SUPPLEMENTARY INFORMATION

Abbreviations:

snoRNA: small nucleolar RNA

snRNA: small nuclear RNA

snRNP: small nuclear ribonucleoprotein

rRNA: ribosomal RNA

tRNA: transfer RNA

miRNA: micro RNA

MEGF8: multiple epidermal growth factor-like-domains 8

SMARCA4: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4

MAD1L1: MAD1 mitotic arrest deficient-like 1

GRIK5: glutamate receptor, ionotropic, kainate 5

HIPPI: huntington-interacting protein protein interactor

FGFR3: fibroblast growth factor receptor 3

mRNA: messenger RNA

pre-mRNA: precursor messenger RNA

GAPDH: glyceraldehyde phosphate dehydrogenase

HBII-180: human brain snoRNA II - 180

MBII: mouse brain snoRNA II

PCR: polymerase chain reaction

CMV: cytomegalovirus

HDAC: histone deacetylase 1

HSP70: heat shock protein 70

CBP: CREB binding protein isoform a

HPRT: hypoxanthine phosphoribosyltransferase 1

PSP-1: paraspeckle protein 1

LEGENDS TO SUPPLEMENTARY FIGURES

Supplementary Figure 1. Characterization of HBII-180C and chimera snoRNAs. (a) Fluorescence In Situ Hybridisation showing nucleolar localization of wild type HBII-180C and chimera snoRNAs (Cy3, Red). Fibrillarin localisation pattern is also shown on the right panel (FITC, green). DNA is stained by DAPI (Blue). Scale Bar is 10 μm. (b) Characterization of HBII-180C and chimera snoRNAs. RNA blot showing sub cellular distribution of HBII-180C and Chimera snoRNAs. The same amount of HeLa cell total RNA (lanes 1), cytoplasmic RNA (lanes 2), nucleoplasmic RNA (lanes 3) and nucleolar RNA (lanes 4) were compared with either transiently transfected wild type or chimera snoRNA expression mini-genes by northern RNA blotting analysis using specific probes. Fractionation quality was tested using the tRNA specific probe (lower panel).

Supplementary Figure 2. Analysis for HBII-180 snoRNAs M box. (a) Potential mRNA targets of HBII-180 snoRNAs. Two predicted complementary target gene sequences with highest complementarity are shown for each snoRNA (MEGF8: multiple epidermal growth factor-like-domains 8, SMARCA4: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4, MAD1L1: MAD1 mitotic arrest deficient-like 1, GRIK5: glutamate receptor, ionotropic, kainate 5, HIPPI: huntington-interacting protein protein interactor, FGFR3: fibroblast growth factor receptor 3). (b) Another example of gene knock-down using M box-modified snoRNAs. This is the same experiment as shown in Figure 3b, except the stable cell line was HeLa$^{GFP\text{-}SMN}$. pmCherry-HBII-180C is another negative control vector, which has the wild type HBII-180C snoRNA sequence downstream of mCherry cDNA. Scale bar is 10 μm.The arrows indicate transfected cells and arrowheads show untransfected

cells. Note clear reduction in GFP expression specifically in the cells transfected with the expression plasmids encoding either HBII-180C chimera1 or 2.

Supplementary Figure 3. Western blot analysis for chimera snoRNAs knockdown. Detection of protein levels for G/YFP following transfection of HeLa^{GFP} and HeLa$^{YFP\text{-}Fibrillarin}$ stable cell lines using either wild type HBII-180C with mCherry expression plasmid pmCherry-HBII-180C (control), expression plasmid pHBII-180C chimera1 (Chimera-1) or expression plasmid pHBII-180C chimera2 (Chimera-2). An equivalent amount of HeLa extract was loaded for each lane and the proteins separated by SDS PAGE, electroblotted and probed both with a monoclonal anti-GFP antibody and with anti-tubulin as a loading control. Transfection efficiency was also calculated by counting the number of mCherry positive cells in 100 randomly selected cells (TF efficiency). The graphs show average signal intensity and standard deviation for three independent experiments. G/YFP signal ratio was normalized to the tubulin signal.

Supplementary Figure 4. Analysis of off target effects. Quantitative RT-PCR was performed for 10 different genes to check possible off target effects of Chimera snoMEN vectors. The graph shows average signal intensity and standard deviation for three independent experiments. Each signal was normalised by U3 snoRNA intensity. Left panel shows semi-quantitative PCR results. Each primer set specificity was also made sure by semi-quantitative PCR as a single band. Abbreviation is available in Supplementary information.

Supplementary Figure 5. Analysis of HBII-180C and chimera snoRNA mutations. Mutant plasmids (CM2m2-1, m2-2 & m3) were transiently transfected into a stable cell line expressing

the GFP-SMN fusion protein (HeLa$^{GFP-SMN}$). Scale bar is 15 μm.Arrow: transfected, Arrowhead: untransfected. Note clear knock-down effect of the mutant plasmids that destroyed the 28S rRNA complementary sequence (CMm2-1 & m2-2).

Supplementary Figure 6. Analysis of HBII-180C and chimera snoRNA mutations. Mutant plasmids that have either insertions in the M box to increase complementarity to GFP (CM2In-1 to -3), or deletion in the M box (CM2Del-1 to –3), were transiently transfected into the stable cell line expressing GFP-SMN fusion protein (HeLa$^{GFP-SMN}$). Arrow: transfected, Scale bar is 15 μm. Arrowhead: untransfected.

Supplementary Figure 7. Analysis of HBII-180C and chimera snoRNA mutations. Mutant plasmids that have point mutations in M box (CM2X-1 to –6) were transiently transfected into the stable cell line expressing GFP-SMN fusion protein (HeLa$^{GFP-SMN}$). Scale bar is 15 μm. Arrow: transfected, Arrowhead: untransfected.

Supplementary Figure 8. Dose response analysis comparing GFP suppression level with expression level of snoMEN vectors. Graph shows the fluorescence signal ratio between mCherry and GFP-SMN (n>10). Chimera2 and CM2-In3 showed a negative correlation between mCherry signal and GFP signal, while, CM2m7 (boxD mutant) and HBII-180C (Control) showed similar GFP levels independent of the mCherry level.

Supplementary Figure 9. Another example of gene knock-down using triple modified snoRNAs. This shows the same experiment as in Fig. 7b, except that the stable cell line was HeLa ^{GFP}. The arrows indicate transfected cells and arrowheads show untransfected cells. Scale

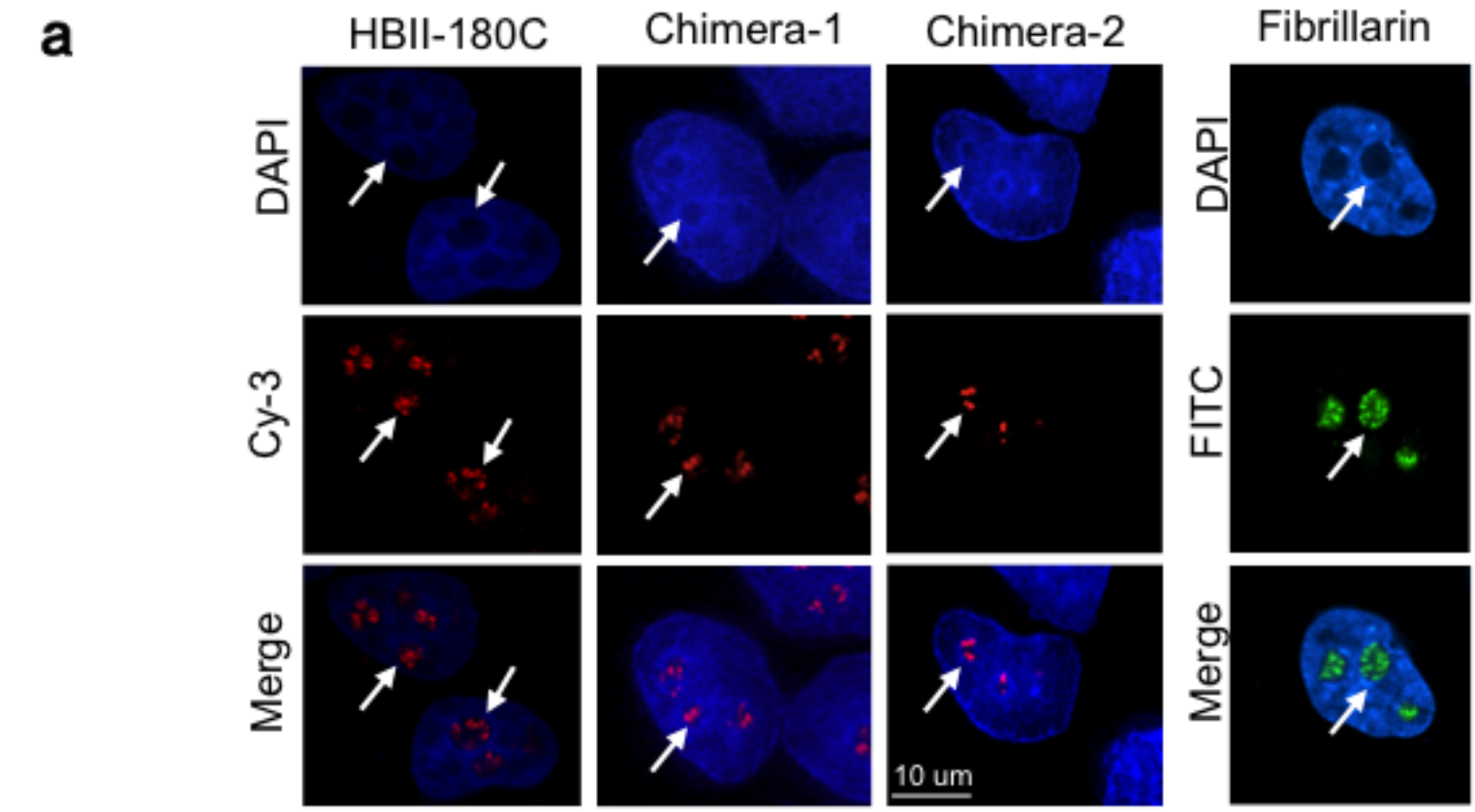
bar is 10 μm. Note clear reduction in GFP expression specifically in the cells transfected with the expression plasmid encoding HBII-180C triple chimera.

Supplementary Figure 10. Western blot analysis for triplet snoMEN knockdown. (a) Knock-down efficiency of each chimera plasmid targeted to G/YFP was tested by western blotting (lane 1-5). Detection of protein levels for YFP-Fibrillarin following transfection of $HeLa^{YFP\text{-}Fibrillarin}$ stable cell lines using either mCherry and wild type HBII-180C expression plasmid pmCherry-HBII-180C (control: lane1), expression plasmid chimera1 (lane2), chimera2 (lane3), chimera3 (lane4), HBII-180C triple chimera (lane5). An equivalent amount of HeLa extract was loaded for each lane and the proteins separated by SDS PAGE, electroblotted onto membrane and probed both with a monoclonal anti-GFP antibody and with anti-B23 as a loading control. Graph shows YFP-Fibrillarin signal intensity ratio normalised using the B23 signal. Transfection efficiency was determined by counting the number of mCherry positive cells in 100 randomly selected cells (TF efficiency). (b) Detection of protein levels for endogenous Coilin and LaminA/C following transfection of HeLa cells using either mCherry triple chimera (Control: lane1) and pCoilin/Lamin-snoMEN (lane2). An equivalent amount of HeLa extract was loaded for each lane and the proteins separated by SDS PAGE, electroblotted and probed both with a monoclonal anti-Coilin and anti-Lamin antibody and with anti-tubulin as a loading control. Graph shows Coilin and Lamin signal intensity normalised to the tubulin signal.

Supplementary Figure 11. SiRNA oligoribonucleotides targeted to pre-mRNA regions. (a) This is the same experiment as in Fig. 13a, except the target gene is Coilin. The same pre-mRNA sequences targetd by the snoMEN vector were targeted by siRNA oligoribonucleotides. Scrambled siRNA (Control siRNA) and Coilin siRNA (Dharmacon) were transfected as a negative and a positive control, respectively. Coilin Mbox siRNA-1 and Coilin Mbox siRNA-2 have the same target sequence with Coilin snoMEN set1 and Coilin snoMEN set2, respectively

(Fig. 14). Scale bar is 5 μm. Arrow: cells not showing knock-down, Arrowhead: cells showing knock-down. (b) Western blot analysis for siRNA experiments. Detection of protein levels for endogenous Coilin following transfection of HeLa cells using either Scrambled siRNA (Control: lane1), Coilin siRNA (lane2), Coilin M box siRNA-1 (lane3) and Coilin M box siRNA-2 (lane4). An equivalent amount of HeLa extract was loaded for each lane and the proteins separated by SDS PAGE, electroblotted onto membrane and probed both with a monoclonal anti-Coilin antibody and with anti-tubulin as a loading control. Graph shows Coilin signal intensity normalised to the tubulin signal.

