

## **The 7SK snRNP associates with the little elongation complex to promote snRNA gene expression**

Sylvain Egloff<sup>1\*</sup>, Patrice Vitali<sup>1</sup>, Michael Tellier<sup>2</sup>, Raoul Raffel<sup>1,4</sup>, Shona Murphy<sup>2</sup> and Tamás Kiss<sup>1,3\*</sup>

<sup>1</sup>Laboratoire de Biologie Moléculaire Eucaryote, Centre de Biologie Intégrative (CBI), Université de Toulouse, CNRS, UPS, 118 route de Narbonne, 31062 Toulouse Cedex 9, France

<sup>2</sup>Sir William Dunn School of Pathology, University of Oxford, UK

<sup>3</sup>Biological Research Centre, Hungarian Academy of Sciences, Szeged, Hungary

<sup>4</sup>Current address: Institut de Génétique Humaine CNRS, Montpellier, France

\*Corresponding authors: Tamás Kiss

e-mail: [tamas@ibcg.biotoul.fr](mailto:tamas@ibcg.biotoul.fr)

phone: (33)561335907

fax: (33)561335886

Sylvain Egloff

e-mail: [Sylvain.Egloff@ibcg.biotoul.fr](mailto:Sylvain.Egloff@ibcg.biotoul.fr)

phone: (33)561335991

Running title: 7SK snRNP promotes snRNA synthesis

## Abstract

The 7SK small nuclear RNP (snRNP), composed of the 7SK small nuclear RNA (snRNA), MePCE and Larp7, regulates the mRNA elongation capacity of RNA polymerase II (RNAPII) through controlling the nuclear activity of positive transcription elongation factor b (P-TEFb). Here, we demonstrate that the human 7SK snRNP also functions as a canonical transcription factor that, in collaboration with the little elongation complex (LEC) comprising ELL, Ice1, Ice2 and ZC3H8, promotes transcription of RNAPII-specific spliceosomal snRNA and small nucleolar RNA (snoRNA) genes. The 7SK snRNA specifically associates with a fraction of RNAPII hyperphosphorylated at Ser5 and Ser7, which is a hallmark of RNAPII engaged in snRNA synthesis. Chromatin immunoprecipitation (ChIP) and chromatin isolation by RNA purification (ChIRP) experiments revealed enrichments for all components of the 7SK snRNP on RNAPII-specific sn/snoRNA genes. Depletion of 7SK snRNA or Larp7 disrupts LEC integrity, inhibits RNAPII recruitment to RNAPII-specific sn/snoRNA genes and reduces nascent snRNA and snoRNA synthesis. Thus, through controlling both mRNA elongation and sn/snoRNA synthesis, the 7SK snRNP is a key regulator of nuclear RNA production by RNAPII.

**Keywords** regulatory RNA; 7SK snRNA; 7SK snRNP; snRNA gene transcription; RNA polymerase II; little elongation complex

## Introduction

The human 7SK snRNA is an abundant RNA polymerase III (RNAPIII)-transcribed nucleoplasmic RNA (Diribarne & Bensaude, 2009). Together with Larp7 (La-related protein 7) and MePCE (methylphosphate capping enzyme) (He et al., 2008, Jeronimo et al., 2007, Krueger et al., 2008, Markert et al., 2008), it forms the 7SK snRNP that is a central regulator of RNAPII-mediated pre-mRNA elongation (Peterlin et al., 2011, Zhou et al., 2012). Shortly after transcription initiation, RNAPII is frequently arrested by negative elongation factors (NELFs). To convert promoter-proximally paused RNAPII into elongation-competent polymerase, the positive transcription elongation factor b (P-TEFb), a cyclin-dependent protein kinase comprising Cdk9 and cyclin T1 (CycT1) or CycT2, phosphorylates NELFs and the YSPTSPS heptapeptide repeat of the C-terminal

domain (CTD) of RNAPII at Ser2 (Peterlin & Price, 2006, Price, 2000, Smith & Shilatifard, 2013, Zhou et al., 2012). The nuclear level of active P-TEFb is controlled primarily by the 7SK snRNP (Nguyen et al., 2001, Yang et al., 2001). In collaboration with HEXIM (hexamethylene bisacetamide-inducible protein), the 7SK snRNP sequesters P-TEFb into the large kinase-inactive 7SK/HEXIM/P-TEFb RNP (Barboric et al., 2007, Michels et al., 2004). Dynamic association of 7SK snRNP with P-TEFb adjusts the level of active P-TEFb to the actual transcriptional condition of the cell. Upon transcriptional stimuli, the 7SK/HEXIM/P-TEFb negative transcriptional regulatory RNP disassembles and the free 7SK snRNP associates with a set of heterogeneous nuclear ribonucleoprotein (hnRNP) proteins, including hnRNP A1, A2/B1, R and Q (Barrandon et al., 2007, Van Herreweghe et al., 2007). The released active P-TEFb is recruited to the transcriptionally active chromatin by several factors including the super elongation complexes (SECs) (He et al., 2010, Lin et al., 2010, Mueller et al., 2009, Sobhian et al., 2010, Yokoyama et al., 2010). SECs promote suppression of promoter proximal RNAPII pausing and increase the catalytic rate of transcription elongation (Luo et al., 2012, Zhou et al., 2012). Mammalian SECs are related multiprotein complexes composed of different combinations of the four AF4/FRM2 family proteins (AFF1, AFF2, AFF3 or AFF4), the eleven-nineteen lysine-rich leukemia family RNAPII elongation factors (ELL1, ELL2 or ELL3), the ELL-associated factors (EAF1 or EAF2) and finally, the eleven-nineteen leukemia (ENL) protein or the highly similar ALL1-fused gene from chromosome 9 (AF9) protein.

RNAPII also participates in the synthesis of the major (U1, U2, U4 and U5) and minor (U11, U12 and U4atac) spliceosomal snRNAs, the U7 snRNA involved in histone mRNA 3'-end processing and the U3, U8 and U13 ribosomal processing snoRNAs (Hernandez, 1992). Transcription of RNAPII-specific sn/snoRNA genes is governed by three major regulatory elements, the distal sequence element (DSE) that functions as a classical transcriptional enhancer, the proximal sequence element (PSE) that coordinates transcription initiation through binding the snRNA activating protein complex (SNAPc, also called PSE-binding transcription factor, PTF) and finally, the 3'-box that directs co-transcriptional RNA 3'-end processing (Egloff et al., 2008, Henry et al., 1998, Hernandez, 2001, Jawdekar & Henry, 2008). RNAPII engaged in snRNA gene transcription shows elevated CTD phosphorylation at Ser5 and Ser7 and weak phosphorylation at Ser2 (Glover-Cutter et al., 2009). Transcription elongation on snRNA genes does not depend on P-TEFb activity and Ser2 phosphorylation (Medlin et al.,

2005). Instead, together with Ser7-P, Ser2-P coordinates the recruitment of the Integrator complex directing the 3'-end cleavage of nascent snRNA transcripts (Baillat et al., 2005, Egloff et al., 2007, Egloff et al., 2010).

Recently, it has been demonstrated that efficient sn/snoRNA synthesis by RNAPII requires an additional snRNA gene-specific transcription factor termed the little elongation complex (LEC) that shares some structural and functional properties with SECs (Hu et al., 2013, Smith et al., 2011). In addition to the LEC-specific Ice1 and Ice2 proteins, *Drosophila* LEC also contains the ELL1 and EAF1/2 subunits of SEC. Instead of EAF1/2, the human LEC associates with the ZC3H8 zinc finger protein that lacks an apparent *Drosophila* homologue. LEC has a dual function in snRNA gene expression; it facilitates RNAPII loading onto the promoters of snRNA genes through an Ice1-dependent mechanism and promotes the elongation capacity of RNAPII in an ELL-dependent manner (Hu et al., 2013).

In this study, we demonstrate that the human 7SK/MePCE/Larp7 snRNP is an essential component of the snRNA gene-specific transcription factor LEC. The 7SK snRNP is required for the integrity and correct subnuclear localization of LEC and it supports efficient loading of LEC and RNAPII onto sn/snoRNA genes. Thus, besides controlling the elongation phase of mRNA transcription, the 7SK snRNP also functions as a canonical transcription factor promoting synthesis of RNAPII-transcribed sn/snoRNAs.

## Results

### 7SK snRNA co-purifies with RNAPII phosphorylated at Ser5 and Ser7

To test whether the 7SK transcriptional regulatory snRNP can directly interact with RNAPII, HeLa RNAPII complexes were immunoprecipitated with antibodies directed against the N-terminal region or against the CTD of the largest subunit of RNAPII, Rpb1, phosphorylated at Ser2, Ser5 or Ser7. Immunoprecipitation (IP) of RNAPII by the four different antibodies was confirmed by western blot analysis (Fig 1, lower panel, lanes 3-6). RNAs co-precipitated with the four RNAPII samples were analyzed by northern blot hybridization with a mixture of oligonucleotide probes complementary to the 7SK, U1 and U2 snRNAs (upper panel). In the anti-Ser5-P and anti-Ser7-P IP reactions, the 7SK snRNA was highly enriched compared to the U1 and U2 control RNAs (lanes 5 and 6). Since the anti-Rpb1 antibody that reacted mostly with the unphosphorylated form of RNAPII (lower band) recovered 7SK snRNA with very low

efficiency (lane 3, and data not shown), we concluded that the human 7SK snRNP preferentially associates with a small fraction of RNAPII phosphorylated at Ser5 and Ser7.

### **7SK snRNP is enriched at RNAPII-specific small nuclear and nucleolar RNA genes**

Next, we attempted to determine the complement of human HeLa genes transcribed by the newly detected 7SK-associated fraction of RNAPII. We performed chromatin IP and high-throughput DNA sequencing (ChIP-seq) experiments using a polyclonal population of HeLa cells stably expressing a Flag-tagged version of the 7SK snRNP protein Larp7 (FL-Larp7) and an anti-Flag antibody, because commercially available Larp7 antisera performed poorly in ChIP analysis (our unpublished data). HeLa cells expressing FL-Larp7 showed no detectable global Larp7 overaccumulation, although the majority of the cells showed FL-Larp7 expression by immunostaining with an anti-Flag antibody (Fig EV1A, lane 4 and Fig EV1B). Larp7 was selected for our ChIP-seq analysis because the other 7SK core protein MePCE also interacts with the U6 spliceosomal snRNA (Muniz et al., 2013). We also determined the genome-wide distribution of RNAPII with the anti-Rpb1 antibody that recognizes the N-terminus of Rpb1. Mapping of the obtained ChIP-seq datasets to the human genome sequence (UCSC Genome Browser) identified a set of genomic sites (<100) with significant Larp7 enrichments. Plotting of Larp7 occupancy against RNAPII distribution revealed that the majority of genomic sites marked by solid Larp7 and RNAPII co-occupancy corresponded to RNAPII-transcribed spliceosomal snRNA (U1, U2, U4, U5 and U4atac) and ribosomal processing snoRNA (U3, U8 and U13) genes (Fig 2A, red dots). Apart from the obvious exception of the NEAT1 long non-coding RNA gene, the remaining Larp7-bound genomic sites which corresponded to known or predicted protein- or non-protein-coding transcription units showed no (data not shown) or low RNAPII occupancy (grey dots), as compared to the RNAPII-specific sn/snoRNA genes. Genes with significant Larp7 and RNAPII co-occupancy are listed in Table EV1.

As representative examples, Larp7 and RNAPII occupancies on a few snRNA (U1 and U2) and snoRNA (U3 and U13) genes as well as on the TTI2 and c-Myc protein-coding genes are shown in Figs 2B and EV5). The similar coverage profiles of RNAPII and Larp7 observed on sn/snoRNA genes support the idea that Larp7 interacts with RNAPII engaged in sn/snoRNA transcription. In contrast, Larp7 showed no

apparent co-occupancy with promoter-proximally paused RNAPII on the inducible TTI2 gene or with RNAPII distributed on the actively transcribed P-TEFb-dependent c-Myc gene.

To validate our Larp7 ChIP-seq data, we performed ChIP-quantitative PCR (ChIP-qPCR) experiments to directly assess the interaction of Larp7 with several selected HeLa snRNA, snoRNA and protein-coding genes (Fig 3A). All of the tested RNAPII-transcribed snRNA (U1, U2 and U4) and snoRNA (U3, U8) genes showed strong enrichments for Larp7, as compared to a randomly selected intergenic region that was used as negative control. In contrast, no significant Larp7 occupancy was measured at the promoters of the c-Myc, PCNA, GAPDH, c-Fos, Cyc D1 and TGM2 protein-coding genes or at the intron-encoded SNORA55 and SNORA74 snoRNA and the RNAPIII-transcribed U6 snRNA genes.

Next, we further interrogated RNAPII, Larp7 and MePCE occupancy at various regions (R1 to R5) of the human U2 snRNA gene by ChIP-qPCR (Fig 3B). Due to the lack of MePCE antibody competent for ChIP, polyclonal HeLa cell populations stably expressing flag-tagged MePCE (FL-MePCE) and Larp7 (FL-Larp7) were used. The R2 promoter region of the U2 gene was highly enriched for RNAPII, FL-Larp7, FL-MePCE as well as for ELL, an essential component of the snRNA gene-specific transcription factor LEC (Hu et al., 2013, Smith et al., 2011). DNA fragments immediately following the U2 coding region (R3) or situated further downstream (R4) showed decreasing RNAPII, FL-Larp7, FL-MePCE and ELL occupancy. None of the tested proteins were enriched at the R1 and R5 intergenic control regions located 1422 bp upstream and 737 bp downstream of the U2 RNA encoding region. In contrast to the FL-Larp7 and FL-MePCE 7SK core proteins, we failed to detect FL-HEXIM1 enrichment on the U2 gene, suggesting that the 7SK core snRNP associates with the U2 snRNA gene.

