Inventory of Supplemental Information

Inventory of Supplementary Figures and Tables

Figure S1 Characteristics of the long ncRNA transcripts. Related to Figure 1.

Figure S2 Annotated long ncRNAs are independent transcripts. Related to Figure 2.

Figure S3 Characterization of the analyzed ncRNA-activators. Related to Figure 3.

Figure S4 ChIP analysis of the analyzed ncRNA-a loci. Related to Figure 4.

Table S1 Annotation and genomic coordinates for the identified 3,019 long ncRNAs from the GENCODE annotation. Related to Figure 1.

Table S2 qPCR validation of the custom microarray expression data for the 1 percent ENCODE region for HeLa cells, human primary keratinocytes and HEK-293 cells. Related to Figure 2.

Table S3 99 ncRNAs with a fold change >|1.5| in keratinocytes after induction of differentiation with TPA treatment using customized expression array. Related to Figure 2.

Table S4 Expression data for knock-down analyses of ncRNA-a loci. Related to Figure 3.

Table S5 ID of the long ncRNAs tested in knock-down experiments. Related to Figure 4.

Table S6 Microarray results from Snai1 and ncRNA-a7 knock-down in A549 cells. Related to Figure 5.

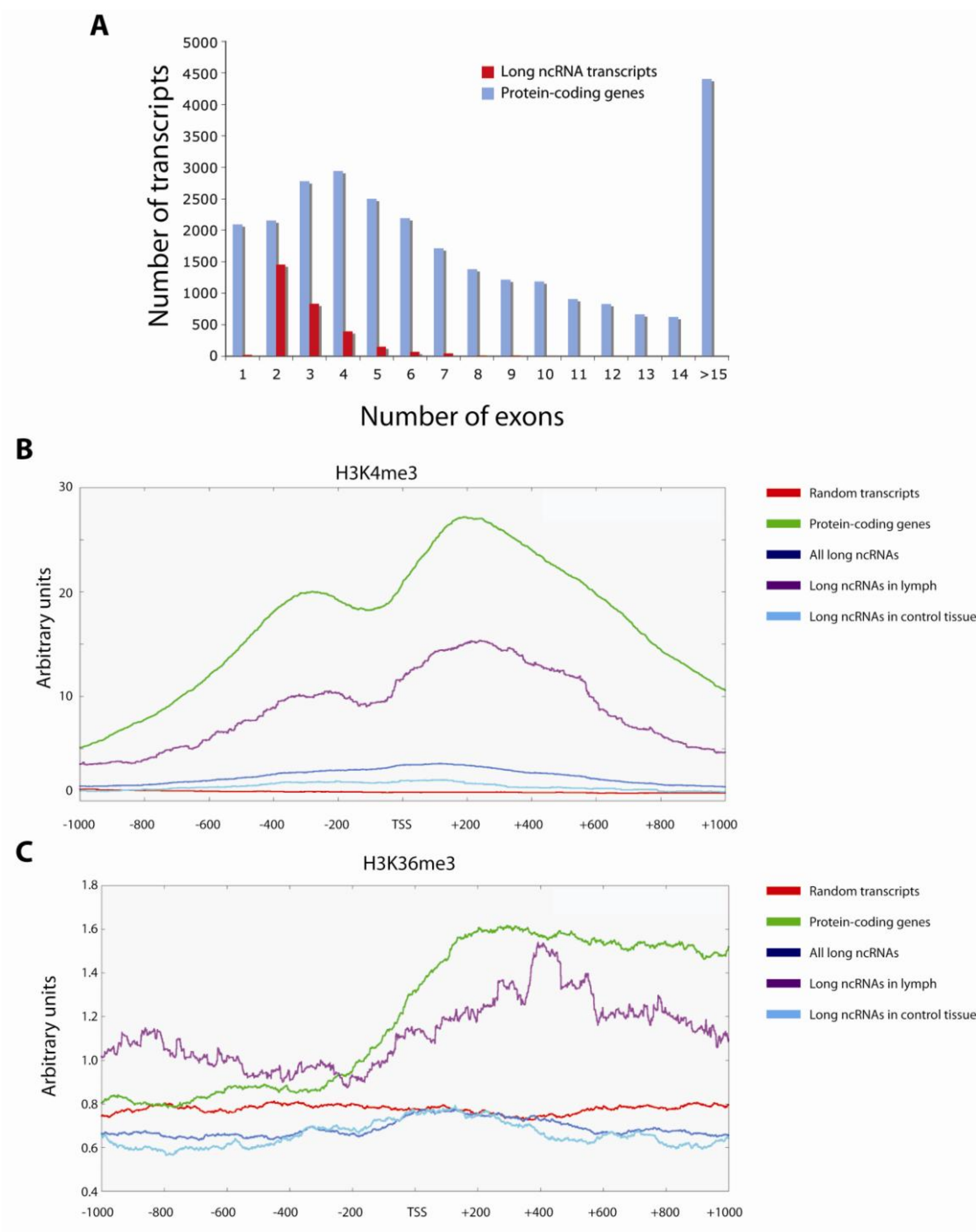

Figure S1. Characteristics of the long ncRNA transcripts. Related to Figure 1.

