Novel insights into human RNA polymerase III transcription: non canonical termination and biogenesis of potential regulatory RNAs

Phd in Biochemistry and Molecular Biology
XXII cycle: 2007-2009

Andrea Orioli

Department of Biochemistry and Molecular Biology,
University of Parma, Italy

PhD Coordinator: Prof. Gian Luigi Rossi
PhD Tutor: Prof. Giorgio Dieci

Chapter 1

Human RNA polymerase III transcription

OVERVIEW OF EUKARYOTIC TRANSCRIPTION

In higher eukaryotes, the task of recognizing nuclear gene promoters and then transcribing the genes is accomplished by three highly related enzymes, known as RNA polymerase I, II and III. Each of these polymerases is dedicated to the transcription of a specific set of genes and, quite obviously, the transcriptional output of a genome is orchestrated by a complex network of functional, structural and regulatory elements.

RNA polymerase I has the characteristic role of transcribing the large, tandemly repeated, ribosomal RNA genes encoding the 18S, 5.8S, and 25S-28S rRNAs that form the catalytic core of ribosomes [1]. To keep up with the cell's metabolic activity and demand for ribosomes, these genes are transcribed with high efficiency in a cell growth co-regulated fashion and give birth to the largest sub-nuclear structure, the nucleolus, where ribosomes are assembled [2].

RNA polymerase II is dedicated to the transcription of protein-coding mRNA genes as well as many other non-coding RNA species, such as snRNAs, snoRNAs and miRNAs. During the RNA transcription cycle, the enzyme is recruited by transiently associated general transcription factors (GTFs) into a core promoter region encompassing the transcription start site. The core promoters come in different types, but they generally contain a TATA-box, an initiator and a downstream promoter element arranged in various combinations. Once a pre-initiation complex (PIC), composed of RNA polymerase II and GTFs, is bound to the core promoter, the subsequent events leading to efficient transcription are summarized as initiation, promoter escape, abortive transcription, elongation and finally transcription termination (for reviews, see [3],[4]). Gene expression is regulated by an extensive array of factors which spread from subtle differences in chromatin state and accessibility to distinct sets of regulators able to bind proximal and distal control regions to facilitate or repress PIC recruitment [5]. Proteins are the main effectors, but also ncRNAs have been recently shown to control the transcription reaction by targeting general factors, co-regulators, RNA polymerase II and chromatin [6]. The emerging picture is that of an articulated regulatory pathway which can be considered as one of the landmarks of higher eukaryotes complexity.

RNA polymerase III is the largest among the three polymerases and synthesizes a varied collection of short, untranslated transcripts that rarely exceeds 400nt in length, such as 5S rRNA, tRNAs, AluRNAs and many others (class III genes). The RNA polymerase III promoters are more varied in structure than the uniform RNA polymerase I ones, yet not as diverse as the RNA polymerase II promoters. Their features, together with our current understanding of RNA polymerase III transcripts and the various pathways leading to the recruitment and termination of the enzyme on the relative genes will be summarized in the following chapters.

Besides the three DNA-dependant RNA polymerases described above, higher plants possess two additional RNA polymerases, named IVa and IVb, specialized in small RNA mediated gene silencing pathways [7]. Curiously, an RNA-directed RNA polymerase is also involved in *S.pombe* co-transcriptional gene silencing [8], suggesting that the enzyme might be a common feature, yet to be discovered, of most eukaryotes.

RNA POLYMERASE III

The different composition of the three eukaryotic polymerases, including some shared and some specific subunits, likely reflects their specialized functional requirements at the initiation, elongation and termination stages of the transcription cycle. RNA polymerase III appears the most complex of the nuclear RNA polymerases, as it consists of 17 subunits ranging from 7.0 to 155.6 KDa and has a total molecular weight of 600-680 KDa [8]. Ten out of the 17 subunits are unique to *S. cerevisiae* RNA polymerase III and are designated C subunits, two are common to RNA polymerase I and III and are designated AC subunits, and five are common to all eukaryotic polymerases and are designated ABC subunits; five of the C subunits are unique to RNA polymerase III and have no paralogs in RNA polymerase I and II (Tab 1). Recently, human RNA polymerase III has been purified from a stable cell line expressing a doubly tagged HsRPC4 subunit, and homologs of all 17 subunits of yeast RNA polymerase III have been identified by mass spectrometry [11].

RNA polymerase III subunits in *S.cerevisiae* (*H.sapiens*)	**MW in *S.cerevisiae* (*H.sapiens*), KDa**	**Paralogs in RNA polymerase II/I**	**Accession number *S.cerevisiae (H.sapiens)***
C160 (HsRPC1)	162.1 (155.6)	RPB1/RPA190	P04051 (AAB86536)
C128 (HsRPC2)	129.3 (127.6)	RPB2/RPA135	AAB59324 (AY092084)
C82 (HsRPC3)	73.6 (60.5)		CAA45072 (NP_006459)
C53 (HsRPC4)	46.6 (44.4)		P25441 (AY092086)
C37 (HsRPC5)	32.1 (79.8)		NP_012950 (AY092086)
C34 (HsRPC6)	36.1 (35.6)		P32910 (NP_006457)
C31 (HsRPC7)	27.7 (25.9)		P17890 (AAB63676)
C25 (HsRPC8)	24.3 (22.9)	RPB7/RPA43	P35718 (AY092087)
C17 (HsRPC9)	18.6 (16.8)	RPB4/RPA14	P47076 (AAC25992)
C11 (HsRPAC10)	12.5 (12.3)	RPB9/RPA12	AAD12060 (NP_057394)
AC40 (HsRPAC1)	37.6 (38.6)		P07703 (NP_004866)
AC19 (HsRPAC2)	16.1 (15.2)		P28000 (NP_057056)
ABC27 (HsRPABC1)	25.1 (24.6)		P20434 (P19388)
ABC23 (HsRPABC2)	17.9 (14.5)		AAA34989 (P41584)
ABC14.5 (HsRPABC3)	16.5 (17.1)		CAA37383 (P52434)
ABC10α (HsRPABC4)	7.7 (7.0)		AAA64417 (P53803)
ABC10β (HsRPABC5)	8.2 (7.6)		P22139 (P52436)

Tab 1. *Subunit composition of S.cerevisiae and H.sapiens RNA polymerase III, adapted from Schramm and Hernandez, 2002. C subunits are unique to RNA polymerase III; AB subunits are common to RNA polymerase I and III; ABC subunits are common to RNA polymerase I, II and III. Subunits in bold have no paralogs in RNA polymerase I and II.*

Simple logic suggests that RNA polymerase III specific subunits should contribute to the enzyme specific properties, but much remains to be found out since most of our current knowledge comes from genetic and two-hybrid screens in S.cerevisiae [9,10]. Information on RNA polymerase III architecture are relatively poor to date, nonetheless, in 2006 the first structural data became available, including a homology model for the core enzyme [13] and a 17 Å EM structure [14, Fig. 1A]. Later on, with the use of DMSO treatment and tandem mass spectrometry, Cramer and colleagues dissociated and identified peripheral subcomplexes from the core enzyme, providing important insights into the subunit architecture of RNA polymerase III [15]. Apparently, the enzyme comprises a 10 subunit core that resembles the RNA polymerase II core and likely includes subunit C11, even though a trimeric subcomplex, C53/37/11, cannot be completely ruled out; the subunit C11 has a region of homology to the elongation factor TFIIS and plays an important role in transcription termination as well as in the RNA polymerase III switch from polymerase to

endonuclease activity [12]. Arrayed around the periphery of the core are three heterodimeric subcomplexes, C53/37, C82/34, and C17/25. Subunit C31 forms a stable trimer with the initiation related C82/C34 subcomplex [16], but also binds to the RNA polymerase III core and subunit C17 [8,12], suggesting that it is involved in bridging between C82/C34, the core and C17/C25. The C37/C53 heterodimer likely adopt a position on the outside of the core enzyme [14], where it slows down nascent chain elongation [17], thus playing a major role in correct recognition of termination signals of class III genes. C17/C25 is the first RNA polymerase III complex whose structure was determined by X-ray crystallography [15, Fig. 1B]. The heterodimer contributes to initiation complex assembly, since C17 interacts with the TFIIB-related domain of TFIIIB70 [18]; nonetheless, its affinity for class III genes is greater than that of the RNA polymerase II related complex Rpb4/7 [15], suggesting that specific properties of the subunits surrounding the core might account for differences in promoter/termination recognition among the three eukaryotic polymerases.

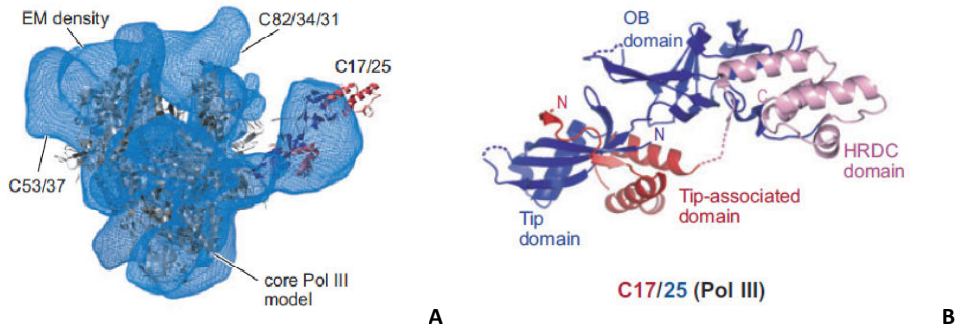


Figure 1 **A**. *RNA polymerase III 17Å EM structure, adapted from Fernandez-Tornero et.al, 2007.* **B**. *Crystal structure of C17/C25, Cramer et. al, 2007.*

TRANSCRIPTION FACTORS

TFIIIA

Much of our current knowledge on TFIIIA comes from studies in *X.laevis*, considering that long ago its purification was greatly facilitated by its considerable abundance in early Xenopus oocytes [126]. Miller *et al*., proposed that *X.laevis* TFIIIA is composed of nine tandem repeats arranged in compact and independent domains; each of them enclose a central zinc ion coordinated to the invariant cysteines and histidines within a loop-like structure, termed a zinc finger, that directly contacts DNA [127]. Later on, extended X-ray analysis confirmed that a zinc ion, essential for the DNA binding activity, is coordinated at each C2H2 domain [129]. More recently, a 42 kD protein that co-fractionates with TFIIIA activity has been purified from HeLa cells [128]; the polypeptide has the ability to direct specific 5S gene transcription, cross-reacts with antibodies raised against *X.laevis* TFIIIA and produce a footprinting pattern on a human 5S gene closely resembling that produced by the *X.laevis* homolog. A unique feature of TFIIIA is its ability not only to specifically recognize the 5S rDNA promoter sequence, but also to bind the 5S rRNA [178]. The observation that TFIIIA can specifically bind only eukaryotic 5S rRNA suggests its involvement in a eukaryotic-cells related process. Indeed, TFIIIA acts as a factor providing for the export of 5S rRNA from the nucleus and its stabilization in the cytosol as a 7S RNP until its binding with ribosomal protein L5;

the exchange of 5S rRNA from 7S RNP with ribosomal protein L5 results in the formation of 5S RNP, which can enter the nucleus and then the nucleolus, where it is used in ribosome biosynthesis [179].

TFIIIB (TFIIIB-α and TFIIIB-β)

The composition of mammalian TFIIIB is far less well characterized than that of yeast TFIIIB. The first efforts aimed at a more comprehensive definition of human TFIIIB subunits led to the identification of two chromatographically separable components named HsTFIIIBα and HsTFIIIBβ [134,136,143]. HsTFIIIBα is active in the transcription of type 3 genes, whereas HsTFIIIBβ is specific for the transcription of type 1 and 2 genes [136]. Like ScTFIIIB, HsTFIIIBβ is minimally comprised of a stable HsTBP-HsBrf1 complex that associates reversibly with a human homologue (HsBdp1) of ScBdp1 [137,140]. HsBrf1 was absolutely required for transcription at type 1 and 2 promoters, since depletion of extracts with antibodies raised against the C-terminal or N-terminal domain resulted in debilitated transcription of a VA1 or 5S template [136,138]. On the other hand, only antibodies directed against the N-terminal half of HsBrf1, but not the C-terminal half, inhibited transcription at the U6 snRNA promoter [136], suggesting that transcription at type 3 promoters required a novel protein related to the N-terminal domain of HsBrf1, now referred to as HsBrf2 [139]. HsTFIIIBα activity thus reflects a loose association of HsTBP, HsBrf2, and HsBdp1 that can be assembled on a U6 promoter in a stepwise fashion in vitro. It should be noted that U6 transcription in HsBrf2-depleted extract can be restored either by addition of HsBrf2-containing extracts immunopurified from HeLa cells [140] or by addition recombinant HsBrf2 purified from *E.coli* [141], meaning that HsBrf2 associated polypeptides might contribute to the overall reaction efficiency rather than to promoter specificity. Despite their difference at the C-terminal end, homology modeling based on the NMR structure of PfTFIIB [142] indicate that both HsBrf1 and HsBrf2 share a common N-terminal zinc-binding domain and a core domain, consisting of two imperfect repeats, likely involved in TBP binding. Human Bdp1 was found through a combination of database searches of sequence similarity to a specific Bdp1 domain of the yeast homolog [141]. In HeLa cells, several alternative forms of Bdp1 have been found, but it has still to be determined which one is responsible for the transcriptional activity. As previously shown for ScBdp1, all the three forms of human Bdp1 contain a characteristic SANT domain [8], a putative DNA-binding domain also found in the SWI-SNF, ADA complexes and the transcriptional co-repressor N-CoR [145]. Outside of the SANT domain, the yeast and human protein are not conserved, with HsBdp1 displaying a striking C-terminal extension containing a number of repeats with potential phosphorylation sites. Immunodepletions of HeLa extracts with antibodies raised against regions upstream and downstream to the SANT motif within the N-terminal domain of HsBdpI [141], as well as against the C-terminal repeats (Schramm and Hernandez, unpublished data), cause loss of RNA polymerase III transcription from both type 2 and type 3 promoters. Transcription can be restored by addition of recombinant Bdp1, both full length or truncated downstream to the SANT domain; these experiments suggest that the C-terminal repeats are not required for basal in vitro transcription, yet present in the functional protein and perhaps involved in undiscovered regulatory roles.

TFIIIC (TFIIIC-1 and TFIIIC-2)

Human TFIIIC was originally resolved by Mono Q [146] or sequence-specific DNA affinity chromatography [147] into two functional entities, TFIIIC1 and TFIIIC2, that were both required for transcription of the VA1 gene [146]. TFIIIC2 was shown to be required for sequence-specific DNA binding to type 2 and, in conjunction with TFIIIA, to type 1 RNA polymerase III promoters [146,148]. Footprinting experiments revealed that TFIIIC2 interacts strongly with the B-box of type 2 promoters, and this interaction was reinforced, as well as extended over the A-box to sequences close to the transcription initiation site, by the addition of TFIIIC1. What is known to date is that sequence or functional homologs of human TFIIIC2 have been identified in human cells for all six *S.cerevisiae* TFIIIC subunits, namely Tfc1 (HsTFIIIC63), Tfc3 (HsTFIIIC220), Tfc4 (HsTFIIIC102), Tfc6 (HsTFIIIC110), Tfc8 (HsTFIIIC90) (for a review see [8]) and, more recently, Tfc7 (HsTFIIIC35) [154]. It is thus likely that human TFIIIC2 share a common structure with *S.cerevisiae* TFIIIC, which is organized in two DNA binding domains separated by a flexible linker [149] that can accommodate variously spaced A-box and B-box elements. Besides the active TFIIIC2a complex, the TFIIIC2 fraction extracted from HeLa cells contains a transcriptionally inactive TFIIIC2 complex designated TFIIIC2b [153] which represents 10%–20% of total TFIIIC2 in actively dividing cells; TFIIIC2b lacks the TFIIIC110 subunit and appears to contain a 77-kD subunit, absent in TFIIIC2a, of unclear function. In addition to TFIIIC2 requirements for expression of class III genes, in vitro transcription with partially purified TFIIIC1 demonstrated that this factor was absolutely required for transcription of type 2 promoter-containing genes [146,150]. Later, TFIIIC1 was also shown to be required for transcription of type 1 and type 3 promoters by RNA polymerase III [148]. Although the exact composition of TFIIIC1 remains unclear, it was shown to comprise at least human Bdp1, a component of human TFIIIBα and TFIIIBβ [151]. TFIIIC2, TFIIIC1, and TFIIIBβ, in conjunction with RNA polymerase III, are able to specifically initiate and also terminate transcription from type 2 promoters in vitro, but the efficiency of transcription is strongly enhanced by the addition of the RNA polymerase II co-activators PC3 (topoisomerase 1) or PC4 [152]. Requirements of type 3 promoter containing genes are less well characterized, but the general understanding is that RNA polymerase III recruitment by the PSE-binding transcription factors/TFIIIBα complex [136] is somehow helped by TFIIIC1 [148].