To demonstrate that the 7SK snRNA also associates with RNAPII-transcribed snRNA genes, we performed ChIRP (chromatin isolation by RNA purification)-qPCR experiments (Chu et al., 2011). After *in vivo* cross-linking and chromatin fragmentation, the 7SK and as a negative control, the U1 snRNA were affinity-selected with biotinylated oligonucleotides complementary to three internal regions of the 7SK (7SK-1, 7SK-2 and 7SK-3) and to the 5'-terminal region of U1 snRNA. Northern blot analysis demonstrated that the 7SK and U1 target RNAs were recovered with great efficiency and specificity (Fig EV2A). Analysis of the co-purified genomic DNAs by qPCR detected no significant U1 snRNA occupancy on the U2, 7SL and hnRNP H1 genes (Fig 3C).

However, the 7SK snRNA showed increased association with the U2 gene in the three independent ChIRP-qPCR experiments [t-statistics 6.65 (7SK-1), 4.68 (7SK-2) and 4.58 (7SK-3) with p-values of  $\leq 0.01$ ]. We failed to detect significant 7SK snRNA association with the 7SL and hnRNP H1 control genes. We also measured an increased 7SK occupancy on the U1 snRNA gene that also showed an extremely strong association with U1 snRNA sequences, likely representing nascent U1 transcripts (Fig 2B).

We concluded that all components of the 7SK core snRNP, including Larp7, MePCE and the 7SK snRNA, accumulate at the U2 and probably other RNAPII-specific snRNA genes (see Discussion). These observations, together with the finding that 7SK snRNA associates with RNAPII phosphorylated at Ser5 and Ser7 and non-phosphorylated at Ser2, a CTD phosphorylation pattern similar to that observed for RNAPII engaged in snRNA synthesis (Glover-Cutter et al., 2009), strongly suggested that the human 7SK snRNP participates in sn/snoRNA gene expression.

### **Depletion of 7SK snRNP inhibits transcription of RNAPII-specific snRNA genes**

To test whether the 7SK snRNP is required for expression of RNAPII-transcribed sn/snoRNA genes, Larp7 was depleted by RNA interference (RNAi). Transfection of HeLa cells with a Larp7-specific small interfering RNA (siRNA) knocked down Larp7 expression about 90% as compared to cells treated with control siRNA (Fig 4A). As expected, depletion of Larp7 largely diminished the level of 7SK snRNA, but had no detectable effect on the accumulation of HEXIM1, Cdk9, CycT1 and hnRNP A1 (He et al., 2008, Krueger et al., 2008, Markert et al., 2008). Accumulation of several RNAPII- and RNAPIII-synthesized small RNAs as well as a few mRNAs was measured in the Larp7-depleted cells by northern blot analysis and/or by quantitative RT-PCR (Fig 4A,B). Depletion of 7SK snRNP largely reduced accumulation of the RNAPII-synthesized U1, U2, U4 and U5 spliceosomal snRNAs and the U3 and U8 snoRNAs. Larp7 depletion also inhibited accumulation of the newly-synthesized precursor (pre-)U2 snRNA that carries a short 3'-terminal tail. This indicated that the 7SK snRNP promotes snRNP biogenesis at the level of snRNA synthesis, rather than facilitating pre-snRNA maturation and/or snRNP assembly. In contrast, Larp7 depletion had no significant effect on the accumulation of the RNAPIII-transcribed signal recognition particle RNA (7SL RNA), the U6 spliceosomal snRNA and the 5S rRNA. Likewise, expression of the tested GAPDH,  $\beta$ -actin, c-Myc, c-Fos, CD44 and Gadd45b mRNAs or the U19 and U64 snoRNAs which are processed from pre-mRNA introns showed no significant alteration

upon Larp7 and 7SK snRNA depletion, indicating that the nuclear P-TEFb activity is not limiting for expression of protein-coding genes in HeLa cells.

The facts that Larp7 is required for 7SK snRNA stability (Fig 4B) and that all components of the 7SK core snRNP associate with the U2 snRNA gene (Fig 3) strongly suggested that Larp7 promotes sn/snoRNA synthesis as an integral component of the 7SK snRNP. To unambiguously verify this conclusion, we also investigated the effect of targeted 7SK snRNA depletion on sn/snoRNA expression (Fig 4C). Suppression of 7SK snRNA accumulation with a 7SK-specific siRNA greatly inhibited accumulation of the RNAPII-transcribed U1, U2 and U3 RNAs, but failed to compromise Larp7 and HEXIM1 accumulation, excluding the formal possibility that Larp7 promotes snRNA synthesis in a 7SK-independent fashion. As expected, 7SK depletion did not alter the level of the RNAPIII-synthesized U6 snRNA.

To corroborate the conclusion that 7SK snRNP promotes the synthesis of RNAPII-transcribed snRNAs, instead of facilitating snRNP assembly, we performed nuclear run-on assays to compare the rates of nascent snRNA synthesis in Larp7-depleted and control HeLa cells (Fig 4D). Probing of antisense U1, U2, U6, 5S and 7SL oligonucleotide blots with internally-labeled nuclear run-on test RNAs showed that Larp7 depletion inhibited the synthesis of nascent U1 and U2 snRNAs by about 50%, but had no significant effect on the production of the RNAPIII-specific U6, 5S and 7SL RNAs (for quantitation, see Fig EV3). Immobilized sense U1 and U2 control oligonucleotides failed to react with the nuclear run-on test RNAs. As reported before (Hu et al., 2013, Smith et al., 2011), depletion of the Ice1 subunit of the snRNA gene-specific transcription factor LEC also inhibited U1 and U2 snRNA transcription by about 50%. Thus, the 7SK snRNP, like the previously described LEC, is required for efficient transcription of RNAPII-specific sn/snoRNA genes.

### **The 7SK snRNP associates with the little elongation complex**

Efficient expression of RNAPII-transcribed snRNA genes requires at least three snRNA gene-specific multiprotein transcription factors, the snRNA activating protein complex (SNAPc), the integrator complex and the little elongation complex (LEC) (Egloff et al., 2008, Henry et al., 1998, Hernandez, 2001, Jawdekar & Henry, 2008). To test whether the 7SK snRNP promotes snRNA transcription through cooperation with any of these factors, HeLa LEC, SNAPc and integrator complex were immunoprecipitated and co-purification of 7SK snRNA was monitored by northern blot analysis (Fig 5). IP of the



Ice1 and Ice2 subunits of human LEC with specific antibodies recovered the 7SK snRNA, but failed to pull down the U2 and U6 snRNAs (Fig 5A, lanes 4 and 5). Ice1 and Ice2 co-purified with each other in both IP reactions, indicating that they were integrated into LEC. IP of two other components of human LEC, ELL and ZC3H8 also recovered 7SK RNA, demonstrating that the human 7SK snRNA interacts with LEC (Fig 5B). In contrast, no detectable 7SK RNA was recovered upon IP of SNAPc with an antibody directed against its SNAPc2 subunit (Fig 5A, lane 3). Finally, purification of the integrator complex by IP of transiently-expressed Flag-tagged integrator subunits Int7 and Int11 failed to recover 7SK RNA (Fig 5C, lane 2 and 3), although control IP of transiently expressed Flag-tagged HEXIM1 pulled down 7SK snRNA (lane 4). We concluded that the 7SK snRNA interacts with the snRNA gene-specific transcription factor LEC.

To demonstrate that 7SK snRNA interacts with LEC in the form of 7SK snRNP, the 7SK core protein Larp7 was immunoprecipitated from a HeLa cell extract (Fig 5D). Western blot analysis of the recovered proteins confirmed that in addition to the other 7SK core protein MePCE, the ELL and Ice2 LEC proteins also co-purified with Larp7, while GAPDH, an unrelated control protein, was not detectable in the IP pellet (lane 3). Likewise, IP of Larp7 from a HeLa extract prepared from cells expressing Flag-tagged ZC3H8 LEC protein efficiently recovered FL-ZC3H8, demonstrating that the 7SK snRNP rather than naked 7SK RNA interacts with LEC (Fig 5E, lane 3).

Besides existing in the form of free snRNP, a fraction of the human 7SK/MePCE/Larp7 snRNP is sequestered into large 7SK/HEXIM/P-TEFb and 7SK/hnRNP particles (see Introduction). Our results presented thus far are most consistent with the idea that the 7SK core snRNP, rather than any of the large 7SK complexes, promotes snRNA gene expression. To confirm this assumption, HeLa LEC was immunoprecipitated with an anti-Ice2 antibody and the co-purified proteins were analyzed by western blotting (Fig 5F). As expected, IP of Ice2 pulled down the ELL component of LEC as well as the 7SK snRNA and the two 7SK core proteins Larp7 and MePCE (lane 3). In contrast, the major 7SK cofactor directing P-TEFb recruitment and inactivation, HEXIM1 (Barboric et al., 2007, Michels et al., 2004), failed to co-purify with the 7SK/LEC RNP. Likewise, hnRNP A1, the most prominent component of 7SK/hnRNP particles (Barrandon et al., 2007, Van Herreweghe et al., 2007), also failed to co-precipitate with the immunopurified 7SK/LEC RNP, corroborating the idea that the 7SK core snRNP interacts with LEC to promote RNAPII-mediated snRNA synthesis.

Transcription inhibition provokes rapid disassembly of 7SK/HEXIM/P-TEFb to increase the nuclear level of active P-TEFb in order to alleviate RNAPII transcription blockade (Nguyen et al., 2001, Yang et al., 2001). Consistent with this, ActD treatment of HeLa cells largely abolished HEXIM1 interaction with 7SK snRNA, but at the same time, it significantly increased the Ice1, ZC3H8 and ELL association with 7SK, suggesting that at least a fraction of 7SK snRNP released from 7SK/HEXIM/P-TEFb can be incorporated into 7SK/LEC RNP (Fig 5G). We propose that RNAPII inhibition increases the levels of active P-TEFb and 7SK/LEC to promote both pre-mRNA and sn/snoRNA synthesis.

### **The 7SK snRNA is required for LEC integrity and recruitment to snRNA genes**

To test whether integrity of the LEC protein complex is RNA-dependent, extracts prepared from HeLa cells transiently expressing either FL-Ice2 or FL-ZC3H8 were subjected to RNase A treatment before IP of LEC with an anti-ELL antibody (Fig 6A). As demonstrated by western blot analysis of the precipitated proteins, the RNase treatment largely reduced ELL association with both FL-ZC3H8 and FL-Ice2 (lanes 9 and 12). From the untreated control extracts, both FL-ZC3H8 and FL-Ice2 co-purified with ELL (lanes 8 and 11), indicating that an RNA component is required for the structural integrity of human LEC.

To confirm that the 7SK snRNA supports LEC stability, we investigated the integrity of LEC in 7SK-depleted HeLa cells. Transfection of a 7SK-specific interfering RNA largely diminished 7SK snRNA accumulation, but had no effect on the level of 7SL RNA (Fig 6B, left panel). IP of ELL from the 7SK-depleted extract failed to recover ZC3H8 (lane 6), although ZC3H8 co-purified with ELL immunoprecipitated from the control cell extract (lane 5). Besides demonstrating that Larp7 and MePCE are unable to support *in vivo* LEC stability in the absence of 7SK snRNA, these results also suggest that the 7SK snRNP plays a central role in structural organization of the 7SK/LEC RNP.

Depletion of 7SK snRNP had no effect on the accumulation of LEC proteins including ELL, ZC3H8 and Ice1 that serves as the central scaffold for LEC assembly (Fig 6B and Fig EV4). This suggests that the compromised snRNA synthesis observed in 7SK-depleted HeLa cells was due to reduced LEC recruitment to snRNA genes. To test this idea, we investigated the ELL occupancy at the R2 promoter region of the U2 snRNA gene (see Fig 3B) in 7SK snRNP-depleted HeLa cells by ChIP-qPCR (Fig 6C).

As compared to control cells, Larp7 depletion by RNAi highly reduced ELL occupancy at the U2 promoter. A less robust, but significant reduction was observed for RNAPII occupancy at the U2 promoter in the absence of 7SK snRNP. To test whether in the absence of 7SK snRNP, compromised polymerase occupancy is a common feature of RNAPII-specific snRNA genes, we compared RNAPII association with the promoter regions of the U1 and U2 snRNA and the U3 snoRNA genes in Larp7-depleted and control cells by ChIP-qPCR (Fig 6D). Disruption of 7SK snRNP decreased RNAPII occupancy on the U1, U2 and U3 genes by similar extents, but failed to inhibit RNAPII recruitment to the promoter region of the GAPDH gene. These results demonstrate that the 7SK snRNP, as a component of the 7SK/LEC transcription factor, stimulates sn/snoRNA synthesis through supporting RNAPII loading onto the promoters of RNAPII-specific sn/snoRNA genes.

### **The 7SK snRNP is required for localization of LEC components to the Cajal body**

Cajal bodies (CBs) are distinct subnuclear domains where assembly and maturation of large multiprotein and RNP complexes take place (Machyna et al., 2013). HeLa LEC proteins frequently concentrate within CBs in a transcription-dependent manner, raising the possibility that CBs provide the nuclear locale for the assembly of 7SK/LEC RNP (Hu et al., 2013, Polak et al., 2003, Smith et al., 2011). Previous fluorescent *in situ* hybridization (FISH) experiments revealed a diffuse punctuate nucleoplasmic localization pattern for HeLa 7SK snRNA (Egloff et al., 2006, Prasanth et al., 2010) (Fig 7A, upper panel). However, due to the high concentration of 7SK snRNA in the nucleoplasm, FISH failed to confirm 7SK accumulation within CBs. Therefore, to visualize 7SK snRNA associated with CBs, HeLa cells were pre-extracted with detergent to partially remove the soluble fraction of nucleoplasmic 7SK before cell fixation. Co-staining of the pre-extracted cells with a fluorescently labeled 7SK-specific oligonucleotide probe and an antibody against the CB marker protein coilin revealed a small, but significant fraction of 7SK snRNA that co-localized with coilin in CBs (lower panel, indicated by arrowheads).

