cassette exons are bound by DGCR8. (**b**) UV cross-linking assay determines binding of DGCR8 to cassette exons that were identified by CLIP (upper panels). WT indicates intact RNA secondary structure, whereas Mut, indicates mutations introduced to disrupt this structure (lanes 2, 4, 6 and 8). Lower panels show Coomassie blue stainings of the same gels to show equal presence of T7-DGCR8 protein (**c**) The ZNF644 isoform including exon 3 is upregulated in human cells depleted of Drosha. The lower panel illustrates the presence of small RNAs identified in this region[32]. (**d**) The mouse TCF3 isoform including exon 18 is upregulated in mouse cells lacking DGCR8 or Drosha but not Dicer, indicating that is a miRNA-independent effect (**Supplementary Fig. 5 f,g**). The lower panel illustrates the presence of small RNAs identified in this region[32] in human transcripts and the presence of DGCR8 binding sites by CLIP. (**e**) The mouse NEDD4L isoform including exon 18 is upregulated in mouse cells lacking DGCR8. (**f)** The mouse CSNK1D isoform including exon 9 is upregulated in the absence of DGCR8 (left) but not Dicer (right).

ONLINE METHODS

RNA extraction and quantitative RT-PCR

Total RNA was isolated from cells by Trizol, following manufacturer's instructions. RNA was DNAse treated (RQ1 DNAse, Promega, M601A) and checked for DNA contamination. 500ng of total RNA was used for quantitative RT-PCR analyses with SuperScript III Platinum SYBR Green One-Step qRT-PCR Kit (Invitrogen, 11736-051). Primers for qRT-PCR analysis are listed in **Supplementary Table 2**. Data was analyzed with Bio-Rad CFX Manager software. All the experiments show the average and standard deviations of at least three independent biological replicas.

Northern blot for snoRNAs

Northern blots were performed as previously described[49]. Briefly, 5-15 μg of total RNA was separated in 1xTBE-urea 10% poly-acrylamide gel. Gel was transferred to nylon membrane (Hybond N^+, GE Healthcare) for 1 h using Genie Blotter in 1xTBE. Membrane was crosslinked twice 1200J/m^2. Pre-hybridization was performed overnight at 40°C in hybridization buffer (1xSSC, 1%SDS and 200 μg/ml ssDNA from Sigma (D7656)). Templates and probes for hybridization were prepared using miRvana Probe Construction kit (Ambion, AM1550). Probe hybridization was performed overnight at 40°C and 4 washes, 30 min each, with wash buffer (0.2xSSC and 0.2%SDS). Radioactive signals were analyzed using Phosphorimager.

Radioactive RNA labeling and *in vitro* processing

Templates for RNA synthesis and labeling were obtained by genomic PCR (primers are listed in Supplementary Table 2) adding a T7 promoter sequence at the 5'end. Pri-miR-24-2 was cloned in pGEMt-easy (Promega) and *Nde*I digested overnight to be transcribed. Transcription reactions were performed with T7 polymerase from Ambion (AM2082) in the presence of 40 μmols of ^{32}P-αUTP. Probes were gel-purified, phenol-extracted and ethanol precipitated. 50,000 cpms of each probe were incubated either with immunoprecipitated T7-DGCR8, Flag-Drosha or control immunoprecipitates in the presence of buffer A (0.5mM ATP, 20mM creatine phosphate and 3.2 mM $MgCl_2$). Reactions were incubated for 30 min at 30°C, followed by standard phenol/chloroform extraction and ethanol precipitation. RNAs were separated in an 8-10% 1xTBE poly-acrylamide urea gel. Gels were analyzed using Phosphorimager.