PSE and DSE binding proteins

As mentioned above, type 3 promoter genes differ from other class III gens in that they contain proximal and distal sequence elements (PSE and DSE) which recruits specific multisubunit complexes and somehow overcome the lack of TFIIIC2 binding; indeed, depletion of a transcription extract with antibodies directed against the TFIIIC220 subunit, affects transcription from the 5S, tRNA, and VA1 genes but not from the 7SK and U6 snRNA genes [155]. The human DSE contains a variety of protein binding sites, but two of them are almost invariably an SPH (also called NONOCT) element, which recruits *in vitro* the transcription factor STAF (also called SPH binding factor or SBF), and an octamer sequence, which recruits the transcription factor Oct-1 [156]. STAF was originally identified in *X.laevis* as a zinc finger protein containing seven zinc fingers of the C2-H2 type, different sets of which can be used to bind to different DNA targets; the human homolog ZNF143 shares similar DNA binding specificities, and can activate RNA polymerase II and III snRNA gene transcription [157,158]. Oct-1 contains two structurally independent domains (POU-domains) that cooperate functionally as a DNA-binding unit; the bipartite POU domain leads to flexible

interactions with DNA and other proteins and, thus, functional versatility in transcriptional regulation [159]. On the other hand, the factors binding to the human PSE are variously known as PBP, PTF, or $SNAP_c$ (snRNA activating protein complex). $SNAP_c$ is a complex containing five types of subunits, SNAP190, SNAP50 (PTFβ), SNAP45 (PTFδ), SNAP43 (PTFγ), and SNAP19 [160]. SNAP190 forms the backbone of the complex, with SNAP19 and SNAP45 associating toward the N- and C-terminus, respectively, of the molecule. SNAP43 can associate with the same region of SNAP190 as SNAP19, while SNAP50 joins the complex by associating with SNAP43. Details on $SNAP_C$ recruitment can be found in the section on type 3 promoters.

La protein

Human La auto-antigen is a highly abundant protein, consisting of a La motif (LM) and a RNA recognition motif (RRM) arranged in a fixed configuration [162], that interacts with a variety of newly synthesized RNA polymerase III transcripts [161]. In its best characterized activity, La binds to tRNAs, 5S rRNA, U6 snRNA, 7SL RNA, 7SK RNA, Alu RNAs and other RNAs (for a review see [161] and references therein) via the common UUU-OH motif at their 3'ends [163-165], which results from transcriptional termination at the oligo(dT) termination signal. As a result of UUU-OH binding, La stabilizes and protects these transcripts from 3'-exonucleolytic digestion and thus promotes their processing to the mature forms [166]. A recent paper reporting the ability of La to discriminate between precursor and mature tRNAs [168] and to protect various precursor from 5'-processing, as well as the observation that digestion of the RNA with phosphatases decreased interaction with La [167], suggest that La also interacts with the 5'-end of the above precursor transcript [167], which presumably share a triphosphate moiety at their 5'-terminal. Beside its RNA-binding activity, La has been also reported to have effects on RNA polymerase III transcription. Immunodepletion of La from mammalian cell extracts was found to reduce RNA polymerase III output in vitro, which led to the suggestion that La could act as a transcriptional termination factor that mediates nascent transcript release [169,170]; indeed, addition of La to purified mammalian RNA polymerase III transcription complexes lead to increased transcriptional activity, apparently due to enhanced recycling and re-initiation [171,172]. Nonetheless, these data appear to be controversial, since other studies reported that immunodepletion of La from HeLa extracts did not result in any significant decrease in RNA polymerase III transcription [173]. Novel insights came from a recent paper reporting that human La can be at all three types of class III templates in living HeLa cells, as shown by ChIP experiments [174]. Interestingly, none among TFIIA, NF1 and Ro, which had been previously found to be part with La of a putative human RNA polymerase III holoenzyme [175], were shown to bind the same class III genes and contrasts with reports that these factors can stimulate RNA polymerase III transcription from purified systems in vitro [176]. A possible explanation is that La might contribute to transcription only under specific circumstances, related to cell cycle or cell state. Notably, gene occupancy depends on the phosphorylation status of La, with the less prevalent, unphosphorylated form [177] being found selectively on class III genes, suggesting that La might act differently depending on its phosphorylation status.

RNA POLYMERASE III TRANSCRIPTION

Transcription of class III genes requires accessory factors that bind internal or external promoter elements, genomic sequences with a various degree of conservation displayed from yeast to

humans. The organization of these promoter elements directs RNA polymerase III on its target genes via different pathways. Genes with type 1 and type 2 promoters share a common mechanism for TFIIIB and RNA polymerase III recruitment, with the notable difference that type 2 promoters are directly recognized by TFIIIC in the previous step, while type 1 promoters require the assembly of a TFIIIA-DNA complex for the subsequent binding of TFIIIC. Type 3 promoter-containing genes are set apart in that they contain external elements that bind various factors upstream of the transcriptional start site, required to incorporate TFIIIB into the pre-initiation complex. Details on promoter organization and the pathways followed for recruitment of RNA polymerase III on its different target genes are described in the following paragraphs.

RNA POLYMERASE III RECRUITMENT ON ITS TARGET PROMOTER

Type 1 promoters

5S gene promoter is the only example of an RNA polymerase III type 1 promoter. The intragenic element consist of an A-box, an intermediate element and a C-box [180,181], that is conserved in the 5S genes of different species and in *S.cerevisiae* represent the only element actually required for transcription [182]. These three elements arranged together constitute the internal control region (ICR), which is recognized and bound by the zinc finger domains of the specificity factor TFIIIA [183,184]. Once assembled, this binary complex is enlarged by the 600 kDa multisubunit TFIIIC2 [148], which, in a strictly TFIIIA-mediated pathway, binds the 5S rRNA gene upstream of the transcriptional start site with other domains than those required for interaction with the B-box sequence. [148,185]. The low affinity of TFIIIC2 alone for the 5S gene promoter is perfectly conceivable, considering the gene lacks a B-box motif. However, it was also reported that, at high concentrations, TFIIIC2 is able to directly contact the 5S promoter [186], but the data could be biased by traces amount of TFIIIA in the fractions used, considering that TFIIIA is known to interact with TFIIIC2 in the absence of promoter elements [155,187]. In conjunction with TFIIIA, TFIIIC2 recruits the TBP-Brf1-Bdp1 complex TFIIIBβ, which allows the subsequent recruitment of RNA polymerase III [16]. While in yeast, TFIIIB is sufficient for the recruitment of RNA polymerase III for multiple rounds of transcription [187], human cells have evolved a novel activity, named TFIIIC1, which is required for transcription of all types of RNA polymerase III promoters and enhances the binding of TFIIIA,TFIIIC2 and PBP [148]. More recently, a novel separation method enabled Weser *et al*. to isolate intermediate stages of the assembly pathway on the 5S rRNA gene and helped to determine the sequence in which factors enter the nascent complex [188]. Considering that previous investigations concerning the order of complex assembly on the 5S gene [189,190] had been performed before TFIIIC1 was discovered [146], the successful isolation of a ternary 5S rDNA-TFIIIA- TFIIIC2 complex served as the starting point for further investigations. According to a new model, after the primary association of TFIIIA with the ICR and the subsequent enlargement of the complex by TFIIIC2, TFIIIBβ is definitely the third transcription factor to enter the growing complex without any requirement for TFIIIC1 for this process. Nonetheless, the isolated 5S-TFIIIA-TFIIIC2-TFIIIBβ quaternary complex alone does not suffice for transcription, unless TFIIIC1 is integrated into the complex as the last transcription factor. The combined action of TFIIIBβ and TFIIIC1 is then sufficient for the recruitment of RNA polymerase III and the transcription of the 5S gene [188, Fig.2A].

Type 2 promoters

tRNA genes have been traditionally the preferred substrates for type 2 promoter studies, but the same reasoning can be applied to adenovirus 2 VA genes, Alu elements and other short interspersed repeats containing the same internal control elements. The promoter of these genes consist of two highly conserved sequence elements [203], a proximal A-box and a more distal B-box within the transcribed region, corresponding to the tRNA D- and T-loops [191-193]. While the A-box strictly reside 10-20bp downstream of the start site, the spacing between the A-box and B-box varies greatly, in part to accommodate introns. The ability to accommodate a variety of distances between the A-box and B-box usually found in tRNA genes is likely to be mediated by TFIIIC-90 [195]; the 90kDa subunit, in a way comparable to Tfc8 which serves as the τa-τb connector in *S.cerevisiae*, probably provides the flexible linker required to connect the TFIIIC2 subassembly 110-220kDa, bound to the B-box [153,194], with the TFIIIC2 subassembly 35-63-102kDa, bound to the A-box [154,196]. In addition to the intragenic control elements, 5' flanking sequences are likely to have some influence upon the transcription of most tRNA genes and numerous examples have been reported in *H.sapiens* [197-200]. The 5' elements that affect tRNA transcription are generally located within about 80 bp of the initiation site, but are often close to the transcription start site. In general, there is little or no extensive sequence homology conserved between the 5' flanking regions of different tRNA genes, even between the same isoacceptor species. One exception to this is the presence of short TATA motifs or A/T-rich regions which occur upstream of some tRNA genes in many organisms, but rarely in vertebrates [201].

The recruitment of RNA polymerase III on type 2 promoter genes follow the same pathway described above for the type 1 promoter, with the difference that TFIIIC2 is directly recruited to the target genes thanks to its high affinity for the B-box element that the 5S gene lacks (Fig. 2B). Initially, type 2 promoters are recognized by TFIIIC2, which binds strongly to the B-box and weakly to the A-box [146,155,194]. This DNA-protein interaction then allows the recruitment of TFIIIBβ which in turns allows the recruitment of RNA polymerase III. The role of TFIIIC1 in this process is still controversial. It was proposed that TFIIIC1 enlarges the TFIIIC2 footprint over the A-box of the promoters of tRNA-like genes [146], contributing to create an initiation complex that extends to approximately –40bp upstream of the start site [204]. Moreover, *in vitro* transcription with partially purified TFIIIC1 demonstrated that the factor was absolutely required for transcription of tRNA and VA1 genes [146,148,150,194]. The precise step at which TFIIIC1 is involved has still to be determined, but based on several studies some speculations can be brought forward. Analysis of the 5S gene revealed that further addition of TFIIIC1 to an isolated 5S-TFIIIA-TFIIIC2-TFIIIBβ complex led to a significant increase of transcription. Considering that, when TFIIIBβ primarily associates to an artificial TATA box-containing promoter under certain conditions, additional TFIIIC1 is required for a functional recruitment of RNA polymerase III [202], the same reasoning could be applied to type 2 promoters. Collectively, these data suggest that human TFIIIBβ cannot be regarded as the sole initiation factor for transcription, as had been firmly established for TFIIIB in yeast by Kassavetis *et al*. [187]; future investigations will have to reveal the exact function of TFIIIC1 during the final phase of the assembly of the transcription complex.

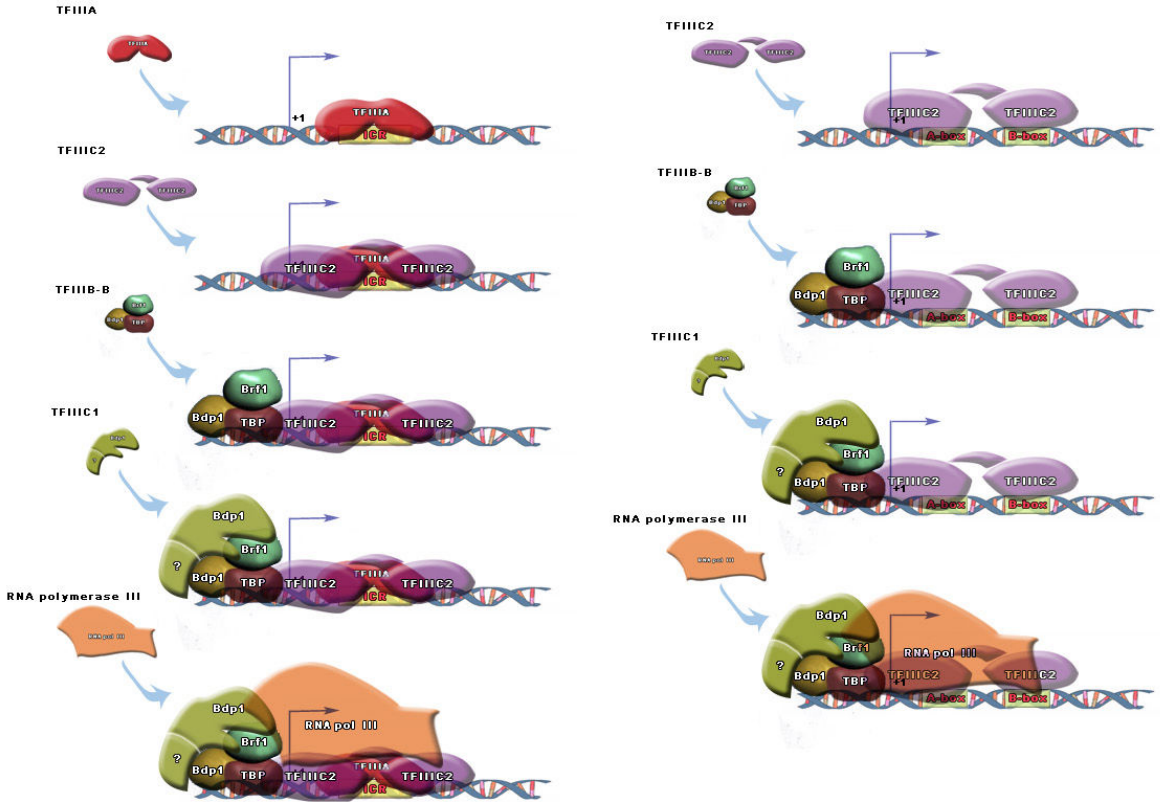


Figure 2. *Up to date representation of PIC (pre-initiation complex) assembly on human type 1 and type 2 promoters*. **A**. *TFIIIC2 recruitment on the type 1 promoter is mediated by the specificity factor TFIIIA, which recognize and binds the 5S rDNA at the ICR (internal control region), made up of an A-box, IE (internal element) and C-box. TFIIIC2 recruits the other transcription factors and RNA polymerase III in the order shown above: TFIIIC1 plays a key role in enhancing RNA polymerase III activity on the target promoter.* **B.** *PIC assembly on type 2 promoters is likely to follow the same route described above (only indirect data on TFIIIC1 recruitment by TFIIIB-β are available), with the difference that TFIIIC2 is directly recruited on the target genes thanks to its interaction with the A-box and B-box intragenic promoter.* ***?*** *indicates that the subunits composition of TFIIIC1 has still to be fully resolved.* ***+1*** *and the arrow indicate the transcription start site.*

Type 3 promoters

Set apart from type 1 and type 2 promoters, human type 3 promoters are characteristically located in the 5'-flanking region of the gene and consist of a proximal sequence element (PSE), a TATA-box and an upstream distal sequence element (DSE) which activate transcription from the PSE-TATA core promoter; the human U6 snRNA gene represent the most studied example of these kind of promoters, which were also found in the H1, 7SK and other genes of unknown function. A special feature in this promoter organization is that some elements are shared by genes transcribed by RNA polymerase II and III; both the DSE and the PSE can be interchanged between RNA polymerase II and III snRNA promoters with no effect on RNA polymerase specificity, which is determined by the presence or absence of the TATA-box [205]. As mentioned above, the core promoter of the human U6 snRNA gene consist of a TATA-box sequence (-29 to -24) and a PSE sequence (-66 to -47)

[207]. The distance between the two elements is under rigid constraints; when moved together while maintaining their original spacing, new start sites are efficiently selected, but none of them alone is able to specify a RNA polymerase III transcription start site [208]. On the other hand, the DSE is composed of various protein binding sites, but one of them is almost invariably the octamer sequence ATGCAAAT [209]. In addition, the DSE contains a second element referred to as the SPH element (originally identified in the enhancer of the *X.laevis* selenocysteine tRNA gene, whose promoter contains a PSE and TATA box [210]), which corresponds to the NONOCT element [211]. In the human U6 snRNA promoter, both the octamer and SPH elements stimulate the formation of pre-initiation complexes [212], as described in the next paragraph.