(A) Number of exons for the entire transcript and the corresponding number of transcripts shown for all 3,019 long ncRNAs and all protein-coding genes.

(B and C) Analysis of histone marks H3K4 (B) or H3K36 (C) trimethylation from -1,000 to +1,000 from the TSS of the 3,019 long ncRNAs for the various sets shown in the legend. Data are extracted from publicly available databases (Barski et al., 2007).

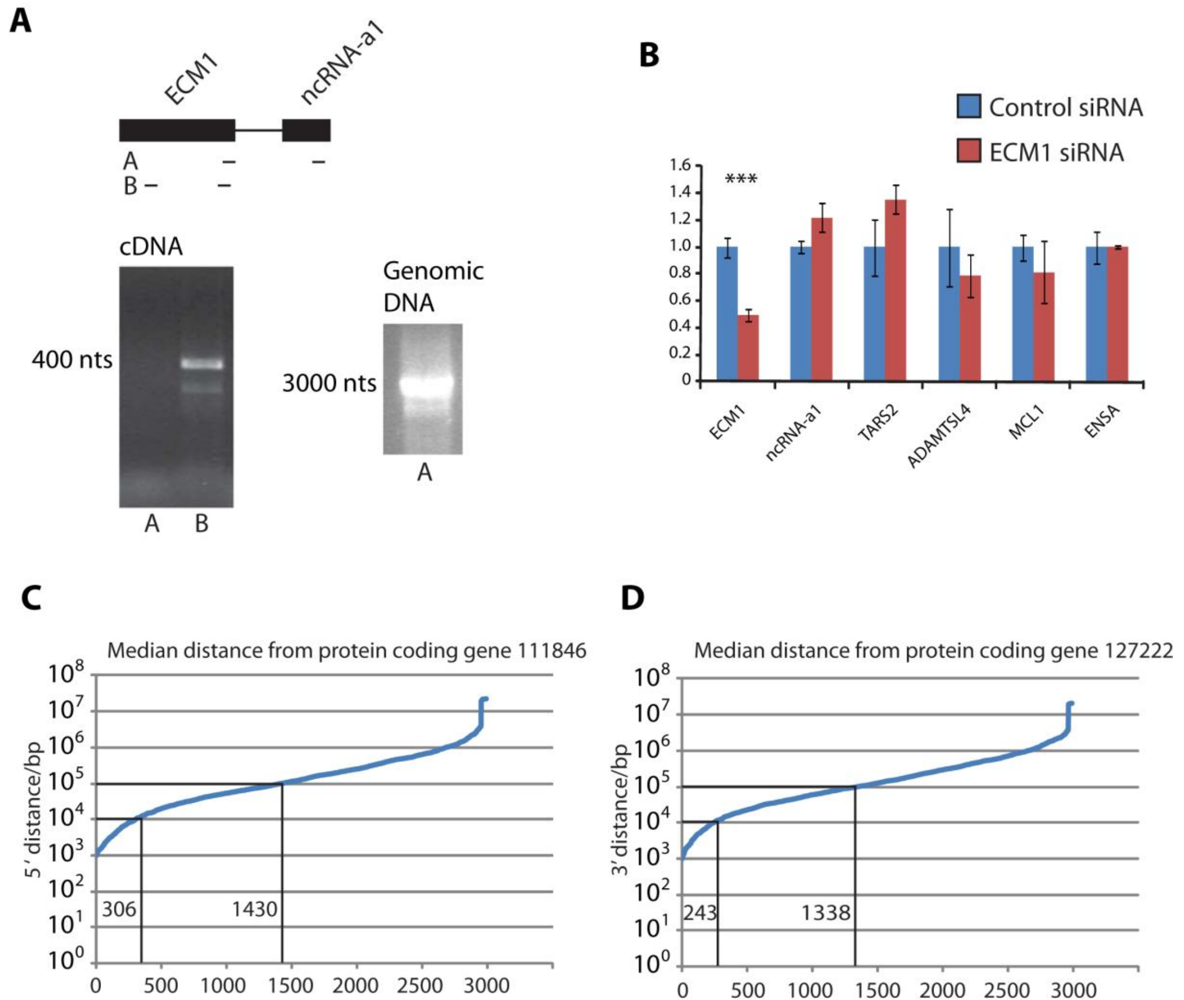

Figure S2. Annotated long ncRNAs are independent transcripts. Related to Figure 2.

(A) PCR amplification using the primer pairs shown in the upper panel on cDNA from HEK293 cells (left panel) or human genomic DNA (right panel).

(B) Analysis of the ECM1 locus upon siRNA knock-down of ECM1 using qPCR. y-axis show RNA level compared to Gapdh and normalized to control siRNA values set to an average of 1.

(C and D) Plot showing the distance of the long ncRNAs from its nearest protein coding gene at the 5' end (C) and the 3' end (D). Y-axes show the distance in bp on a logarithmic scale. Numbers on X-axes represent the 3,019 long ncRNAs ordered by increasing distance to their nearest protein coding gene. Indicated are numbers of transcripts closer than 10 kb or 100 kb, respectively, to their nearest protein coding genes.