After having been demonstrated that 7SK snRNA is present in CBs, we tested the importance of 7SK snRNP for CB-specific accumulation of LEC. Consistent with previous reports, immunofluorescence microscopy performed with Ice2- and coilin-specific antibodies detected Ice2 co-localization with coilin in the CBs of a fraction of HeLa cells (Fig 7B, upper panel). Upon closer inspection of 1274 randomly selected

HeLa cells derived from three independent experiments, we detected CB-specific Ice2 accumulation in 6.44% of the cells (Fig 7C). However, depletion of 7SK snRNA abolished Ice2 accumulation in CBs (Figs 7B and 7C). Conscientious examination of 1278 cells derived from three independent 7SK knock-down experiments identified only 3 cells (0.2%) with detectable Ice2 accumulation in CBs. The same results were obtained upon investigation of the CB-specific localization of ZC3H8 in wild-type and 7SK snRNA-depleted HeLa cells (data not shown). Likewise, RNAi-mediated depletion of the 7SK core proteins Larp7 and MePCE largely inhibited Ice2 accumulation in CBs (Fig EV4B). Thus, we concluded that the 7SK snRNP is required for localization of LEC into CBs.

## Discussion

Expression of mammalian RNAPII-specific spliceosomal snRNA and snoRNA genes is supported by the little elongation complex (LEC) that promotes both transcription initiation and elongation on snRNA genes (Smith & Shilatfard, 2013, Smith et al., 2011). Here, we have demonstrated that the human 7SK snRNP, composed of the 7SK snRNA, Larp7 and MePCE, specifically associates with LEC to promote RNAPII-mediated snRNA and snoRNA synthesis. As an essential component of the 7SK/LEC sn/snoRNA transcriptional stimulatory RNP, the 7SK snRNP functions as a canonical RNAPII transcription factor. Earlier, the U1 snRNP has been reported to interact with the general RNAPII initiation factor TFIIF to stimulate transcription initiation, indicating that snRNPs may function in various steps of the RNAPII transcription cycle (Kwek et al., 2002).

Instead of serving as an auxiliary cofactor of LEC, the 7SK snRNP seems to form a stable complex with the four LEC proteins, ELL, Ice1, Ice2 and ZC3H8. 7SK is an abundant snRNA that is present in about 200,000 copies in mammalian cells (Diribarne & Bensaude, 2009). Our IP experiments detected only a small portion (less than 2-3%) of cellular 7SK associated with LEC, indicating that the nuclei of HeLa cells contain maximum 5,000-6,000 7SK/LEC particles. The structural organization of the 7SK/LEC particle remains unknown, but the fact that the 7SK snRNA is required for LEC integrity suggests that the 7SK snRNP provides the structural scaffold for the assembly of the 7SK/LEC RNP. The 7SK snRNP and the four LEC proteins accumulate within CBs raising the possibility that assembly and/or recycling of the 7SK/LEC particles takes place in these nucleoplasmic domains (Hu et al., 2013, Polak et al., 2003, Smith et al., 2011). The major function of CBs is in promoting assembly of macromolecular

complexes involved in transcription, splicing, ribosome biogenesis and telomere synthesis (Hu et al., 2013, Smith et al., 2011). The Ice2 and ZC3H8 LEC proteins accumulate in CBs in a 7SK snRNP-dependent manner supporting the notion that 7SK snRNP plays a central role in the assembly and intracellular trafficking of the 7SK/LEC RNP. CBs are highly mobile subnuclear organelles which are frequently found in the proximity of actively-transcribed snRNA genes, raising the possibility that CBs might also deliver the 7SK/LEC transcription factor to snRNA genes (Ogg & Lamond, 2002). However, further efforts are required to understand the structural organization, biogenesis and trafficking of the 7SK/LEC RNP. It is also possible that in addition to the 7SK snRNP and the four LEC proteins, transcriptionally fully active 7SK/LEC RNP contains other, yet unidentified components.

Similar to the RNAPII-specific sno/snRNA genes, the NEAT1 long non-coding RNA gene showed strong Larp7 and RNAPII co-occupancy. The NEAT1 gene encodes a 4 kb unspliced polyadenylated non-coding RNA that is required for paraspeckle integrity (Clemson et al., 2009). Our preliminary experiments have suggested that Larp7 is required for efficient NEAT1 RNA synthesis, although in contrast to the RNAPII-specific snRNA genes, expression of the NEAT1 gene is governed by a classical TATA box-containing promoter (our unpublished data). Our genome-wide ChIP analysis also revealed high Larp7 occupancy on a set of known or predicted protein-coding and long non-coding RNA genes which showed very low or no apparent RNAPII occupancy (data not shown). These observations raise the possibility that the 7SK snRNP and/or Larp7 play a complex role in controlling expression of various classes of RNAPII-transcribed genes.

Our ChIP-seq analysis failed to detect significant Larp7 accumulation on the overwhelming majority of protein-coding genes. A recent study, however, reported widespread promoter occupancy of Larp7 on protein-coding genes and in contrast to our results, did not detect Larp7 enrichments on sn/snoRNA genes (McNamara et al., 2016). The reason for these contradictory observations remains uncertain. It is possible that the ectopically expressed FL-Larp7 is competed out by the excess of endogenous Larp7, reducing FL-Larp7 signals on protein-coding genes (Fig EV1). On the other hand, the anti-Larp7 epitope of endogenous Larp7 might be masked in the RNAPII-7SK/LEC complex, impeding 7SK snRNP detection on sn/snoRNA genes. Nevertheless, to confirm our conclusion that 7SK/LEC specifically associates with RNAPII-specific sn/snoRNA genes, we reanalyzed the recently published 7SK ChIRP-

seq and Ice1, Ice2 and ZC3H8 ChIP-seq data sets and compared them with our FL-Larp7 and RNAPII ChIP-seq results (Flynn et al., 2016, Hu et al., 2013). We found that all components of the 7SK/LEC RNP analyzed so far, including the 7SK snRNA, Larp7, Ice1, Ice2 and ZC3H8 preferentially associate with RNAPII-specific sn/snoRNA genes (Fig EV5A-EV5C).

Promoting the release of promoter-proximally paused RNAPII by P-TEFb is a prominent regulatory step of human gene expression (Adelman & Lis, 2012, Smith & Shilatifard, 2013, Zhou et al., 2012). The nuclear level of active P-TEFb and consequently, the global level of pre-mRNA production are controlled primarily by the 7SK snRNP that, together with HEXIM1/2, sequesters P-TEFb into the transcriptionally inactive 7SK/HEXIM/P-TEFb RNP. Demonstration that the 7SK snRNP also functions in snRNA gene expression implies that the 7SK/HEXIM/P-TEFb RNP is a bifunctional negative transcriptional regulatory particle that inhibits both pre-mRNA and snRNA synthesis. In the 7SK/HEXIM/P-TEFb complex, the 7SK snRNP and P-TEFb, with the help of HEXIM1/2, mutually inhibit the transcriptional stimulatory capacity of each other. Consistent with this idea, disassembly of the 7SK/HEXIM/P-TEFb RNP in response to transcription inhibition by ActD increased the level of active P-TEFb and at the same time augmented 7SK/LEC accumulation to alleviate the blockade of both pre-mRNA and spliceosomal snRNA synthesis (Fig 5G). Thus, we propose that sequestering P-TEFb and 7SK snRNP into the same inhibitory complex provides an elegant regulatory mechanism to co-ordinate cellular pre-mRNA synthesis with spliceosomal snRNP production.

Larp7 provides metabolic stability for the 7SK snRNA and directly participates in P-TEFb recruitment and inactivation (He et al., 2008, Krueger et al., 2008, Muniz et al., 2013). Larp7 depletion increases nuclear P-TEFb activity, promotes cell growth, transformation and cancer progression (Biewenga et al., 2008, Cheng et al., 2012, He et al., 2006, Ji et al., 2014). Primordial dwarfism (PD) is a rare human genetic disorder characterized with global growth failure, multiple developmental defects and finally, death before the age of three years (Alazami et al., 2012). Mutations causing primordial dwarfism have been identified in the U4atac minor spliceosomal snRNA gene, indicating that compromised minor spliceosomal activity is responsible for the disease (Edery et al., 2011, He et al., 2011). Paradoxically, recent studies linked a form of primordial dwarfism to loss-of-function mutations in the human *LARP7* gene, contrary to the fact that Larp7 depletion is expected to promote cell growth through disrupting 7SK snRNP

and thereby, increasing P-TEFb activity (Alazami et al., 2012, Najmabadi et al., 2011). In the light of our current results, however, we can propose that the lack of functional Larp7 in primordial dwarfism patients compromises the synthesis of RNAPII-transcribed spliceosomal snRNAs. Since the minor spliceosomal snRNAs accumulate at low levels, it is conceivable that proper function of the minor splicing machinery is highly sensitive to snRNA production and compromised snRNA accumulation may lead to splicing defects similar to those caused by U4atac mutations.

In summary, demonstration that the human 7SK snRNP is an essential component of the 7SK/LEC snRNA gene-specific transcription factor provides support to the recently emerging view that snRNPs, instead of participating in a single nuclear reaction, frequently possess multiple nuclear roles. Identification of 7SK snRNP as a snRNA gene-specific transcription factor has further expanded the complex functional repertoire of eukaryotic snRNPs.

## **Materials and Methods**

### **General procedures, oligonucleotides**

Standard laboratory protocols were used for manipulating RNA, DNA and proteins. Unless indicated differently, all DNA modifying enzymes were purchased from Promega. All oligodeoxynucleotides used in this study were purchased from Eurofins MWG Operon and they are either specified in this chapter or listed in Table EV2. Human HeLa cells were grown in DMEM media supplemented with 10% (v/v) fetal bovine serum and 100 µg/ml penicillin and streptomycin (GIBCO, Invitrogen).

### **Expression plasmids**

To generate pFL-Int7 (utilized restriction sites, *EcoRI/XbaI*), pFL-Int11 (*HindIII/EcoRI*), pFL-ZC3H8 (*HindIII/EcoRI*), pFL-Ice2 (*NotI/XbaI*) and pFL-Hexim1 (*HindIII/XbaI*) expression plasmids, the corresponding full-length cDNAs were amplified by RT-PCR and inserted into the p3XFLAG-CMV-10 expression vector (Sigma). The identity of all constructs was verified by sequence analysis. Expression plasmids were introduced into HeLa cells using the Lipofectamine 2000 (Invitrogen) transfection reagent according to the manufacturer's instructions. For ChIP-seq and ChIP-qPCR analysis, HeLa cells were transfected with pFL-Larp7, pFL-MePCE or pFL-HEXIM1 expression plasmids and the transfected cells were selected by administration of 1 mg/ml of G418 48 hours after transfection to generate polyclonal cell populations stably expressing FL-Larp7, FL-

MePCE or FL-HEXIM1. Accumulation of FL-MePCE, FL-Larp7 or FL-HEXIM1 was verified by Western blot analysis and immunofluorescence microscopy with monoclonal anti-Flag M2 antibody conjugated with anti-mouse IgG-FITC (Sigma-Aldrich, F1804 and F9237, respectively) (Fig EV1).

### **siRNA-mediated knock down**

Control siRNA (D-001210-01) and siRNAs targeting Larp7 (M-020996-01) and Ice1 (M-024272-01) mRNAs were purchased from Dharmacon and transfected into HeLa cells with Lipofectamine 2000 (Invitrogen). 7SK-specific siRNA complementary to the A230-U250 region of human 7SK snRNA was purchased from Euromedex.

### **Chromatin immunoprecipitation and deep sequencing**

Formaldehyde cross-linking and IP of chromatin was performed essentially as described by (Boyd et al., 1998). Briefly, HeLa cells were cross-linked with 1% formaldehyde for 10 min at RT and quenched for 5 min by adding glycine to 125 mM final concentration. Cells were resuspended in Cell Lysis Buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% Nonidet P40 supplemented with protease inhibitor cocktail, Roche) and incubated on ice for 10 min. After collection by centrifugation at 5000 rpm for 5 min, nuclei were lysed in Nuclei Lysis Buffer (50 mM Tris-HCl pH 8.1, 10 mM EDTA, 1% SDS) for 20 min. Chromatin was disrupted with Bioruptor Plus (Diagenode) sonifier (high power setting, 20 cycles for 30 s on and 30 s off) on ice. The extract was cleared by centrifugation and incubated with 2 µg of appropriate antibody overnight at 4°C. Immunocomplexes were immobilized by reacting with 10 µl of BSA-saturated Pansorbin cells (Calbiochem) for 15 min at 4°C. IP of RNAPII was performed with, or as a control without, anti-PolIII (sc-899X, Santa Cruz) antibody. To recover FL-Larp7, the anti-Flag M2 (F1804, Sigma) antibody was used in combination with a rabbit anti-mouse IgG (Jackson ImmunoResearch). In the control reaction, the anti-Flag antibody was omitted. After extensive and stringent washing and elution, protein-DNA cross-links were reversed by incubation at 67°C for 5 hours and DNA was deproteinized with proteinase K treatment at 50°C for 30 min. DNA was purified by the Nucleospin PCR clean-up kit of Macherey-Nagel. ChIP samples were directly analyzed by qPCR or by deep sequencing using Illumina HiSeq 2500 (Wellcome Trust Centre for Human Genetics, University of Oxford). Sequences were mapped to the Hg18 human genome build using the Burrows-Wheeler Aligner (BWA) software and the Larp7-enriched regions of ChIP-seq signal were



analyzed by the Model-based Analysis of ChIP-seq 2 (MACS2) program (Li & Durbin, 2009, Zhang et al., 2008) (Fig. 2B). Larp7 peaks were annotated using the Ensembl project ([http://ftp.ensembl.org/pub/release-75/gtf/homo\\_sapiens](http://ftp.ensembl.org/pub/release-75/gtf/homo_sapiens)). RNAPII co-occupancy at Larp7-enriched regions was measured by calculating the numbers of extended read fragments per million (rpm) within +/- 100 bp of the Larp7 peaks with a cut-off 50 rpm for the RNAPII signal to select for the most significant Larp7/RNAPII-enriched regions.