Assembly of a stable pre-initiation complex on the U6 snRNA gene can be summarized as the cooperative binding of TBP, $SNAP_C$ and Oct-1 (Figure 3). Oct-1 binds to its DSE site relatively weakly, but is stabilized at this site by interaction with DNA-bound $SNAP_C$. In turn, $SNAP_C$ binding to the PSE is strongly dependent on DNA-bound Oct-1, since a sort of "flap" in the C-terminal domain of $SNAP_C190$ occludes the PSE-binding domain located in the N-proximal third of the 190kDa subunit. By interacting with the POUs domain of Oct-1, the occluding flap is displaced and the DNA-protein complex at the PSE site is consequently stabilized [156,213]. Interestingly, a positioned nucleosome between the octamer sequence and the PSE brings into close proximity the two elements (separated by about 150bp), such that SNAPc and the Oct-1 POU domain can contact and recruit each other to the DNA [214]. Similarly, TATA-box binding to the TBP component of TFIIIBα is limited by its N-terminal part and deletion of this segment greatly increases the ability of the truncated protein to bind to TATA boxes [215]. Interaction with PSE-bound SNAPc relieves the autoinhibition of TBP which can form the characteristic sharply bent DNA complex through its C-terminal core [215].

Given that type 3 promoters do not contain binding sites for TFIIIA and TFIIIC2, they do not require none of these factors for the assembly of a pre-initiation complex; indeed, depletion of a nuclear extract with antibodies raised against TFIIIC220, affects transcription from the 5S, tRNA, and VAI genes but not from the 7SK and U6 snRNA genes [155] (it should be mentioned that a U6 snRNA variant, that assembles properly in the corresponding snRNP, is transcribed from a gene with type 2 promoter elements [216]). Conversely, TFIIIC1 has been shown to be required for transcription of the 7SK [217] and U6 [148] as it can generally strengthen the specific association of PSE binding proteins and promote RNA polymerase III recruitment.

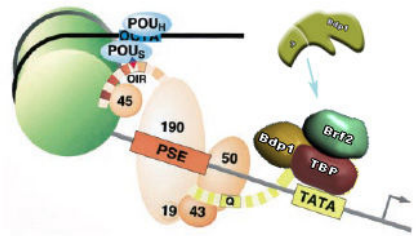


Figure 3. *Assembly of a stable pre-initiation complex on the U6 snRNA gene. Oct-1, SNAPc and TFIIIBα are recruited cooperatively on the respective DNA binding sites; a positioned nucleosome (shown in green) spanning a 150bp DNA region brings in close proximity the PSE and DSE promoting a tighter interaction between the respective binding factors. TFIIIC1 (composed as previously described in figure 2) positively promotes RNA polymerase III recruitment, but its interaction with the other components of the PIC is unknown.*

Type 4 promoters

A fourth type of RNA polymerase III transcribed genes is characterized by a combination of internal and external promoter elements and includes the *X.laevis* selenocysteine tRNA gene [222], the Epstein-Barr virus EBER genes [223], vertebrate vault RNA genes [224] and the human 7SL RNA gene [225] which is probably the best studied case of type 4 promoter genes. The three originally identified genes (Hs7SL1-3, [225]) contain an A-box and B-box -like promoter elements that do not suffice for fully efficient expression, but are absolutely required for human 7SL RNA transcription [226] and bound by TFIIIC2 [227]. To keep up with the high expression levels of 7SL RNA observed in human cells (about 1 million copies, expressed by the sole Hs7SL1 and Hs7SL-3, considering that Hs7SL-2 is apparently not expressed in vivo), additionally regulatory elements in the 5'-flanking region of these genes are required. First, a binding site for the transcriptional activator ATF/CRE between -43 and -50 was found to be involved in the expression of the human 7SL RNA [228]. The element was originally refereed as ATF as it binds the Activating Transcription Factor observed to stimulate transcription of multiple E1-A inducible adenovirus early promoters [229]; later on, sequence analysis demonstrated that the ATF and the cAMP response element (CRE) of the somatostatin promoter were almost identical [230] giving a more comprehensive picture of the ATF/CRE binding site. Secondly, a TATA-like sequence located approximately between -20 and -30 [226] was found to be crucial, completing the picture of a relative complicate arrangement of promoter elements, shared by the EBER and vRNA genes, that cooperate for efficient RNA polymerase III transcription.

THE TRANSCRIPTION CYCLE

Similarly to the other RNA polymerases, the transcription cycle of RNA polymerase III consists of four major steps: PIC assembly and transcription initiation, elongation, and termination. To date, there is little to no information on the mechanisms of initiation and elongation in human cells, but considering that components of TFIIIB, TFIIIC and the 17 subunit RNA polymerase III are highly conserved in structure and function from yeast to human [16,218], it is likely that they are broadly shared among these organisms (for a review and up to date details on yeast RNA polymerase III promoter opening and elongation, see [218-221]). More detailed studies are available in humans for the termination and re-initiation steps. First in *X.laevis* [231] and later for several human genes [205,232-236], transcription by RNA polymerase III has been shown to terminate within stretches of at least four consecutive dT nucleotides located in the 3 '-flanking sequences of class III genes. Albeit RNA polymerase III itself is able to terminate accurately by recognizing the simple dT-rich consensus, several auxiliary factors have been found to affect the efficiency of recognition in *H.sapiens*. The general RNA polymerase III transcription factor TFIIIC, the RNA polymerase II transcription coactivator PC4 and topoisomerase I interact with the terminator region of VA1 and tRNA genes and appear to facilitate RNA polymerase III termination and reinitiation [150,152]. TFIIIC0, a component of the human TFIIIC fraction functionally and chromatographically distinct from TFIIIC1 and TFIIIC2, shows strong DNase I protection of the terminator region of the $tRNA^{met}$, VAI and 5S rRNA genes [148]. Analysis of the VAI 3'-trailer identified a series of sites bound by a member of the NF1 protein family. NF1 interacts with the VAI terminator region and with two subunits of hTFIIIC2, promoting both accurate termination by RNA polymerase III and multiple round transcription (re-initiation) on the VAI template [176]; however, neither the NF1-binding sequence nor NF1 itself are present at the terminator region of many class III genes and the factor

might only be stimulatory for specific genes [65]. The La protein, which binds pre-tRNAs at their UUU-3'OH ends facilitating their maturation, has been proposed to play a role in transcript release and termination/re-initiation [161,169], an idea reinforced by the observation that non phosphorylated La is detectable at type 1-3 promoter genes in living HeLa cells [65]. Taken together, these data suggest that RNA polymerase III termination might be more complicated than what originally assumed by Bogenhagen and colleagues [231] almost 30 years ago.

Efficient production of proper transcripts requires efficient release of RNA polymerase III at the terminator and fast recruitment of the enzyme at the promoter elements, to keep up with the high levels of certain class III RNAs normally found in living cells. Template competition experiments in *S.cerevisiae* have shown that after the first cycle of transcription, the polymerase is committed to reinitiate on the same gene, a phenomenon referred to as facilitated recycling, which contributes to the high rate of RNA polymerase III transcription in yeast [245,246]. The process requires proper termination and PIC assembly, since promoter-independent transcription of tailed linear templates does not allow efficient recycling [245,247]. In humans, several activities have been shown to contribute to RNA polymerase III re-initiation, as mentioned above about NF1, PC4 and topoisomerase I. A recent study in human, which used a stalled ternary complex assembled on an immobilized VAI template, demonstrated that facilitated recycling is conserved from yeast to humans [248]. The same study, sorted out a novel role for Maf1 in RNA polymerase III transcription. Originally pointed as a tumor suppressor [249], human Maf1 is a potent inhibitor of RNA polymerase III transcription in HeLa cell extracts and also represses a subset of RNA polymerase II transcribed genes, including the one encoding for the TBP [249], thus regulating indirectly transcription by RNA polymerase I. Recombinant human Maf1 interacts with both Brf1 and RNA polymerase III in *in vitro* binding assays [250], but, rather than affecting re-initiation, it prevents RNA polymerase III recruitment to promoter-bound TFIIIBβ and TFIIIC and, to a lesser extent, binding of TFIIIBβ to promoter-bound TFIIIC [248]. Once RNA polymerase III has been engaged in transcription, facilitated recycling protects somehow the complex from Maf1 repression [248]. This recent discovery point out that several progresses are being achieved in understanding RNA polymerase III termination, but also that much remains to be found out in mammalian systems.

OUTPUTs OF RNA POLYMERASE III TRANSCRIPTION

tRNA

Transfer RNAs (tRNAs) are by far the best known RNA polymerase III transcripts. As adaptors between mRNA and proteins, tRNAs play an indispensable function during protein synthesis, converting the genetic information stored in nucleotides into the amino acid sequence. tRNAs form a characteristic L-shaped tertiary structure, organized in five domains known as acceptor stem, D-arm, anticodon-arm , TψC-arm and the variable loop: the acceptor stem and TψC-arm stack to each other to form a continuous α-helix, while the D-arm and anticodon-arm stack to form another continuous helix (Figure 4, for a review see [19]). Aminoacyl-tRNA synthetases and modifying enzymes recognize the overall shape of tRNAs, but strictly select their cognate tRNA or catalyze definite reactions by features present in localized domains of the tertiary structure.

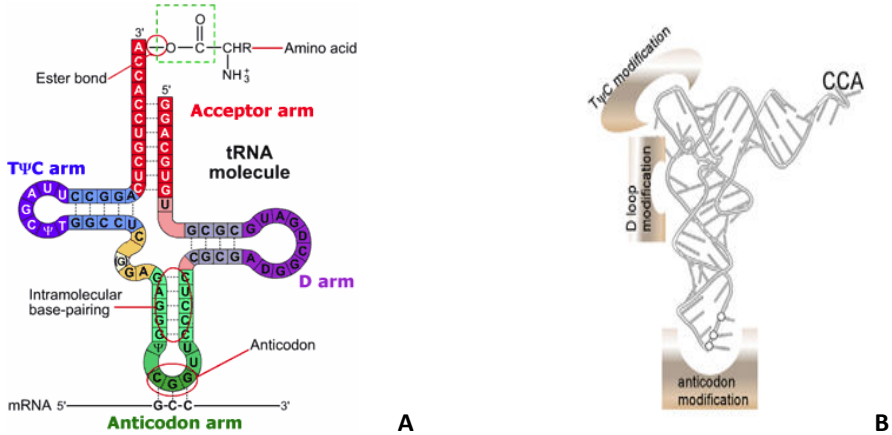


Figure 4. *tRNA secondary (**A**) and tertiary structure (**B**), "Current opinion in structural biology". Arranged around the cloverleaf structure, the brown blocks indicate the major sites of precursor tRNA recognized by processing and modification enzymes.*

Almost 50 years ago, during a meeting at the King's College in Cambridge, Jacob, Brenner and Crick laid the basic concepts of genetic control of protein synthesis. Given that DNA has 61 codons for 20 natural aminoacids and three stop signals, the genetic code is degenerate, in that most aminoacids are encoded by more than one codon. How tRNAs can recognize more than one codon is given by Francis Crick's Wobble Hypothesis [20]. Crick proposed that the two first base pairs between tRNA nucleosides 36 and 35 and the coding triplet in mRNA would be canonical Watson-Crick pairings of a purine with a pyrimidine and *viceversa*. The third base pair could also be canonical, but to expand tRNA recognition of codons in protein synthesis, five additional non-canonical base pairs were possible: two are G-U and U-G base pairs, while the remaining three involved the nucleoside inosine (I), able to base pair with A, U and C. Nowadays, Crick's attention to the importance of the canonical geometry for the first two base pairs has proven absolutely correct. In 1991, P.F. Argis complemented the original hypothesis in suggesting that some tRNA base modifications had evolved to counter or facilitate the wobble [58], rather than the ribosome inducing a correct architecture for the anticodon to fit the decoding site as originally supposed by Crick. Today we

know that precise aminoacylation of cognate tRNA and selection of the correct anticodon by the ribosome, vouch for high accuracy in protein synthesis, with one wrong incorporation in 10^3-10^4 aminoacids [59,60].

tRNA processing and maturation

Eukaryotic tRNA is transcribed by RNA polymerase III as a precursor molecule (pre-tRNA) with additional sequences at both 5' and 3' ends. To fulfill its adapter function, the tRNA must be processed into the standard length and be post-transcriptionally modified; this modification step is essential for the tRNA to maintain the canonical L-shaped structure, which is required for its decoding function.

The first step in pre-tRNA processing is the endonucleolytic cleavage of the 5'-leader by RNaseP, a ribozyme which catalyzes the hydrolysis of a specific phosphodiester bond, leaving a phosphate group at the 5'-end of mature tRNA and a hydroxyl group at the 3'-end of the leader [38]. RNase P has a catalytic RNA core (Rpr1 in *S.cerevisiae* [40], H1 in *H.sapiens* [41]) that has adapted to prokaryotic and eukaryotic cellular environments with the addition of various protein subunits. Given the poor conservation of the sequences surrounding the cleavage site, both elements of the ribozyme have evolved to recognize common structural elements of the tRNA cloverleaf [38], but cleavage efficiency may also be affected by the surrounding 5'-leader and 3'-trailing sequences [42], or by accessory proteins bound to the pre-tRNA, such as the human La autoantigen [39]. While most of the catalytic mechanism has been solved out, little is known about the subcellular localization of RNaseP; pre-tRNA biosynthesis had been assumed to be distributed throughout the nucleoplasm, but emerging data suggest this might not be the case. In *S.cerevisiae*, most pre-tRNAs and RNaseP RNA are localized primarily to the nucleolus, with additional foci in the nucleoplasm [43]. Similarly, H1 RNA and protein subunits of human RNaseP have been shown to shuttle among distinct cellular compartments, including the nucleoplasm, Cajal bodies and nucleolus, suggesting that RNaseP is also linked to active gene transcription [44]. Remarkably, as determined by ChIP analysis, a recent study has shown that many subunits of RNaseP can be found associated with chromatin of tRNA, 5S rRNA and U6 snRNA genes in rapidly dividing HeLa cells [45], indicating that RNaseP subunits could act as general transcription factors for class III genes.
Compared to 5'-end processing, removal of the 3'-trailer is poorly understood to date. It is known from earlier work in *E.coli* that CCA-containing and CCA-less pre-tRNAs are selectively processed by exonucleases or endonucleases, respectively. In the exonucleolytic pathway, RNaseII and PNPase trim the longer 3'-trailers leaving the last few nucleotides for RNasePH and RNaseT processing [46]. In the second pathway, RNaseE is likely the initial endonuclease [47], while tRNase Z cleaves after the discriminator nucleotide of CCA-less pre-tRNAs [48]. The 3'-end processing machinery in eukaryotes is much more elusive; in contrast with the situation in bacteria, eukaryotic nuclear tRNA genes do not encode the 3'CCA sequence and a few factors cooperate in promoting the endonucleolytic processing. One of these is the yeast homologous of the human La autoantigen (Lhp1p in S.cerevisiae), an abundant eukaryotic RNA-binding protein that associates with the 3' oligo(U) residues of many RNA polymerase III derived RNAs, impeding the accession of 3' to 5' exonucleases to the 3'-end of pre-tRNAs [49]. Since there is no indication that La possesses endonuclease activity, a candidate has been recently identified in the yet uncharacterized eukaryotic tRNAse Z. At least one long (in *S.cerevisiae* and *D.melanogaster*) and one long and one short (in *H.sapiens*) forms of tRNase Z have been identified so far [50]. Both enzymes show

processing activity on a variety of substrates, but the majority of the data originates from *in vitro* experiments [51]; recent data in *S.pombe* suggests that tRNaseZ^{Long} is functionally conserved among *S.pombe*, *S.cerevisiae* and *H.sapiens*, but the limited amount of *in vivo* information impedes a comprehensive understanding of its biological functions [52].
Once the tRNA has been cleaved after the discriminator nucleotide, a unique nucleotidyl transferase catalyzes the post-transcriptional addition of CCA onto the 3' terminus of immature tRNAs [53]. The ability of the CCA-adding enzyme to add specific nucleotides in the absence of a nucleic acid template makes it one of the most intriguing polymerases, but the central question regarding the specificity of the CCA-adding enzyme in eukaryotes remains unanswered. Nonetheless, sequence homology among different kingdoms suggests a conserved mechanism that has been recently defined thanks to the crystal structure of a few archaea/bacteria CCA-adding enzymes [54,55]. The enzyme has a L-shaped nucleotide binding site where the terminal nucleoside of tRNA is projected to stack with the upcoming NTP; discrimination between CTP and ATP at a given addition step and at termination arises from conformational changes of the catalytic pocket that is progressively altered by the elongating 3' end of the tRNA [54].