Fig. 1 (Sup)


a
HBII-180C
Chimera-1
Chimera-2
Fibrillarin
DAPI
Cy-3
Merge
DAPI
FITC
Merge
10 um

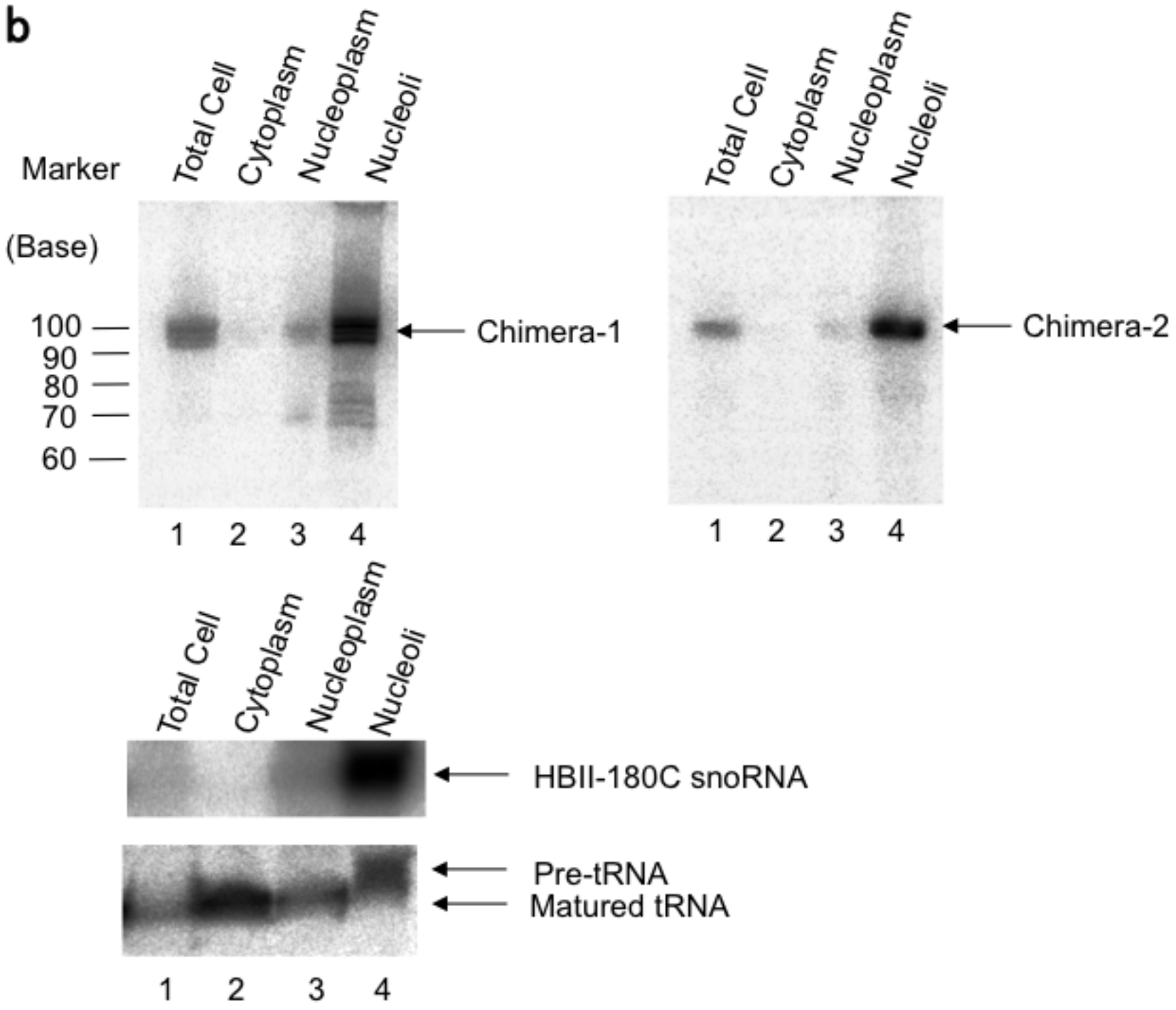

b
Marker
Total Cell
Cytoplasm
Nucleoplasm
Nucleoli
(Base)
100
90
80
70
60
Chimera-1
1 2 3 4
Total Cell
Cytoplasm
Nucleoplasm
Nucleoli
Chimera-2
1 2 3 4
Total Cell
Cytoplasm
Nucleoplasm
Nucleoli
HBII-180C snoRNA
Pre-tRNA
Matured tRNA
1 2 3 4

Fig. 2(Sup)

a

HBII-180A

MEGF8

```
5' TGCCACTGAGGACACGGTG 3'
   ||||||||||||||||||  ||
3' ACGGTGACTCCTGTGCAAC 5'
```

SMARCA4

```
5' TGCCACTGAGGACACGGTGC 3'
   ||||||||||||| ||||  ||
3' ACGGTGACTCCTCTGCCCCG 5'
```

HBII-180B

MAD1L1

```
5' TGCCCCTGAGGAC-ACGGTG 3'
   |||||||||||||| ||||||
3' ACGGGGACTCCTGGTGCCAC 5'
```

GRIK5

```
5' GCCCCTGAGGACACGGTG 3'
   ||||||||||||||||||
3' CGGGGACTCCTGTGTCAC 5'
```

HBII-180C

HIPPI

```
5' ACCCCTGAGGACACAGTGCA 3'
   ||||||||||||||||||||
3' TGGGGACTCCTGTGTCACGT 5'
```

FGFR-3

```
5' CACCCCTGAGGACACAGTGCA 3'
   |||||||||||||| |||||  |
3' GTGGGGACTCCTGGGTCACCT 5'
```

b

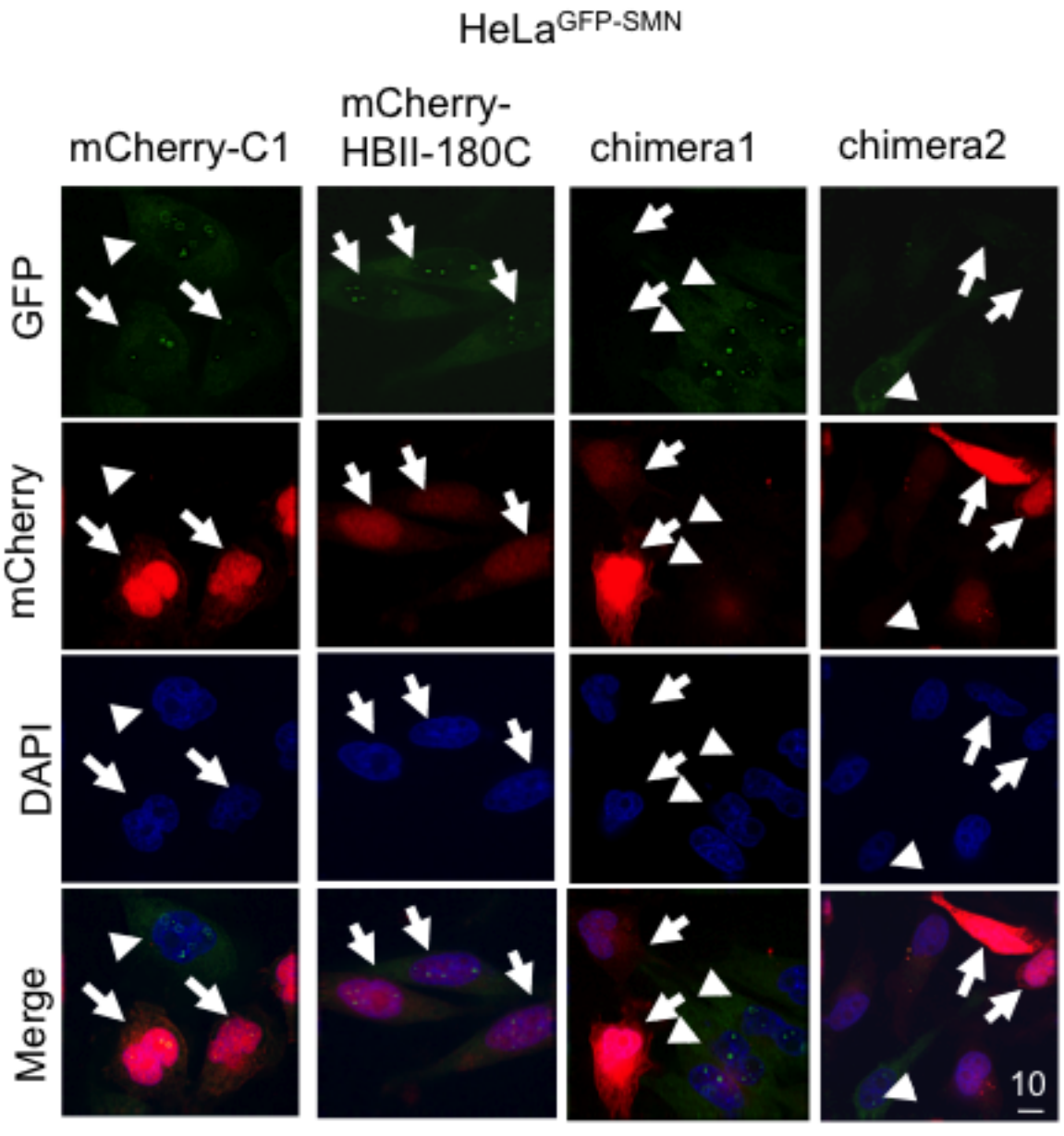


Fig. 3 (Sup)

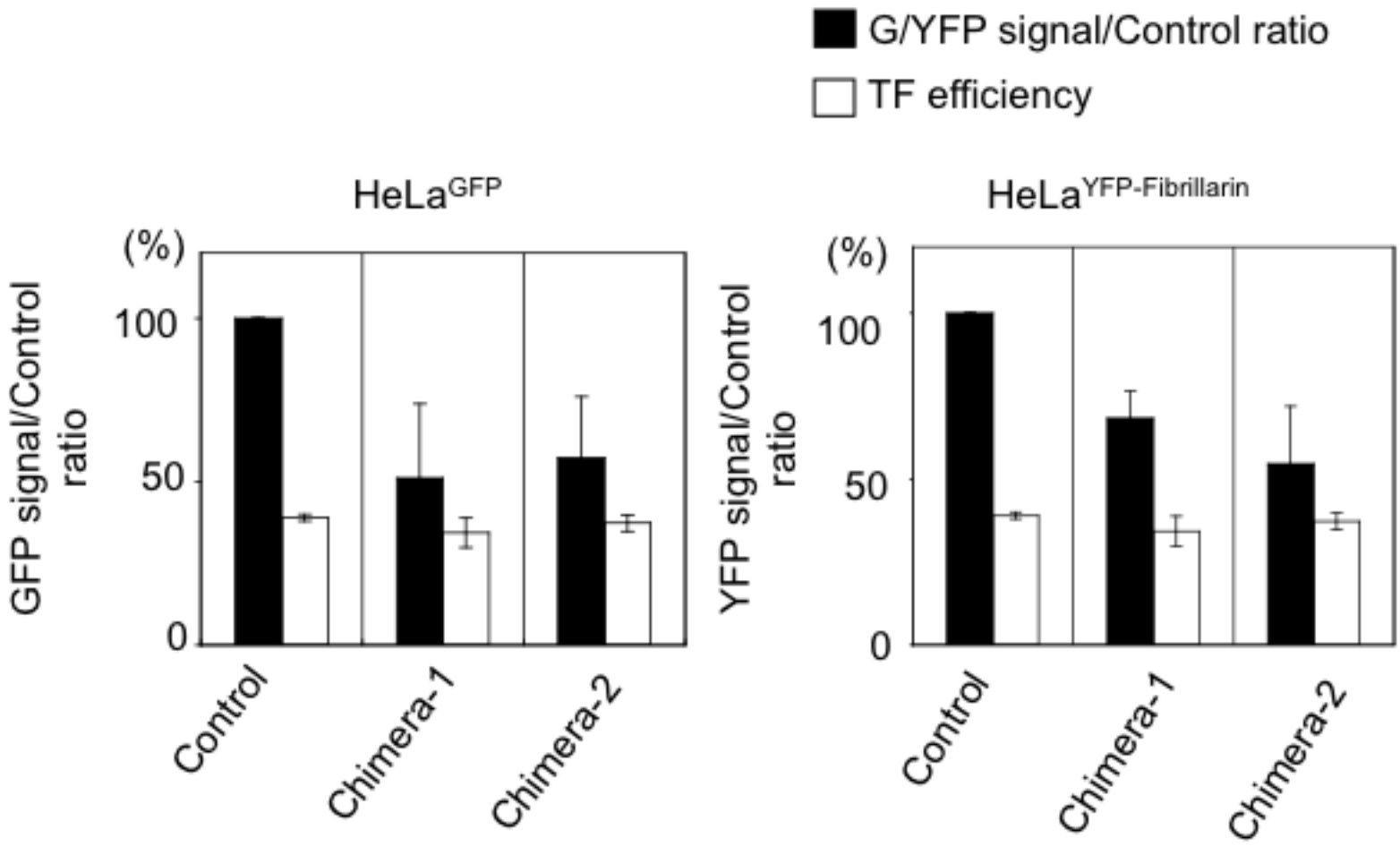

G/YFP signal/Control ratio
TF efficiency
HeLa^GFP
GFP signal/Control ratio
(%)
100
50
0
Control
Chimera-1
Chimera-2
HeLa^YFP-Fibrillarin
YFP signal/Control ratio
(%)
100
50
0
Control
Chimera-1
Chimera-2