A

5’ RACE 3’ RACE PCR

5’ RACE
100 bp 3’ RACE
ncRNA-a1
400 nts
1,000 nts
500 nts

5’ RACE
1000 bp
ncRNA-a2
3’ RACE
400 nts
1,000 nts
500 nts

5’ RACE
3’ RACE
100 bp
ncRNA-a3
400 nts
1,000 nts
500 nts

5’ RACE
3’ RACE
100 bp
ncRNA-a4
400 nts
1,000 nts
500 nts

5’ RACE
3’ RACE
100 bp
ncRNA-a5
400 nts
1,000 nts
500 nts

5’ RACE
3’ RACE
1000 bp
ncRNA-a6
400 nts
1,000 nts
500 nts

5’ RACE
3’ RACE
100 bp
ncRNA-a7
400 nts
1,000 nts
500 nts

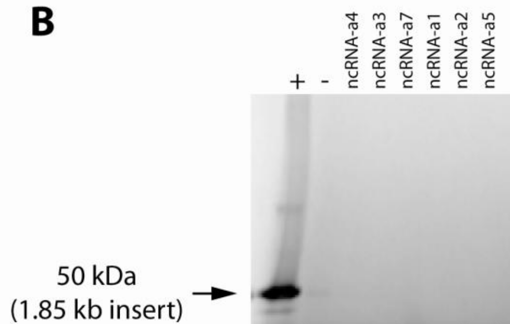

Figure S3. Characterization of the analyzed ncRNA-activators. Related to Figure 3.

(A) RACE analysis of the identified ncRNA-a using the primers indicated (left panel) for 5' and 3' RACE analysis (middle panel). PCR analysis of exon-exon junctions of the transcripts are shown in the right panel. Expected product size according to the HAVANA annotation is around 400 bp. The two bands that appear at a larger size (ncRNA-a3 and ncRNA-a6) were sequenced and showed different processing that annotated, as indicated above these two panels in red

(B) In vitro translation assay. Lane + is the pTRI-REF positive control, lane - is a negative control and following lanes are the indicated ncRNA-a cloned downstream of a T7 promoter and transcribed and used in a translation assay in vitro.