Despite their complexity, it is generally accepted that the mechanisms described above are restricted to eukaryotic cells nuclei. According to the current view, splicing of intron-containing tRNAs occurs before the resulting mature molecule is exported into the cytoplasm, where it receives the appropriate aminoacid. However, recent studies have shown that several of the conventional wisdoms are incorrect; tRNA aminoacylation does not occur solely in the cytoplasm and splicing does not occur solely in the nucleus. In yeast, the splicing of intron-containing tRNAs occurs on the cytoplasmic surface of mitochondria [56], where the four subunits of the tRNA-splicing endonuclease are found [29]. In humans, pre-tRNAs are spliced in the nucleus before excision of the 5′ and 3′ extensions and it is generally accepted that the splicing endonuclease is restrained in this subcellular location; since the four subunits of the splicing endonuclease were found associated with pre-mRNA 3′ end processing factors [57], it is not surprising that the enzyme is restrained into the nucleus, where it is likely to function in multiple RNA processing events.

As the last step in tRNA maturation, post-transcriptional modification of tRNA residues involves a significant amount of genes, enzymes and energy. In the past years, progress has been more rapid and complete for the yeast *S.cerevisiae*; nonetheless, the tRNA evolutionary conservation and the fact that many modifications have been described in various organisms, suggest that most tRNA modifications are broadly shared among higher eukaryotes [37]. Of the seventy-five modifications found in tRNAs, those occurring in the physically and functionally separable anticodon stem and loop domain (ASL) are best understood. Position 34 and 37 are by far the two most modified nucleosides, and have the highest variety of modification chemistries found among all RNAs. At wobble position 34, post transcriptional deamination of A_{34} to I_{34} is the rule, rather than the exception [130], as I_{34} improves wobble base-pairing and, unlike A_{34}, would not weaken ribosomal A-site binding when in the P-site. Other modifications at nucleotide 34 appear to expand codon recognition by tRNA; though Crick was correct in that a tRNA U_{34} has not been observed to bind to U3 of an NNU codon on the ribosome (where N is any of the four major nucleosides), mcm^5U_{34} (2-thio-5-methoxycarbonylmethyluridine) and its 2-thio-derivatives [37] are commonly found at the wobble position of eukaryotic tRNAs, as they expand or facilitate the wobble [58]. Similarly, modifications of Pur_{37} are very important as they negate translational frameshifting. For example, addition of hydrophobic or hydrophilic moieties to Pur_{37} facilitates codon/anticodon interactions

and negates intra-loop hydrogen bonding, retaining the open loop structure required for codon binding [131,132]. Many other chemistries have been described so far and are extensively listed in [133].

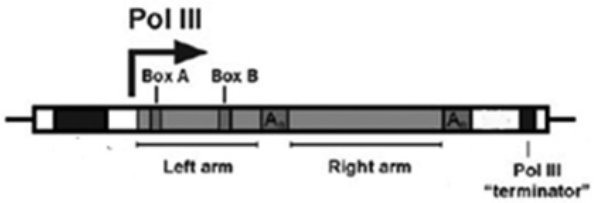
tRNA loading and its role in translation.

All aminoacyl-tRNA synthetases catalyze a high fidelity two-step aminoacylation reaction. This entails condensation of the amino acid with ATP to form an activated aminoacyl adenylate intermediate, followed by transfer of the amino acid to the 3'-terminal ribose of tRNA to generate the aminoacylated product [22]. Despite their conserved mechanisms of catalysis, aminoacyl-tRNA synthetases can be broadly divided into two classes (I and II), each of which has developed exclusive sequence motifs at the active site and tRNA-binding site [23]. Briefly, it has been shown by several groups that AARSs mostly interact specifically with one or more of the discriminator base (N73), the acceptor stem, and the anticodon [21], so that each amino acid becomes associated with one or more anticodon sequences in the cognate tRNA isoacceptor set. Aminoacylated tRNAs are thence needed in the cytoplasm for ribosomal translation and exported from the nucleus through a specific receptor, exportin-t, which binds directly to tRNAs in a RanGTP-dependent manner [24,25]. The receptor properly recognizes secondary and tertiary structural elements, so that fully processed, mature tRNAs are preferentially selected by the tRNA export machinery [26,27]. However, exportin-t is not able to discriminate between spliced and unspliced tRNAs [27,28], in accordance with a study reporting that the yeast tRNA-splicing endonuclease is associated with the outer mitochondrial membrane [29]. Surprisingly, these mature tRNAs can be re-imported into the nucleus [30], suggesting that the export machinery is involved in some sort of tRNA quality control.
Loaded tRNAs are ready to initiate the translation process, which involves the coordinated action of ribosomes, mRNA and a plethora of initiation factors. Translation has been studied extensively during the last half century, but most of our current knowledge culminated in the X-Ray crystal structures of prokaryotic 70S ribosome functional complexes [31]. A glimpse of the mechanism of tRNA recognition by an eukaryotic ribosome comes from the observation of the yeast 80S structure [32]. This and other studies [33,34] support the idea that most of the fundamental activities of the ribosome are the same from prokaryotes to higher eukaryotes, and are in agreement with the common model which considers the tRNA to bind the ribosome at the three different sites A (aminoacyl), P (peptidyl) and E (exit), respectively [35]. Principles of ribosome assembly and eukaryotic translation are reviewed at [36], but won't be addressed here since they are beyond the purpose of this thesis' work.

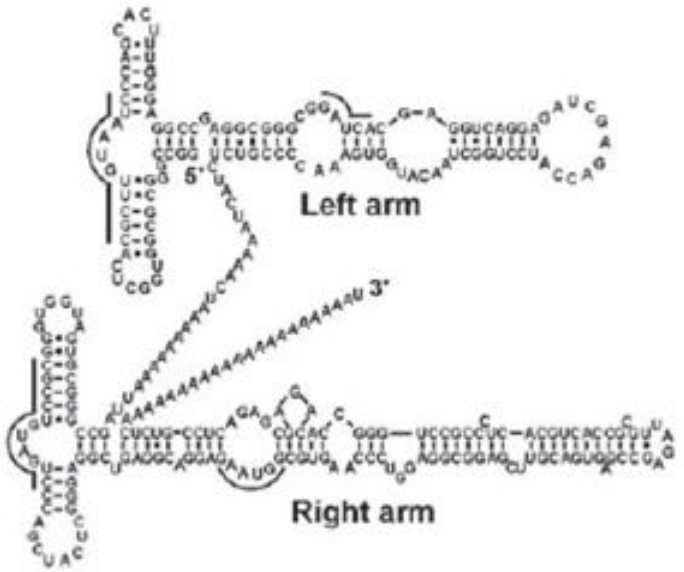
Alu RNA

Repetitive elements account for at least 50% of the human genome [96] and can be broadly subdivided into those that are tandemly arrayed or interspersed. The last ones mainly derive from four types of transposable elements, of which one transposes directly as DNA (DNA transposons) and three transpose through RNA intermediates (LINEs, SINEs and LTR retroposons) [97-100]. SINEs (Short Interspersed Nuclear Elements) are the most extensively studied human retroposons and include three distinct monophyletic families; the active Alu, and the inactive MIR and Ther2/MIR3. With more than one million copies, Alu repeats are the most abundant repetitive elements in the human genome [96]; altogether they cover about 10% of the whole genome and

occur at higher frequency within non coding regions of gene-rich areas [101]. Alu elements are approximately 300bp in length and consist of two similar but distinct monomeric units derived from the 7SL RNA gene [102]. The two monomers are held together by an A-rich linker and terminated by a poly(A) tail of variable length which probably played an important role in retroposition [103]. Although the mechanism still remains to be elucidated, the amplification of Alu elements is thought to occur by reverse transcription of an Alu-derived RNA polymerase III transcript [104,105]. Indeed, Alu elements inherited from the 7SL RNA gene an internal promoter closely resembling the type 2 promoter of tRNA genes [106] and are thus capable of autonomous transcription by RNA polymerase III (Figure 5B, [107]). An RNA polymerase III terminator is not included in the sequence itself, even if a termination signal is often present in the downstream genomic region allowing the synthesis of a non-coding Alu RNA which probably borrows the factors required for retroposition from LINE-1 [108].



5A



5B

Figure 5. A. *Representation of an Alu element transcribed by RNA polymerase III. The left arm and right arms derive from rearrangements of the 7SL RNA gene. Transcription is terminated at an RNA polymerase III termination signal, located in 3'trailer at a variable distance from the second poly(A) tail.* **B.** *Secondary structure model of the HAFP Alu RNA based on digestion analysis, from JBC,1991, 266:8675-8678. The two related monomers are bound by a poly(A) tail of variable length.*

7SL RNA

The signal recognition particle (SRP) is a multisubunit ribonucleoprotein responsible for the cotranslational targeting of secretory and membrane proteins to the endoplasmic reticulum in eukaryotes or to the plasma membrane in prokaryotes. The cycle is initiated when the SRP recognizes the N-terminal sequence of nascent polypeptide chains as they emerge from the ribosome; the binding of the signal sequence transiently arrest elongation and targets the complex to the ER membrane via a GTP-dependent interaction with the SRP receptor. Subsequent GTP hydrolysis dissociates the SRP–SRP receptor complex, allowing SRP to initiate a new round of targeting, and leaving the nascent chain directed through the translocon or integrated into the membrane (for a comprehensive review, see [120]). The SRP consists of a single 7SL RNA molecule bound to two heterodimeric protein subunits (SRP 68/72 and SRP 9/14) and two monomeric polypeptides (SRP 19 and SRP 54), as depicted in Figure 6A. The roughly 300nt long sequence of the human 7SL RNA can be divided into an Alu domain and an S domain. The S-domain comprise the central sequence, which is unique to the 7SL RNA, bound to SRP 19, SRP 54 and the heterodimer SRP 68/72; SRP 19 significantly enhances the attachment of SRP 54, which has a key role both in the recognition of the signal sequence and in the interaction with the SRP receptor [122]. The Alu domain consist of the heterodimer SRP 9/14 bound to the 5' and 3'-ends of the Alu related sequence of the 7SL RNA and it plays a major role in mediating the elongation arrest function of the particle [121]. Given that each monomer of the Alu RNAs is highly similar to the 7SL-Alu domain (compare Fig. 5B and 6B), it was supposed that AluRNAs might share some of the SRP functions in mammalian cells; indeed, it was shown that each monomer of AluRNAs can bind the cognate SRP protein SRP 9/14 *in vitro* [123] and *in vivo* [124], possibly explaining that SRP 9/14 are present in a large excess over SRPs in mammalian cells [125].

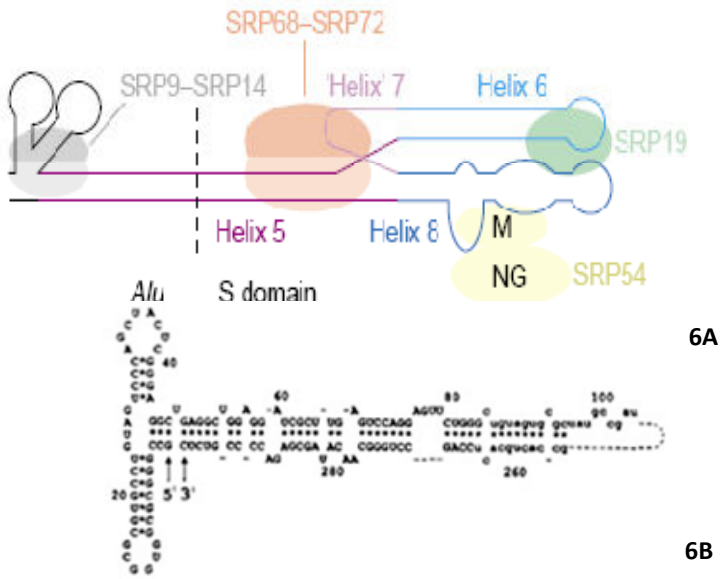


Figure 6. A. *Schematic representation of eukaryotic SRP, adapted from Current opinion in Structural Biology, 2003, 13:64-70. The 300nt long 7SL RNA is bound by six protein components indicated as SRP68/72, SRP9/14, SRP19 and SRP54.* **B.** *Secondary structure model of the 7SL RNA based on digestion analysis, from JBC,1991, 266:8675-8678. The figure shows only the 7SL RNA sequence segments homologous to the AluRNA.*

5S ribosomal RNA

The smallest RNA component of the large ribosomal subunits is a 120nt long RNA polymerase III transcript known as 5S rRNA. Regardless of its origin, 5S rRNA can be folded into a common secondary structure consisting of five helices (I-V), two hairpin loops (C and E), two internal loops (B and D) and a hinge region (A) [for a review, see 79]. In eukaryotes, 5S rRNA is exported from the nucleus to the cytoplasm in a complex with TFIIIA as a 7S RNP and then re-enters the nucleus in a complex with protein L5; notably, 5S is bound to a 16-aminoacid region on the nascent chain of L5 which is strictly species-specific [80]. The exchange of 5S rRNA between 7S RNP and L5 releases TFIIIA, which can return to the nucleus and activate the synthesis of more 5S rRNA, establishing a tight link between 5S rRNA and L5 biosynthesis [81]. To date, the full picture of interactions of 5S rRNA within the ribosome of higher eukaryotes has still to be solved. The standard reference for ribosome studies is the crystal structure of the 50S subunits of the *H.marismortui* ribosome [82]; the emerging picture is that 5S rRNA makes only direct contacts with 23S rRNA and the three ribosomal proteins L18, L5 and L21, but data in yeast or human cells still need to be confirmed.

U6 snRNA

The removal of introns from mRNA precursors has an indispensable role in eukaryotic gene expression; this two-step transesterification reaction is performed by the spliceosome, one of the most complicated cellular machines comprising five small nuclear RNA molecules (snRNAs) and a hundred or more proteins [61]. Spliceosomes appear to assemble de novo on each pre-mRNA in a stepwise fashion, with spliceosomal components joining the assembly either singly or in pre-formed complexes. Initially, the 5' splice site of the pre-mRNA is recognized by the U1 snRNP. Then, U2 snRNP binds to the branchpoint sequence located near the 3'-end of the intron. Finally, the spliceosome is completed by incorporation of the U4/U6/U5 tri-snRNP [61]. The splicing reaction requires an extensive network of RNA–protein and RNA–RNA interactions, including the unwinding of the U4/U6 snRNA duplex and the formation of a U2/U6/pre-mRNA structure [62]. Among the five small nuclear RNAs, U6 snRNA has the remarkable feature to be the only snRNA transcribed by RNA polymerase III, as shown by several *in vitro* assays performed in Novikoff hepatoma whole cell extracts [63] or isolated HeLa cells nuclei [64]; *in vivo* data on RNA polymerase III transcription of human U6 snRNA are still missing today, but several indirect analysis, such as ChIP analysis [65] and efficient immunoprecipitation of U6 snRNA by an antibody against the La antigen [64], support this hypothesis (it should be noted that *S. cerevisiae* U6 snRNA is efficiently transcribed *in vivo* by RNA polymerase III, since a mutant strain with a temperature-sensitive defect in the large subunit of the enzyme shows a defect in U6 RNA levels [66]). Another peculiar feature of U6 snRNA is its unusual CAP structure, characterized by a γ-monomethyl phosphate at the 5'-end [67]. Finally, it must be mentioned that U6 snRNA molecules have a remarkable heterogeneity, resulting from extensive posttranscriptional modification of their 3'-termini catalyzed by a U6-specific 3'-exonuclease [68] and a terminal uridylyltransferase (TUTase) [69]. All these peculiarities support the structural significance of U6 snRNA and it is conceivable that they are essential for the assembly and catalytic function of spliceosomes.