### **Chromatin isolation by RNA purification (ChIRP) experiments**

ChIRP experiments were performed essentially as described earlier (Chu et al., 2012). Briefly, HeLa cells (~20 million) grown in monolayer cultures were harvested and crosslinked in PBS containing 1% of glutaraldehyde for 10 min at RT with continuous gentle agitation. The reaction was stopped by adding glycine to 125 mM final concentration. Cells were collected, washed with PBS, resuspended in Lysis Buffer (50 mM Tris-Cl pH 7.0, 10 mM EDTA, 1% SDS, protease inhibitors), and sonicated for 2 hours in a 4°C water bath (Bioruptor Plus, Diagenode, 30 sec pulses with 30 sec intervals at high setting). Hybridization of 7SK- and U1-specific and control biotinylated oligonucleotides (~100 pmol) to cellular RNAs was performed at RT for 4 hours in 50 mM Tris-Cl pH 7.0, 750 mM NaCl, 1% SDS, 1 mM EDTA, 15% formamide supplemented with protease inhibitors. Biotinylated oligonucleotides were affinity-selected by streptavidin-agarose beads (S1638, Sigma) for 30 min at RT. Beads were extensively washed in 2x SSC, 0.5% SDS supplemented with protease inhibitors and the bound chromatin was eluted in DNA Elution Buffer (50 mM NaHCO<sub>3</sub>, 1% SDS). DNA was deproteinized with proteinase K treatment and phenol extraction and analyzed by quantitative PCR.

### **Nuclear run-on assay**

Nuclear run-on analysis was performed according to (Medlin et al., 2003). For evaluation of U1 and U2 synthesis, two pairs of 80-mer oligonucleotides complementary to the U5-G164 and C28-C187 regions of U1 and U2 snRNAs, respectively, were used. For analysis of U6, 5S and 7SL RNA transcription, we used oligonucleotides complementary to the A27-U106, C9-A88 and C63-U142 sequences of the U6, 5S and 7SL RNAs, respectively. Control oligonucleotides (U1 sense and U2 sense) corresponded to the U1 and U2 snRNA sequences between U5-G164 and C28-C187, respectively.

### Immunoprecipitation and protein analysis

HeLa cells grown in monolayer cultures were rinsed with phosphate-buffered saline (PBS), scraped and collected by centrifugation. Cells were placed on ice, resuspended in 0.5 ml of cold NET-2 buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl and 0.1% Nonidet P40) and disrupted with a Bioruptor Plus Sonicator (Diagenode) by sonication five times for 30 s with 30 s intervals at high setting. Cell debris were removed by sedimentation at 16 000 x g for 10 min and the clarified cell extracts were incubated with 2 µg of antibodies coupled to 20 µl of packed protein A agarose beads (Sigma). For IP of Flag-tagged proteins, 20 µl of anti-Flag M2 Affinity Gel (A2220, Sigma) was used. After immunoprecipitation, beads were washed six times in NET-2 buffer and the associated proteins were eluted by 1 x SDS-PAGE loading buffer. For IP, western blot analysis and immunofluorescence microscopy, the following antibodies were used: anti-SNAPC2 (Yoon et al., 1995), anti-Ice1, anti-Ice2 and anti-ELL (provided by Dr. A. Shilatifard), anti-Flag M2 (F1804, Sigma), anti-Ser2P (3E10, ChromoTek), anti-Ser5P (3E8, ChromoTek), anti-Ser7P (4E12, ChromoTek), anti-PolIII (sc-899X, Santa Cruz), anti-ZC3H8 (ab97821, Abcam), anti-HEXIM1 (ab25388, Abcam), anti-Larp7 (ARP40847, Aviva System Biology), anti-MePCE (ab185991, Abcam) and anti-coilin (kind gift of Dr. M. Carmo-Fonseca). Unless indicated differently by the manufacturers, antibodies were diluted to 1:1000 for Western blot analyses. As secondary antibodies, horseradish peroxidase-conjugated anti-rabbit and anti-mouse IgGs (W401B and W402B, Promega) were used.

### RNA analysis

RNAs from HeLa cells, cell extracts and protein A agarose beads were purified by the guanidine thiocyanate/phenol-chloroform RNA extraction procedure. For northern blot analysis, RNAs were size-fractionated on a 6% sequencing gel, electroblotted onto a Hybond-N nylon membrane and probed with labelled oligonucleotides. Real-time quantitative RT-PCR was performed on 2 µg of total RNA isolated from wild-type or Larp7-depleted HeLa cells. DNA first strands were synthesized by AMV reverse transcriptase (Promega) and qPCR reactions were performed on a Bio-Rad C1000 *Touch thermal cycler* using the iQ SYBR Green Supermix of Bio-Rad. The delta-delta Ct method was used for relative quantification of RNA accumulation. The oligonucleotides used for northern blot analysis and RT-qPCR are listed in Supplementary Material.

## Fluorescence in situ hybridization

Fluorescence *in situ* hybridization, image acquisition and processing have been described (Darzacq et al., 2002). Removal of the nucleoplasmic soluble fraction of 7SK snRNA from living HeLa cells and fixation of the cells was performed according to the protocols of the Lawrence laboratory (<http://www.umassmed.edu/cellbio/labs/lawrence/protocols/>). The following oligonucleotide probe was used to detect HeLa 7SK snRNA: (5'-G\*TGTCTGGAGTCTTGAAGC\*-3'). The aminoallyl-nucleotides labelled with FluoroLink Cy3 monofunctional reactive dye (GE Healthcare) are indicated by asterisks. Nuclear DNA was visualized with 1 µg/ml DAPI. Coilin and Ice2 were detected with polyclonal mouse anti-coilin and rabbit anti-Ice2 antibodies provided by Drs M. Carmo Fonseca and A. Shilatifard, respectively.

## Data availability

FL-Larp7 and RNAPII ChIP-seq datasets have been deposited into the GEO genomics data repository (accession number: GSE87470).

## Acknowledgements

We are indebted to Dr A. Shilatifard (Stowers Institute for Medical Research, Kansas City, USA) and Dr M. Carmo-Fonseca (Universidade de Lisboa, Portugal) for providing antibodies. We thank Dr. O. Cuvier (Laboratoire de Biologie Moléculaire Eucaryote du CNRS, Toulouse, France) for his advices on ChIP-seq data analyses. Our work was supported by grants from la Fondation ARC pour la recherche sur le cancer (to S.E.), la Fondation pour la Recherche Médicale and Agence Nationale de la Recherche (to T.K.).

## Author contributions

S.E., S.M. and T.K. conceived and designed the study. S.E. carried out most of the experiments. P.V. performed FISH experiments. R.R. and M.T. analyzed the ChIP-seq data. T.K. wrote the manuscript.

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

- Adelman K, Lis JT (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* 13: 720-31
- Alazami AM, Al-Owain M, Alzahrani F, Shuaib T, Al-Shamrani H, Al-Falki YH, Al-Qahtani SM, Alsheddi T, Colak D, Alkuraya FS (2012) Loss of function mutation in LARP7, chaperone of 7SK ncRNA, causes a syndrome of facial dysmorphism, intellectual disability, and primordial dwarfism. *Hum Mutat* 33: 1429-34
- Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R (2005) Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* 123: 265-76
- Barboric M, Yik JH, Czudnochowski N, Yang Z, Chen R, Contreras X, Geyer M, Matija Peterlin B, Zhou Q (2007) Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription. *Nucleic Acids Res* 35: 2003-12
- Barrandon C, Bonnet F, Nguyen VT, Labas V, Bensaude O (2007) The transcription-dependent dissociation of P-TEFb-HEXIM1-7SK RNA relies upon formation of hnRNP-7SK RNA complexes. *Mol Cell Biol* 27: 6996-7006
- Biewenga P, Buist MR, Moerland PD, Ver Loren van Themaat E, van Kampen AH, ten Kate FJ, Baas F (2008) Gene expression in early stage cervical cancer. *Gynecol Oncol* 108: 520-6
- Boyd KE, Wells J, Gutman J, Bartley SM, Farnham PJ (1998) c-Myc target gene specificity is determined by a post-DNA binding mechanism. *Proc Nat Acad Sci USA* 95: 13887-92
- Cheng Y, Jin Z, Agarwal R, Ma K, Yang J, Ibrahim S, Olaru AV, David S, Ashktorab H, Smoot DT, Duncan MD, Hutcheon DF, Abraham JM, Meltzer SJ, Mori Y (2012) LARP7 is a potential tumor suppressor gene in gastric cancer. *Lab Invest* 92: 1313-1319.
- Chu C, Qu K, Zhong FL, Artandi SE, Chang HY (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell* 44: 667-78
- Chu C, Quinn J, Chang HY (2012) Chromatin isolation by RNA purification (ChIRP). *Journal of visualized experiments : JoVE* 61

- Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* 33: 717-26
- Darzacq X, Jady BE, Verheggen C, Kiss AM, Bertrand E, Kiss T (2002) Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *The EMBO J* 21: 2746-56
- Diribarne G, Bensaude O (2009) 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. *RNA Biol* 6: 122-128.
- Edery P, Marcaillou C, Sahbatou M, Labalme A, Chastang J, Touraine R, Tubacher E, Senni F, Bober MB, Nampoothiri S, Jouk PS, Steichen E, Berland S, Toutain A, Wise CA, Sanlaville D, Rousseau F, Clerget-Darpoux F, Leutenegger AL (2011) Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* 332: 240-3
- Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S (2007) Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* 318: 1777-9
- Egloff S, O'Reilly D, Murphy S (2008) Expression of human snRNA genes from beginning to end. *Biochemical Society Transactions* 36: 590-4
- Egloff S, Szczepaniak SA, Dienstbier M, Taylor A, Knight S, Murphy S (2010) The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J Biol Chem* 285: 20564-9
- Egloff S, Van Herreweghe E, Kiss T (2006) Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and HEXIM1 binding. *Mol Cell Biol* 26: 630-42
- Flynn RA, Do BT, Rubin AJ, Calo E, Lee B, Kuchelmeister H, Rale M, Chu C, Kool ET, Wysocka J, Khavari PA, Chang HY (2016) 7SK-BAF axis controls pervasive transcription at enhancers. *Nat Struct Mol Biol* 23: 231-8
- Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, Bentley DL (2009) TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol Cell Biol* 29: 5455-64
- He H, Liyanarachchi S, Akagi K, Nagy R, Li J, Dietrich RC, Li W, Sebastian N, Wen B, Xin B, Singh J, Yan P, Alder H, Haan E, Wieczorek D, Albrecht B, Puffenberger E, Wang H, Westman JA, Padgett RA et al. (2011) Mutations in U4atac snRNA, a

- component of the minor spliceosome, in the developmental disorder MOPD I. *Science* 332: 238-40
- He N, Jahchan NS, Hong E, Li Q, Bayfield MA, Maraia RJ, Luo K, Zhou Q (2008) A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* 29: 588-99
- He N, Liu M, Hsu J, Xue Y, Chou S, Burlingame A, Krogan NJ, Alber T, Zhou Q (2010) HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol Cell* 38: 428-38
- He N, Pezda AC, Zhou Q (2006) Modulation of a P-TEFb functional equilibrium for the global control of cell growth and differentiation. *Mol Cell Biol* 26: 7068-76
- Henry RW, Ford E, Mital R, Mittal V, Hernandez N (1998) Crossing the line between RNA Polymerases: Transcription of human snRNA genes by RNA Polymerases II and III. *Cold Spring Harbor symposia on quantitative biology LXIII*: 111-120
- Hernandez N (1992) Transcription of vertebrate snRNA genes and related genes. In *Transcriptional regulation*, Yamamoto KRaM, S.L (ed) pp 281-313. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press
- Hernandez N (2001) Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription. *J Biol Chem* 276: 26733-6
- Hu D, Smith ER, Garruss AS, Mohaghegh N, Varberg JM, Lin C, Jackson J, Gao X, Saraf A, Florens L, Washburn MP, Eissenberg JC, Shilatifard A (2013) The little elongation complex functions at initiation and elongation phases of snRNA gene transcription. *Mol Cell* 51: 493-505
- Jawdekar GW, Henry RW (2008) Transcriptional regulation of human small nuclear RNA genes. *Biochim biophys Acta* 1779: 295-305
- Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J, Chabot B, Poirier GG, Hughes TR, Blanchette M, Price DH, Coulombe B (2007) Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 27: 262-74
- Ji X, Lu H, Zhou Q, Luo K (2014) LARP7 suppresses P-TEFb activity to inhibit breast cancer progression and metastasis. *eLife* 3: e02907
- Krueger BJ, Jeronimo C, Roy BB, Bouchard A, Barrandon C, Byers SA, Searcey CE, Cooper JJ, Bensaude O, Cohen EA, Coulombe B, Price DH (2008) LARP7 is a

- stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res* 36: 2219-29
- Kwek KY, Murphy S, Furger A, Thomas B, O'Gorman W, Kimura H, Proudfoot NJ, Akoulitchev A (2002) U1 snRNA associates with TFIIF and regulates transcriptional initiation. *Nat Struct Biol* 9: 800-5
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9: 357-9
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* 25: 1754-60
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* 25: 2078-9
- Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L, Washburn MP, Conaway JW, Conaway RC, Shilatifard A (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol Cell* 37: 429-37
- Luo Z, Lin C, Shilatifard A (2012) The super elongation complex (SEC) family in transcriptional control. *Nature Reviews* 13: 543-7
- Machyna M, Heyn P, Neugebauer KM (2013) Cajal bodies: where form meets function. *Wiley Interdisciplinary Reviews* 4: 17-34
- Markert A, Grimm M, Martinez J, Wiesner J, Meyerhans A, Meyuhas O, Sickmann A, Fischer U (2008) The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep* 9: 569-75
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* 17: 10-12.
- McNamara RP, Reeder JE, McMillan EA, Bacon CW, McCann JL, D'Orso I (2016) KAP1 Recruitment of the 7SK snRNP Complex to Promoters Enables Transcription Elongation by RNA Polymerase II. *Mol Cell* 61: 39-53
- Medlin J, Scurry A, Taylor A, Zhang F, Peterlin BM, Murphy S (2005) P-TEFb is not an essential elongation factor for the intronless human U2 snRNA and histone H2b genes. *The EMBO J* 24: 4154-65
- Michels AA, Fraldi A, Li Q, Adamson TE, Bonnet F, Nguyen VT, Sedore SC, Price JP, Price DH, Lania L, Bensaude O (2004) Binding of the 7SK snRNA turns the

- HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *The EMBO J* 23: 2608-2619
- Mueller D, Garcia-Cuellar MP, Bach C, Buhl S, Maethner E, Slany RK (2009) Misguided transcriptional elongation causes mixed lineage leukemia. *PLoS Biology* 7: e1000249
- Muniz L, Egloff S, Kiss T (2013) RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. *Nucleic Acids Res* 41: 4686-98
- Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zecha A, Mohseni M, Puttmann L, Vahid LN, Jensen C, Moheb LA, Bienek M, Larti F, Mueller I, Weissmann R et al. (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* 478: 57-63
- Nguyen VT, Kiss T, Michels AA, Bensaude O (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* 414: 322-5
- Ogg SC, Lamond AI (2002) Cajal bodies and coilin--moving towards function. *J Cell Biol* 159: 17-21
- Peterlin BM, Brogie JE, Price DH (2011) 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdisciplinary Reviews* 3: 92-103
- Peterlin BM, Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 23: 297-305
- Polak PE, Simone F, Kaberlein JJ, Luo RT, Thirman MJ (2003) ELL and EAF1 are Cajal body components that are disrupted in MLL-ELL leukemia. *Mol Biol Cell* 14: 1517-28
- Prasanth KV, Camiolo M, Chan G, Tripathi V, Denis L, Nakamura T, Hubner MR, Spector DL (2010) Nuclear organization and dynamics of 7SK RNA in regulating gene expression. *Mol Biol Cell* 21: 4184-96
- Price DH (2000) P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol Cell Biol* 20: 2629-2634.
- Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* 26: 841-2



- Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44: W160-5
- Smith E, Shilatifard A (2013) Transcriptional elongation checkpoint control in development and disease. *Genes & development* 27: 1079-88
- Smith ER, Lin C, Garrett AS, Thornton J, Mohaghegh N, Hu D, Jackson J, Saraf A, Swanson SK, Seidel C, Florens L, Washburn MP, Eissenberg JC, Shilatifard A (2011) The little elongation complex regulates small nuclear RNA transcription. *Mol Cell* 44: 954-65
- Sobhian B, Laguet N, Yatim A, Nakamura M, Levy Y, Kiernan R, Benkirane M (2010) HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol Cell* 38: 439-51
- Van Herreweghe E, Egloff S, Goiffon I, J  dy BE, Froment C, Monsarrat B, Kiss T (2007) Dynamic remodelling of human 7SK snRNP controls the nuclear level of active P-TEFb. *The EMBO J* 26: 3570-80
- Yang Z, Zhu Q, Luo K, Zhou Q (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* 414: 317-322.
- Yokoyama A, Lin M, Naresh A, Kitabayashi I, Cleary ML (2010) A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer Cell* 17: 198-212
- Yoon JB, Murphy S, Bai L, Wang Z, Roeder RG (1995) Proximal sequence element-binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. *Mol Cell Biol* 15: 2019-27
- Zhang Y, Liu T, Meyer CA, Eickhout J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biology* 9: R137
- Zhou Q, Li T, Price DH (2012) RNA polymerase II elongation control. *Annu Rev Biochem* 81: 119-43

## Figure Legends

**Figure 1.** Human 7SK snRNA interacts with RNAPII phosphorylated at Ser5 and Ser7. HeLa RNAPII was immunoprecipitated with antibodies specific for the N-terminal part of

the largest RNAPII subunit Rpb1 ( $\alpha$ -Rpb1), and for the CTD of Rpb1 phosphorylated at Ser2 ( $\alpha$ -Ser2-P), Ser5 ( $\alpha$ -Ser5-P) or Ser7 ( $\alpha$ -Ser7-P). Recovery of RNAPII was confirmed by western blot analysis with the  $\alpha$ -Rpb1 antibody. Co-IP of 7SK, U1 and U2 snRNAs was measured by northern blot analysis. Total cell extract (Ext) representing 1% of the input was also analyzed. Control IP performed without antibody (no AB) were also analyzed. Lane M, DNA size markers.

**Figure 2.** Larp7 and RNAPII are enriched at RNAPII-transcribed snRNA and snoRNA genes in human HeLa cells.

A Scatter plot depicting Larp7 and RNAPII co-occupancy determined by ChIP-seq analyses. RNAPII-specific snRNA and snoRNA genes showing strong Larp7 and RNAPII co-occupancy are indicated in red. The RNAPII occupancy levels indicate the numbers of extended read fragments per million (rpm) within  $\pm 100$  bp of the Larp7 peaks. The complete list of Larp7-associated genes with significant RNAPII co-occupancy represented by at least 50 reads (indicated by grey and red dots) is shown in Table EV1.

B Representative genome browser track examples displaying read coverage profiles for Larp7 and RNAPII on HeLa snRNA (U1, U2), snoRNA (U3, SNORD13) and protein-coding (TTI2 and c-Myc) genes.

**Figure 3.** ChIP-qPCR validation of Larp7 interaction with RNAPII-specific snRNA genes in HeLa cells.

A Association of FL-Larp7 with various classes of human genes. The bar graphs indicate the relative enrichment of genomic DNA association with FL-Larp7 over the input control. Data are the means of three independent measurements with error bars representing standard deviation.

B Quantitative ChIP analyses of endogenous RNAPII, ELL and FL-Larp7, FL-MePCE and FL-HEXIM1 occupancy of HeLa U2 snRNA gene. HeLa polyclonal cell populations stably expressing FL-Larp7, FL-MePCE and FL-HEXIM1 showed no detectable Larp7, MePCE and HEXIM1 over-accumulations (Fig EV1). Schematic diagram of the human U2 snRNA gene including the distal and proximal sequence elements (DSE and PSE), the coding region (U2), the 3' box (3') and the downstream CT-rich region is shown.

Positions of the PCR-amplified R1 to R5 fragments relative to the transcription initiation site (indicated by an arrow) are shown. Other details are identical to panel A.

C ChIRP-qPCR analyses of 7SK and U1 snRNAs occupancy of U2, 7SL and hnRNP H1 genes. Biotinylated oligonucleotides complementary to the C178-A197 (7SK-1), C221-G241 (7SK-2) and G272-C291 (7SK-3) sequences of the 7SK snRNA or to the A1-G20 sequences of the U1 snRNA were used for RNA affinity selection from a cross-linked and sonicated HeLa cell extract. Chromatin DNAs co-purified with 7SK and U1 snRNAs were analyzed by qPCR using U2, 7SL and hnRNP H1 gene-specific primers. Bar graphs represent mean values of three replicates. Error bars represent standard deviations.

**Figure 4.** Knock down of 7SK snRNP inhibits RNAPII-mediated synthesis of snRNAs and snoRNAs.

A Larp7 depletion reduces the accumulation of RNAPII-synthesized snRNAs and snoRNAs. HeLa cells were transfected twice (at 0 and 48 hours) with Larp7-specific or non-targeting control siRNAs and incubated for additional 24 hours. RNAs were isolated and size-fractionated on a sequencing gel, electroblotted onto a nylon membrane and probed with sequence-specific oligodeoxynucleotide probes as indicated. Accumulation of Larp7, HEXIM1, Cdk9, CycT1 and hnRNP A1 was monitored by western blot analysis with specific antibodies.

B Relative expression of selected protein-coding genes (black bars), RNAPII- (grey bars) and RNAPIII-specific (white bars) small RNA genes was determined by RT-qPCR in HeLa cells in response to siRNA-mediated Larp7 depletion. The signals were normalized to GAPDH mRNA levels that showed no significant change after siRNA treatment. Bars indicate average values of three independent experiments with error bars representing standard deviations.

C Knock down of 7SK snRNA by RNA interference inhibits accumulation of RNAPII-synthesized small nuclear RNAs. Other details are identical to panel A.

D Nuclear run-on assay. Transcriptionally active nuclei isolated from HeLa cells transfected with Larp7- and Ice1-specific or with non-targeting control siRNA were incubated with radiolabeled uridine 5'-[ $\alpha$ -<sup>32</sup>P]-triphosphate. After RNA extraction, the newly synthesized internally labeled RNAs were partially hydrolyzed and reacted with

immobilized oligodeoxynucleotides complementary to the U1, U2, U6, 5S and 7SL RNAs. As negative controls, sense U1 and U2 oligonucleotides were used.

**Figure 5.** The human 7SK core snRNP interacts with the snRNA gene-specific transcription factor LEC.

A, B, C HeLa endogenous SNAPC2, Ice1, Ice2 (A) ELL and ZC3H8 (B) and transiently expressed FL-Int7, FL-Int11 and FL-HEXIM1 proteins (C) were immunoprecipitated and co-purification of 7SK, U2, U6 or 7SK RNAs was monitored by northern blot analysis. The immunoprecipitated proteins were analyzed by western blotting. Lanes NT, control IPs from extracts prepared from non-transfected cells. Ext, total cell extracts representing 1% of the IP input.

D Co-IP of ELL, Ice2 and MePCE with Larp7.

E Transiently expressed FL-ZC3H8 co-immunoprecipitates with Larp7.

F LEC interacts with the 7SK core snRNP. HeLa LEC was immunopurified with an anti-Ice2 antibody. Co-precipitation of the ELL subunit of LEC, the Larp7 and MePCE 7SK core proteins, as well as the HEXIM1 and hnRNPA1 proteins which transiently interact with 7SK snRNP was tested. Co-IP of 7SK snRNA was monitored by western blotting.

G Transcription inhibition promotes 7SK/LEC accumulation. Ice2, ZC3H8, ELL and HEXIM1 were immunoprecipitated from extracts prepared from HeLa cells either treated (+) or not treated (-) with ActD. IP of Ice2, ZC3H8, ELL and HEXIM1 was confirmed by western blotting and co-precipitation of 7SK and 7SL RNAs was measured by northern blot analysis.

**Figure 6.** The 7SK snRNA is required for the assembly and recruitment of LEC to snRNA genes.

A RNase treatment disrupts LEC. Extracts (Ext) prepared from HeLa cells transfected with the pFL-ZC3H8 or pFL-Ice2 expression plasmid were treated (+) or not treated (-) with RNase A. ELL was immunoprecipitated and co-precipitation of transiently expressed FL-Ice2 and FL-ZC3H8 was monitored by western blotting. Lane no AB, control IP without antibody. Ext, total cell extract representing 1% of the IP input.

**B** Targeted depletion of 7SK snRNA disrupts LEC in living cells. ELL was immunoprecipitated from extracts prepared from HeLa cells transfected with 7SK-specific (+) or control (-) interfering RNAs. Co-purification of endogenous ZC3H8 was assayed by western blotting. Depletion of 7SK and 7SL RNAs was monitored by northern blotting.

**C** The 7SK snRNP is required for efficient recruitment of RNAPII and ELL to the U2 snRNA gene. ChIP-qPCR was performed with Rpb1 and ELL antibodies in control and Larp7-depleted (Larp7 KD) HeLa cells. Depicted are means from three independent experiments, error bars correspond to standard deviation. For other details, see Fig 3.

**D** 7SK snRNP promotes RNAPII loading onto the promoters of the U1 snRNA and U3 snoRNA genes. ChIP was performed with anti-Rpb1 antibody from control and Larp7-depleted cells. Co-purification of U1, U3, U2 and GAPDH DNA was measured by qPCR using sequence-specific primers. Other details are identical to panel C.

**Figure 7.** 7SK snRNP is required for Ice2 localization to CBs.

**A** HeLa 7SK snRNA is present in CBs. Human HeLa cells with or without pre-extraction were probed with a fluorescently labelled oligonucleotide complementary to the 7SK snRNA and an anti-coilin antibody as indicated above the panels. Closed arrowheads indicate CBs with 7SK accumulation. Scale bar, 10  $\mu$ m.

**B** Depletion of 7SK snRNA disrupts co-localization of Ice2 with coilin in CBs. HeLa cell either non-treated (WT) or treated (7SK RNAi) with a 7SK-specific interfering RNA were immunostained with antibodies against Ice2 and coilin. CBs with or without Ice2 are indicated by closed and open arrowheads, respectively. Scale bar, 25  $\mu$ m.

**C** Statistical analysis of Ice2 localization to CBs in control (WT) and 7SK-depleted HeLa cells. Depicted are means (percentage of HeLa cells accumulating Ice2 in CBs) derived from three independent immunostaining (WT) and 7SK snRNA depletion (7SK RNAi) experiments. Error bars correspond to standard deviation. For other details, see the text.