Fig. 4 (sup)

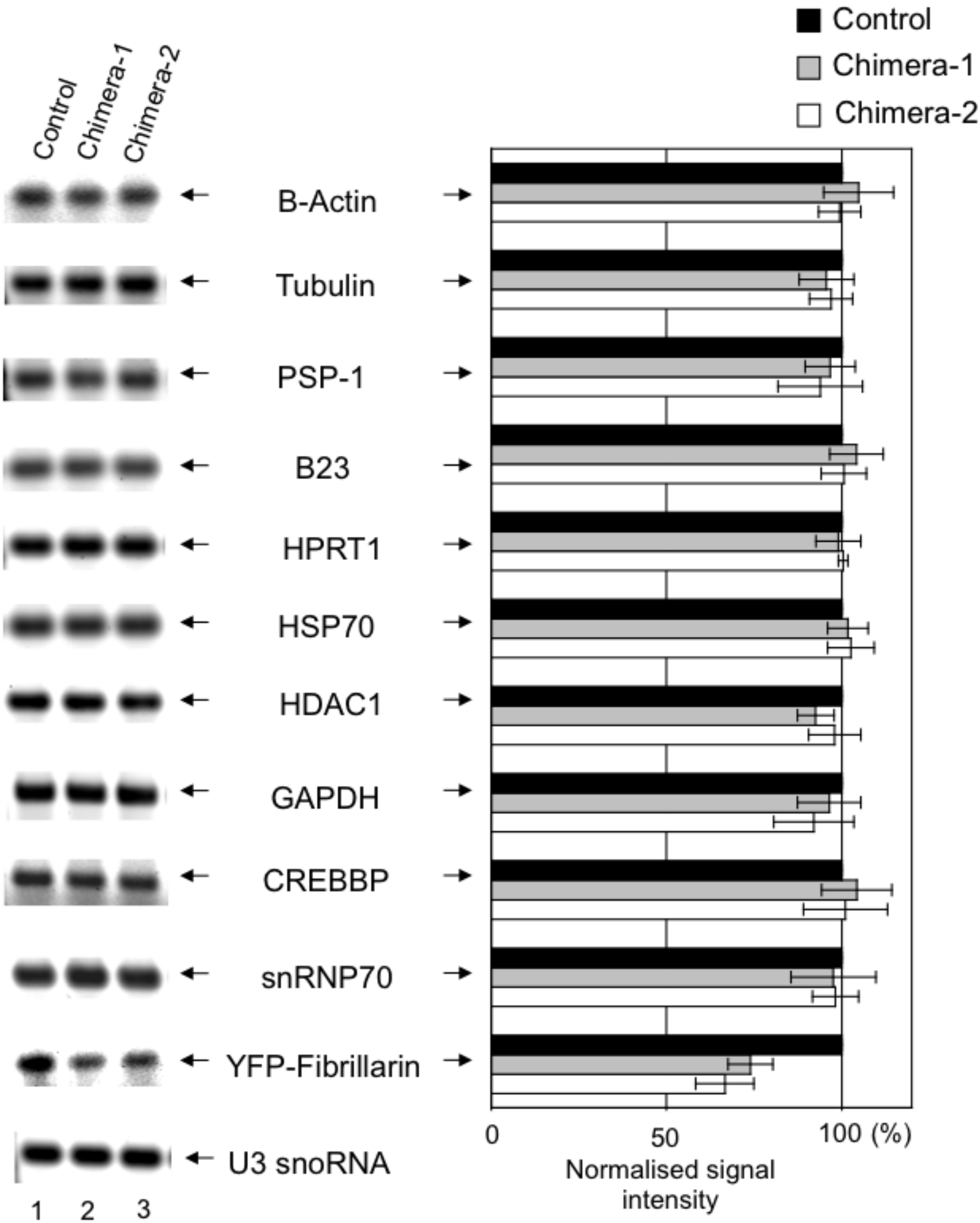
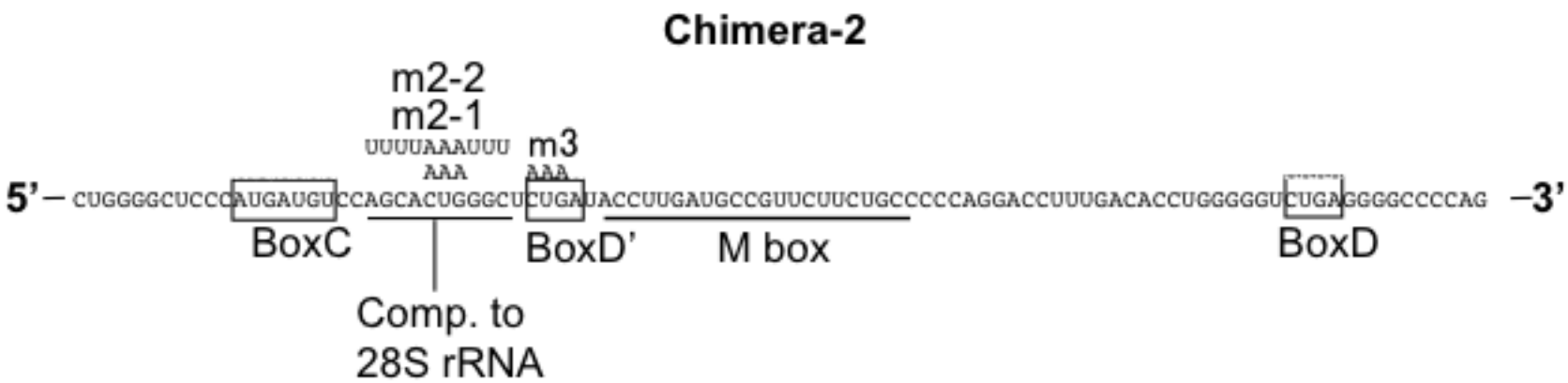

Control
Chimera-1
Chimera-2
B-Actin
Tubulin
PSP-1
B23
HPRT1
HSP70
HDAC1
GAPDH
CREBBP
snRNP70
YFP-Fibrillarin
U3 snoRNA
1
2
3
Control
Chimera-1
Chimera-2
0
50
100 (%)
Normalised signal intensity

Fig. 5 (Sup)


Chimera-2
m2-2
m2-1
UUUUAAAUUU
AAA
m3
AAA
5'– CUGGGGCUCCCAUGAUGUCCAGCACUGGGCUCUGAUACCUUGAUGCCGUUCUUCUGCCCCCAGGACCUUUGACACCUGGGGGUCUGAGGGGCCCCAG –3'
BoxC
BoxD'
M box
BoxD
Comp. to
28S rRNA

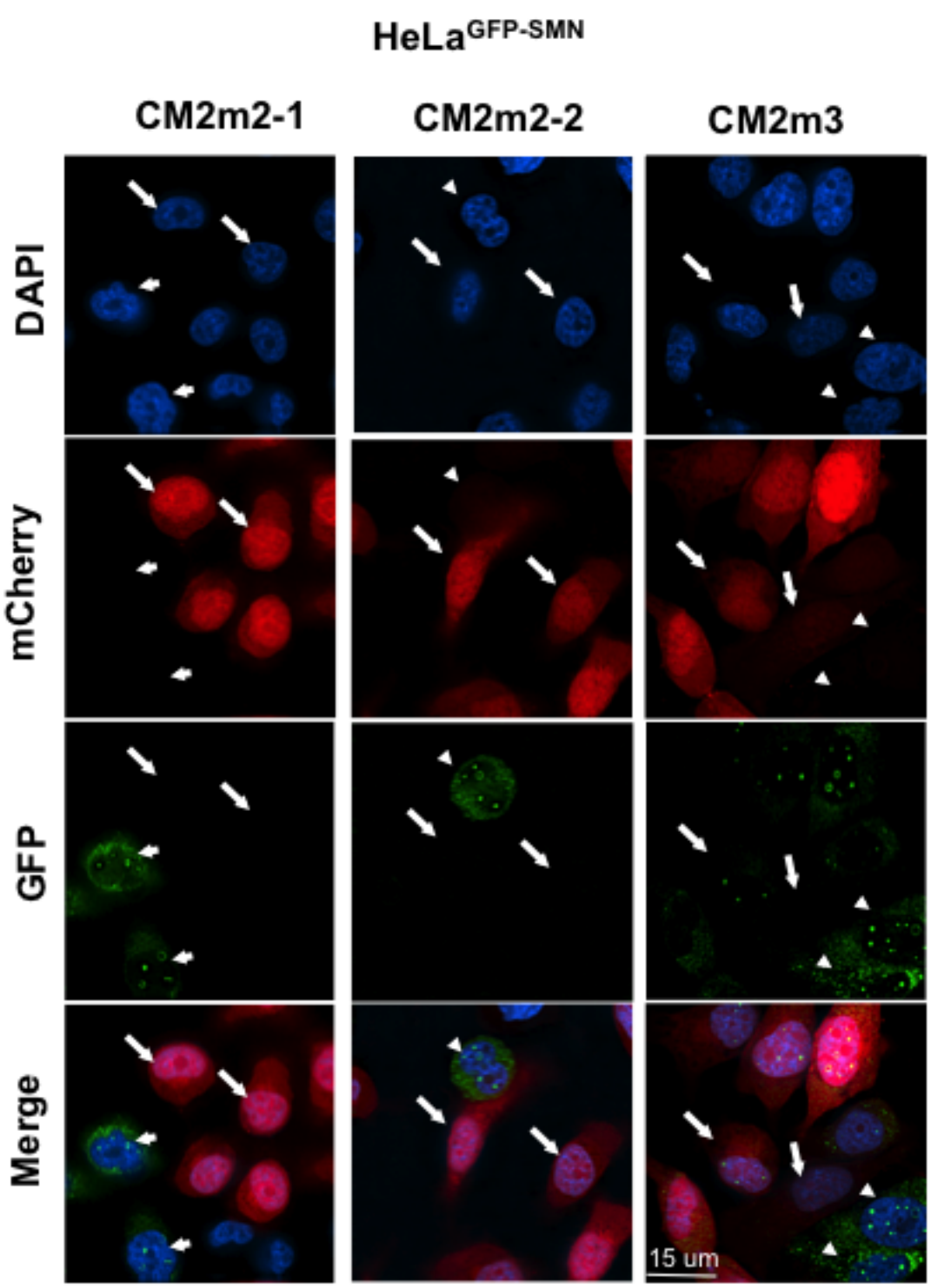

HeLaGFP-SMN
CM2m2-1
CM2m2-2
CM2m3
DAPI
mCherry
GFP
Merge
15 um

Fig. 6 (sup)


Chimera-2
In−3
In-2
In-1
GUUCACCUUGAUGCCGUUCUUCUGUUGT
UCACCUUGAUGCCGUUCUUCUGUU
UCACCUUGAUGCCGUUCUUCUG
5'−
CUGGGGCUCCCAUGAUGUCCAGCACUGGGCUCUGAUACCUUGAUGCCGUUCUUCUGCCCCCAGGACCUUUGACACCUGGGGGUCUGAGGGGCCCCAG
−3'
BoxC
BoxD'
M box
BoxD
Comp. to
28S rRNA
--CUUGAUGCCGUUCUUCUG
--CUUGAUGCCGUUCUUCU-
----UGAUGCCGUUCUU---
Del-1
Del-2
Del-3