H1 RNA

Purification of human RNase P from HeLa cells has revealed the existence of an RNA subunit, called H1 RNA [70]. The RNA moiety of some bacteria and archaeal RNaseP has been demonstrated to be directly involved in catalysis and, as such, can be considered to be remnants from an RNA world. In 2007, it was shown that also H1 RNA mediate cleavage of four tRNA precursors and a model RNA hairpin loop substrate in the absence of protein. Compared with bacterial RNase P RNA, the rate of cleavage was five to six orders of magnitude lower, whereas the affinity for the substrate was reduced 20- to 50-fold [71], suggesting that the catalytic activity is conserved in all domains of life.

7SK RNA

In eukaryotic cells, phosphorylation of Ser2 at the carboxy-terminal domain (CTD) of the largest subunit of RNA polymerase II plays a critical role in transition from transcriptional initiation to elongation [72]. The reaction is catalyzed by the positive transcription elongation factor b (P-TEFb) and is essential for the production of full-length mRNAs of most protein-coding genes [73]. Biochemical experiments suggest that cellular P-TEFb exists in two forms, the active heterodimeric CDK9/cyclin T and the inactive, reversible form containing the 7SK small nuclear RNA and HEXIM1 or HEXIM2 proteins [74,75]. The 7SK RNA is an abundant RNA polymerase III transcript that shows high sequence conservation in vertebrates [76,77]. In HeLa cells, about 30% of 7SK snRNA is sequestered into the 7SK/HEXIM/P-TEFb, where it has been proposed to play a central role in providing the structural platform for the coordinated assembly of HEXIM1/2 and P-TEFb. A recent study demonstrated that two distinct regions of 7SK RNA directly recruit HEXIM1 and P-TEFb in living cells; the G24-C48/G60-C87 distal segment of the 5'-terminal hairpin is essential for recruitment of both P-TEFb and HEXIM1, while the G306-C324 apical domain of the 3'-terminal hairpin contains elements that function only in P-TEFb binding [78]. The same group also shown that, in spite of being present in large excess compared to HEXIM1 and P-TEFb, nuclear 7SK RNA is limiting for the formation of the 7SK/HEXIM1/P-TEFb complex; a possible interpretation is that the major portion of 7SK RNA is sequestered into another RNP complex, thus playing a more complex role in transcriptional regulation.

Vault RNA

Vault complexes are large ribonucleoparticles highly conserved among higher and lower eukaryotes [86,87], yet probably not essential for eukaryotic cells in general as they were not detected in *S.cerevisiae* [88] and *D.melanogaster*. Analysis of mammals vaults revealed that the complex comprise three distinct high molecular weight proteins and several small vault RNAs (vRNAs) that account for about 5% of the total mass [86]. Despite the characterization of individual vault components and the development of a detailed structural modeling recent years, the cellular function of vaults has still not been fully elucidated. Most recent studies link vaults to nucleocytoplasmic transport and cellular detoxification; in favor of this second hypothesis is the fact that the major vault protein is highly expressed in tissues that are chronically exposed to elevated levels of xenobiotics (e.g. bronchus epithelium), metabolic active tissue and macrophages [89]. Intriguingly, variations in the relative amount of vRNAs associated with vaults have been recently observed in multidrug-resistance cancer cell lines suggesting their possible involvement in tumor development [90]. The role of these small RNA polymerase III transcripts (hvg1, hvg2 and

hvg3 in *H.sapiens*) has still to be grasped, but it is likely that they play a functional rather than structural role in the complex as degradation of vRNAs does not affect vaults morphology [91].

Viral RNAs

Together with the notable RNA molecules listed above, there is a wide range of less known RNA polymerase III transcripts poorly studied or selectively expressed by certain organisms. The list include some viral RNAs, such as the VA-I and VA-II RNAs from adenovirus and EBER1 and EBER2 RNAs from EBV [83,84], which are transcribed by cellular RNA polymerase III with the aid of associated factors TFIIIB and TFIIIC. Curiously, these RNAs contain a triphosphate at their 5'-end [85] which makes them clearly distinguishable from endogenous RNA polymerase III transcripts.

RNA polymerase III-derived miRNAs

A growing and still poorly developed field is that of viral/human microRNAs (miRNAs) derived from RNA polymerase III transcription. In 2004, the discovery of miRNAs in the herpesvirus family indicated that viruses have evolved the ability to exploit RNA silencing for regulation of both host and viral genes [92]. Primary transcripts of miRNAs generally originate as RNA polymerase II transcripts, but in 2005 it was discovered that tRNA promoters in the genome of murine gammaherpesvirus 68 (MHV68) are able to drive miRNA expression in the host cell [93]. The upstream tRNAs are surprisingly expressed by the virus, but not aminoacylated [94], suggesting that these tRNA-like molecules are non-functional remnants required more for their promoter capabilities than for their classical role in translation. More recently, it was shown that miRNAs can arise from RNA polymerase III also in the human genome. C19MC Alu can function as a promoter for biosynthesis of four downstream miRNAs [95], suggesting a connection between miRNAs and transposable elements.

Chapter 2

Revisiting the minimal requirements for human RNA polymerase III termination at non canonical sites

Since the original finding that RNA polymerase III is able to terminate transcription at a simple stretch of consecutive dT residues [231,252], little progresses have been done to further elucidate the significance of these termination sequences in the human genome. The observation that ~¼ of human tRNA genes are enriched in canonical terminators located at ≥50 bp from the 3'-end of the tRNA coding sequence led us to the identification of several novel non canonical terminators located further upstream at which RNA polymerase III terminates with different efficiency, generally lower than those observed for termination at the canonical $\geq T_4$ sites. Considering that more than one million of type 2 promoter elements are scattered through the human genome (tDNAs, Alu, MIR, etc), read-through at non-canonical sites could end up in the biogenesis of primary RNAs, with potential regulatory functions, embedded in the 3'trailer of predicted type 2 promoter transcripts. Relaxed RNA polymerase III terminators could thus represent the evolution of a more sophisticated regulatory mechanism for the biogenesis of small RNAs.

Introduction

As the final step in the transcription cycle, proper termination is particularly critical for successful gene expression. Primarily, it allows nascent RNA release and facilitates the recycling of RNA polymerases for further rounds of transcription, but also ensures that other promoters are not perturbed by read-through polymerases that have failed to terminate at upstream genes. In mammals, the three nuclear RNA polymerases use different strategies to keep up with these requirements. RNA polymerase II termination can occur anywhere from a few base pairs to several kilobases downstream from the mature mRNA. Considering that the precursor RNA has to undergo several processing events to produce the fully translatable mRNA, a plethora of coordinated *trans*-acting factors and *cis*-acting elements, such as an intact polyadenylation signal at the 3'-end of protein coding genes, is required to couple the transcription events with termination and 3'-end processing [251]. RNA polymerase I and III termination is much less complicated and apparently involves significant fewer factors than RNA polymerase II. The key event in transcription termination by RNA polymerase I is the binding of the transcription termination factor TTF-1 to an 11/18 bp sequence element (Sal-box) located downstream of the pre-rRNA coding region [253,254]. TTF-1 stops elongating RNA polymerase I when bound to the terminator, but it is not sufficient for the dissociation of paused ternary complexes; an additional dT-rich signal and the transcript release factor PTRF are required to fulfill this function, as well as for proper 3'-end processing of pre-rRNAs [255,256].

Transcription termination by RNA polymerase III seems to be even simpler than the previous models, as it was shown to occur within a stretch of consecutive dT residues located in the 3'-flanking sequence of class III genes [231,252]. Apparently, RNA polymerase III itself possesses transcription termination activity without further requirements of associated factors [243]. Studies in *S.cerevisiae* suggest that the termination activity resides in the trimeric subcomplex C53/37/11: the heterodimer C53/C37 is likely to be directly involved in termination, allowing for increased pausing time at the terminator elements [17], while C11 appears to play a major role in re-initiation [12,17]. It is not yet clear the exact mechanism by which RNA polymerase III releases the nascent RNA once found a proper terminator, but early studies in Xenopus by Campbell and Setzer suggest a two-step pathway. According to the authors, RNA polymerase III has the intrinsic property to pause at several termination signals during the elongation step, but it is unable to release the transcript as far as the latter is anchored by a persistent DNA:RNA hybrid; once the polymerase finds a proper terminator that allows for a weaker DNA:RNA hybrid, and eventually proper surrounding sequences, the nascent RNA strand can be displaced, allowing for polymerase and transcript release [244]. In humans, RNA polymerase III still retains its ability to autonomously recognize a termination signal, but several auxiliary factors have been found to affect the efficiency of recognition, such as the RNA polymerase II transcription coactivators PC4 and topoisomerase I, NF1 and the La protein. Taken together, these data point to possible different requirements for yeast and mammalian RNA polymerase III termination. *S.cerevisiae* RNA polymerase III requires a somewhat longer stretch of five/six consecutive dT residues for efficient termination [241,257] than does Xenopus RNA polymerase III, which can terminate with high efficiency at a sequence with as few as four consecutive dT residues [231]. The latter is true in humans, even if it shall be noted that, despite the efficiency of the TTTT sequence as a termination signal, RNA polymerase III termination might be affected by the sequence context in which the dT stretch is embedded [242] and read-through at canonical terminators have been observed in a few cases. For example, transcription termination of the adenovirus 2 VAI gene occurs principally at a

run of four dTs located 160 nt from the start site, but a small proportion of transcripts proceed to a backup termination signal of five dTs located 35nt further downstream [232,233]. The emerging picture is that at the yeast terminator, the RNA polymerase III/DNA/RNA complex may be less stable than the human one, as it possesses a longer U:A heteroduplex which could lead to an easier displacement of the nascent transcript. In contrast, the shorter mammalian terminators might require the addition of accessory factors, which could eventually facilitate polymerase and RNA release, to terminate transcription efficiently. In brief, it is clear that there is still room to gain further insights into that matter.

Human tRNA genes have unusually distant RNA polymerase III canonical terminators

In lower eukaryotes a cluster of dT residues is usually present within the first 20 bp following a tRNA gene, thereby giving rise to precursor molecules with relatively short 3' trailer sequences that must be removed during processing [257]. The elusiveness of this definition was further assessed by analyzing the distribution of RNA polymerase III terminators in the 3'-end of the 274 *S.cerevisiae* tRNA genes [259,260]. Canonical terminators made up of five or more consecutive dT residues were sorted in 5nt bins covering the whole 1kb downstream flanking region of each tRNA gene. For each bin, the percentage of tRNA genes containing a properly located canonical terminator was plotted as a function of the distance from the first nucleotide downstream of the mature tRNA sequence and represented with a cumulative distribution function. As shown in **Figure 1A** (blue line), more that 97% of *S.cerevisiae* tDNAs terminate within the first 15 bp, in agreement with the prevailing view on RNA polymerase III termination. Considering that eight tRNAs in the genome of *S.cerevisiae* are tandemly arrayed in bicistronic units which have lost the intergenic terminator [258,259], the data point to strict termination requirements immediately downstream of the tRNA coding region. To further extend our analysis, we focused on the 675 tRNA genes identified in the human genome by Pol3Scan [259], now considering as a canonical terminator a stretch of four or more consecutive dT residues. As depicted in **Figure 1A** (red line), the percentage of genes with a canonical terminator within the first 15 bp drops to 50-60% and the curve approximates the plateau much slower than previously shown for *S.cerevisiae*. The difference is striking when the fraction of tRNA genes containing a canonical terminator in each 5nt bin is plotted in the histogram of **Figure 1B**. With respect to *S.cerevisiae*, human tRNA genes are surprisingly enriched in canonical terminators located at ≥50 bp from the 5'-terminal nucleotide of the 3'-trailer (~23% of the 675 tDNAs), suggesting that human RNA polymerase III transcription spans a broader region downstream of type 2 promoter genes.
Considering that RNA polymerase III transcripts of well known function are usually longer than 70-100nt and that short regulatory transcripts, such as microRNAs, are normally processed into the standard length from ~60-80 nucleotides hairpin precursors by the cytoplasmic RNase III Dicer, we considered tRNA genes harbouring a 3'-trailer longer than 50nt of potential interest. The ability of RNA polymerase III to transcribe several of these genes was thus experimentally assessed in *in-vitro* transcriptional assays.

Human RNA polymerase III terminates at novel non-canonical terminators *in-vitro*

In-vitro transcription of eight human tRNA genes representing a wide variety of ≥50 bp long 3'-trailers (**Figure 2**) was assayed in HeLa Nuclear Extract under standard conditions (*see materials*

and methods). As shown in **Figure 3A**, while two of the selected tRNAs gave a barely detectable signal or a varied pattern without a clear main termination event (lane 1-2), the remaining genes

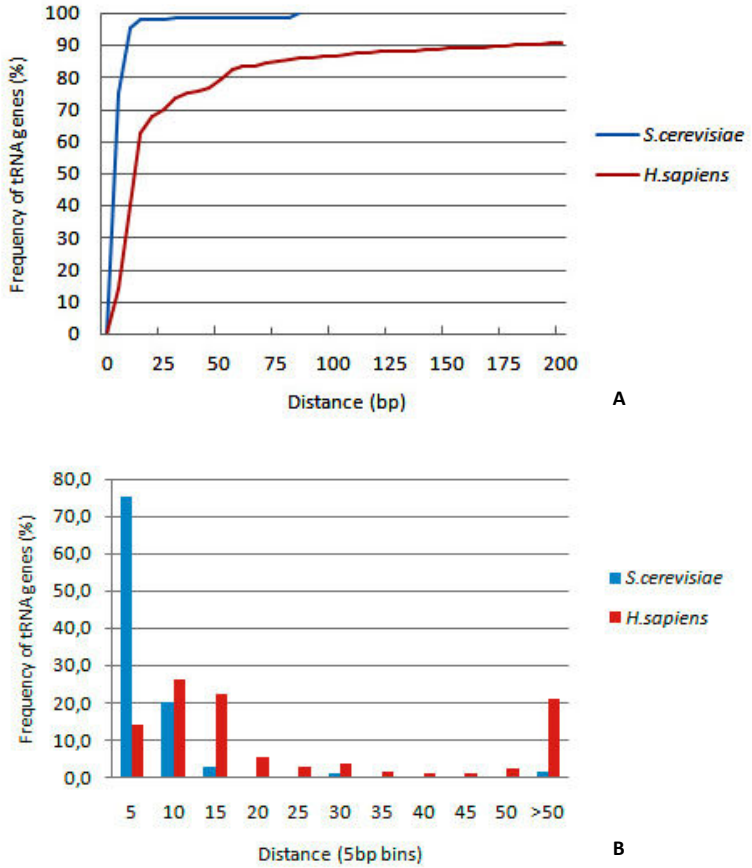


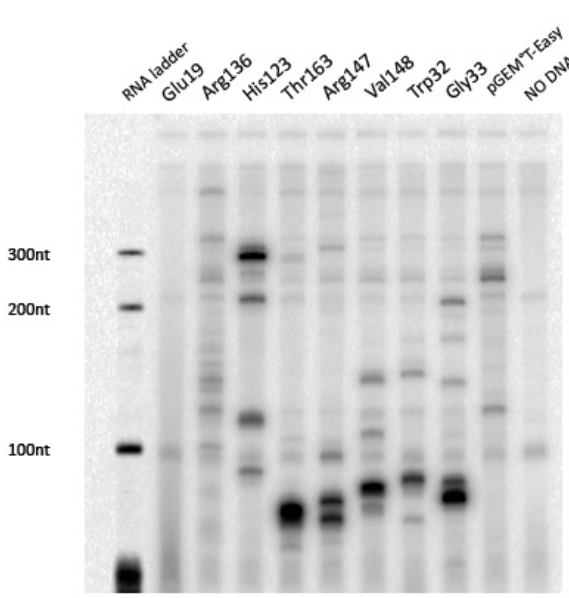
Figure 1. Distribution of RNA polymerase III canonical terminators in the 3'-trailer of tRNA genes. **A**. Cumulative distribution of canonical terminators in the 3'trailer of 274 *S.cerevisiae* and 675 *H.sapiens* tRNA genes, identified with the software Pol3Scan [259] as described in *materials and methods*. Numbers on the x-axis indicate the relative positions of the terminator element with respect to the first nucleotide of the 3'-trailer. Distances were clustered in 5bp bins and each terminator was classified in the relative bin depending on the position of the first dT. Canonical terminators are defined as a stretch of at least four (*H.sapiens*) or five (*S.cerevisiae*) consecutive dT residues. The chart does not take into account terminators located at >200bp, which represent ~9% of the total number (the original Pol3Scan analysis considered terminators as far as 1kb from the reported 3'-end of the tRNA coding sequence, see *materials and methods*). **B**. Data described in A are chartered in a histogram which depicts the frequency distribution of canonical terminators in each 5bp bin. In the last bin denoted by >50, all termination sites located at a distance greater than 50bp and smaller than 1kb are subsumed. 142 out of the 675 human tRNA genes belong to this category, while in yeast only the 5'-moiety of the four tandemly arrayed tRNA genes terminates farther than 50bp. Among the 142 tRNA genes with long trailers, 69 were further validated as tRNAs by tRNAScan analysis [262].