### Expanded View Figure Legends

**Figure EV1.** HeLa polyclonal cell populations stably expressing Flag- (FL)-tagged Larp7, MePCE or HEXIM1 show no significant over-accumulation of Larp7, MePCE or HEXIM1.

A Western blot analysis. Extracts prepared from cells expressing FL-HEXIM1 (lane 2), FL-MePCE (lane 3), FL-Larp7 (lane 4) or from non-transformed (NT) control cells (lane 1) were analyzed by western blotting with anti-Larp7, anti-MePCE, anti-HEXIM1 and as a loading control, with anti-hnRNP A1 antibodies. Please note that the Flag tag is not large enough to differentiate the tagged version of the proteins from the untagged under the gel conditions used. Accumulation of the ectopically expressed FL-tagged MePCE, Larp7 and HEXIM1 proteins was detected by an anti-Flag antibody (lower panel).

B Immunofluorescence microscopy. Expression of FL-Larp7 in stably transfected HeLa cells was verified by immunofluorescence staining with monoclonal anti-Flag M2 antibody followed by incubation with anti-mouse IgG conjugated to fluorescein (Sigma-Aldrich, F1804 and C2181, respectively). Scale bar, 20  $\mu$ m.

**Figure EV2.** Association of 7SK and U1 snRNAs with HeLa U1, 7SK, and 7SL genes was analyzed by ChIRP- (chromatin isolation by RNA purification)-qPCR (see also Fig 3C in the main text).

A Affinity selection of U1 and 7SK snRNAs. Biotinylated oligonucleotides complementary to the U1 or to the 7SK (7SK-1, 7SK-2 and 7SK-3) snRNA were used for streptavidin affinity selection of U1 and 7SK snRNAs from an extract (Ext) prepared from *in vivo* cross-linked HeLa cells. Recovery of U1 and 7SK snRNAs was monitored by Northern blot analysis with a mixture of U1- and 7SK-specific oligonucleotide probes. Lane Cont, RNA affinity selection with a biotinylated control oligonucleotide.

B ChIRP-qPCR analyses of U1 and 7SK snRNA occupancy of HeLa U1, 7SK and 7SL genes. The bar graphs indicate the relative enrichment of genomic U1, 7SK and 7SL DNA association with U1 and 7SK snRNAs over the input control, upon correction with background values obtained by control selections with non-specific oligonucleotides. The observed extremely strong association of U1 and 7SK snRNAs with their own genes is most probably due to transient accumulation of the nascent snRNA transcripts at the site of transcription. Please note that 7SK snRNA shows similar levels of

accumulation on the U1 and U2 snRNA genes (see Fig 3C). Bars represent the mean from three independent experiments, error bars correspond to standard deviation.

**Figure EV3.** Nuclear run on. Nuclear run-on analysis was performed to determine the relative levels of RNAPII-mediated synthesis of nascent HeLa U1 and U2 snRNAs in Ice1- and Larp7-depleted and in control cells (see main text, Figure 4D). The transcription rates of the RNAPIII-specific U6, 5S and 7SL RNA genes were also measured. The obtained slot blot hybridization signals were quantitated by PhosphorImager. The nascent RNA levels detected in control cells treated with a non-specific interfering RNA were considered as 100%.

**Figure EV4.** 7SK snRNP is required for LEC localization to CBs.

A 7SK snRNP is not required for Ice1 expression. Total cell extracts prepared from HeLa cells transfected with Larp7- (Larp7 RNAi), 7SK- (7SK RNAi) or non-specific (Cont RNAi) interfering RNAs were analyzed by western blotting with antibodies specific for Ice1 and Larp7. Accumulation of 7SK snRNA was monitored by northern blotting. As controls, the levels of the 7SL signal recognition particle RNA and hnRNP A1 were also measured. While transfection with Larp7-specific interfering RNA diminished accumulation of both Larp7 and 7SK snRNA, RNAi-mediated depletion of 7SK snRNA has no effect on Larp7 expression (see also Fig 4A and 4C in the main text). More importantly, neither Larp7 nor 7SK snRNA depletion affected Ice1 accumulation.

B Depletion of Larp7 and MePCE inhibits Ice2 accumulation in Cajal bodies. HeLa cells either non-treated (WT) or treated with Larp7- or MePCE-specific interfering RNAs were immunostained with antibodies specific for Ice2 and coilin. Interfering RNAs targeting Larp7 (M-020996-01) and MePCE (M-016636-00) were purchased from Dharmacon. Scale bar, 10  $\mu$ m.

**Figure EV5.** Larp7, 7SK and the little elongation complex are enriched on RNAPII-specific sn/snoRNA genes.

A Representative genome browser track examples displaying read coverage profiles for 7SK, FL-Larp7, RNAPII, Ice1, Ice2, ZC3H8 and their respective negative controls on human U3A and U3C and U13 snoRNA genes. The 7SK ChIRP-seq (two independent datasets) and the Ice1, Ice2 and ZC3H8 ChIP-seq datasets have been published (Flynn et al., 2016, Hu et al., 2013). Adapters were trimmed with Cutadapt v. 1.9.1 with the

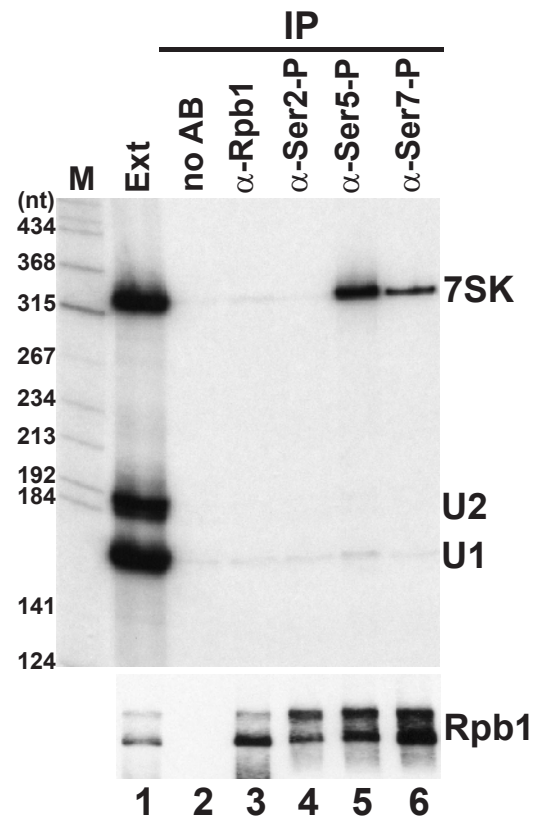
following constant parameters: --minimum-length 10 -q 15, 10 --max-n 1 (Martin, 2011). The obtained sequences were mapped to the human Hg19 reference sequence with Bowtie2 v. 2.2.5 (Langmead & Salzberg, 2012). Unmapped reads were removed with SAMtools v. 1.3.1 (Li et al., 2009). Mapped reads were then de-duplicated using Picard to remove PCR-duplicates. Bam files were sorted and indexed with SAMtools. Trackhubs in the UCSC browser were created by employing deepTools2 v. 2.2.4 bamCoverage tool with the following parameters: -bs 10 --normalizeUsingRPKM -e 200 -p max (Ramirez et al., 2016).

B Profiles of Larp7, RNAPII (left panel, this study) and 7SK (right panel, Flynn, 2016 #578) and their respective negative controls across the U2 snRNA repeat genes (RNAU2). The number of mapped reads to the RNU2 gene for each ChIP/ChIRP-seq were normalized to reads per million (RPM). Adapters trimmed sequences were mapped to the human RNU2 repeat (GenBank: U57614.1) with Bowtie2. Unmapped reads were removed with SAMtools. Mapped reads were then de-duplicated using Picard to remove PCR-duplicates. Bam files were sorted and indexed with SAMtools. Mapped reads were scaled to one million and read coverages across the RNU2 repeat were visualized with Bedtools v. 2.23 genomecov tool (Quinlan & Hall, 2010). Positions of the enhancer (DSE), promoter (PSE), and 3' box are indicated on the plot.

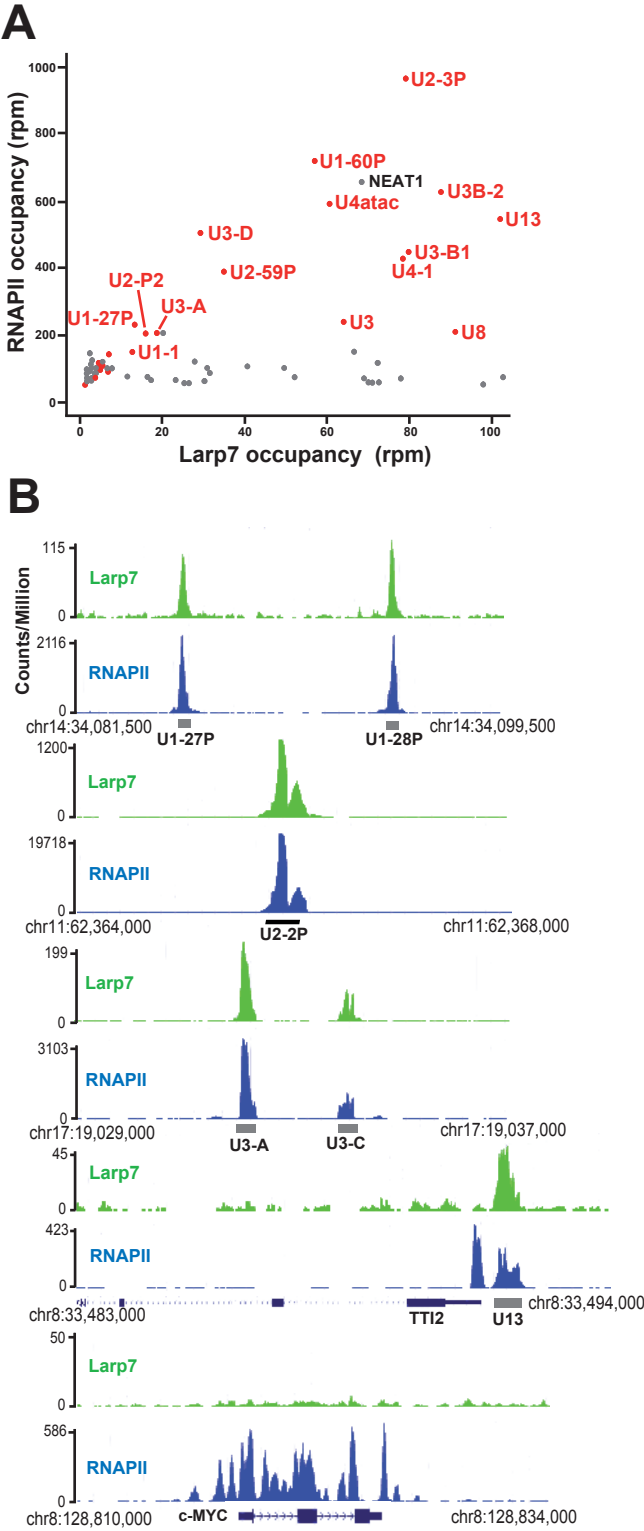
C Boxplots depicting Larp7, 7SK, Ice1, Ice2, and ZC3H8 occupancy on non-overlapping protein- and sn/snoRNA-coding genes. All these factors show a strong preference for sn/snoRNA genes compared to protein-coding genes. For retrieving non-overlapping transcription units, the GENCODE V19 annotation was parsed using a custom Python script to keep only non-overlapping transcription units within a region of 2.5 kb upstream of the TSS and 2.5 kb downstream of the TES. Protein occupancy was then calculated on non-overlapping protein- and sn/snoRNA genes. Input and IP coverage were computed with SAMtools bedcov tool across the gene body (GB: TSS to TES) and normalized to reads per million (RPM). The Input value was subtracted to the IP value and then divided by the transcription unit size. Only transcription unit with a protein occupancy above 0 were kept. The protein occupancy is therefore defined as:  $\log_{10}((GB \text{ IP} * RPM) - (GB \text{ Input} * RPM)) / (length_{GB})$ . Boxplots were created with GraphPad Prism 6. The number of genes in each category is indicated. P-values were computed by a Wilcoxon rank sum test in R.