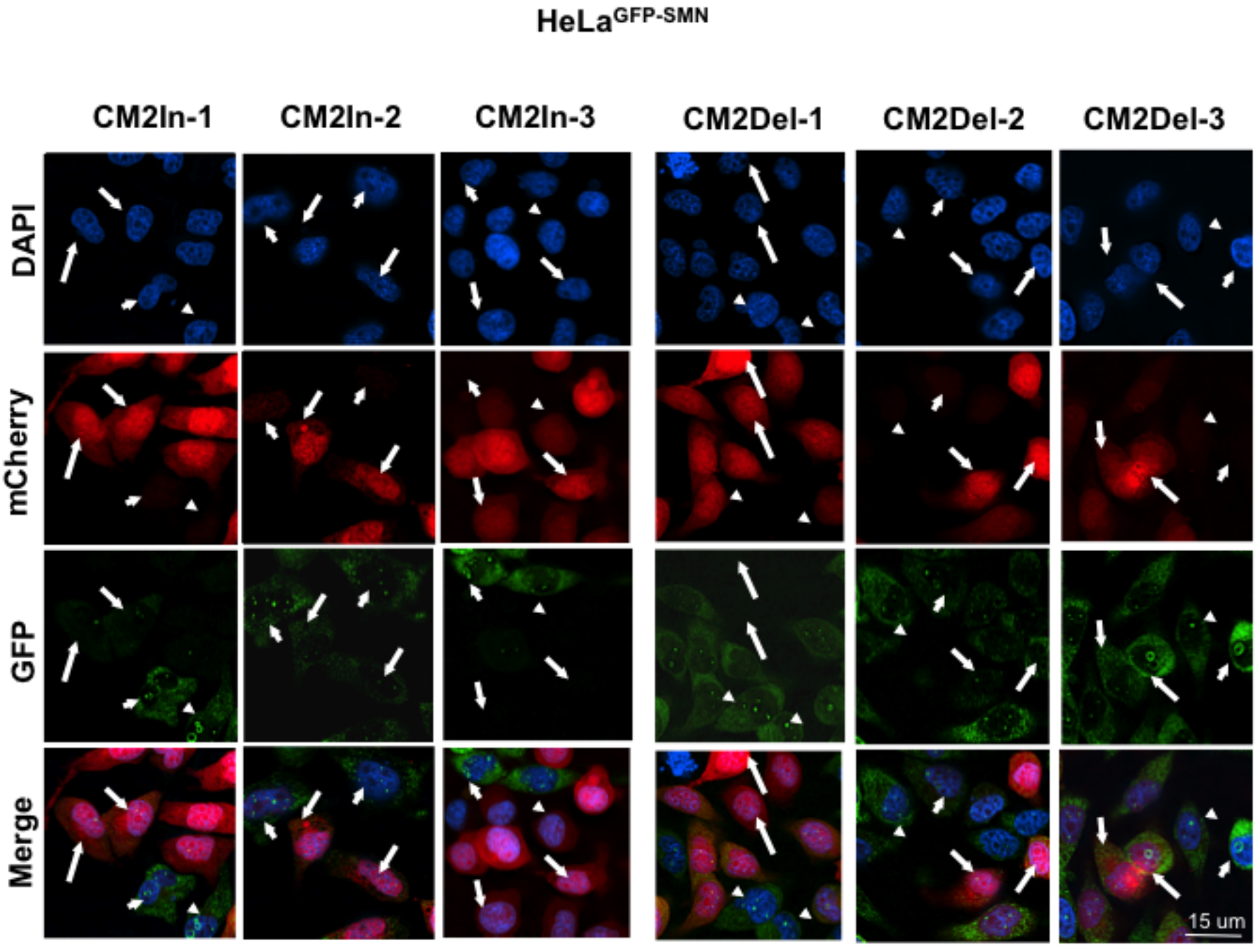
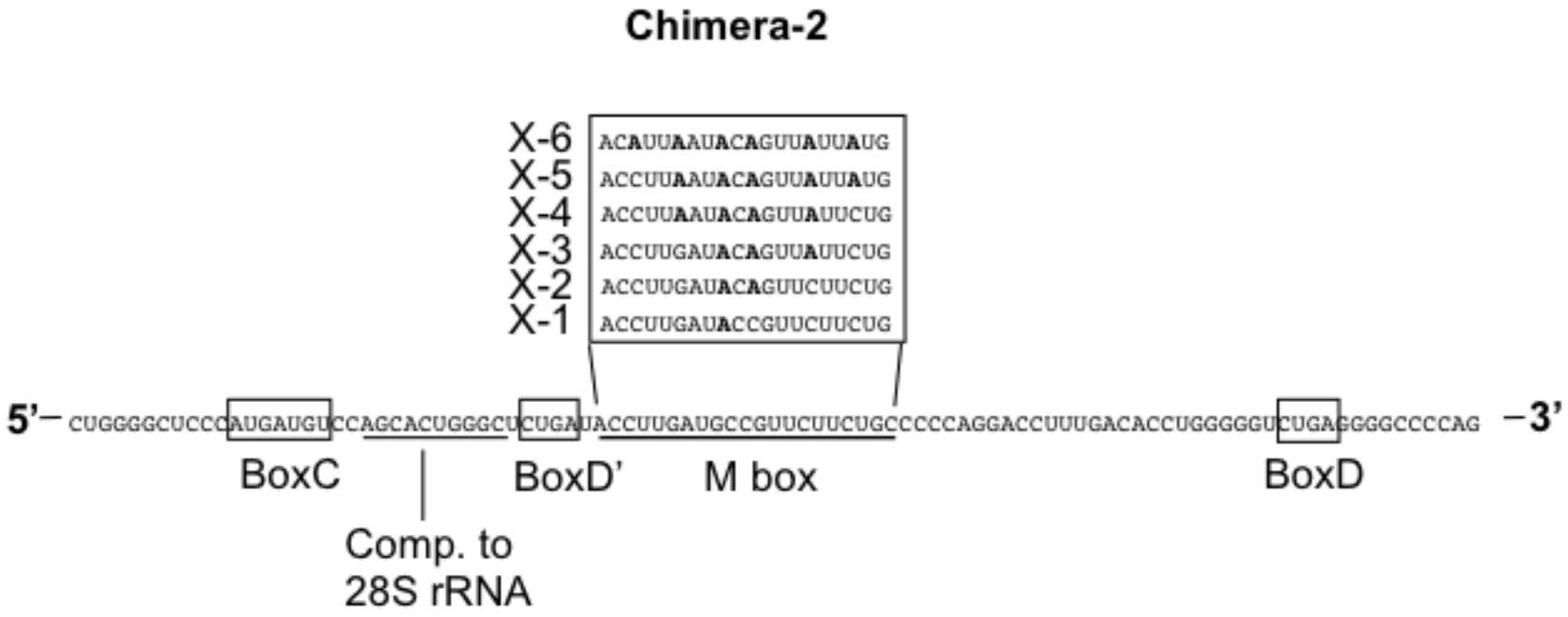

HeLaGFP-SMN
CM2In-1
CM2In-2
CM2In-3
CM2Del-1
CM2Del-2
CM2Del-3
DAPI
mCherry
GFP
Merge
15 um

Fig. 7 (sup)


Chimera-2
X-6 ACAUUAAUACAGUUAUUAUG
X-5 ACCUUAAUACAGUUAUUAUG
X-4 ACCUUAAUACAGUUAUUCUG
X-3 ACCUUGAUACAGUUAUUCUG
X-2 ACCUUGAUACAGUUCUUCUG
X-1 ACCUUGAUACCGUUCUUCUG
5'– CUGGGGCUCCCAUGAUGUCCAGCACUGGGCUCUGAUACCUUGAUGCCGUUCUUCUGCCCCCAGGACCUUUGACACCUGGGGGUCUGAGGGGCCCCAG –3'
BoxC
BoxD'
M box
BoxD
Comp. to
28S rRNA

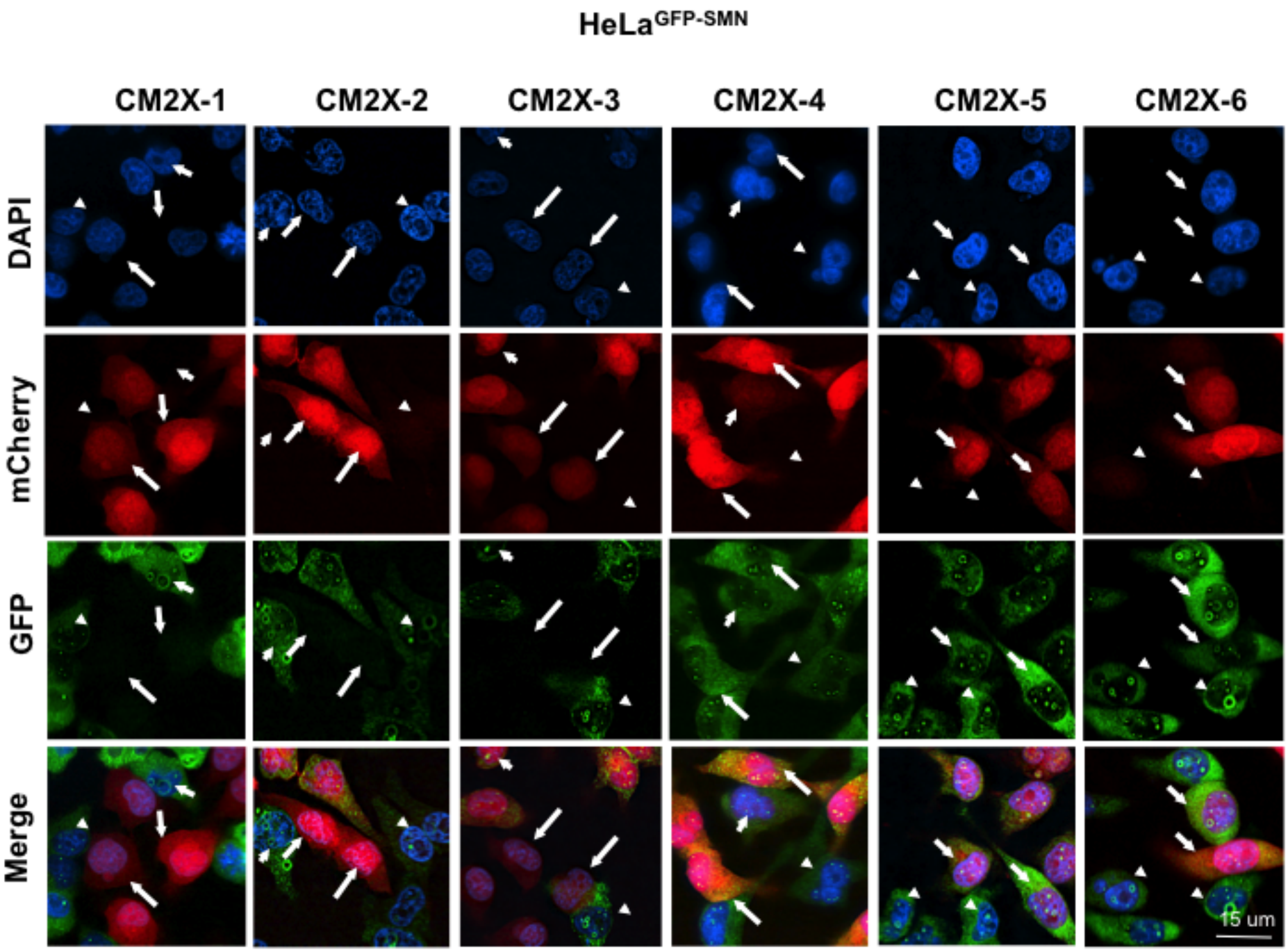
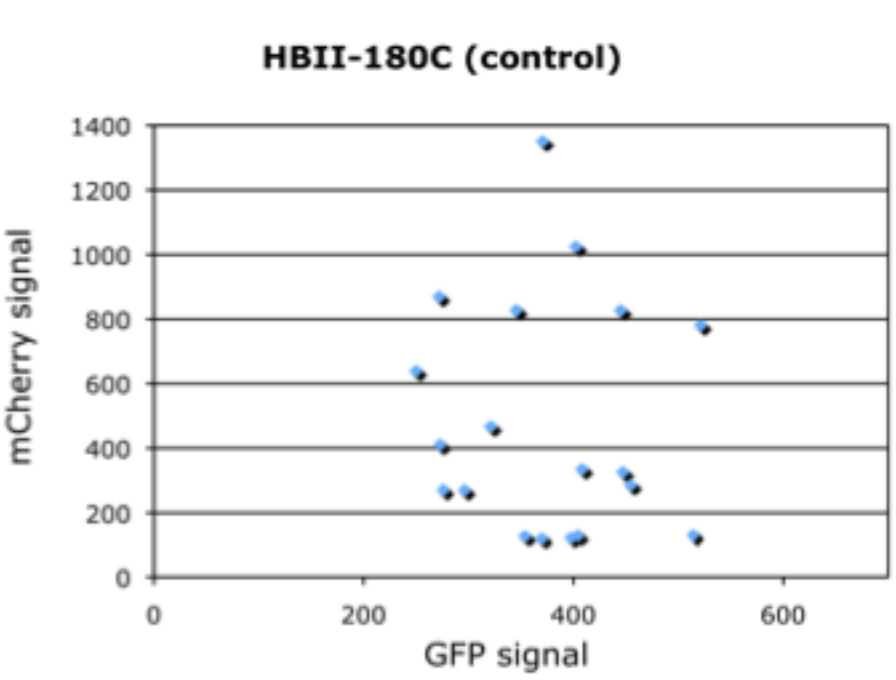
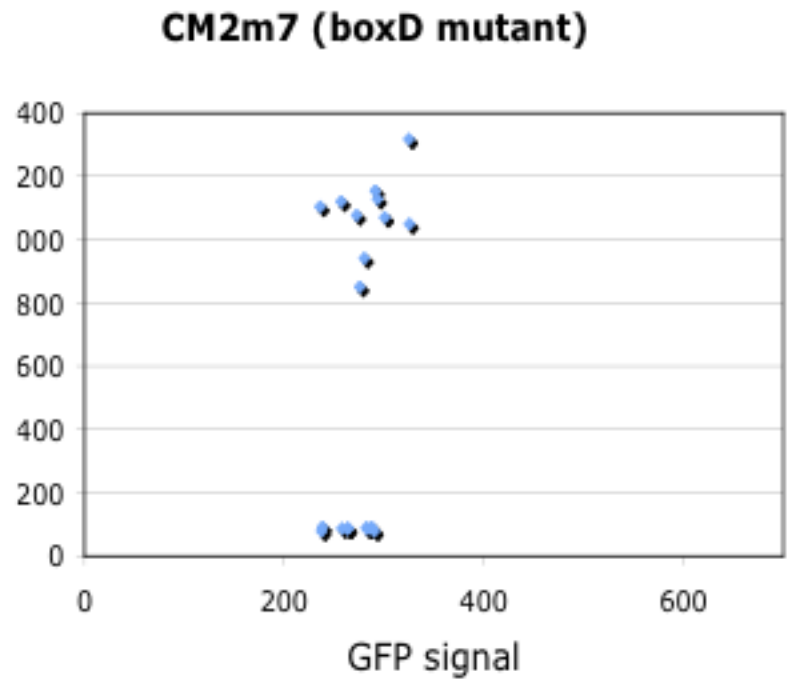
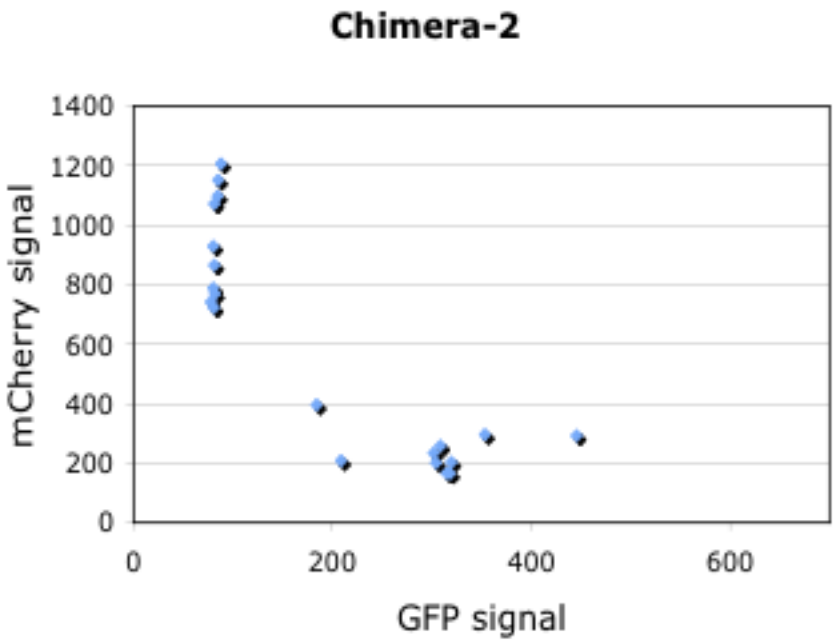