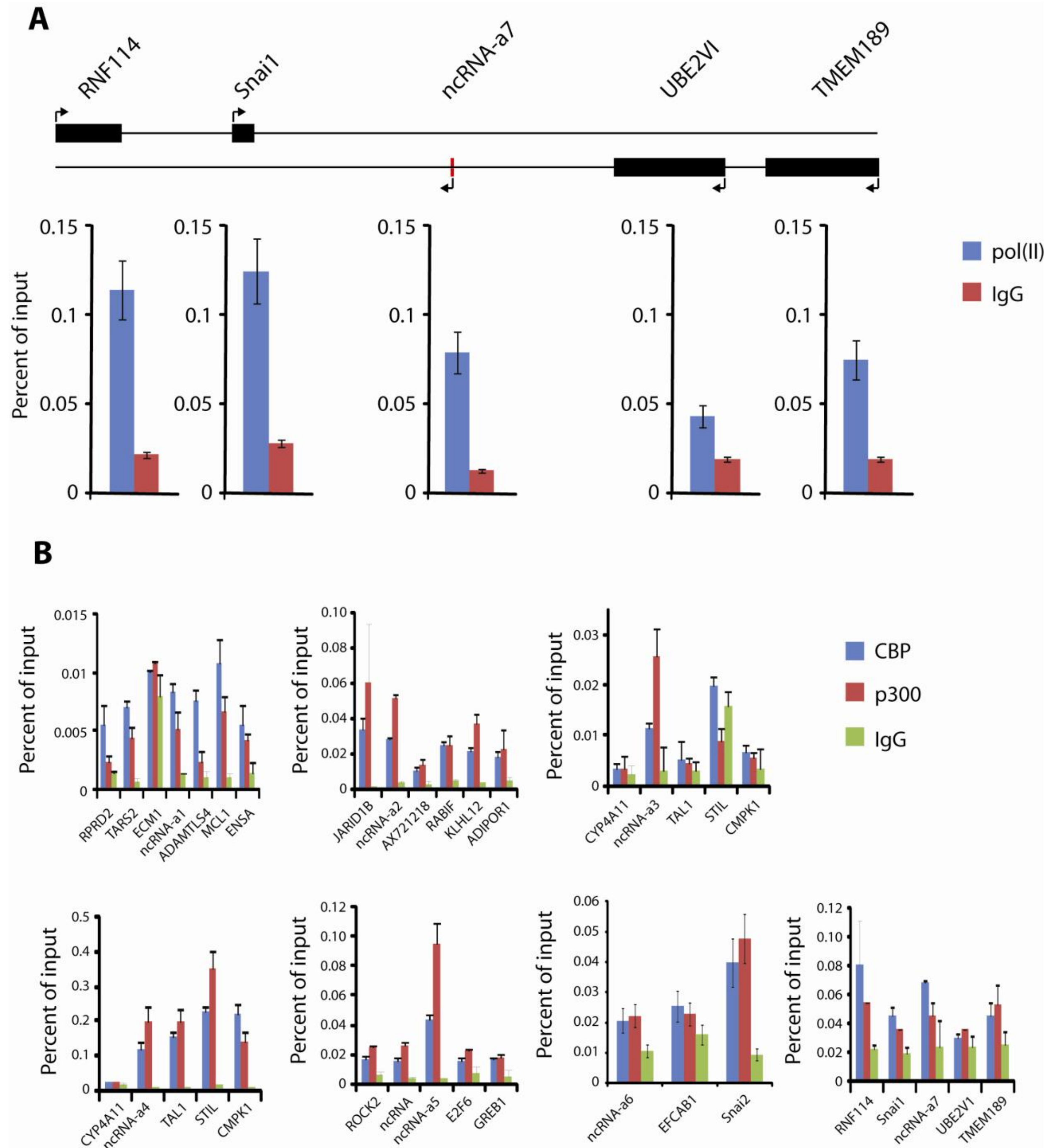

Figure S4. ChIP analysis of the analyzed ncRNA-a loci. Related to Figure 4.

(A) Pol(II) ChIP of the Snai1 locus.

(B) ChIP for the analyzed ncRNA-a loci using antibodies against p300 or CBP. Y-axes show amount of material in the IP relative to the Input.

Transcript ID	HeLa	Keratinocytes	HEK-293
AC000123.2-001	V	V	V
AC000123.3-001	V	V	
AC000124.1-001		V	
AC005538.3-002	V	V	V
AC005592.1-001		V	
AC063976.3-001	V	V	
AC104389.16-001			V
AC113617.1-002		V	
AL163953.3-002		V	
AP002856.7-001		V	
RP1-149A16.12-001	V		
RP11-629E24.2-001	V		
RP11-65J3.1-001	V	V	V
RP1-90G24.6-001			V
RP3-477O4.15-001	V	V	V
RP1-149M18.4-001	V		V

Table S2. qPCR validation of the custom microarray expression data for the 1 percent ENCODE region for HeLa cells, human primary keratinocytes and HEK-293 cells. Related to Figure 2.

Gene	RPRD2	TARS2	ECM1	ncRNA-a1	ADAMTSL4	MCL1	ENSA	
siRNA 1	1.03	0.95	0.51	0.44	1.24	1.27	0.77	
siRNA2	1.31	1.04	0.51	0.37	1.40	2.93	1.10	
control 1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
control 2	0.91	0.92	1.15	1.08	0.90	0.50	0.66	
Gene	JARID1B	ncRNA-a2	LOC641515	AX711218	NR_002929	RABIF	KLHL12	ADIPOR1
siRNA 1	0.80	0.28	N.D.	1.12	N.D.	0.99	0.79	1.05
siRNA2	0.80	0.41	N.D.	0.96	N.D.	1.26	0.78	0.92
control 1	1.00	1.00	N.D.	1.00	N.D.	1.00	1.00	1.00
control 2	1.05	0.87	N.D.	1.00	N.D.	0.97	1.05	1.07
Gene	CYP4A11	ncRNA-a3	ncRNA-a4	PDZK1IP1	TAL1	STIL	CMPK1	
siRNA 1	0.41	0.18	N.D.	N.D.	0.23	1.34	1.22	
siRNA2	0.67	0.37	N.D.	N.D.	0.31	1.01	1.38	
control 1	1.00	1.00	N.D.	N.D.	1.00	1.00	1.00	
control 2	0.99	1.33	N.D.	N.D.	1.29	1.11	1.46	
Gene	CYP4A11	ncRNA-a3	ncRNA-a4	PDZK1IP1	TAL1	STIL	CMPK1	
siRNA 1	0.94	0.95	0.55	N.D.	1.14	0.88	0.60	
siRNA2	0.85	0.96	0.53	N.D.	0.72	1.31	0.60	
control 1	1.00	1.00	1.00	N.D.	1.00	1.00	1.00	
control 2	0.75	1.00	0.98	N.D.	0.92	0.88	1.36	
Gene	PQLC3	ROCK2	OTTHUMT00000324050	ncRNA-a5	E2F6			
siRNA 1	1.12	0.71	N.D.	0.16	0.91			
siRNA2	1.04	0.73	N.D.	0.38	0.79			
control 1	1.00	1.00	N.D.	1.00	1.00			
control 2	0.85	0.98	N.D.	0.94	0.90			
Gene	ncRNA-a6	EFCAB1	Snai2					
siRNA 1	0.20	1.02	0.44					
siRNA2	0.38	0.95	0.50					
control	1.00	1.00	1.00					
Gene	RNF114	Snai1	ncRNA-a7	UBE2VI	TMEM189	CEBPB		
siRNA 1	1.08	0.34	0.30	1.12	1.05	0.96		
siRNA2	0.96	0.42	0.30	1.10	1.04	1.00		
control	1.00	1.00	1.00	1.00	1.00	1.00		

Table S4. Expression data for knock-down analyses of ncRNA-a loci. Related to Figure 3.