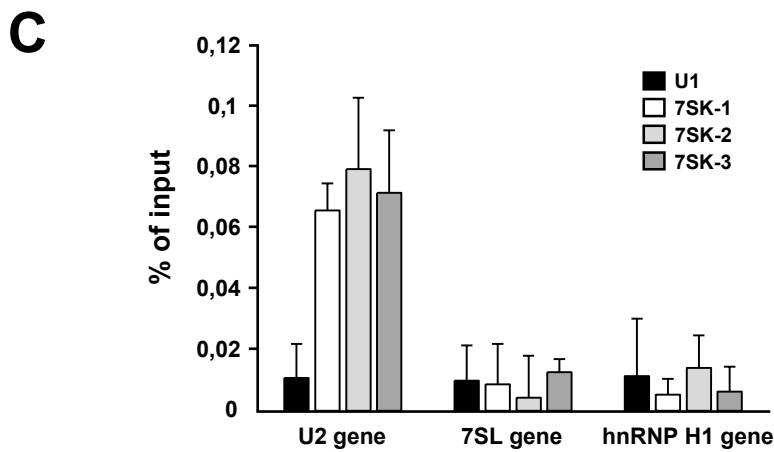
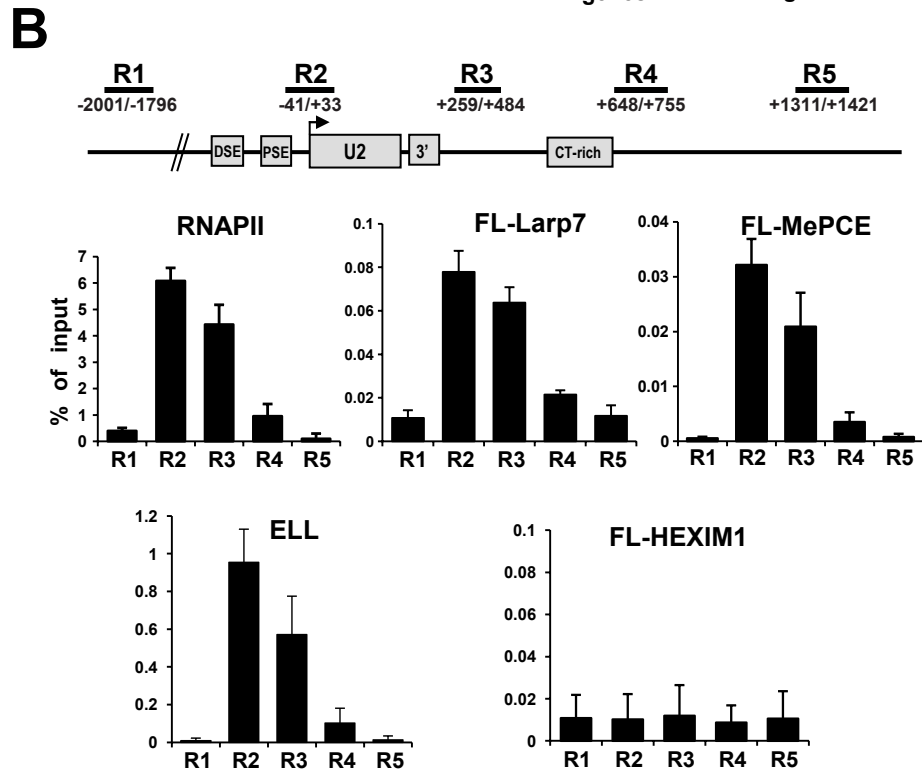
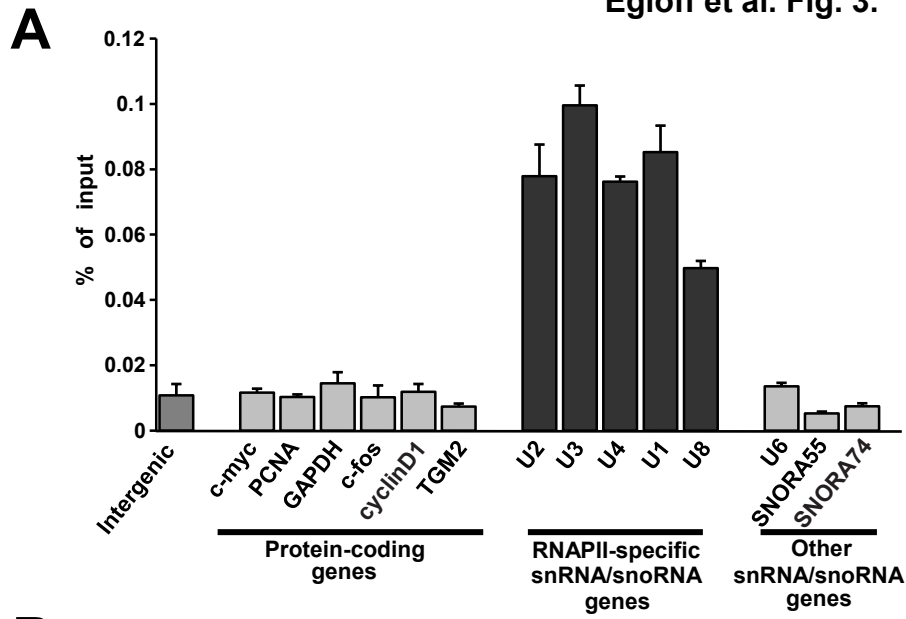


Egloff et al. Fig. 1.

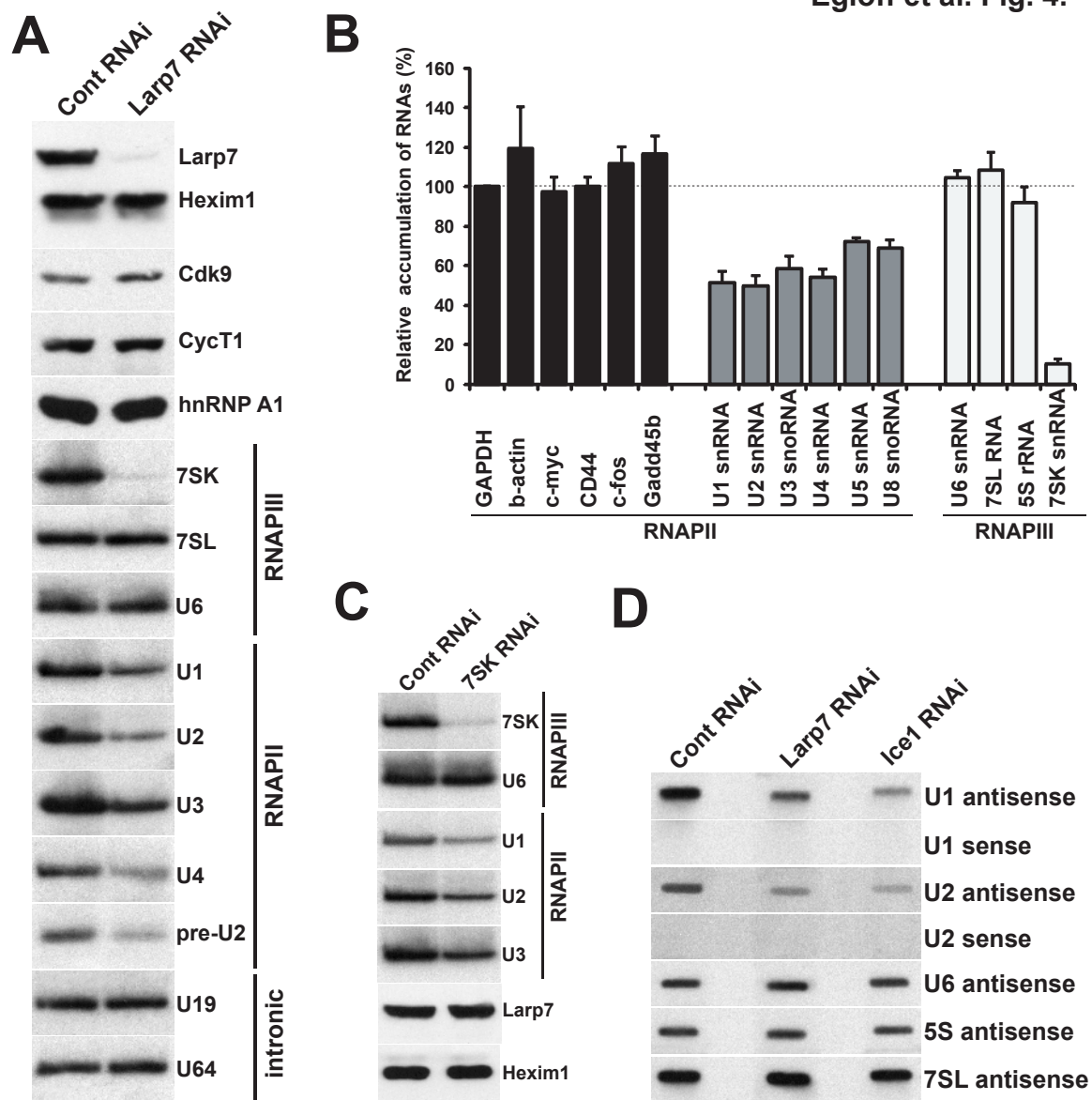


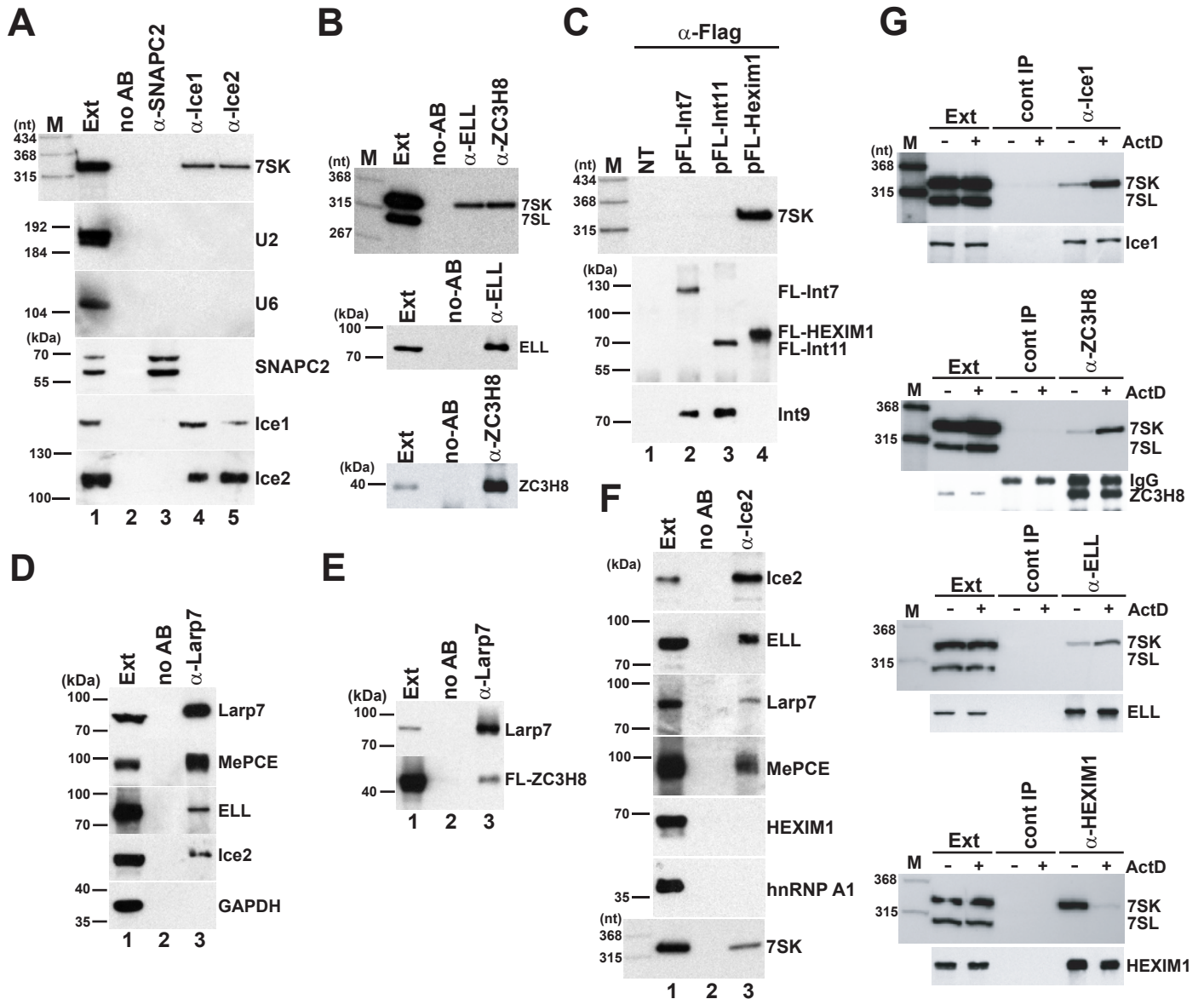
Egloff et al. Fig. 2.

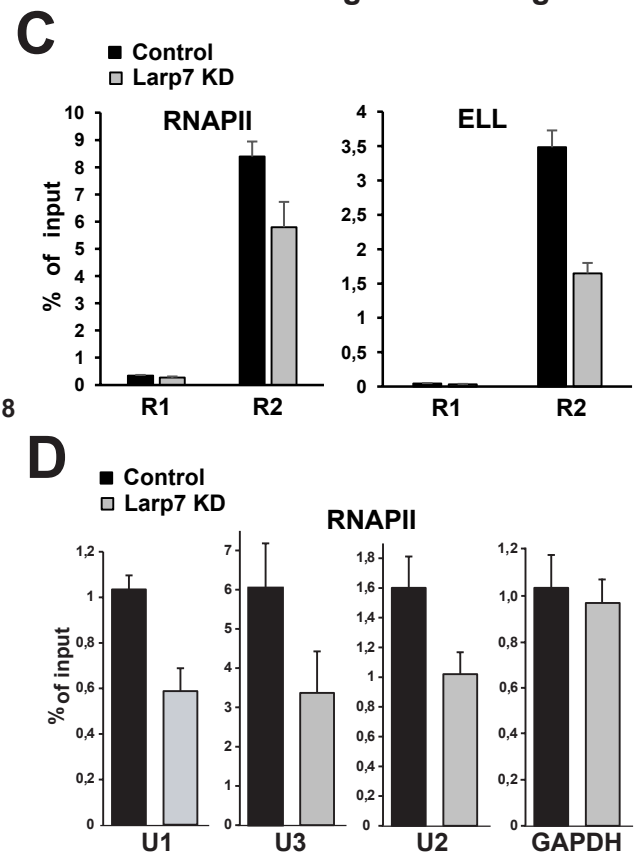
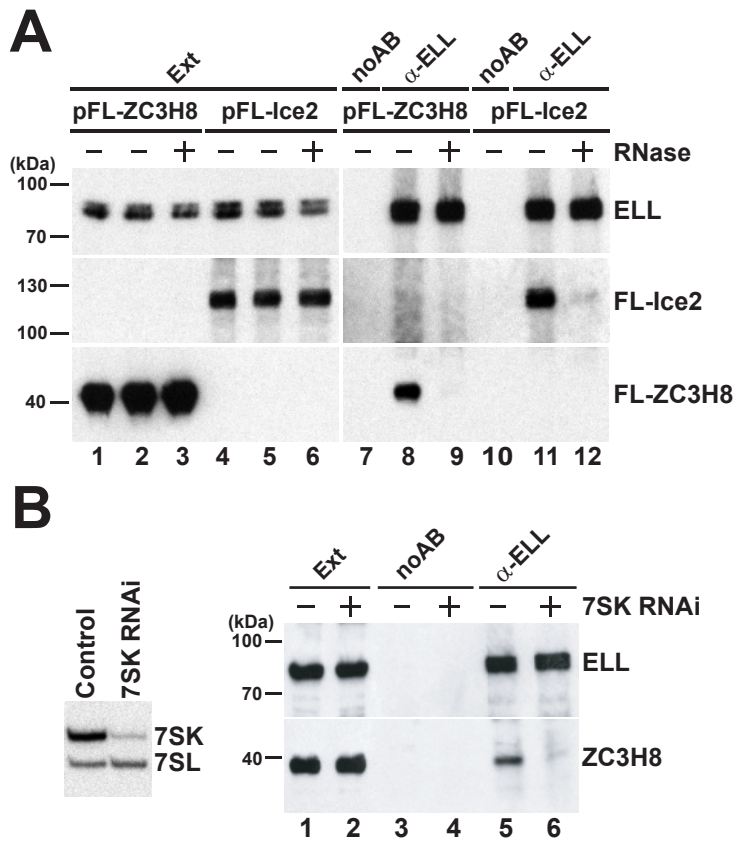