HeLa GFP-SMN
CM2X-1
CM2X-2
CM2X-3
CM2X-4
CM2X-5
CM2X-6
DAPI
mCherry
GFP
Merge
15 um

Fig. 8 (sup)


HBII-180C (control)
mCherry signal
GFP signal
0
200
400
600
800
1000
1200
1400


CM2m7 (boxD mutant)
GFP signal
0
200
400
600


Chimera-2
mCherry signal
GFP signal
0
200
400
600
800
1000
1200
1400

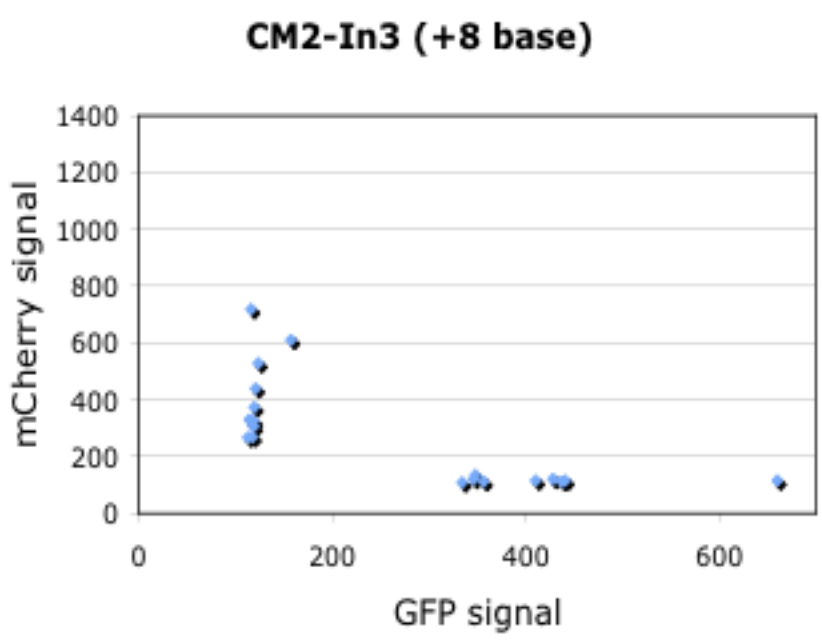

CM2-In3 (+8 base)
mCherry signal
GFP signal
0
200
400
600
800
1000
1200
1400

Fig. 9 (sup)

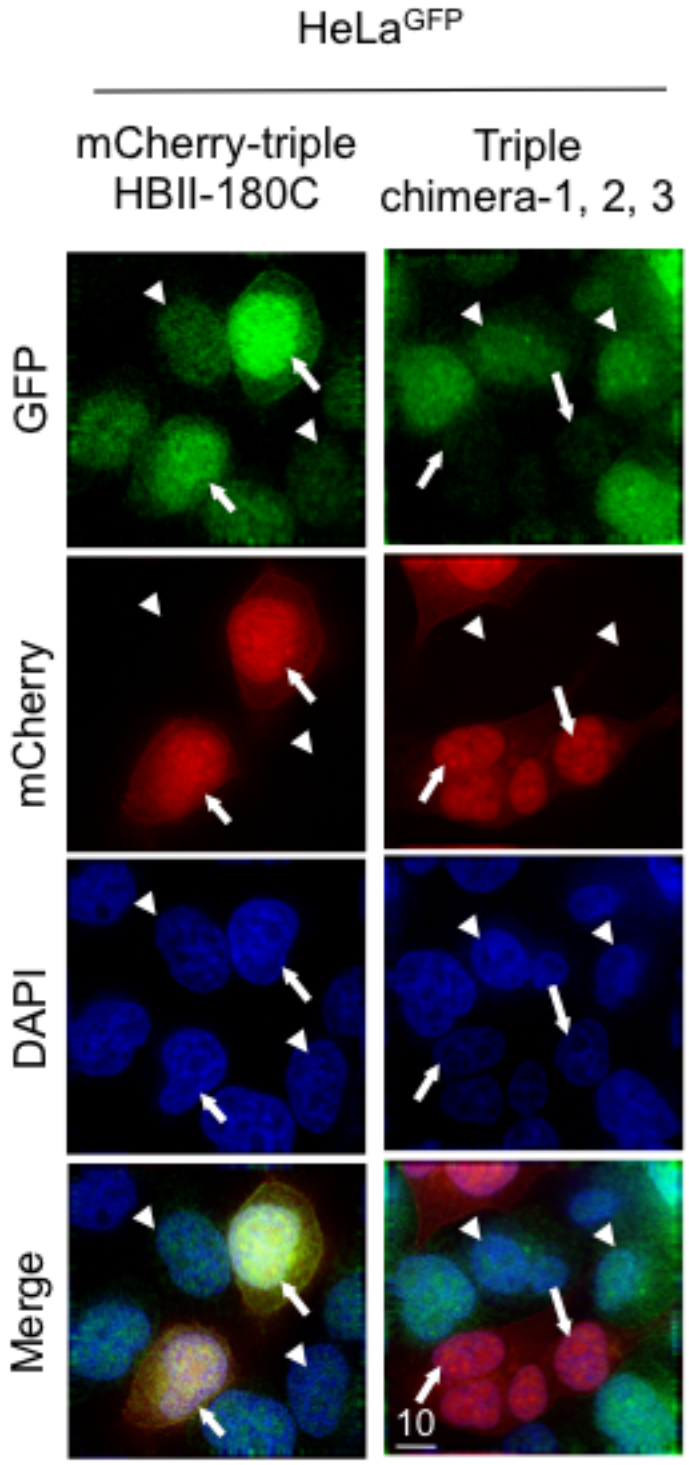
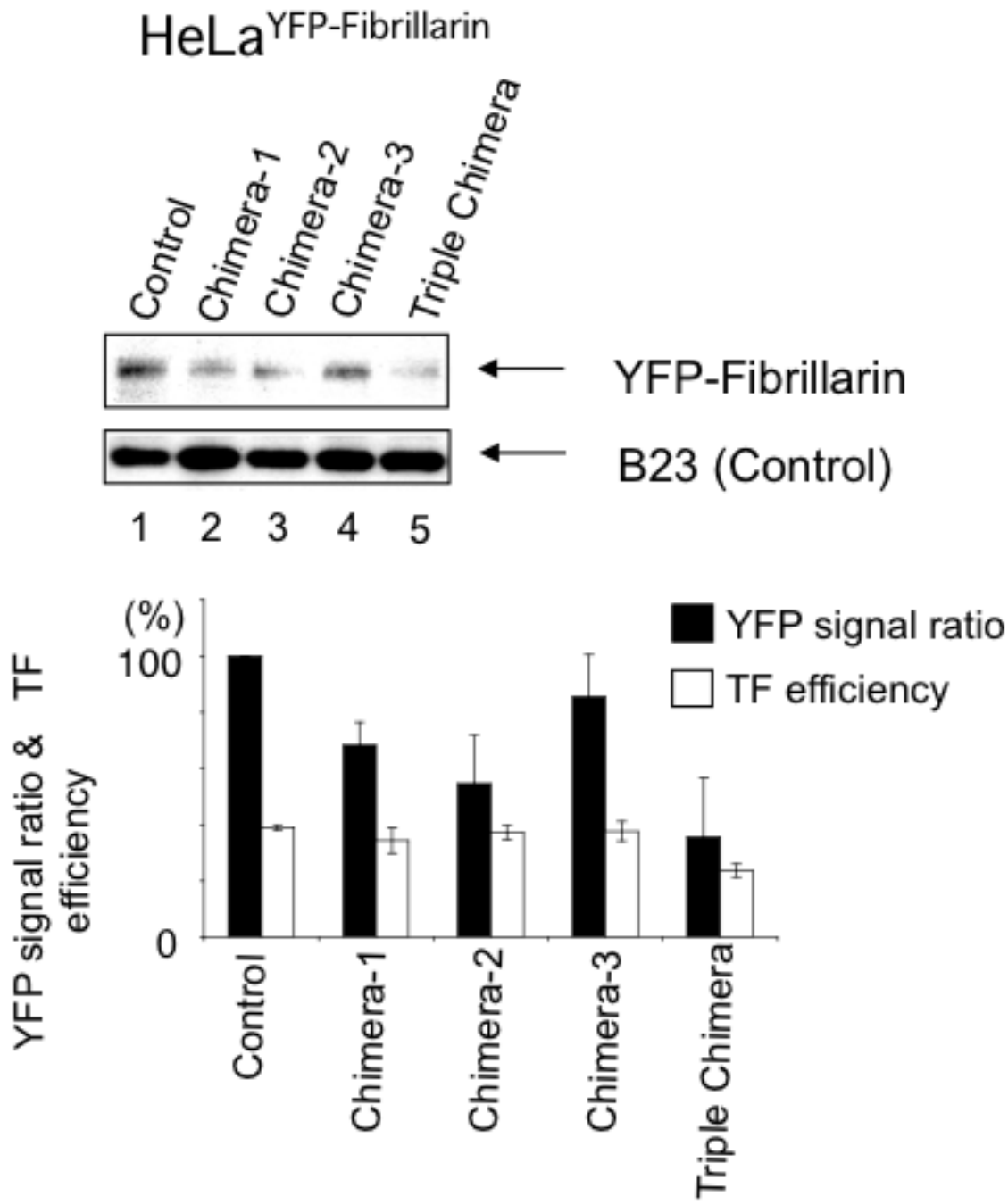

HeLa^GFP
mCherry-triple HBII-180C
Triple chimera-1, 2, 3
GFP
mCherry
DAPI
Merge
10