Name	ID	Affected gene	Cell line
ncRNA-a1	OTTHUMT00000087261	ECM1	HEK293
ncRNA-a2	OTTHUMT00000099182	KLHL12	HeLa
ncRNA-a3	OTTHUMT00000021645	TAL1	MCF-7
ncRNA-a4	OTTHUMT00000021648	CMPK1	Jurkat
ncRNA-a5	OTTHUMT00000324045	ROCK2	HeLa
ncRNA-a6	OTTHUMT00000280498	Snai2	A549
ncRNA-a7	OTTHUMT00000079667	Snai1	A549
	OTTHUMT00000054962	N/A	HEK293
	OTTHUMT00000048112	N/A	Hs68
	OTTHUMT00000040694	N/A	HeLa
	OTTHUMT00000089836	N/A	HeLa
	OTTHUMT00000256004	N/A	HepG2

Table S5. ID of the long ncRNAs tested in knock-down experiments. Related to Figure 4.

Methods

Extracting long ncRNA data

The HAVANA annotation12 has been downloaded using the DAS server26 provided by the Sanger institute (version July,16th 2008). We removed all annotated biotypes or functional elements belonging to specific categories such as pseudogenes or PCGs. We excluded all transcripts overlapping with known protein coding loci annotated by HAVANA, RefSeq or UCSC27. Transcripts falling into a 1kb window (5' and 3' UTRs) of all PCGs were also removed. Finally, we excluded all transcripts covered by known non-coding RNAs such as miRNAs (miRbase version 11.0 April 2008).

Conservation analysis

To estimate the evolutionary constraints among mammalian sequences we constructed the cumulative distribution of PhastCons28 scores for ancestral repeats (ARs), RefSeq genes and long ncRNAs using the following datasets. For Ancestral repeats, we used the RepeatMasker annotation29 for Human genome release Hg18. We excluded repeats annotated as "Simple Repeats", "Low complexity" and primate- or rodents-specific families. For those annotated as RNA or L1, we selected the ones showing more than 25% and 20% divergence, respectively, from consensus in order to ensure that repeats were not recent insertions. Then, we excluded those which overlapped with genes annotated by RefSeq, HAVANA or UCSC, as well as those overlapping miRNAs (miRbase version 11.0, 2008). We kept those which were conserved using the Multiz alignment30. Finally, in order to avoid GC bias content due to Biased Gene Conversion31, we selected repeats in a close vicinity (10Kb) of long ncRNA loci. RefSeq transcripts32 have been retrieved from the UCSC genome browser (July 2008). The cumulative distributions of these transcripts or repeats are plotted using a log-scale on the y-axis.

Cell culture

Human primary keratinocytes from four different biological donors were grown in Keratinocyte medium (KFSM, Invitrogen). Differentiation was induced by 2.5 ng/ml 12-O-tetradecanoylphorbol-13-acetate (TPA) during 48h.

HEK293, A549, HeLa and MCF-7 cells were cultured in complete DMEM media (GIBCO) containing 10 % FBS, and 1X Anti/Anti (GIBCO). Jurkat cells were cultured in complete RPMI media (GIBCO) containing 10 % FBS and 1X Anti/Anti (GIBCO).

siRNA transfections

For transfections of 293, HeLa, A549 and MCF-7 cells we used Lipofectamine 2000 (Invitrogen) according to the manufacturer's recommendations and an siRNA concentration of 50 nM. Jurkat cells were transfected using HiPerFect (Qiagen) according to the manufacturer's recommendations and an siRNA concentration of 100 nM.

siRNA sequences:

Specific siRNAs for ncRNA-a1

Sense 1 agaacauacaggaggaagauu;

Antisense 1 ucuuccuccuguauguucuuu;

Sense 2 guagcaacuucaaacggaauu;

Antisense 2 uuccguuugaaguugcuacuu;

siRNA controls for ncRNA-a1

Sense 3 gcuuccugcuacacagaauuu;

Antisense 3 auucuguguagcaggaagcuu;

Sense 4 ggaauguacuauagagccauu;

Antisense 4 uggcucuauaguacauuccuu;

Specific siRNAs for ncRNA-a2

Sense 1 ccuccuuacucuuggacaauu;

Antisense 1 uuguccaagaguaaggagguu;

Sense 2 cccacaaacgcuugcgauuuu;

Antisense 2 aaucgcaagcguuuguggguu;

siRNA controls for ncRNA-a2

Sense 3 ggaggaaacauuugauauuuu;

Antisense 3 aauaucaaauguuuccuccuu;

Sense 4 ccucagacggacagcauaauu;

Antisense 4 uuaugcuguccgucugagguu;

Specific siRNAs for ncRNA-a3

Sense 1 cuauggaauuuaagcccaauu;

Antisense 1 uugggcuuaaauuccauaguu;

Sense 2 gcaagaucacggugcgucauu;

Antisense 2 ugacgcaccgugaucuugcuu;

siRNA controls for ncRNA-a3

Sense 3 ggaauguacuauagagccauu;