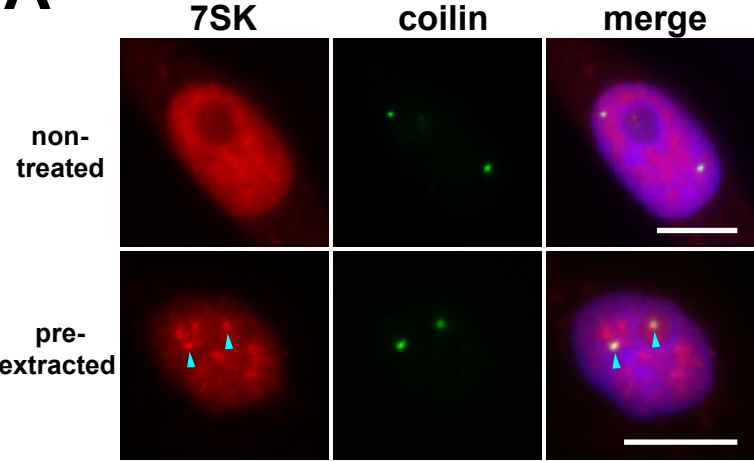
Egloff et al. Fig. 4.



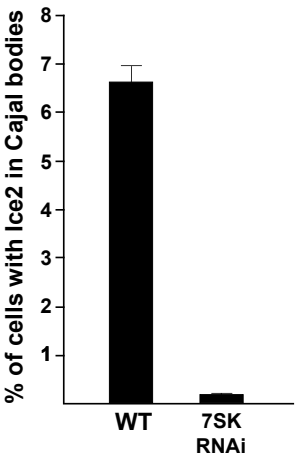




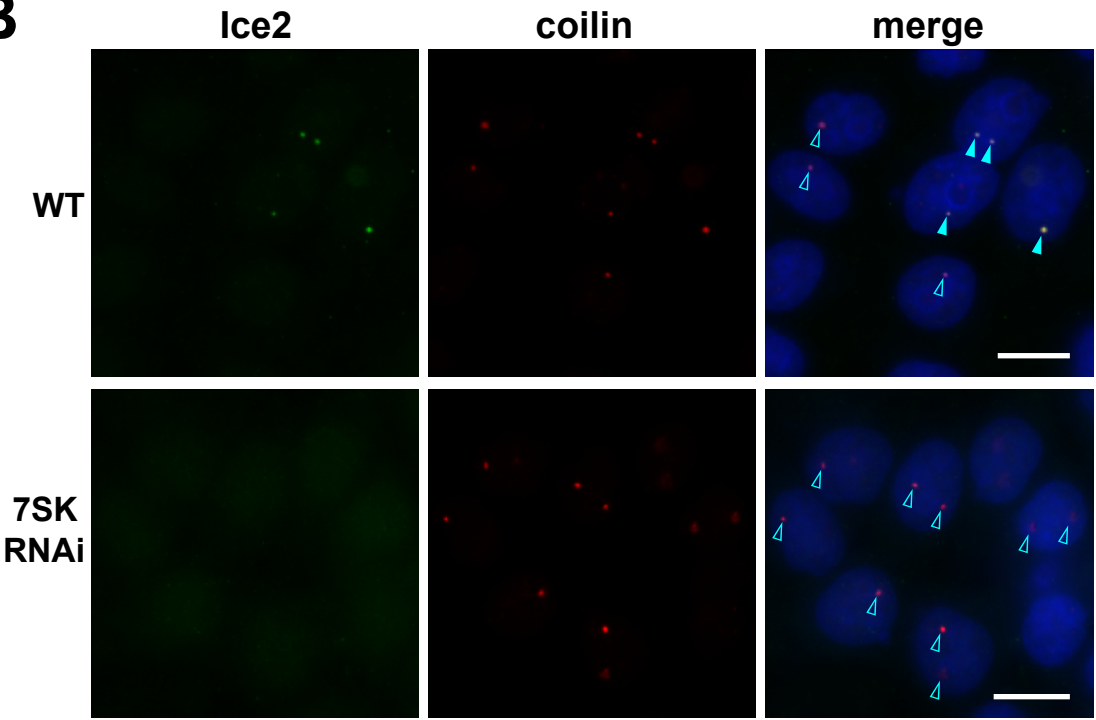
**A**



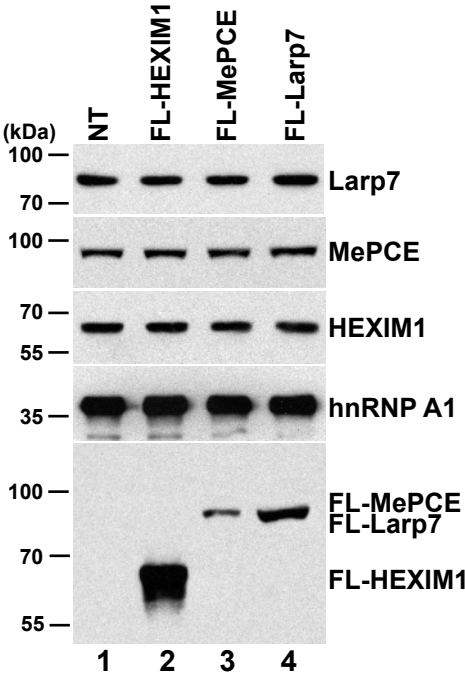
**C**



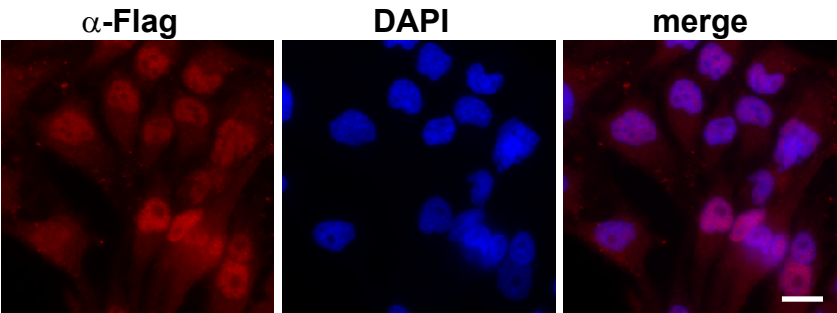
**B**



A

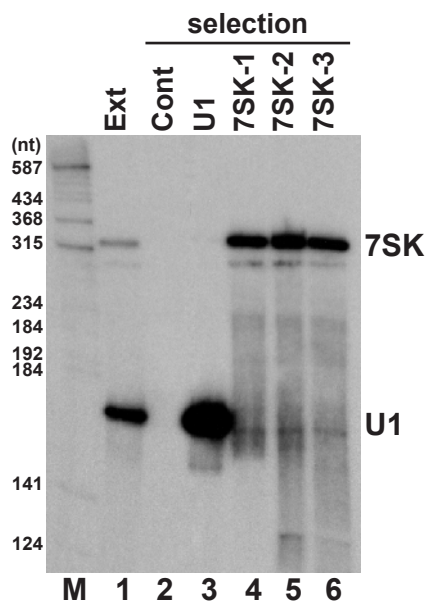


B

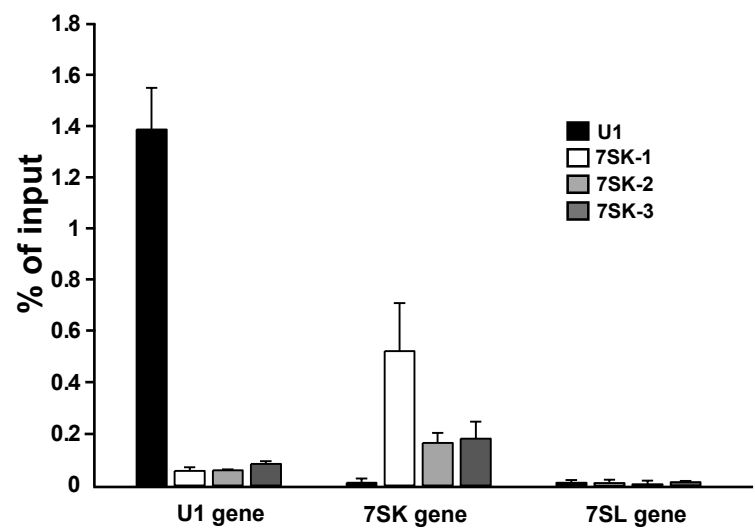




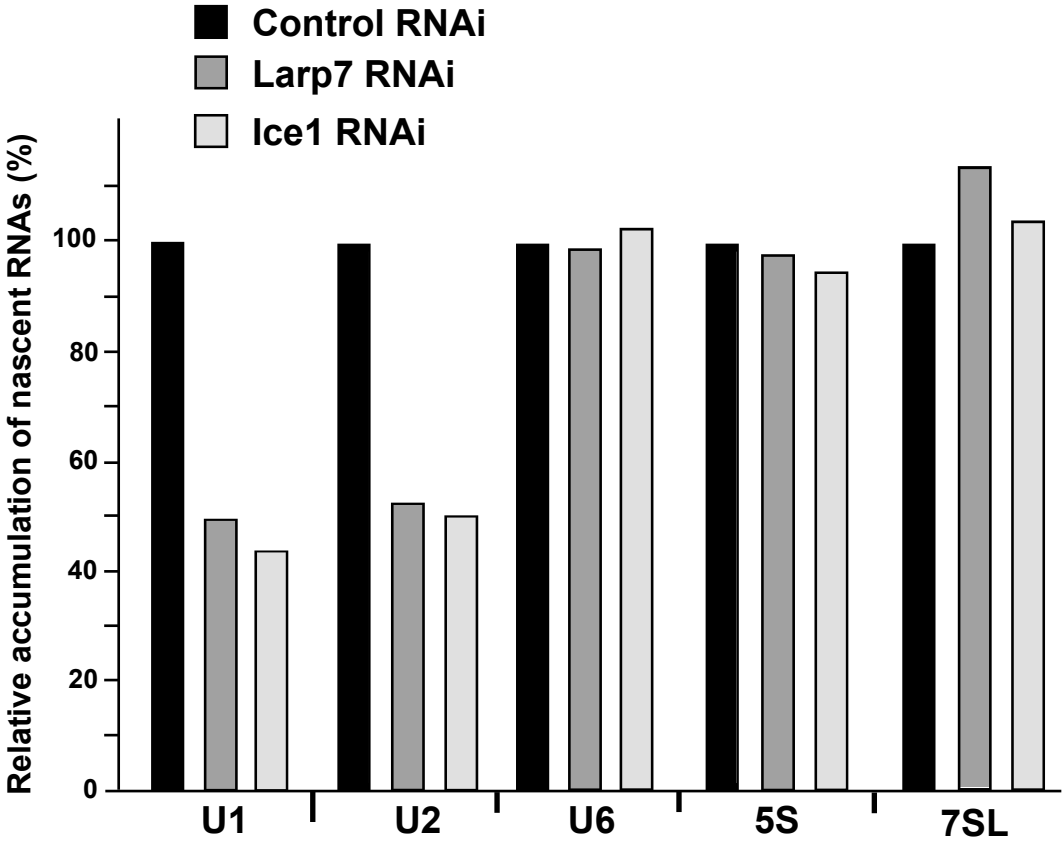
**A**



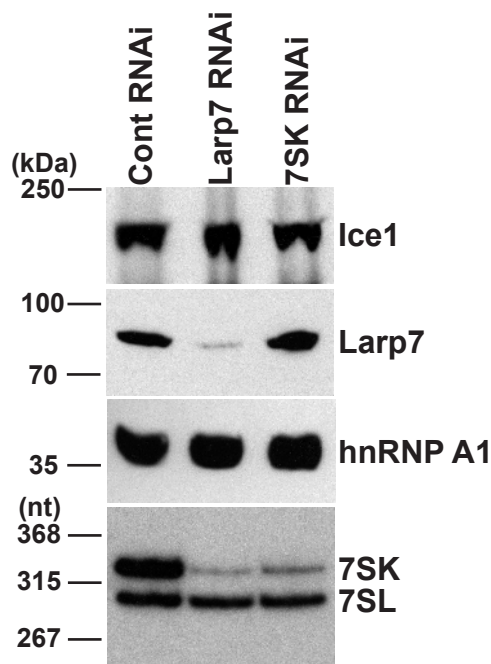
**B**



Egloff et al. Fig. EV3



**A**



**B**