Fig. 10 (Sup)

a


HeLa YFP-Fibrillarin
Control
Chimera-1
Chimera-2
Chimera-3
Triple Chimera
YFP-Fibrillarin
B23 (Control)
1 2 3 4 5
(%)
100
0
YFP signal ratio & TF efficiency
YFP signal ratio
TF efficiency
Control
Chimera-1
Chimera-2
Chimera-3
Triple Chimera

b

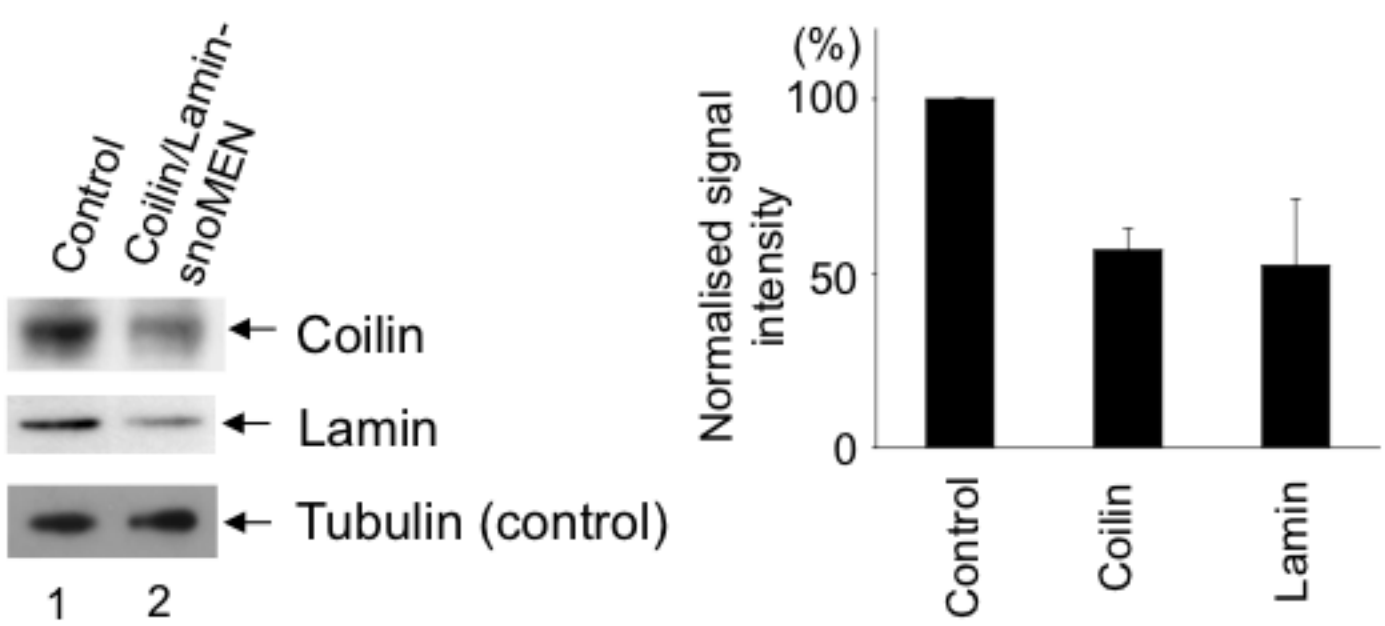
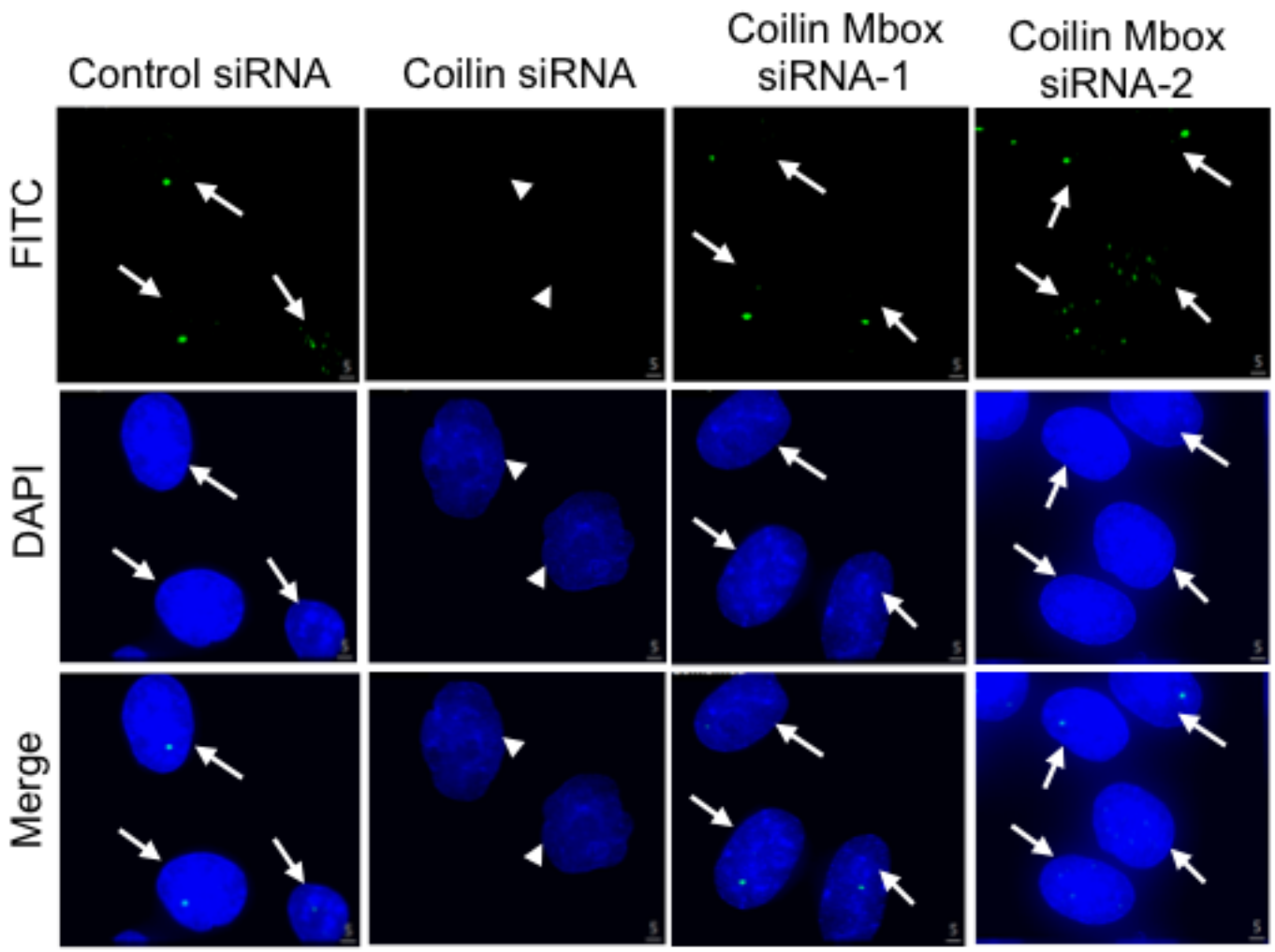
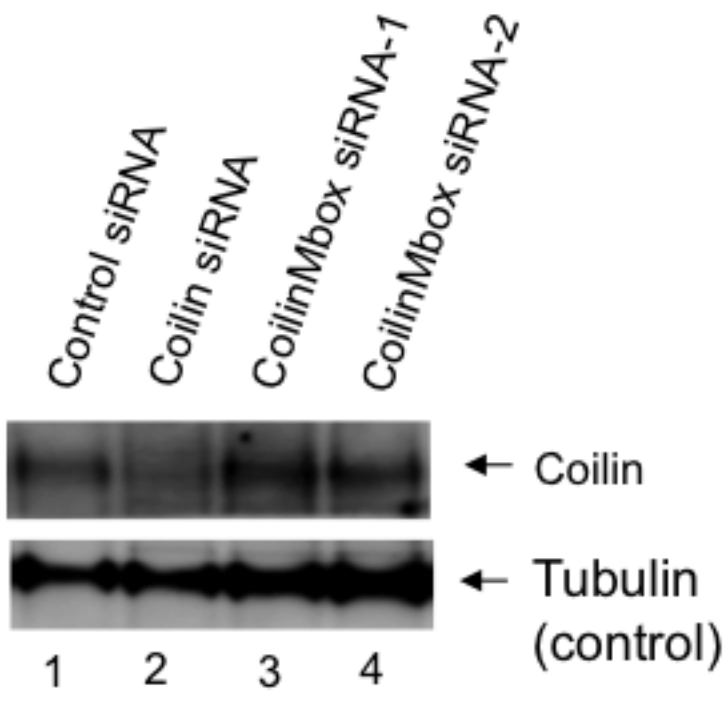

Control
Coilin/Lamin-snoMEN
Coilin
Lamin
Tubulin (control)
1 2
(%)
100
50
0
Normalised signal intensity
Control
Coilin
Lamin

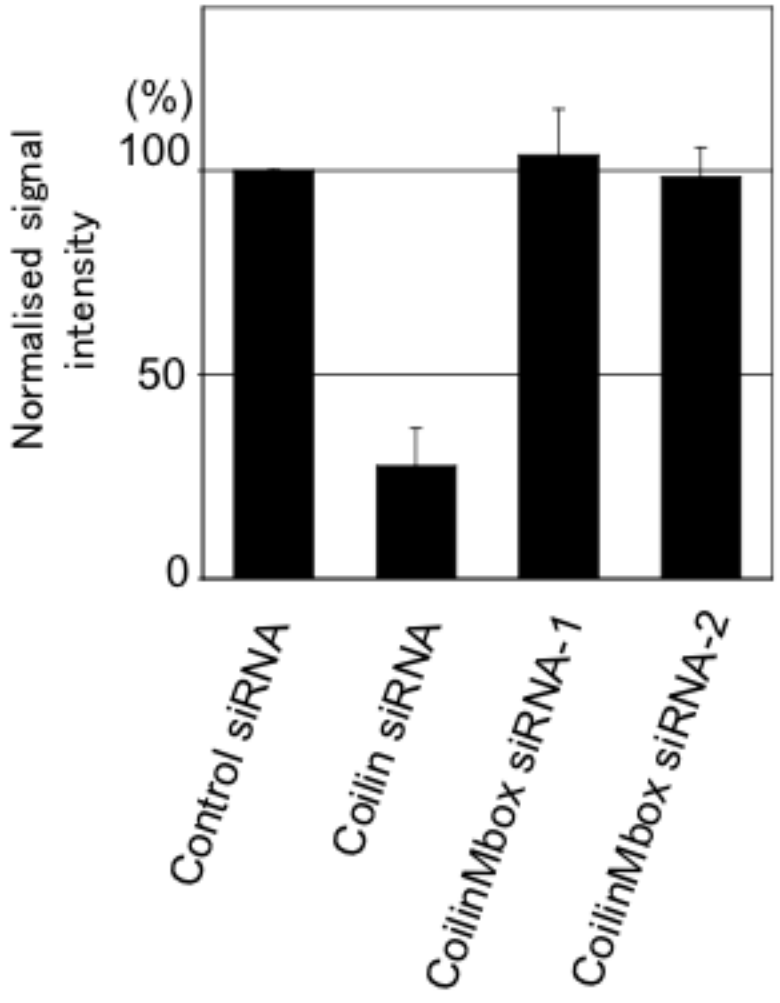
Fig. 11 (Sup)

a


Control siRNA
Coilin siRNA
Coilin Mbox siRNA-1
Coilin Mbox siRNA-2
FITC
DAPI
Merge

b


Control siRNA
Coilin siRNA
CoilinMbox siRNA-1
CoilinMbox siRNA-2
Coilin
Tubulin (control)
1
2
3
4


Normalised signal intensity
(%)
100
50
0
Control siRNA
Coilin siRNA
CoilinMbox siRNA-1
CoilinMbox siRNA-2

Supplementary Table 1A: List of small nucleolar RNAs identified from the unamplified cDNA library