Antisense 3 uggcucuauaguacauuccuu;

Sense 4 ccacaagcaucgaggagcauu;

Antisense 4 ugcuccucgaugcuugugguu;

Specific siRNAs for ncRNA-a4

Sense 1 ggaauguacuauagagccauu;

Antisense 1 uggcucuauaguacauuccuu;

Sense 2 ccacaagcaucgaggagcauu;

Antisense 2 ugcuccucgaugcuugugguu;

siRNA controls for ncRNA-a4

Sense 3 cuauggaauuuaagcccaauu;

Antisense 3 uugggcuuaaauuccauaguu;

Sense 4 gcaagaucacggugcgucauu;

Antisense 4 ugacgcaccgugaucuugcuu;

Specific siRNA sequences for ncRNA-a5

Sense 1 GGAUGUGAGACGUGGAAUCUU;

Antisense 1 GAUUCCACGUCUCACAUCCUU;

Sense 2 CCUUGGAGAAUAAAGCUUAUU;

Antisense 2 UAAGCUUUAUUCUCCAAGGUU;

siRNA controls for ncRNA-a5

Sense 3 gcuuccugcuacacagaauuu;

Antisense 3 auucuguguagcaggaagcuu;

Sense 4 ggaauguacuauagagccauu;

Antisense 4 uggcucuauaguacauuccuu;

Specific siRNA sequences for ncRNA-a6

Sense 1 GGAGAAAUGAGGAAGUAAAUU;

Antisense 1 UUUACUUCCUCAUUUCUCCUU;

Sense 2 CGAAAGAAGCACAGGGUGUUU;

Antisense 2 ACACCCUGUGCUUCUUUCGUU;

siRNA controls for ncRNA-a6

Sense 3 gcuuccugcuacacagaauuu;

Antisense 3 auucuguguagcaggaagcuu;

Specific siRNA sequences for ncRNA-a7

Sense 1 GAAGGGAACCAGUGCUAAAUU;

Antisense 1 UUUAGCACUGGUUCCCUUCUU;

Sense 2 CCGAUUUGAGAGAGUGAGAUU;

Antisense 2 UCUCACUCUCUCAAAUCGGUU;

siRNA controls for ncRNA-a7

Sense 3 gcuuccugcuacacagaauuu;

Antisense 3 auucuguguagcaggaagcuu;

ECM1 siRNA Sense GUGCAGAGGAGGAGAAAUUUU

ECM1 siRNA Antisense AAUUUCUCCUCCUCUGCACUU

Snai1 siRNA Sense GCUUUGAGCUACAGGACAAUU

Snai1 siRNA Antisense UUGUCCUGUAGCUCAAAGCUU

RNA purification, cDNA synthesis and quantitative PCR

Cells were harvested and resuspended in TRIzol (Invitrogen) and RNA extracted according to the manufacturer's protocol. cDNA synthesis was done using MultiScribe reverse transcriptase and random primers from Applied Biosystems. Quantitative PCR was done using SybrGreen reaction mix (Applied Biosystems) and an HT7900 sequence detection system (Applied Biosystems). For all quantitative PCR reactions Gapdh was measured for an internal control and used to normalize the data. Primer sequences are as follows:

ncRNA-a1 FW gcaagcggagacttgtcttt;

ncRNA-a1 RV ggctggtcttgaactcctga;

ECM1 FW gccctctgctgttacctgag;

ECM1 RV ccttggcgttctcagtgtct;

TARS2 FW tggatgcagactctggactg;

TARS2 RV ctttgctctttctggccaac;

ADAMTSL4 FW cccagcctgtggatgtctat;

ADAMTSL4 RV caaggattggaggaggtgaa;

MCL1 FW gaaggcgctggagaccttac;

MCL1 RV atgtccagtttccgaagcat;

ENSA FW ggtcttgggtgtgatgtgtg;

ENSA RV ggcctttagctttgcctctt;

RPRD2 FW gcagctgctctagtgaagga;

RPRD2 RV aaagcttctctcaatgccaca;

ncRNA-a3 FW gtgcgtcagagctagcaaga;

ncRNA-a3 RV tcgtcgggtactctccattc;

ncRNA-a4 FW tcaaccagaaagagacagagca;

ncRNA-a4 RV gcctttcagatcttccattcc;

PDZK1IP1 FW ctgcacacatgatcctgacc;

PDZK1IP1 RV tctcatgctcactggacctg;

STIL FW ttcccttgtgatggcaagat;

STIL RV tcagttcacaacggattgga;

CMPK1 FW ctcctctgctctccacgtct;

CMPK1 RV gcaagatgtcatccctcgac;

CYP4A11 FW ccaactgaggaaggctcaac;

CYP4A11 RV attttggccaagaggaggat;

TAL1 FW agggcctggttgaagaagat;

TAL1 RV aagtaagggcgactgggttt;

JARID1B FW tcagtgcagagagccagaga;

JARID1B RV ggatagatcggcctcgtgta;

ncRNA-a2 FW ccttcgcccctagatacacc;

ncRNA-a2 RV agcacgagcaaggcagag;