Gene and chromosome location	Sequence	A-box score	B-box score
Arg147 chr6 28818780-28818708	**GGCCGCGTGGCCTAATGGATAAGGCGTCTGATTCCGGATCAGAAGATTGAGGGTTCGA GTCCCTTCGTGGTCGT**GTTCTTACTATTGTCAGGAAATATCTTTACTTTCTGCACATC TCCTTGAGCTTTCTGCAAGCCAGTTAAAAAAAAAAAAACCGGCTTGTTCCCAAGACTT TGGGAACTTGAAATGAAAGAAGACTATTGACAAAAAAAACAAGAACCCGGTAAGCAAC CTGTGTATTGACTATTAAAGTGCACTTTGTCCCTGGCATAGGTACTACCGGAGGCATT CGTAATTTGGAATGGGCTTGGTTTTT	-14.03	-3.62
Val148 chr6 28811756-28810685	**GGGGGTGTAGCTCAGTGGTAGAGCGTATGCTTAACATTCATGAGGCTCTGGGTTCGAT CCCCAGCACTTCCA**CAAGTACATTTCCTTATATCTATCGTTGAACATTTCGTTTCAGT CTTTGAATACTTAGTTCCAGTCTTTT	-9.50	-3.21
Gly33 chr16 69380911-69380981	**GCATTGGTGGTTCAGTGGTAGAATTCTCGCCTGCCACGCGGGAGGCCCGGGTTCGATT CCCGGCCAATGCA**GCAGCAGACCTTTAGTTTAATCGTGAGAGACTGAATTTGAGATTT CACACTGCCTAAAGAATTTATAAATCCACTCTGAGAACTGAGAATGACTTGTTTCTAG TTGGATTTCAAATGCCTTTGATGAAAACACTTATTTT	-11.19	-2.17
Trp32 chr7 98905243-98905314	**GACCTCGTGGCGCAACGGCAGCGCGTCTGACTCCAGATCAGAAGGTTGCGTGTTCAAA TCACGTCGGGGTCA**AACATGTGATCATCCCTTTCTTTAGCCAACAGGAGTTATAGCTT GCCGGTAACTCTCAACCTTGGACACTGTATTTT	-15.95	-9.90
Arg136 chr1 145949915-145949985	**GCATTGGTAGTTCAGCGGTGGCATTCTCCCCACCTACGCGGGAGACCTGGGTTCAACT CCCGGCCAAAGCA**GGCCTTCTTACCCTTCATGCTCCAAGTTCATGCTTGAAGGGCTAT TACATTTCTAGCCAACAAAGTTATTGTACTACACAGGCTGTACAGCATTTAGGGGCCT TGAGTCGTGGTGGCTGGAAGCTAAACCTGCCACCTCTGGAATATTCTCTGTGCTGGAT CTAATCTTGAAATTACATTTCCTCTGCTCAAAGCCTGTCTTTGAAAGTGGGGGTGTTT CCGGTGGTTTGCTGCAGTTTGGAGTACTGGGAACTAAATAAAGCCTTTGCTACTTTCA TCTAGTCTCCGGGCACACCCCTGAGTAATCCAAAGGCCATCAGACGGATTCTCGTGCT GGACTTTCCCCTACCTTAGGCCACATCTCACTATGGAGTGAGGATCAAGAAGTAAAGG TCTGGGAAGGAGAGAGCTGCAGGGGCCAGGGAGACTCTTCGTGAGTCGGAGAGCCATC AGCGGCCCAAGGAAAAGGGAGGATCGTGGGGAGACGAGGGGAGCAGTGGAGATGAGGG GAGCAGCGGAGACAGAGTCTCACGCACAGGTATGCCCAGACACCCTGAACCCAGTTAC TTAACCTCGTCAACCACTGTGGCATCCGCGGGGTCCGTGGGCCCATGGGTCTTCCGGG AGGGATTCTGTGTCCCTGTTATTGTGTGTCAATAGGTGTTAGCCTGATATTTATATTT T	-16.80	-4.82
Thr163 chr6 27802452-27802525	**GGCTTCGTGGCTTAGCTGGTTAAAGCGCCTGTCTAGTAAACAGGAGATCCTGGGTTCG AATCCCAGCGAGGCCT**CTTTATTTCTTCCCCTAAACTTAGGTAAATTCTTGTCACTAG TTAAACCTGACTTAGATGTATACCTAATAAAGCAGCAAACCTTGAATCTCGATGTCTG AGGCACTTTGTAGAGGAACGTGGTGATGGGTTGGAAATATTTACGATGAATTAGCTAA ATCTATGCCAGAGACCAACCGTCTAGAACCATCTAGAATTGAGCTGCAATTTT	-12.71	-1.92
His123 chr1 14742312-147421952	**GCCATGATCGTATAGTGGTTAGTACTCTGCGCTGTGGCCGCAGCAACCTCGGTTCGAA TCCGAGTCACGGCA**ATGTCGTTCTTCCAGGCCGTCAGCTATCCACTTTCGCTCCCTGC AACCAGGCGCCTCAGGGAAAGGAAAAGAGACCTCACAGCCCACTAACCTGGCGAGGAA TCCAAGGGAAGCGCTGGACGTGCTCCGTCCATTTCCGAGCCTCACACAGCAAATAACC CAGTCAGCCCTGTTGGTGGCTTACACGAAGGCAATGGTGCCCTCAGAACCTGTTTT	-16.94	-1.92
Glu19 chr2 150927523-150927594	**TTCCCCTTGGTCTAGTGGTTAGGATTCAACACTCTCACCGCCGCAGCCCGGGTTTGAT TCCCAGGCAGGGAA**GCCTTCCTTCTTTGGCTCTGCCTGCCAACTGCCAGGCCCCTTCA CACCTCCCCTTTT	-11.82	-7.60

Figure 2. tRNA genes with an unusually distant canonical RNA polymerase III terminator selected for the *in-vitro* transcription assays. Eight loci from the pool of 142 tRNA genes with a ≥50 bp canonical terminator were selected based on the high score of A-box and B-box promoter elements [259]. Each locus was further validated as a tRNA coding unit by tRNAScanSE [262] and RepeatMasker [263]. The genomic sequence corresponding to the selected genes plus ~300bp upstream of the TSS and ~100bp downstream of the terminator element were PCR amplified from buccal cells DNA and cloned in the commercial vector pGEM-T-*Easy* (*Promega®*). Chromosome positions are indicated on the left column and referred to the human assembly NCBI36. Bold: tRNA coding sequence. Underlined: hypothesized non-canonical terminators.

(lane 3-8) were efficiently transcribed by RNA polymerase III, but did not resolve into a single product as expected for full termination at the canonical $\geq T_4$ sequence. Rather, the observed transcripts pattern could be explained by assuming that termination is modulated *in vitro* by novel dT-rich elements located upstream of the canonical terminator (**Figure 2**, underlined sequences), which give very efficient termination (lane 4,5) or more consistent read-through (lane 3,6-8). In

order to investigate whether the smaller sized transcripts represent processing products of full length pre-tRNAs or derive from transcriptional termination at non-canonical sequences, we introduced point mutations into the first putative termination element of the Val148 and Thr163 tRNA genes (**Figure 3B**) and analyzed the effects on transcription in HeLa nuclear extract. As expected, almost no transcripts were terminated properly when both the non-canonical sequences were mutated in dT-poor elements; instead, the transcription reactions gave rise to several long read-through transcripts terminated at the downstream predicted sites (**Figure 3C**). Interestingly, the overall level of transcription from a linear VA1 template mutated at the canonical T_4 was reported to be much lower than the level observed with the wild-type template, both in single and multiple round transcription [176]. In contrast, the overall transcription efficiency of the mutated and wild-type templates was approximately the same for both Val148 and Thr163. These results suggest that the non-canonical sites in analysis affects mainly termination, but not initiation and reinitiation; however, it cannot be excluded a gene-specific effect or that RNA polymerase III transcription proceeds differently on the supercoiled template employed in the assay.



Val148 wt -GTACATTTCCTTATAT-
Val148 Tmut -GTACATGTCCGTATAT-

Thr163 wt -CCTCTTTATTTCTTCCC-
Thr163 Tmut -CCTCTCTATGGCTCCCC-

B

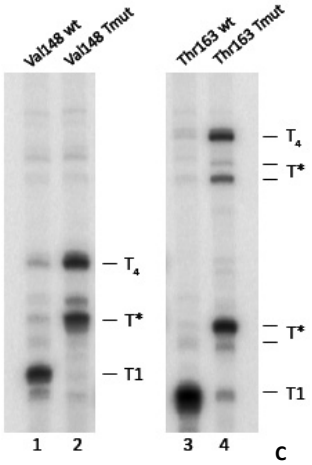


Figure 3. Human tRNA genes with a canonical terminator located at ≥50 bp from the 3'-end of the tRNA coding sequence terminate at novel non canonical terminators. **A.** *In-vitro* transcription reactions were performed in HeLa NE using 500ng of the indicated templates (lanes 1-8), the empty vector (lane 9) or no DNA (lane 10). Reactions were stopped by phenol:chloroform extraction and transcription products were resolved by denaturing polyacrylamide gel electrophoresis. The size of the transcripts is consistent with termination at the non canonical sites underlined in Figure 2. L: 0.1-2Kb RNA ladder (*Invitrogen®*) was end-labeled at the 5'-end with T4 polynucleotide kinase (*Promega®*) and [γ-^{32}P]ATP. **B.** Sequence of wild-type (wt) and mutant (Tmut) termination sites immediately downstream of Val148 and Thr163 tRNA genes. **C.** *In vitro* transcription of Val148 and Thr163 wild-type templates (lane 1,3) and their respective mutants (lane 2,4) was performed as previously described. Transcripts terminated at the first non-canonical site (T1), other predicted non-canonical sites (T*) and the canonical T-run (T_4) are indicated on the right and have the predicted size for transcriptional termination at the sites underlined in Figure 2.

Comprehensive analysis of human RNA polymerase III non-canonical terminators by *in vitro* transcriptional assays.

As noted above, in the human genome there is a considerable enrichment in tRNA genes with a canonical terminator located at ≥50 bp from the 3'-end of the coding sequence. The previous analysis on a limited set of genes suggests that RNA polymerase III transcription terminates at upstream located non-canonical sites, but also revealed the synthesis of several long read-through transcripts, thus not excluding transcription of long 3'-trailers terminated at the canonical $\geq T_4$.
Since the original finding that RNA polymerase III terminates within stretches of consecutive dT residues [231,252], the sequence of a native RNA polymerase III terminator and its significance remain elusive. Despite being one of the first eukaryotic transcriptional regulatory elements identified, the occurrence of RNA polymerase III terminators in the human genome has never been addressed thoroughly, and several reports suggest that things are more complicated than originally proposed. Read-through at supposedly canonical sites have been observed for several genes [232,233] and a few alternative elements, such as various combinations of 2-3 consecutive dT residues [237-239], long stretches of dAs and potential hairpins at the 3'-end of transcribed genes [240], have been sparsely described in literature as possessing termination activity. The sequence contest in which the T-run is embedded was also shown to play a significant role in its recognition as a termination signal [241,242] and, as pointed out by Bogenhagen and colleagues in their original paper, a run of five consecutive T residues is less sensitive to the influence of neighboring sequences [231]. Considering the abundance of type 2 promoters in the human genome, a more comprehensive analysis of the sequence requirements for efficient RNA polymerase III termination is of paramount importance, since termination or read-through at these sites could lead to regulated biogenesis of downstream transcripts. To identify non-canonical terminators among the 675 transcriptional units previously found, our strategy centered around the imposition of several loose constraints under which a minimal terminator is expected to operate. These constraints include a maximum sequence length (8bp) and a minimal consensus sequence ($T_aN_xT_b$), the rationale for which is presented in **Figure 4**. Application of these constraints led to the identification of 163 sequences with potential non-canonical terminators upstream of the $\geq T_4$, while this number dropped to 110 when we applied the survey to the 142 tRNA genes with a canonical terminator located at ≥50 bp. Among the 45 5/8-mer sequences of the minimal consensus that we included in the survey, nine of them were particularly enriched in the 3'-trailer of tRNA genes, representing 73,5% of the overall potential non-canonical terminators (this percentage raised to 86,5% when the analysis was restricted to the first non-canonical terminator occurring in the 3'-trailer, **Figure 4**). According to the rank order, we decided to analyze RNA polymerase III termination at these top-scoring sequences by *in-vitro* transcription. We first engineered an expression plasmid containing the coding region of Val148 by replacing the

TTTNNTTT
TTTNTTT
TTNNTTT
TTTNNTT
TTNTTT
TTTNTT
TNTTT **73,5%**
TTTNT
TTNTT

Figure 4. Minimal consensus sequence for potential non-canonical terminators. Considering that a stretch of four or more dT residues is rated as an efficient terminator for human RNA polymerase III and that the few alternative terminators described in literature [237-239] never have more than three nucleotides between two consecutive stretches of dTs, we assumed a minimal consensus terminator $T_aN_xT_b$ generating 45 possible sequences with the following constraints: $a,b < 4$; $4 \leq a+b \leq 6$; $0 < x < 3$; N denotes any nucleotide. Scanning the 3'trailer of the 675 tRNA and tRNA-like genes identified with Pol3Scan, we found 468 hits for the consensus indicated on left and that the 3'trailers are generally enriched in the sequences highlighted in bold (73,5%).

