63	-2.2	0.842209
		54419	57,553,766	57,553,886	+	4	120	82.63	0.69	-2.89	0.989836
		54416	57,553,726	57,553,846	-	4	120	81.09	0.63	-2.51	0.962321
89617	chr6	161687	85,997,067	85,997,187	+	4	120	81.36	0.8	-2.34	0.9805
		161692	85,997,147	85,997,266	+	4	120	79.89	0.74	-2.51	0.991869
		161690	85,997,107	85,997,226	-	4	120	78.77	0.79	-1.84	0.908923
95813	chr7	172643	101,021,517	101,021,626	+	5	120	76.12	0.54	-2.99	0.999155
		172639	101,021,478	101,021,586	-	5	120	77.81	0.49	-2.27	0.934702
		172638	101,021,478	101,021,586	+	5	120	77.81	0.6	-2.79	0.996624
27222	chr13	53606	43,744,032	43,744,208	-	4	183	79.12	0.56	-2.81	0.982372
90290	chr6	162988	99,701,979	99,702,098	+	5	120	86.35	0.8	-3.84	0.999157
108394	chrX	194774	20,510,387	20,510,503	+	4	120	81.98	0.76	-2.35	0.972823
27091	chr13	53384	39,976,457	39,976,573	+	4	120	82.19	0.83	-2.44	0.985861

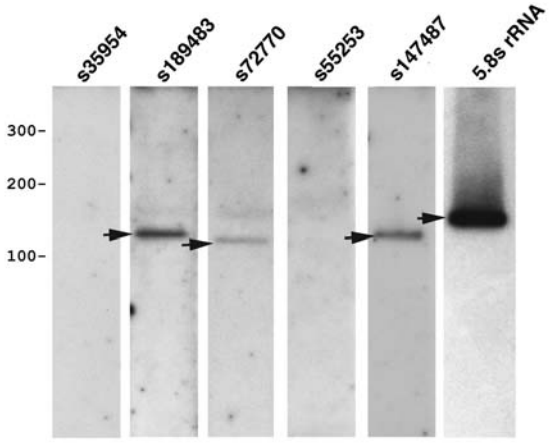
Cluster	Chr.	Structure	From	To	Strand	N	Length	ID	SCI	*z*	*P*
		53386	39,976,497	39,976,587	+	4	94	80.95	0.73	-1.47	0.5
89504	chr6	161517	82,802,692	82,802,812	-	4	120	86.53	0.83	-2	0.909879
109345	chrX	196509	40,146,862	40,146,980	-	4	120	82.54	0.73	-3.31	0.997741
		196511	40,146,901	40,146,982	-	4	82	86.3	0.86	-3.67	0.997816
		196496	40,146,782	40,146,901	+	4	120	78.49	0.58	-2.1	0.82181
		196500	40,146,822	40,146,941	-	4	120	76.68	0.51	-2.77	0.979585
63128	chr20	119069	48,283,591	48,283,708	-	4	118	84.38	0.84	-3.36	0.997966
91875	chr6	165653	149,564,187	149,564,381	+	4	194	83.23	0.71	-3.27	0.995457
109735	chrX	197192	50,141,861	50,142,031	+	4	172	87.9	0.82	-2.71	0.983158
94490	chr7	170234	61,948,610	61,948,729	-	4	120	87.68	0.78	-2.06	0.881409
		170235	61,948,649	61,948,741	+	4	93	86.74	0.69	-2.69	0.924065
63684	chr21	169	15,807,334	15,807,454	-	4	120	89.54	0.93	-2.09	0.974365
61640	chr20	116254	13,295,621	13,295,741	-	4	120	87.36	0.78	-2.17	0.90653
49886	chr19	94864	37,091,436	37,091,556	-	5	120	87.17	0.88	-1.72	0.900379

Cluster, Structure … See Methods how overlapping windows were scored and combined into clusters
Chr, From, To … Chromosome and coordinates based on the hg17 human assembly
Strand … Forward strand (+) or reverse complement (-)
N … Number of sequences
Length … Number of columns
ID … Mean pairwise identity
SCI … Structure conservation index
z … Mean *z*-score of the single sequences
P … RNA class probability calculated by RNAz

The table is linked to our website where all predictions can be viewed and downloaded:

http://www.tbi.univie.ac.at/papers/SUPPLEMENTS/ncRNA

Supplementary Figure 1: Northern Blot analysis of five H/ACA snoRNA candidates



Hybridisation signals are indicated by arrows. Sizes of RNAs are estimated by comparison with an internal RNA marker indicated on the left. As a control, expression of 5.8S rRNA is analyzed in addition. Genomic location and probe sequences of the candidates:

- s35954: chr.11:43,883,889, CTACATTGTGTTCCTGTTTGCAAA
- s189483: chr.9:92,134,300, CCATGCATAAATCACAGGATTGCTAT
- s72770: chr.16:2,786,411, CACCAGTGAATCAAGACCAGTAGATTC
- s55253: chr.13:70,139,481, CATTTGCAAATAACAATACATCAGAAC
- s147487: chr5:82,395,781, AAGGTTGGCAGCCCTGAGAACTCC

Candidates s189483 and s72770 are shown in Fig. 3 c and d, respectively. Candidate s189483 is located in an intron of Isoleucine-tRNA synthetase (D28473). Candidate s72770 is not associated with any known protein coding gene, but near an annotated pseudogene of the ribosomoal protein 27a. The snoRNA and the pseudogene are flanked by LINE elements. Candidate s147487 is located in an intron of a hypothetical protein (MGC23909) from a cDNA sequencing project.

Experimental details: HeLa cells were cultivated in DMEM, containing 10% FCS, 2 mM L-Glutamin, 100 U/ml Penicillin, 100 U/ml Streptomycin at 37^{o}C and 5% CO2. 107 cells were pelleted and resuspended in 1 ml of Trizol (Invitrogen), and total RNA was purified following the manufacturer's instructions. Subsequently, total RNA (20 g) was separated on an 8% denaturing polyacrylamide gel (7M urea, 1 x TBE buffer) and transferred onto a nylon membrane (Quiabrane Nylon Plus, Quiagen) using the Biorad semi-dry blotting apparatus (Trans-blot SD, Biorad). After immobilizing of RNAs using the STRATAGENE crosslinker, we pre-hybridized the nylon membrane for 1 h in 1 M sodium phosphate buffer (pH 6.2), 7% SDS. Oligonucleotides complementary to potentially novel RNA species were end-labelled with $\gamma[^{32}P]$ATP and T4 polynucleotide kinase; hybridization was carried out at 58^{o}C in 1 M sodium phosphate buffer (pH 6.2), 7% SDS for 12 hrs. Blots were washed twice at room temperature in 2 x SSC buffer (20 mM sodium phosphate, pH 7.4; 0.3 M NaCl; 2 mM EDTA), 0.1% SDS for 15 min and subsequently at 58^{o}C in 0.1 x SSC, 0.5% SDS for 1 min. Membranes were exposed to Kodak MS-1 film for 1 to 3 days.