Description	Group	Type	Accesion No.[a]	Chrom.	Feature	Previous status
AIL-1	snoRNA	box C/D	NA	9p13.3	Intron (ubiquitin associated protein 2)	Previously unknown
HBII-234	snoRNA	box C/D	NR_003058	2q33.1	Intron (nucleolar protein NOP5/NOP58)	Predicted by homology to mouse MBII-234 [1]
HBII-429	snoRNA	box C/D	NR_002435	6q23.2	Intron (ribosomal protein S12) not U101	Predicted by homology to mouse MBII-429 [1]
HBII-296B	snoRNA	box C/D	NR_003073	17p13.3	Intron (hypothetical protein FLJ10534)	Predicted by homology to mouse MBII-296 [1]
HBII-180C	snoRNA	box C/D	NR_003069	19q13.33	Intron in multidrug resistance-related protein	Predicted by homology to mouse MBII-180C and MBII-211 [1]
U106	snoRNA	box C/D	AY349607	20q13.13	Intron (similar to RPE-spondin)	Predicted by homology to rat U106 [2]
HBII-135	snoRNA	box C/D	NR_003054	17p11.2	Intron (hypothetical protein MGC40157) not U49	Experimentally identified in human [3]
HBII-202	snoRNA	box C/D	NR_002450	16q24.3	Intron (ribosomal protein L13)	Predicted by homology to mouse MBII-202 [1]
AIL-8	snoRNA	box C/D	NA	20p13	Intron (small nuclear ribonucleoprotein polypeptides B and B1)	Previously unknown
U101	snoRNA	box C/D	NR_002434	6q23.2	Intron (ribosomal protein S12)	Predicted by homology to rat U101 [2]
U104	snoRNA	box C/D	AY349605	17q23.3	Intron (hmm27142)	Predicted by homology to rat U104 [2]
U13	snoRNA	box C/D	NR_003041	3p21.31	Intron (kinesin family member 9)	Previously experimentally identified in human [4]
U14A	snoRNA	box C/D	NR_000022	11q24.1	Intron (heat shock 70kDa protein 8)	Previously experimentally idenytified in human [5]
U15b	snoRNA	box C/D	NR_000025	11q13.3-5	Intron (ribosomal protein S3)	Paralogue of experimentally identified U15A [6]
U18A	snoRNA	box C/D	NR_002441	15q22	Intron (ribosomal protein L4)	Previously experimentally idenytified in human [5]
U20	snoRNA	box C/D	NR_002908	2q12	Intron (nucleolin)	Previously experimentally identified in human [7]
U22	snoRNA	box C/D	NR_000008	11q13	Intron (UHG)	Experimentally identified in human [3]
U24	snoRNA	box C/D	NR_002447	9q34	Intron (ribosomal protein L7a)	Previously experimentally identified in human [8]
U3	snoRNA	box C/D	M14061	17p11.2	Intron (GRB2-related adaptor protein-like) or close 5'region	Previously experimentally identified in human [4]
U32A	snoRNA	box C/D	NR_000021	19q13.3	Intron (ribosomal protein L13a)	Previously experimentally identified in human [9]
U33	snoRNA	box C/D	NR_000020	19q13.3	Intron (ribosomal protein L13a)	Previously experimentally identified in human [9]
U35A	snoRNA	box C/D	NR_000018	19q13.3	Intron (ribosomal protein L13a)	Previously experimentally identified in human [9]
U36b	snoRNA	box C/D	NR_000017	9q34	Intron (ribosomal protein L7a)	Previously experimentally identified in human [9]
U41	snoRNA	box C/D	NR_002751	19p13.13	Intron (transportin 2 (importin 3, karyopherin beta 2b))	Previously experimentally idenytified in human [5]
U42A	snoRNA	box C/D	NR_000014	17q11	Intron (ribosomal protein L23a)	Previously experimentally identified in human [5]
U44	snoRNA	box C/D	NR_002750	1q25.1	Intron (gas5) or hmm10502	Previously experimentally identified in human [5]
U47	snoRNA	box C/D	NR_002746	1q25.1	Intron (gas5) or hmm10502	Previously experimentally identified in human [5]
U48	snoRNA	box C/D	NR_002745	6p21.3	Intron (chromosome 6 open reading frame 48) close to HSP70a1b	Previously experimentally identified in human [5]
U49A	snoRNA	box C/D	NR_002744	17p11.2	Intron (hypothetical protein MGC40157) not AW3	Previously experimentally identified in human [5]
U50B	snoRNA	box C/D	NR_003044	6q14.3	close to LOC441164	Predicted paralogue of experimentally identified U50 [5]

						experimentally identified U50 [5]
U51	snoRNA	box C/D	NR_002589	2q33-34	Intron (eukaryotic translation elongation factor 1 beta 2)	Previously experimentally identified in human [5]
U54	snoRNA	box C/D	NR_002437	8q12	Intron (ribosomal protein S20)	Previously experimentally identified in human [5]
U57	snoRNA	box C/D	NR_002738	20p13	Intron (nucleolar protein 5A (56kDa with KKE/D repeat))	Previously experimentally identified in human [5]
U58A	snoRNA	box C/D	NR_002571	18q21	Intron (ribosomal protein L17)	Previously experimentally identified in human [5]
U59A	snoRNA	box C/D	NR_002737	12q13.13	Intron (ATP synthase, H+ transporting, mitochondrial F1 complex, beta)	Previously experimentally identified in human [5]
U64	snoRNA	box H/ACA	NR_002326	16p13.3	Intron (ribosomal protein S2)	Previously experimentally identified in human [10]
U73a	snoRNA	box C/D	NR_000007	4q31.2-3	Intron (ribosomal protein S3A)	Previously experimentally identified in human [11]
U76	snoRNA	box C/D	AF141346	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U79 [12]
U77	snoRNA	box C/D	AF141346	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U79 [12]
U78	snoRNA	box C/D	AF141346	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U79 [12]
U79	snoRNA	box C/D	AF141346	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U79 [12]
U83	snoRNA	box C/D	AJ243200	6p21.3	Intron (HLA-B associated transcript 1)	Previously experimentally identified in human [13]
U83a	snoRNA	box C/D	NR_000027	22q13	Intron (ribosomal protein L3)	Previously experimentally identified in human [14]
U84	snoRNA	box C/D	NR_003065	6p21.3	Intron (HLA-B associated transcript 1)	Previously experimentally identified in human [13]
U87	snoRNA	box C/D and box H/ACA	NR_003008	8q13	close to hmm36010	Previously experimentally identified in human [15]
U95	snoRNA	box C/D	NR_002591	5q35.3	Intron (guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1)	Predicted by homology to rat U95 [2]
Z15(U80)	snoRNA	box C/D	AJ224022	1q25.1	Intron (gas5) or hmm10502	Previously experimentally identified in human [12]
Z19(U75)	snoRNA	box C/D	AF141346	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U75 [12]
Z23(U81)	snoRNA	box C/D	AJ224028	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U81 [12]
Z39 homolog	snoRNA	box C/D	AJ240057 (mouse)	5q14	Intron (cytochrome c oxidase subunit VIIc)	Previously unknown in human
Z40 homolog	snoRNA	box C/D	AJ240059 (rattus)	11q21	Intron (hypothetical protein MGC5306)	Previously unknown in human
U1	snRNA	box C/D	V00591	1p36.13	close to hypothetical gene supported by BC033316 (3' dawnstream)	Previously experimentally identified in human [16]
U60	snRNA	box C/D	NR_002736	16p13.3	between TNF receptor-associated factor 7 &member RAS oncogene family	Previously experimentally identified in human [5]
ACA62	snoRNA	box H/ACA	NR_002995	17q23.3	close to uncharacterized hypothalamus protein HT008 or EXON of hmm27142	Computationally predicted [17]
U19-2	snoRNA	box H/ACA	NR_002988	10q11.23	Intron (hmm290)	Previously experimentally identified in human [18]
ACA10	snoRNA	box H/ACA	NR_002327	16p13.3	Intron (ribosomal protein S2)	Previously experimentally identified in human [19]
ACA14A	snoRNA	box H/ACA	NR_002955	1q42.3	Intron (similar to Translocase of outer mitochondrial membrane 20 homolog)	Previously experimentally identified in human [19]
ACA14b	snoRNA	box H/ACA	NR_002956	1q42.3	Intron (similar to Translocase of outer mitochondrial membrane 20 homolog)	Previously experimentally identified in human [19]
ACA16/U98	snoRNA	box H/ACA	NR_003035	1P35.3	close to ACA44, ACA61 (PNAS-123, 3'non-coding)	Previously experimentally identified in human [19]
ACA18	snoRNA	box H/ACA	NR_002959	11q21	Intron (hypothetical protein MGC5306)	Previously experimentally identified in human [19]
ACA19	snoRNA	box H/ACA	NR_002917	10q26	Intron (eukaryotic translation initiation factor 3, subunit 10 theta, 150/170kDa)	Previously experimentally identified in human [19]
ACA20	snoRNA	box H/ACA	NR_002960	6q25-27	Intron (t-complex 1)	Previously experimentally identified in human [19]

ACA21	snoRNA	box H/ACA	NR_002576	17q	Intron (ribosomal protein L23)	Previously experimentally identified in human [19]
ACA22	snoRNA	box H/ACA	NR_002961	7q11.21	Intron (similar to chaperonin containing TCP1, subunit 6A (zeta 1))	Previously experimentally identified in human [19]
ACA23	snoRNA	box H/ACA	NR_002962	11p15.4	Intron (importin 7)	Previously experimentally identified in human [19]
ACA24	snoRNA	box H/ACA	NR_002963	4q28.1	close to protease, serine, 12 and unknown gene	Previously experimentally identified in human [19]
ACA28	snoRNA	box H/ACA	NR_002964	14q32.32	Intron (eukaryotic translation initiation factor 5)	Previously experimentally identified in human [19]
ACA3	snoRNA	box H/ACA	NR_002580	11p15	Intron (ribosomal protein L27a)	Previously experimentally identified in human [19]
ACA31	snoRNA	box H/ACA	NR_002967	13q12-14	Intron (tumor protein, translationally-controlled 1)	Previously experimentally identified in human [19]
ACA33	snoRNA	box H/ACA	NR_002436	6q23.2	Intron (ribosomal protein S12) not U101 not marco	Previously experimentally identified in human [19]
ACA34	snoRNA	box H/ACA	NR_002968	12q13.11	Intron (hypothetical protein FLJ20436)	Previously experimentally identified in human [19]
ACA40	snoRNA	box H/ACA	NR_002973	11q21	Intron (hypothetical protein MGC5306)	Previously experimentally identified in human [19]
ACA44	snoRNA	box H/ACA	NR_002976	1P35.3	close to ACA16, ACA61 (PNAS-123, not intron)	Previously experimentally identified in human [19]
ACA46	snoRNA	box H/ACA	NR_002978	12q24.33	Intron (E1A binding protein p400)	Previously experimentally identified in human [19]
ACA48	snoRNA	box H/ACA	NR_002918	17p13	Intron (eukaryotic translation initiation factor 4A, isoform 1)	Previously experimentally identified in human [19]
ACA49	snoRNA	box H/ACA	NR_002979	12q24.33	Intron (E1A binding protein p400)	Previously experimentally identified in human [19]
ACA54	snoRNA	box H/ACA	NR_002982	11p15.5	Intron (nucleosome assembly protein 1-like 4)	Previously experimentally identified in human [19]
ACA6	snoRNA	box H/ACA	NR_002325	3P21.3	Intron (laminin receptor 1 (ribosomal protein SA, 67kDa))	Previously experimentally identified in human [19]
ACA61	snoRNA	box H/ACA	NR_002987	1P35.3	close to ACA16, ACA44 (PNAS-123, not intron)	Previously experimentally identified in human [19]
ACA7	snoRNA	box H/ACA	NR_002582	3p25-24	Intron (ribosomal protein L32)	Previously experimentally identified in human [19]
ACA8	snoRNA	box H/ACA	NR_002920	11q21	Intron (hypothetical protein MGC5306)	Previously experimentally identified in human [19]
ACA9	snoRNA	box H/ACA	NR_002952	7p13	close to hmm34350	Previously experimentally identified in human [19]
E1(U17)	snoRNA	box H/ACA	NR_002907	1p36.1	Intron (chromosome condensation 1)	Previously experimentally identified in human [20]
E1b	snoRNA	box H/ACA	U12210	1p36.1	Intron (chromosome condensation 1)	Previously experimentally identified in human [21]
E3	snoRNA	box H/ACA	NR_002586	3q28	Intron (eukaryotic translation initiation factor 4A, isoform 2)	Previously experimentally identified in human [20]
AIL-2	snoRNA	box H/ACA	NA	17q24.2	Nucleolar protein 11 Intron (DKFZP586L0724 protein)	Previously unknown
AIL-5	snoRNA	box H/ACA	NA	5q31.1	Intron (SEC24 related gene family, member A)	Previously unknown
AIL-7	snoRNA	box H/ACA	NA	5q21.3	Intron (EFNA5)ephrin-A5	Previously unknown
U19	snoRNA	box H/ACA	NR_002915	5q31.2	Intron (hmm33082) close to matrin 3	Previously experimentally identified in human [22]
U67	snoRNA	box H/ACA	NR_002912	17P13	Intron (eukaryotic translation initiation factor 4A, isoform 1)	Previously experimentally identified in human [10]
U68	snoRNA	box H/ACA	NR_000012	19p13	Intron (ribosomal protein L18a)	Previously experimentally identified in human [10]
U70	snoRNA	box H/ACA	NR_000011	Xq28	Intron (ribosomal protein L10)	Previously experimentally identified in human [10]
U71b	snoRNA	box H/ACA	NR_002910	20q11.23	Intron (hypothetical LOC388796)	Previously experimentally identified in human [10]
U72	snoRNA	box H/ACA	NR_002581	8q22	Intron (ribosomal protein L30)	Previously experimentally identified in human [10]
U99	snoRNA	box H/ACA	AY349600	11q12.3	Intron (hypothetical protein MGC2477)	Previously experimentally identified in human [2]
U2	snRNA	box H/ACA	K03023	17q21.31	(hmm6552, 3'region) close to U2 small nuclear pseudogene 2	Previously experimentally identified in human [23]

AIL-3	snoRNA	box H/ACA	NA	1p36.13	Intron (KIAA0476)	Previously unknown
MRP/Th	snoRNA	NA	X51867	9p13.3	close to hypothetical protein MGC31967 (5' upstreame)	Previously experimentally identified in human [24]
U5E	snRNA	NA	M77839	1p36.22	close to hypothetical protein MGC33867	Previously experimentally identified in human [25]

[a] When available, Refseq identifiers were used. Otherwise, other GenBank or EnsEMBL accession numbers were chosen.