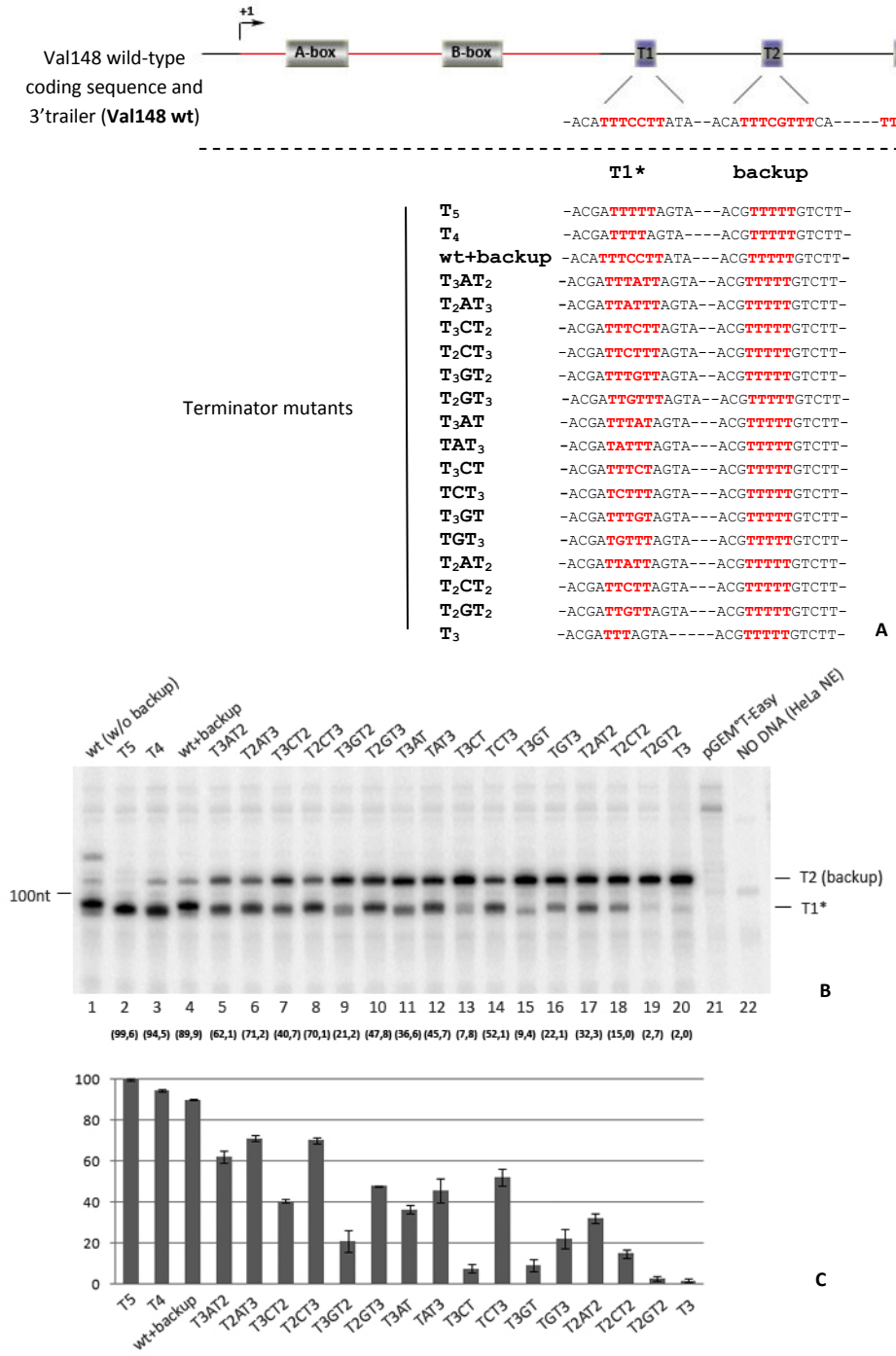


Figure 5. Non canonical terminators support correct RNA polymerase III termination by *in-vitro* transcription in HeLa NE. **A.** Schematic representation of Val148-derived templates. Terminator mutants of the Val148 tRNA gene were generated by site-specific mutagenesis as described in *materials and methods*. The endogenous terminators T1 and T2 were replaced by the non-canonical terminators in analysis (T1*) and a backup terminator, respectively. Minimal termination signals are indicated in red. **B,C**. *In-vitro* transcriptions were performed in HeLa nuclear extract as previously described. Val 148-derived transcripts terminated either at the first non-canonical terminator (T1*) or the backup terminator (T2, backup) are indicated on the right. Termination efficiency is indicated at the bottom and expressed as termination at T1* relative to total termination (T1*+T2). Error bars refer to the standard error of three independent measurement.

terminator T1 with the sequences to be analyzed, plus strong (T_4 and T_5) and leaky (T_3) termination signals. We next introduced a backup terminator in place of T2 to have downstream terminated transcripts in case of read-through at the non-canonical sites (**Figure 5A**) and finally assayed the Val148-derived templates by *in-vitro* transcription in HeLa NE (**Figure 5B**). As expected for transcripts terminating at the canonical T4 and T5 sites, read-through products were barely visible, indicating that a termination signal consisting of a run of four of five T residues can function efficiently as previously reported for several human class III genes [205,232-236]. Consistent with the hypothesis that non-canonical elements induce partial transcription termination, we observed a significant read-through at the sequences in analysis, while RNA polymerase III proceeded straight through the leaky termination signal (T_3) to produce a full-length transcript terminated at the backup site (**Figure 5B,C**).

The limiting amount of UTP in the reaction mix (25μM) made us consider that the shorter products terminated at the non-canonical sites might originate from increased pausing of RNA polymerase III and not from authentic termination, thus reflecting an artifact created by the reaction conditions. In order to allow RNA polymerase III to overcome extended pausing during elongation, we repeated the same assay increasing the UTP concentration to 250uM (**Figure 6**), much closer to the estimated intracellular concentrations of UTP in mammalian cells (0.567 ±0.460 mM, [264]). Under these conditions, the transcription pattern precisely overlapped with the one previously shown in figure 5B, where the reaction was carried on with a 10-fold lower cold UTP concentration. Taken together, these data indicate that RNA polymerase III terminates with different efficiency at *bona fide* non-canonical terminators and also suggest that the nucleotide embedded in the interrupted run of dT residues significantly contributes to the difference in termination efficiency.

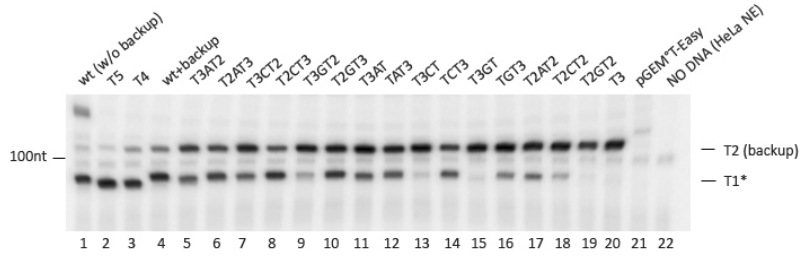
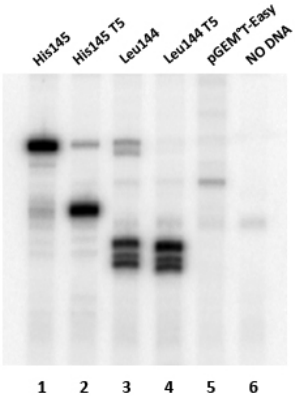


Figure 6. Nascent RNA displacement is not due to the limiting concentration of UTP in the reaction mix, but rather depends on *bona fine* RNA polymerase III termination at the indicated sites. *In*-vitro transcription in HeLa NE were performed as previously described, with the difference that UTP concentration was raised to 250μM.

Given that the sequences in analysis were cloned in the same contest, we wondered whether the non-canonical signals identified so far worked the same in their natural background. Thus, we assessed termination efficiency at supposedly strong (T_3AT_2) and weak (T_3CT) non-canonical terminators, naturally occurring in the 3'-trailer of human tRNA genes. As shown in **Figure 7** (lane 1,3), RNA polymerase III terminated at these natural sites with efficiencies comparable to the ones observed with chimeric Val148 templates. Replacement of the two non canonical terminators with

T_5 restored almost complete termination downstream of the two tRNA genes (lane 2,4), suggesting that type 2 promoter class III genes are likely to undergo alternative transcription termination according to the rules emerged from our analysis. However, more subtle contest effects cannot be excluded, as Leu144 shows a couple of prematurely terminated transcripts despite lacking any dT residue in the immediate surroundings of its terminator.



3'-trailer sequence

His145	ACCTTGATGTCGCCTCAATTTCTCAACGTACTG AGCAGTACCTTGACGTCGCCTCAATTTTTC
His145T$_5$	ACCTTGATGTCGCCTCAATTTTTCAACGTACTG AGCAGTACCTTGACGTCGCCTCAATTTTTC
Leu144	GGTTTATTGCAAAATTAGTAGGGATATCTAAAG CTCATTGAGATTTGCAAGAAAAACATTTTC
Leu144T$_5$	GGTTTTTAGCAAAATTAGTAGGGATATCTAAAG CTCATTGAGATTTGCAAGAAAAACATTTTC

Figure 7. Non-canonical terminators in their natural contest are recognized by RNA polymerase III with different efficiency. *In vitro* transcription of His145 and Leu144 wild-type tRNAs (lane 1,3) and their respective mutants (lane 2,4) was performed as previously described. Endogenous terminators were replaced by a stretch of five consecutive dTs through site-specific mutagenesis. Lane 5 and 6 indicate the empty vector (pGEM-T-Easy) and the reaction performed without DNA, respectively.

It shall be noted that the relative termination frequencies indicated in **Figure 5C** could be biased by the specific processing of full-length products, albeit the observation that leaky transcription at the T_3 terminator couple with traces amount of the shorter transcript should be enough to suggest that this is not the case. To further validate this point, standard transcriptions of three representative templates (Val148_wt, T_3 and T_5) were chased for various times with a large excess of unlabeled UTP (**Figure 8**). The experiment failed to identify any enrichment in the smaller sized products, whereas the amount of label in the long form of Val148-T_3 started to decline 50' after the chase. Such a longer transcript specific decay is unlikely to affect significantly the levels of read-through transcripts observed in the previous assays, as the incubation was carried on for 25'. Consistently with this last data, a similar processing pattern was observed when labeled T7-promoter driven transcripts were incubated with the HeLa nuclear extract in the absence of [α-^{32}P]UTP (compare **Figure 8,** lane 5-8 and **Figure 9**). These data argue that the HeLa nuclear extract used in the previous assays is largely inactive in terms of specific processing, albeit it retains some basal ribonuclease activity. It is unclear why T7-derived transcripts are more sensitive to basal processing, but we can infer that the assembly of an active transcriptional complex on the supercoiled templates somehow protects RNA polymerase III derived transcripts from ribonuclease digestion.

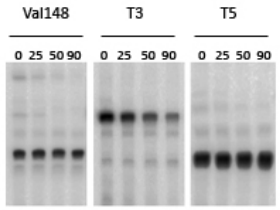


Figure 8. Pulse-chase reaction. In-vitro transcription in HeLa NE of Val148 wt, T3 and T5 were carried out under standard conditions and subsequent chased with UTP 2mM for the indicated times. The processing of the smaller transcripts is negligible and non-specific, since it doesn't lead to the accumulation of any discrete sized band corresponding to the mature tRNA. The longer transcript shows a similar behavior, with a more enhanced processing 50' after the chase.

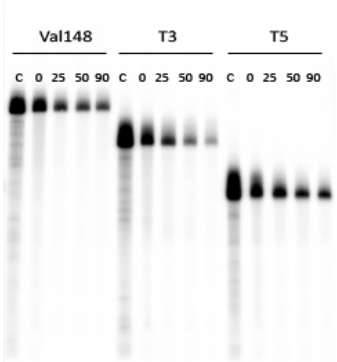
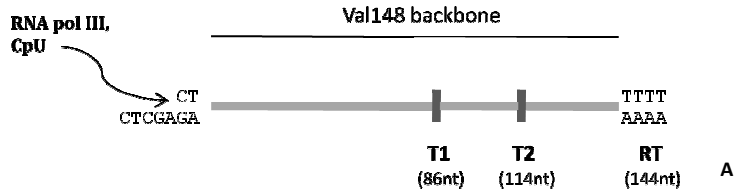


Figure 9. *In vitro* processing of T7 RNA polymerase synthesized RNAs in HeLa NE. Labeled RNAs corresponding to Val148 wt (142nt), T3 (114nt) and T5 (86nt) were synthesized under the control of the T7 promoter as described in *materials and methods*. 20.000 cpm of each run-off transcript were either directly loaded onto the gel (C) or incubated for the indicated times (0,25,50,90 minutes) with HeLa NE. RNA decay reactions were carried out under the same conditions used for the *in-vitro* transcription assays.

RNA polymerase III is able to discriminate among different non-canonical terminators in the absence of accessory factors.

In 1983, Cozzarelli and colleagues suggested that RNA polymerase III alone is sufficient to terminate accurately by recognizing the simple T-rich consensus [243]. Later on, using highly purified RNA polymerase III to transcribe a 3'-overhang template in the absence or presence of its cognate transcription factors, several studies demonstrated factor-dependent mechanisms for efficient termination and reinitiation in mammals [16,152,176]. To further address whether human RNA polymerase III itself is able to discriminate among the different non-canonical terminators, we set up a factor-free assay as depicted in **Figure 10A**. We initially performed titration reactions to define balanced concentrations of salts and RNA polymerase III (**Figure 10B,C**). Once defined optimal reaction conditions, we carried out promoter independent assays on the 3'-overhang templates containing the entire set of previously tested non-canonical terminators (**Figure 10D**). As a result of initiation at the single stranded end, and in agreement with precedent assays on a C-tailed VA-I template [176], RNA polymerase III generated RNA products that started at the 5'-end and stopped at either the first termination site (T1*), the second (T2) or the end of the template (RT). Notably, relative termination efficiency at T1* correlated with the one previously observed in promoter dependent assays, suggesting that RNA polymerase III itself accounts for the difference in terminator recognition. The generally low levels of absolute termination at both non canonical and backup sites can be addressed to either the artificial read-through at tailed templates [176] or to the requirements of additional factors for efficient RNA polymerase III termination, or both.



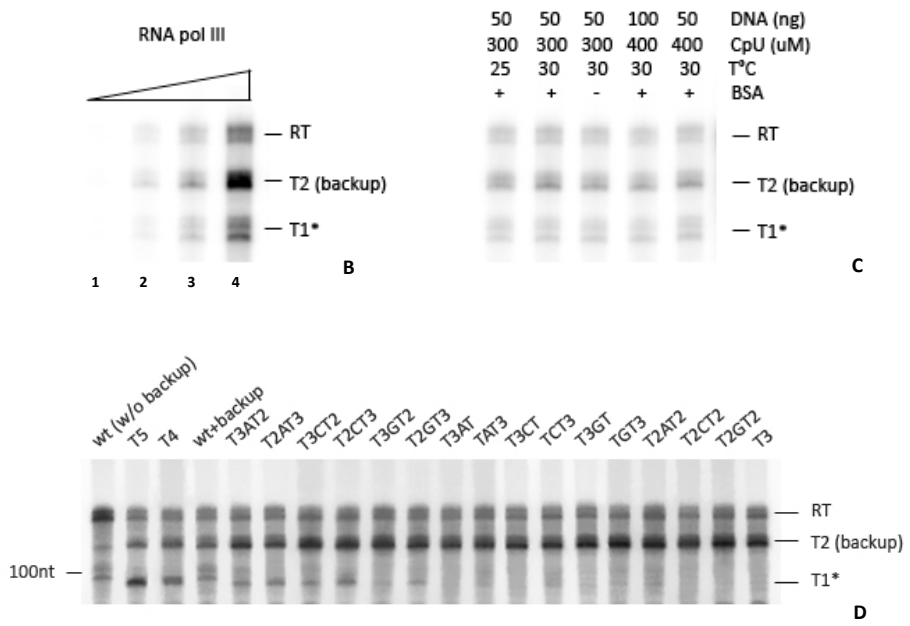


Figure 10. Factor-free transcription assays reveal that purified RNA polymerase III alone is able to discriminate among different non canonical terminators. **A.** Schematic representation of factor-free transcriptional assay. SacI-digestion of a PCR derived linear template generates a 3'-overhang at which RNA polymerase III can initiate transcription in the absence of other accessory factors. Positioned downstream of the tailed end is a 144nt Val148-derived template, designed according to RNA polymerase III initiation at the natural transcriptional start site. Transcripts terminated at either T1* (non-canonical terminator), T2 (backup terminator) and RT (the end of the template) are indicated with their lengths given in nucleotides. **B.** RNA polymerase III titration assay. Reactions containing 50ng of the tailed Val148 wt template were allowed to proceed as described in *materials and methods* with 2.5 (lane 1), 6 (lane 2), 12 (lane 3) or 30 (lane 4) ng of purified human RNA polymerase III. **C.** Promoter independent assays were run on the tailed Val148 wt template using 6ng of purified RNA polymerase III, and different temperatures/amounts of CpU, BSA and DNA as indicated in the figure. Transcripts terminated at T1*, T2 and RT are indicated. **D.** Promoter-independent reactions contained 6ng of purified RNA polymerase III and 50ng of tailed Val148-derived templates modified with the terminators assayed in the previous promoter-dependent reactions in place of T1*. The arrows indicate transcripts terminated at T1*, T2 and RT.

A novel approach to study non-canonical terminators *in-vivo*

The recent discovery of viral miRNAs indicates that viruses have evolved to exploit RNA silencing for regulation of both host and viral genes [92]. In contrast with the RNA polymerase II origin of most miRNAs [93,265], the genome of murine gammaherpesvirus 68 (MHV68) encodes nine pri-miRNA transcripts [94] which are likely to be initiated by RNA polymerase III type 2 promoters and processed in a Drosha-independent stepwise fashion [266,267]; according to a recent report, tRNaseZ cleaves 3' to the tRNA moiety to liberate a canonical pre-miRNA stem-loop that is subsequently processed by Dicer to give the mature miRNA.