Cell culture, antibodies, plasmids and transfection experiments

HEK 293T and HeLa cells were cultured in DMEM medium supplemented with 10% FCS and Pen/Strep and grown in standard conditions. For mouse Embryonic stem cells (mES), gelatin-coated plates were used without MEFs feeders and cultured in DMEM supplemented with 75ml FCS, 5ml 100x P/S, 5ml 100x non-essential amino acid, 5ml 200mM Glutamine, 4µl 100% β-mercaptoethanol and LIF. *Dgcr8* KO mES cells were purchased from Novus Biologicals (NBA1-19349) and the parental strain, v6.5 from Thermo Scientific (MES1402). f/f Dicer and *DicerKO* mES cells were kindly provided by R. Blelloch lab (University of California, San Francisco). The following commercial antibodies were used: anti-DGCR8 (Abcam, ab36865, ab90579), anti-Drosha (Upstate, 07-717), anti-T7 mouse monoclonal antibody (Novagen, 69522), anti-β-tubulin (Sigma (T4026)). Anti-Dcp1 antibody was kindly provided by B. Seraphin (IGBMC, Strasbourg). A mammalian expression vector for DGCR8 was created by inserting a DGCR8 ORF (ENST00000351989, Ensembl) using *Xba*I and *Dpn*II sites into the mammalian expression vector, pCGT7, which has been previously described[50]. Vectors for overexpression of wild-type Drosha and dominant negative forms of DGCR8 and Drosha (Flag-Drosha, mDRBD1&2 DGCR8, ΔC114 Drosha, respectively) were kindly provided by Narry Kim (Seoul National University). Transfections were performed using Lipofectamine 2000 (Invitrogen) and cells were collected 48hr after transfection. Drosha knock down was performed in HeLa cells with Dharmafect 4 (Dharmacon) and two rounds of siRNA transfection. SiRNAs against Drosha, DGCR8 mRNAs and non-targeting siRNAs were purchased from Dharmacon (L-016996-00, L-015713-00 and D-001810-02, respectively). Cells were collected 24hr after the second transfection.

Immunoprecipitations and q-/RT-PCR (IP RT-PCR)

Cells overexpressing DGCR8 or Drosha proteins were collected 48 h after transfection. Cells were trypsinized, centrifuged and washed with PBS. Cell pellets were resuspended in buffer D (20mM Hepes-KOH pH 7.9, 100mM KCl, 0.2mM EDTA, 0.5mM DTT, 0.2mM PMSF, 5% glycerol and RNase inhibitors) and sonicated with Bioruptor for 5 cycles (30 sec on 30 sec off). Lysates were centrifuged for 10 min at maximum speed and the supernatant was saved as a cell lysate. A T7 tag antibody agarose from Novagen was used (69026) to immunoprecipitate T7-DGCR8, whereas anti-Flag M2 affinity from Sigma (A220) was used to immunoprecipitate Flag-Drosha. Beads were washed 3 times with buffer D and sonicated lysates were added to bind overnight at 4°C. Next day, immunoprecipitated material was washed 5 times with buffer D (150 mM KCl). For *in vitro* processing reactions, beads were resuspended in 150µl of buffer D and kept at -80°C until use. For RNA extraction, beads were resuspended in 250µl of sodium acetate 0.3M and 750µl of Trizol LS (Invitrogen). RNA was purified following manufacturer's instructions. The RNA pellet was resuspended in water (20-25µl). As input, 1/20th part of the starting material was saved prior to immunoprecipitation and was resuspended in 20µl following RNA extraction. Prior to RT-PCR analysis, RNA samples were treated with RQ1 DNAse (Promega, M610A) for 1 h at 37°C, phenol/chloroform extracted and ethanol precipitated. For RT-PCR experiments, 1µl of input and immunoprecipitated RNA was used with SuperScript III One-Step RT-PCR with Platinum Taq (Invitrogen, 12574-026) and several cycles of amplification conditions were used (25-40) in order to obtain semi-quantitative results. For qRT-PCR SuperScript III Platinum SYBR Green One-Step qRT-PCR Kit was used (Invitrogen, 11736-051). qRT-PCR analysis of the immunoprecipitated material was normalized to the amount

of material in the input fraction. Primers used for this analysis are listed in Supplementary Table 2. RT-PCR products were analyzed in ethidium stained agarose gels.

Identification of snoRNA cleavage sites

Transcripts were 5' end-labeled with T4 polynucleotide kinase (Roche) and [γ-32P] ATP (4,500 Ci/mmol; Amersham Pharmacia Biotech). Transcripts were purified by electrophoresis in a 10% denaturing polyacrylamide gel and labeled RNAs were visualized by autoradiography. RNAs were incubated in buffer A in the presence of total extracts or immunopurified T7-DGCR8 (as described in the section Radioactive RNA labeling and *in vitro* processing**)**. To determine the cleavage sites, the products of the 5' ^{32}P-labeled RNA fragmentation reaction along with the products of alkaline hydrolysis and limited T1 nuclease digestion of the same RNA molecule were separated on 10% polyacrylamide gels containing 7.5 M urea, 90 mM Tris-borate buffer and 2 mM EDTA. The alkaline hydrolysis ladder was generated by incubation of labeled RNA in formamide containing 0.5 mM MgCl2 at 100°C for 10 min. Partial T1 ribonuclease digestion of RNAs was performed under semi-denaturing conditions (10 mM sodium citrate pH 5.0; 3.5 M urea) with 0.2 U/μl of the enzyme and incubation at 55°C for 15 min. Electrophoresis was performed at 1500 V and was followed by autoradiography at -80°C with an intensifying screen or exposed to a PhosphorImager screen (Molecular Dynamics)