Supplementary Table 1B: List of small nucleolar RNAs identified from the amplified cDNA library

Description	Group	Type	Accesion No.[a]	Chromosome	Feature	Previous status
ACA16/U98	snoRNA	box H/ACA	NR_003035	1P35.3	close to ACA44, ACA61 (PNAS-123, 3'non-coding)	Previously experimentally identified in human [19]
ACA24	snoRNA	box H/ACA	NR_002963	4q28.1	close to protease, serine, 12 and unknown gene	Previously experimentally identified in human [19]
ACA27	snoRNA	box H/ACA	NR_002575	13q12.2	Intron (ribosomal protein L21)	Previously experimentally identified in human [19]
ACA28	snoRNA	box H/ACA	NR_002964	14q32.32	Intron (eukaryotic translation initiation factor 5)	Previously experimentally identified in human [19]
ACA2a	snoRNA	box H/ACA	NR_002950	12q13.11	Intron (hypothetical protein FLJ20436)	Previously experimentally identified in human [19]
ACA34	snoRNA	box H/ACA	NR_002968	12q13.11	Intron (hypothetical protein FLJ20436)	Previously experimentally identified in human [19]
ACA50	snoRNA	box H/ACA	NR_002980	16q21	Intron (CNOT1)	Previously experimentally identified in human [19]
ACA7	snoRNA	box H/ACA	NR_002582	3p25-24	Intron (ribosomal protein L32)	Previously experimentally identified in human [19]
ACA8	snoRNA	box H/ACA	NR_002920	11q21	Intron (hypothetical protein MGC5306)	Previously experimentally identified in human [19]
ACA9	snoRNA	box H/ACA	NR_002952	7p13	close to hmm34350	Previously experimentally identified in human [19]
AIL-4	snoRNA	box H/ACA	NA	1p36.11	Intron (hmm28229)	Previously unknown
HBII-234	snoRNA	box C/D	NR_003058	2q33.1	Intron (nucleolar protein NOP5/NOP58)	Predicted by homology to mouse MBII-234 [1]
E1(U17)	snoRNA	box H/ACA	NR_002907	1p36.1	Intron (chromosome condensation 1)	Previously experimentally identified in human [20]
E2	snoRNA	box H/ACA	NR_002324	3p22.2	Intron (ribosomal protein SA)	Previously experimentally identified in human [20]
E3	snoRNA	box H/ACA	NR_002586	3q28	Intron (eukaryotic translation initiation factor 4A, isoform 2)	Previously experimentally identified in human [20]
HBII-85	snoRNA	box C/D	NR_003316	15q11.2	Intron (hmm4244)	Previously experimentally identified in human [26]
HBII-429	snoRNA	box C/D	NR_002435	6q23.2	Intron (ribosomal protein S12) not U101	Predicted by homology to mouse MBII-429 [1]
hsa-mir-21	microRNA		NT_010783	17q23.1		Previously experimentally identified in human [27]
ACA63	snoRNA	box H/ACA	NR_003019	22q11.21	Intron (RAN binding protein 1)	Computationally predicted and experimentally validated [17]
HBII-296A	snoRNA	box C/D	NR_003072	17p13.3	Intron (hypothetical protein FLJ10534)	Predicted by homology to mouse MBII-296 [1]
HBII-55	snoRNA	box C/D	NR_003078	20p13	Intron (nucleolar protein 5A (56kDa with KKE/D repeat))	Predicted by homology to mouse MBII-55 [1]
HBII-180C	snoRNA	box C/D	NR_003069	19q13.33	Intron in multidrug resistance-related protein	Predicted by homology to mouse MBII-180C and MBII-211 [1]
U105	snoRNA	box C/D	AY349606	19p13	Intron (peter pan homolog (Drosophila))	Predicted by homology to rat U105 [2]
U106	snoRNA	box C/D	AY349607	20q13.13	Intron (similar to RPE-spondin)	Predicted by homology to rat U106 [2]
U14A	snoRNA	box C/D	NR_000022	11q24.1	Intron (heat shock 70kDa protein 8)	Previously experimentally idenytified in human [5]
U18A	snoRNA	box C/D	NR_002441	15q22	Intron (ribosomal protein L4)	Previously experimentally idenytified in human [5]

U21	snoRNA	box C/D	NR_000006	1p22.1	Intron (ribosomal protein L5)	Previously experimentally identified in human [28]
U34	snoRNA	box C/D	NR_000019	19q13.3	Intron (ribosomal protein L13a)	Previously experimentally identified in human [9]
U36c	snoRNA	box C/D	NR_000016	9q34	Intron (ribosomal protein L7a)	Previously experimentally identified in human [9]
U38b	snoRNA	box C/D	NR_001457	1p34.1-p32	Intron (ribosomal protein S8)	Previously experimentally identified in human [9]
U42A	snoRNA	box C/D	NR_000014	17q11	Intron (ribosomal protein L23a)	Previously experimentally identified in human [5]
U44	snoRNA	box C/D	NR_002750	1q25.1	Intron (gas5) or hmm10502	Previously experimentally identified in human [5]
U49A	snoRNA	box C/D	NR_002744	17p11.2	Intron (hypothetical protein MGC40157) not AW3	Previously experimentally identified in human [5]
U50	snoRNA	box C/D	NR_002743	6q14.3	Inside a cluster of ESTs	Previously experimentally identified in human [5]
U52	snoRNA	box C/D	NR_002742	6p21.3	Intron (chromosome 6 open reading frame 48)	Previously experimentally identified in human [5]
U56	snoRNA	box C/D	NR_002739	20p13	Intron (nucleolar protein 5A (56kDa with KKE/D repeat))	Previously experimentally identified in human [5]
U59A	snoRNA	box C/D	NR_002737	12q13.13	Intron (ATP synthase, H+ transporting, mitochondrial F1 complex, beta)	Previously experimentally identified in human [5]
U60	snoRNA	box C/D	NR_002736	16p13.3	3' down stream of CASK interacting protein 1	Previously experimentally identified in human [5]
U64	snoRNA	box H/ACA	NR_002326	16p13.3	Intron (ribosomal protein S2)	Previously experimentally identified in human [10]
U67	snoRNA	box H/ACA	NR_002912	17P13	Intron (eukaryotic translation initiation factor 4A, isoform 1)	Previously experimentally identified in human [10]
U68	snoRNA	box H/ACA	NR_000012	19p13	Intron (ribosomal protein L18a)	Previously experimentally identified in human [10]
U70	snoRNA	box H/ACA	NR_000011	Xq28	Intron (ribosomal protein L10)	Previously experimentally identified in human [10]
U72	snoRNA	box H/ACA	NR_002581	8q22	Intron (ribosomal protein L30)	Previously experimentally identified in human [10]
U95	snoRNA	box C/D	NR_002591	5q35.3	Intron (guanine nucleotide binding protein (G protein))	Predicted by homology to rat U95 [2]
mgU6-77	snoRNA	box C/D	NR_002604	17p13	Intron (eukaryotic translation initiation factor 4A, isoform 1)	Predicted by homology to mouse mgU6-77 [29]
Z15(U80)	snoRNA	box C/D	AJ224022	1q25.1	Intron (gas5) or hmm10502	Previously experimentally identified in human [12]
Z18(U74)	snoRNA	box C/D	AJ224028	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U74 [12]
Z20(U76)	snoRNA	box C/D	AJ224022	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U76 [12]
Z23(U81)	snoRNA	box C/D	AJ224028	1q25.1	Intron (gas5) or hmm10502	Predicted by homology to mouse U81 [12]

[a] When available, Refseq identifiers were used. Otherwise, other GenBank or EnsEMBL accession numbers were chosen.

1. Huttenhofer, A., Kiefmann, M., Meier-Ewert, S., O'Brien, J., Lehrach, H., Bachellerie, J.P., and Brosius, J. (2001). RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse. Embo J *20*, 2943-2953.
2. Vitali, P., Royo, H., Seitz, H., Bachellerie, J.P., Huttenhofer, A., and Cavaille, J. (2003). Identification of 13 novel human modification guide RNAs. Nucleic Acids Res *31*, 6543-6551.
3. Tycowski, K.T., Shu, M.D., and Steitz, J.A. (1994). Requirement for intron-encoded U22 small nucleolar RNA in 18S ribosomal RNA maturation. Science *266*, 1558-1561.
4. Tyc, K., and Steitz, J.A. (1989). U3, U8 and U13 comprise a new

class of mammalian snRNPs localized in the cell nucleolus. Embo J *8*, 3113-3119.
5. Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell *85*, 1077-1088.
6. Tycowski, K.T., Shu, M.D., and Steitz, J.A. (1993). A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3. Genes Dev *7*, 1176-1190.
7. Nicoloso, M., Caizergues-Ferrer, M., Michot, B., Azum, M.C., and Bachellerie, J.P. (1994). U20, a novel small nucleolar RNA, is encoded in an intron of the nucleolin gene in mammals. Mol Cell Biol *14*, 5766-5776.
8. Qu, L.H., Henry, Y., Nicoloso, M., Michot, B., Azum, M.C., Renalier, M.H., Caizergues-Ferrer, M., and Bachellerie, J.P. (1995). U24, a novel intron-encoded small nucleolar RNA with two 12 nt long, phylogenetically conserved complementarities to 28S rRNA. Nucleic Acids Res *23*, 2669-2676.
9. Nicoloso, M., Qu, L.H., Michot, B., and Bachellerie, J.P. (1996). Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2'-O-ribose methylation of rRNAs. J Mol Biol *260*, 178-195.
10. Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997). The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. Genes Dev *11*, 941-956.
11. Rebane, A., Tamme, R., Laan, M., Pata, I., and Metspalu, A. (1998). A novel snoRNA (U73) is encoded within the introns of the human and mouse ribosomal protein S3a genes. Gene *210*, 255-263.
12. Smith, C.M., and Steitz, J.A. (1998). Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. Mol Cell Biol *18*, 6897-6909.
13. Jady, B.E., and Kiss, T. (2000). Characterisation of the U83 and U84 small nucleolar RNAs: two novel 2'-O-ribose methylation guide RNAs that lack complementarities to ribosomal RNAs. Nucleic Acids Res *28*, 1348-1354.
14. Duga, S., Asselta, R., Malcovati, M., Tenchini, M.L., Ronchi, S., and Simonic, T. (2000). The intron-containing L3 ribosomal protein gene (RPL3): sequence analysis and identification of U43 and of

two novel intronic small nucleolar RNAs. Biochim Biophys Acta *1490*, 225-236.

15. Darzacq, X., Jady, B.E., Verheggen, C., Kiss, A.M., Bertrand, E., and Kiss, T. (2002). Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. Embo J *21*, 2746-2756.
16. Murphy, J.T., Burgess, R.R., Dahlberg, J.E., and Lund, E. (1982). Transcription of a gene for human U1 small nuclear RNA. Cell *29*, 265-274.
17. Schattner, P., Barberan-Soler, S., and Lowe, T.M. (2006). A computational screen for mammalian pseudouridylation guide H/ACA RNAs. Rna *12*, 15-25.
18. Gu, A.D., Zhou, H., Yu, C.H., and Qu, L.H. (2005). A novel experimental approach for systematic identification of box H/ACA snoRNAs from eukaryotes. Nucleic Acids Res *33*, e194.
19. Kiss, A.M., Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human box H/ACA pseudouridylation guide RNA machinery. Mol Cell Biol *24*, 5797-5807.
20. Ruff, E.A., Rimoldi, O.J., Raghu, B., and Eliceiri, G.L. (1993). Three small nucleolar RNAs of unique nucleotide sequences. Proc Natl Acad Sci U S A *90*, 635-638.
21. Selvamurugan, N., Nag, M.K., and Eliceiri, G.L. (1995). Intron-encoded small nucleolar RNAs: new RNA sequence variants and genomic loci. Biochim Biophys Acta *1260*, 230-234.
22. Kiss, T., Bortolin, M.L., and Filipowicz, W. (1996). Characterization of the intron-encoded U19 RNA, a new mammalian small nucleolar RNA that is not associated with fibrillarin. Mol Cell Biol *16*, 1391-1400.
23. Van Arsdell, S.W., and Weiner, A.M. (1984). Human genes for U2 small nuclear RNA are tandemly repeated. Mol Cell Biol *4*, 492-499.
24. Topper, J.N., and Clayton, D.A. (1990). Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. Nucleic Acids Res *18*, 793-799.
25. Sontheimer, E.J., and Steitz, J.A. (1992). Three novel functional variants of human U5 small nuclear RNA. Mol Cell Biol *12*, 734-746.
26. Cavaille, J., Seitz, H., Paulsen, M., Ferguson-Smith, A.C., and Bachellerie, J.P. (2002). Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. Hum Mol Genet *11*, 1527-1538.

27. Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853-858.
28. Qu, L.H., Nicoloso, M., Michot, B., Azum, M.C., Caizergues-Ferrer, M., Renalier, M.H., and Bachellerie, J.P. (1994). U21, a novel small nucleolar RNA with a 13 nt. complementarity to 28S rRNA, is encoded in an intron of ribosomal protein L5 gene in chicken and mammals. Nucleic Acids Res *22*, 4073-4081.
29. Tycowski, K.T., You, Z.H., Graham, P.J., and Steitz, J.A. (1998). Modification of U6 spliceosomal RNA is guided by other small RNAs. Mol Cell *2*, 629-638.