LOC641515 FW ggcagaacctcaccttcaaa;

LOC641515 RV tgtacaagaggctgcttcca;

AX721218 FW aacaccttcttcagggagacc;

AX721218 RV cactgtgggggactggtagt;

NR_002929 FW aaacccatggagttgctcac;

NR_002929 RV tgcttgagttcaggagcttg;

RABIF FW agggaccgctctcttctctc;

RABIF RV agtgttcctggaggagatcg;

KLHL12 FW tcaagtgcgacgaaattcag;

KLHL12 RV gctctttcttggcatgcttc;

ADIPOR1 FW acatctggacccatctgctt;

ADIPOR1 RV cccaaaaaccaccttctcct;

PQLC3 FW1 CCT CAG CCT TCC GAG TTT AC;

PQLC3 RV1 GAG GAT GGG GTA CTCCAGGT;

E2F6 FW1 CTT CGC CAT GAA TCC TTC TC;

E2F6 RV1 TGG AGC CCA TTC CTA CAT TC;

ncRNA-a5 FW1 AGC CTG CAG CCT TTC TTT C;

ncRNA-a5 RV1 CAT CCC TTT CCT GGG GTA GT;

OTTHUMT00000324048 FW1 ATT CCT TGC GAA GCT GAGAG;

OTTHUMT00000324048 RV1 ACT GCT TAA ATG CGG TTG CT;

OTTHUMT00000324050 FW1 CAA GGA CAC AGC CAA GTTCA;

OTTHUMT00000324050 RV1 TCT CTG TTT CCC TGA CCT CCT;

ROCK2 FW1 TGA AGC CTG ACA ACA TGC TC;

ROCK2 RV1 TAC CAT GCC TGT TTC ATC CA;

ncRNA-a7 FW1 CCGTTGGCTCCACAAACCT;

ncRNA-a7 RV1 CAGTGACAGTAGCAGGCATCCT;

ncRNA-a6 FW1 CAAGGCATTTTTGCACTCAGTAA;

ncRNA-a6 RV1 AACACGGCTCAGCTATGGAAA;

Snai1 FW1 TTCAACTGCAAATACTGCAACAAG;

Snai1 RV1 TGTGGCTTCGGATGTGCAT;

Snai2 FW1 CCATTCCACGCCCAGCTA;

Snai2 RV1 TCACTCGCCCCAAAGATGAG;

TMEM189 FW1 GCTTCGTCTTCTGCCTGATCA;

TMEM189 RV1 AAAGTACGTGTGCGACCACTTG;

UBE2V1 FW1 AAGCAAGAGCGACGCAAGAT;

UBE2V1 RV1 CGGAAATTGCGAGGGACTT;

RNF114 FW1 GCAGCGCCAACTTCAGAGA;

RNF114 RV1 CATCAACATCATAATCCACAAAAGTG;

AURKA FW1 TTGGAAGACTTGGGTCCTTG;

AURKA RV1 TCAAATATCCCCGCACTCTG;

CDH1 FW1 CCTGGGACTCCACCTACAGA;

CDH1 RV1 TGTGAGCAATTCTGCTTGGA;

PKP2 FW1 ATCCAGGAGCAGGTGCAG;

PKP2 RV1 CTGCTGGTTCGGTGAAGATT;

PLOD2 FW1 GGAACATTACGGCAAATGGT;

PLOD2 RV1 TTCTCCAGATCAACTTGCTTCA;

Cloning of pGL3-TK reporters and luciferase assay

pGL3-Basic was digested with BglII and HindIII and the TK promoter from pRL-TK was inserted into these sites. ncRNA inserts were cloned from genomic DNA including a 1,500 bp upstream and a 500 bp downstream region using the primers GGG GAT CCA GCG GTG TGG AAT AAA CTG G and GGG TCG ACC CTG CCT CCT AAG CTT CCT T (for ncRNA-a3/4), GGG GAT CCC ATG CCT CCT TGA GTT CCA T and GGG TCG ACG TGT GTG TCC ATG TGG CTT C (for ncRNA-a5), C CTG GCT CAC AGA TCC TAGC and CC TCTCAGCAGAGGCTGCTTT with an added BamHI or SalI site dependent on the direction of the ncRNA-a7 insert, CTC CAG CAC GTC ATC GAC TA and ATC GGT CTT GTT CTC CCT CA (for ID1 insert) and TTT GGC AGC TCC TCA TCT CT and AAT TAG CGC CAT CGA CAT TC (for GTSF1L). For the control vector (pGL3-TK-control) GGG GAT CCA GCA GGC TTA GGG GAA AGA G and GGG TCG ACC CCT TGC TCA GTC AAG AAG C were used to amplify a 4 kb fragment from genomic DNA not showing any evidence of transcriptional activity. Luciferase assays were performed in 96-well white plates using Dual-Glo (Promega) according to the manufacturer’s protocol.