UV Cross-linking

Uniformally labeled RNA probes were incubated with immunopurified T7-DGCR8 from cells in 20 μl volume reactions containing 10mM Tris-HCl pH 7.5, 50mM KCl

and 0.5 mM DTT and RNAse inhibitors, at 4°C for 30 min. After incubation, reactions were placed in a 96-well plate on ice and irradiated with 254 nm UV light for 5 min. One microliter of RNAse A/T1 cocktail was then added to each reaction and incubated at 37°C for 15 min to degrade unprotected RNA. Samples were loaded in 10% Tris-Glycine protein gels and after electrophoresis, stained with Colloidal Blue Staining kit (Invitrogen, LC6025) for loading purposes, and dried and exposed overnight on a phosphorimager screen to visualize cross-linked DGCR8 protein.

mES *DGCR8KO* rescue experiments

Mouse DGCR8 was amplified from total mouse RNA using oligos containing *EcoR*I and *Sal*I restriction sites. The cDNA from DGCR8 was cloned in the plasmid pEF1α-IRES-DsRed-Express2 vector (Clontech, Cat No. 631980). mES WT cells were transfected with pEF1α-IRES-DsRed-Express2 empty plasmid and *DGCR8KO* with the plasmid expressing mDGCR8 cDNA using Lipofectamine 2000 (Invitrogen). Six hours post-transfection the media was replaced with fresh media, and one day post-transfection G418 was added at 200 μg/ml for 5 days to select for transfected cells. After selection, RNA from transfected cells was extracted to proceed with the analysis.

Luciferase reporter assays

The assays were performed in three independent replicas. HeLa cells were first depleted of Drosha using siRNAs (as indicated) and 24 h later were co-transfected with Luciferase reporter plasmids and a second round of Drosha siRNA using Dharmafect DUO (Dharmacon). An empty plasmid (LUC) containing a TK promoter fused to Firefly luciferase based on pGL3-basic vector (a gift from James Reddington,

MRC HGU, Edinburgh) was used to clone a 500bp fragment from MALAT1 sequence upstream of the Firefly ORF using *Hind*III sites. A miRNA-reporter plasmid (SV40-miR18a) was also transfected to monitor Drosha knock-down efficiency[51]. In all assays, a plasmid expressing Renilla luciferase was co-transfected as an internal control. Twenty-four hours post-transfection cells were lysed and the levels of Firefly and Renilla luciferase activity were measured using Promega Dual Luciferase Reaction System. The data are expressed as a ratio of Firefly/Renilla luciferase activity and normalized to mock (siCtrl) value. Luminescence was measured using a Monolight 3010 luminometer (Pharmingen).

Identification of homologous mouse exons

The Ensembl 54 annotation was used to extract the human-mouse one-to-one orthologs and the mouse exons[52]. For each of the gene pairs, the best-reciprocal exon pairs were calculated by performing all against all local exon alignments with exonerate[53].

Identification of alternative exons

EST alignments to hg18 from UCSC were used for this purpose[54]. For each internal exon the inclusion level I was calculated as the proportion of ESTs verifying both splice sites (EST_i) over all ESTs including or excluding (EST_e) both splice sites:

$$I = \frac{EST_i}{EST_i + EST_e}$$

Only those exons with more than 10 ESTs ($EST_e + EST_i$) were considered. An exon was classified as CASSETTE if the inclusion was lower than 0.9 and as CONSTITUTIVE otherwise. For mouse, the EST alignments to mm9 from UCSC

were used[54]. The inclusion levels for mouse exons were calculated as above and only internal exons were considered.

49. Michlewski,G. & Caceres,J.F. Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nat. Struct. Mol. Biol.* **17**, 1011-1018 (2010).

50. Caceres,J.F., Misteli,T., Screaton,G.R., Spector,D.L., & Krainer,A.R. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.* **138**, 225-238 (1997).

51. Guil,S. & Caceres,J.F. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.* **14**, 591-596 (2007).

52. Flicek,P. *et al.* Ensembl's 10th year. *Nucleic Acids Res.* **38**, D557-D562 (2010).

53. Slater,G.S. & Birney,E. Automated generation of heuristics for biological sequence comparison. *BMC. Bioinformatics.* **6**, 31 (2005).

54. Fujita,P.A. *et al.* The UCSC Genome Browser database: update 2011. *Nucleic Acids Res.* **39**, D876-D882 (2011).

For the HITS-CLIP protocol and for the processing and mapping of sequencing reads, please see **Supplementary Note**.