RACE

For RACE we used the ExactStart kit from Epi Biotechnologies and the following primers:

ncRNA-a3 5RACE TGG GTC TGG CAC ATT CTG TA

ncRNA-a3 3RACE GAA TTG GAT GCA AGA TCA CG

ncRNA-a4 5RACE GTT CAA ATT CCA GCC TGC TC

ncRNA-a4 3RACE CAA TTT TAC CGG CAA AGG AA

ncRNA-a2 5RACE GCT CAA GGC TTG CTA CTT GG

ncRNA-a2 3RACE CCT TTC CTC CCC AAC AAA CT

ncRNA-a1 5RACE CTC CTG ACC TCG TGA TCC A

ncRNA-a1 3RACE CCT GGC CAA GAA GCT CAT AC

ncRNA-a5 5RACE AAG GCC TAG ATT TTT GCA ACC

ncRNA-a5 3RACE AGC CTG CAG CCT TTC TTT C

ncRNA-a7 5RACE ACT GGC TGA GGT TTG TGG AG

ncRNA-a7 3RACE GGC CGA TTT GAG AGA GTG AG

ncRNA-a6 5RACE CGC AGC CTC TTC TAA ACA CC

ncRNA-a6 3RACE CCT TGA GAC CCA TTC ACG TT

ncRNA-a processing PCR

ncRNA-a7 processing FW GGC AGA GAC GGA GAA GAT GA

ncRNA-a7 processing RV ACT GGC TGA GGT TTG TGG AG

ncRNA-a3 processing FW TCT TCC CCG TCA CAC TGG

ncRNA-a3 processing RV AGC GGT GTG GAA TAA ACT GG

ncRNA-a4 processing FW CAG AAA AGC TCC CGA AAC TG

ncRNA-a4 processing RV ACT CCC TAG CTT GGC TCT CC

ncRNA-a2 processing FW CCC CCT CCT TAC TCT TGG AC

ncRNA-a2 processing RV TGA TGG CAT TGA ATT GGA GA

ncRNA-a1 processing FW GCA AGC GGA GAC TTG TCT TT

ncRNA-a1 processing RV TGC TGC CTC CTT TTC TTC AT

ncRNA-a6processing FW AAC GGA AGG GGG AAT TAT TG

ncRNA-a6 processing RV CGC AGC CTC TTC TAA ACA CC

ncRNA-a5 processing FW AGA CGT AAT GGC ACC CTG CT

ncRNA-a5 processing RV CAA GGC TTG CAA GAA TAG GA

Microarrays

Custom-made microarrays (Agilent) were designed based on the library of 3,019 long ncRNA sequences, with on average six probes targeting each transcript. Human whole genome mRNA arrays were from Agilent (G4112F). Total RNA samples were converted to cDNA using oligo-dT primers. Labeling of the cDNA and hybridization to the microarrays were performed according to Agilent standard dye swap protocols. Data analysis was done using the AFM 4.0 software. All microarrays were done in 3 biological replicates.

Transwell migration assay

In vitro cell migration assays were performed as described previously (Gumireddy et al., 2009) using Trans-well chambers (8μM pore size; Costar). Briefly, A549 control or human snail LnRNA siRNA treated cells were allowed to grow to subconfluency (~75–80%) and were serum-starved for 24 h. After detachment with trypsin, cells were washed with PBS, resuspended in serum-free medium and 250 μl cell suspensions (2 × 105 cells ml–1) was added to the upper chamber of trans-well insert. Complete medium was added to the bottom wells of the chambers. After 10h, the cells that had not migrated through the filter were removed from the upper face of the filters using cotton swabs, and the cells that had migrated to the lower face of the filters were fixed with 5% glutaraldehyde solution and stained with 0.5% solution of Toluidine Blue. Images of three random ×10 fields were captured from each membrane and the number of migratory cells was counted. The mean of triplicate assays for each experimental condition was used.

In vitro translation

In vitro translation was done using the ProteinScript II kit from Ambion according to the manufacturer's protocol. ncRNA inserts were cloned from cDNA from RNA harvested

from cells expressing the ncRNAs and cloned downstream of the T7 promoter in pCR2.1 (Invitrogen).

Primers used were as follow:

ncRNA-a4 FW AGC TCA ACT TCC GAG GCT CT

ncRNA-a4 RV GGC TGA AGC CAG CAG AAC T

ncRNA-a3 FW GGC TCT TCC CCG TCA CAC T

ncRNA-a3 RV GCC TCA GCG GTG TGG AAT A

ncRNA-a1 FW CAT CTC CTA CGG CCT CCA G

ncRNA-a1 RV CCA AGT GTC CTG TGT AAT AGG C

ncRNA-a2 FW CAA CCC CAG CCT GGA GGT

ncRNA-a2 RV TAA GTG TTG TAA TAG GGG GTA GTA ACA A

ncRNA-a5 FW ACT GTG CCC TGG ATG TG

ncRNA-a5 RV ATG CAT AAG CTT TAT TCT CCA AGG

ncRNA-a7 FWGGC AGA GAC GGA GAA GAT GA

ncRNA-a7 RV GCC TTA GAC TGT GAA ATT TTA TTG A