To analyze non-canonical terminators *in-vivo*, we prepared a synthetic tRNA-shRNA dicistronic unit resembling the tRNA4-miR-M1-5 locus of MHV68 [93]. The transcriptional unit was engineered by inserting the Upstream Activating Sequence of the 7SL RNA gene upstream of tRNA4 and a 5-6 nt sequence, corresponding to the terminators to be analyzed, between the tRNA-like and the pre-

miRNA moiety. The resulting template was cloned in pSuper (*Oligoengine®*) in place of the H1 promoter and transfected in HeLa cells to assess non-canonical terminators efficiency based on miR-M1-5 biogenesis/processing (**Figure 11A**). Northern analysis using probes matching the mature miRNA sequence failed to detect significant levels of miR-M1-5 in HeLa cells transfected with tRNA4-miR-M1-5(T_5), consistent with the hypothesis that a stretch of five dT residues induce efficient termination downstream of tRNA4. On the other hand, in cells transfected with tRNA4-miR-M1-5(T_3) a ~71 nt and a ~150 nt bands, which are likely to represent the miRNA processing intermediate (tRNA4-miR-M1-5 and pre-miR-M1-5, respectively), were clearly detectable, as expected for read-through at the leaky T_3 signal (**Figure 11B**). It is not clear why miR-M1-5 alone was not detectable, but it is possible that insertion of five nucleotides at the 5'-end of the pre-miRNA moiety would impair Dicer processing. The identity of these precursor RNAs was further confirmed using a probe specific for the tRNA moiety, which also confirmed specific expression of tRNA4 in cells transfected with tRNA4-miR-M1-5(T_5) and indicated the presence of a cross-reacting shorter species in HeLa cells. Expression levels were further normalized using a probe against U6 snRNA. Taken together, these data account for a novel and efficient system to study RNA polymerase III termination in HeLa cells. Transfections and *in vitro* transcriptional assays with the vectors carrying non-canonical terminators are currently on the way.

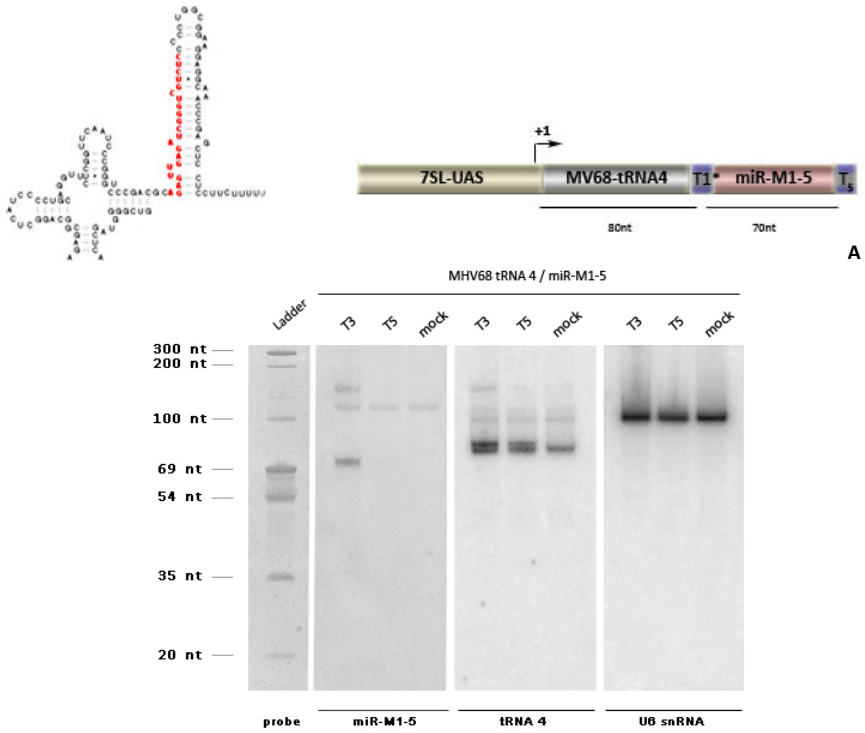


Figure 11. Northern analysis of tRNA4-miR-M1-5(T1*) RNAs derived from RNA polymerase III read-through/termination at T1*. **A.** Schematic representation of an engineered MHV68 tRNA4-miR-M1-5 transcriptional unit. The transcripts have a tRNA-like fold followed by a stem-loop structure in the RNA molecule. Transcripts terminated at either T1* (T_3 and T_5) or T_5 (five dT backup terminator) are indicated with their lengths given in nucleotides. +1 indicates the transcriptional start site. **B.** Northern analysis of total RNA isolated from HeLa cells 36h post transfection with tRNA4-miR-M1-5(T_3), tRNA4-miR-M1-5(T_5) and an empty vector (mock). Northern blots were analyzed using probes complementary to miR-M1-5, tRNA4 and U6 snRNA. An ethidium bromide-stained homemade RNA ladder with the corresponding length in nucleotides is indicated on the left panel.

DISCUSSION

Despite some recent progresses, RNA polymerase III transcription termination in mammalian cells is still poorly understood. Apparently, human RNA polymerase III possesses the ability to autonomously recognize termination signals made up of at least four dT residues, but the efficiency of recognition is largely affected by several auxiliary factors, such as PC4, NF1, TFIIIC2 and La.

In the present report, we redefined the minimal requirements for human RNA polymerase III termination at type 2 promoter class III genes. Our analysis show that many novel terminators comprised of less than four consecutive dT residues are recognized by human RNA polymerase III and that differences in termination efficiency largely rely on the nucleotide embedded in the dT-rich stretch. These terminators are generally weaker than the canonical ≥T4 and consistent read-through can be observed at most of them. Moreover, by employing a factor-free assay we demonstrate that RNA polymerase III itself account for these differences and also confirm that additional factors are likely to be required for fully efficient termination at both canonical and non-canonical sites. The role on these non-canonical elements in-vivo is not known and it's currently being investigated through a system that allows detections of read-through RNAs in transfected HeLa cells. Interestingly, the observation that more than one million of type 2 promoter elements are scattered through the human genome (tDNAs, Alu, MIR, etc) raises the possibility that read-through at non-canonical sites would end up in the biogenesis of primary RNAs, with potential regulatory functions, embedded in the 3'trailer of predicted type 2 promoter transcripts. Early annotation of the genomic position of miRNAs indicated that most miRNAs are located in intergenic regions (>1 kb away from annotated/predicted genes), although a sizeable minority was found in the intronic regions of known genes in the sense or antisense orientation [268,269]. This led to postulate that most miRNA genes are transcribed as autonomous transcription units and that RNA polymerase II transcription account for the majority of these pri-miRNAs. However, recent studies have shown that also RNA polymerase III promoters are able to drive the expression of microRNAs in mammals [266] and more specifically in humans [270,271]. Remarkably, viral miRNAs that are initially transcribed as a tRNA-shRNA dicistronic unit have been shown to be processed by a tRNA processing enzyme instead of the pri-miRNA processing pathway: tRNaseZ cleaves 3' to the tRNA moiety and liberate a non-functional tRNA and a canonical pre-miRNA stem-loop that is subsequently processed by Dicer in a mature miRNA [267]. This finding explains the puzzling detection of uncharged viral tRNAs in MHV68 lately infected mice and suggest that other tRNA or tRNA-like genes could be non-functional remnants that conserved their original promoter elements to drive downstream miRNA, or more generally ncRNAs, expression. Consistent with this hypothesis is the observation that ~23% of the human tRNA genes have a ≥T4 terminator at more than 50bp from the 5'-end of the 3'-trailer and that most of them contains a partially efficient non-canonical terminator within this region. Considering that nascent RNA displacement during RNA polymerase III termination appears to rely on the strength of the DNA:RNA hybrid [244], leaky terminators upstream of a pre-miRNA would account for signals that are more sensitive to regulation by accessory factors. Notably, some important human tRNAs such as $tRNA^{i}$(Met), bear a long dT tract at the 3' boundary of the coding sequence which was possibly conserved to ensure efficient transcription termination, limiting the possibility of read-through [272]. Similarly, most tRNA gens in *S.cerevisiae* terminate at longer dT stretches [241] usually located within the first 15 nucleotides downstream of the coding sequence. Considering that homologs of human RNA polymerase III termination factors have not been found in yeast, the finding of relaxed constraints

for human RNA polymerase III termination with respect to lower eukaryotes could possibly represent the evolution of a more sophisticated regulatory mechanism for the biogenesis of small RNAs.

MATERIALS AND METHODS

Bioinformatic analysis for identification of human tRNA genes

The 274 *S.cerevisiae* tRNA genes were previously identified in [259]. The 675 human tRNA genes with a canonical $\geq T_4$ terminator within the 1kb downstream flanking region of the tRNA coding sequence were identified with the software Pol3Scan [259], which was run on the human release 36 of the EMBL database. The algorithm is based on the statistical analysis of 231 tRNA promoter regions and makes use of weight matrices and weight vectors for scoring. The program discriminates between tRNA genes and related class III elements on the basis of the presence of a transcriptional terminator signal and of the base-pairing within the aminoacyl stem. To include in the analysis tRNA genes with distant terminators (up to 1kb) looser constraints were imposed for the distance of a canonical termination site. Together with the 675 tRNA genes, the software identified 13681 tRNA-like sequences (e.g. tRNA derived SINEs). Among the 675 human tRNA genes, 484 were identified as tDNAs by tRNAscanSE [262]. Among the 675 human tRNA genes, 142 have a canonical $\geq T_4$ terminator located at more than 50bp from the end of the tRNA coding region; among these 142 sequences, 82 and 75 were further confirmed as tRNAs by tRNAScanSE [262] and RepeatMasker [263] analysis, respectively. All the tRNA genes selected for the *in vitro* transcription assay were further validated as tDNAs by tRNAScanSE and RepeatMasker analysis.

Construction of plasmids and templates containing human tRNA genes

Using oligonucleotides defined in appendix 1, eight human tRNA genes indicated in the text and their 3'trailers were PCR amplified from buccal cells genomic DNA with recombinant Taq DNA polymerase (*Fermentas*) as described by the manufacturer. DNA fragments were purified with the Nucleospin Extract II columns (*Fermentas*) and cloned blunt end into the EcoRV site of pGEM-T-Easy® (*Promega*): plasmids were transfected and propagated in XL1-B cells and recovered with the *QIAGEN* MIDI plasmid kit. Non-canonical terminators, inactive terminators (Val148Tmut and Thr163Tmut) and the backup terminator were introduced in the tDNA genes 3'trailer by site specific mutagenesis with Pfu DNA polymerase, using primers containing mismatching nucleotides to replace the sequences described in the text. The MHV68 tRNA4-miR-M1-5-like transcriptional unit indicated in appendix 2 was purchased from *MrGene* and subcloned into the NotI/XhoI sites of an $Amp^R \rightarrow Kan^R$ version of pSuper (*Oligoengine*). The T3 mutant at the T1* site indicated in the text were generated by site-specific mutagenesis with Phusion DNA polymerase and equimolar amount of primers UAST3_for and UAST3_rev as described by the manufacturer.
For the construction of the 3'-overhang template used in factor-free transcription experiments, equimolar amounts of the two oligonucleotides Val148overhang_for and T73'Val148_rev were used to PCR amplify from pGEM-T-Easy-Val148 a linear fragment containing the SacI restriction site at its 5'-end followed by the Val148 coding sequence. PCR reactions were done with Phusion DNA polymerase (*Finnzymes*) on plasmids containing all the non canonical terminators in analysis. The resulting templates were digested with the SacI restriction endonuclease, yielding a 146-bp 3'-overhanged fragment that was gel-purified with QIAquick gel extraction kit (*Qiagen*) and used for transcription experiments.

Linear templates for T7 RNA polymerase mediated transcription were generated by PCR reaction on the Val148, Val148-T3 and Val148-T5 templates using a forward primer containing the T7 RNA polymerase promoter (T75'Val148, [293]) and reverse primers (T73'Val148, T73'T3, T73'T5) corresponding to the 3'-end of the indicated templates.

In vitro transcription and processing reactions

For promoter-dependent transcription in HeLa NE, reaction mixtures in a final volume of 25 µl contained 500 ng of plasmid DNA carrying the templates in analysis, 70 mM KC1, 1 mM $MgCl_2$, 1,5 mM DTT, 11,5% glycerol, 20mM Tris, 5mM phosphocreatine, 2 µg/ml alpha-amanitin, 0.4 U/µl Superase (*Ambion*), 10µl of HeLa Nuclear Extract, 0.5 mM ATP, CTP, and GTP, 10 µCi of [α-32P]-UTP and either 0.025 mM or 0.25 mM UTP. HeLa nuclear extract was prepared according to [294] and kindly provided by the Martin Teichmann group, Bordeaux. Control reactions containing the empty vector or no DNA were assembled in parallel. Reactions were allowed to proceed for 25 minutes at 30°C before being stopped by addition of 75µl H_2O and 1 volume of phenol:chloroform 1:1. Labeled RNAs were phenol:chloroform extracted and precipitated with 1 volume of NH_4Acetate, 2.5 volumes of 100% ethanol and 12 µg of Torula RNA carrier. Labeled RNA pellets were washed twice and resolved on a 6% polyacrylamide, 7M urea, TBE 1x gel. RNA bands were visualized and quantified with the phosphorimager (*Bio-Rad*) and the software Quantity One (*Bio-Rad*).
For 3'overhang template factor-free assays, reaction mixtures in a final volume of 25 µl contained 50 ng of linear DNA carrying the Val148 templates in analysis, 60 mM KC1, 2 mM $MgCl_2$, 1,5 mM DTT, 10% glycerol, 10mM Tris, 50 µg/ml BSA, 0.32 U/µl Superase (*Ambion*), 0.5 mM ATP, CTP, and GTP, 0.025 mM UTP, 10 µCi of [α-32P]-UTP and 10ng of purified human RNA polymerase III, kindly provided by Chiara Pascali from the Martin Teichmann group, Bordeaux (purified human RNA polymerase III was extracted from a stable HeLa cell line expressing an FHM-tagged version of subunit RPC32 and recovered with M2-agarose beads from Sigma: the multisubunit enzyme was then eluted with 250 ng/µl of a FLAG peptide) . Reactions were preincubated for 15 minutes at 30°C before addition of CpU and NTPs and allowed to proceed for 25 minutes at 30°C before resolving and visualizing the labeled RNAs. Different amounts of purified human RNA polymerase III and different concentrations of salts were assayed as described in the text before defining the optimal reaction conditions described above.
For the pulse-chase experiment, the reaction mixture was as described above except that after 25 minutes at 30°C the total UTP concentration was raised to 2mM. Reactions were then allowed to proceed for either 25, 50 or 90 minutes and stopped at the end of each time interval by phenol:chloroform extraction. Labeled RNAs were ethanol precipitated prior to polyacrylamide gel electrophoresis (PAGE) and phosphorimaging detection as described above.
T7 promoter-directed transcriptions were performed on a final volume of 25 µl with 200ng of linear DNA and 15U of T7 RNA polymerase (*USB*) as described by the manufacturer with the following exception: the T7 reaction buffer (*USB*) was supplemented with 5mM $MgCl_2$, 5mM DTT and 0.4U/µl Superase (*Ambion*). Reactions were carried out with 10 µCi of [α-32P]-UTP, 0.5 mM ATP, GTP, UTP and CTP for 1 hour at 37°C and stopped with 0.5M EDTA. Labeled RNAs were purified with Micro Bio-Spin Chromatography columns (*BIO-RAD*) as described by the manufacturer. *In vitro* processing of T7 RNA polymerase transcribed pre-synthesized RNAs was carried out for 0, 25, 50 and 90 minutes at 30 °C in 25µl reaction mixtures, which contained 300.000 cpm of the labeled RNAs and the components already described in the promoter-

dependent transcription assay. The reactions were quenched by phenol:chloroform (1:1) extraction and RNA was analyzed as described above for the pulse chase experiments.

Northern analysis

HeLa cells were maintained in DMEM (Dulbecco's Modified Eagle Medium, *Invitrogen*) containing 5% calf serum and 5% fetal bovine serum and propagated in 5% CO2 at 37°C. One day prior to transfection, cells were seeded at a density of 5 x 10^4 on 150 mm plates. DNA (12µg of plasmids containing the tRNA4-miR-M1-5 expression cassettes and 500ng of the EGFP expressing vector pEGFP-C1) was transiently transfected with the JETPEI kit as described by the manufacturer. 36h our after transfection, total RNA extraction was performed with Trizol (*Invitrogen*) as described by the manufacturer. Northern analysis was performed as described in [295], except that 20µg of total RNA were resolved on a 11% denaturing polyacrylamide gel in TBE 1x and transferred onto Hybond-N+ membrane (Amersham Pharmacia biotech). Pre-hybridization and hybridization were performed at 30 °C with Denhardt reagent 5X, SSC 5X, SDS 0.5% and salmon sperm DNA 0.1 mg/ml and the membrane was probed with ^{32}P-radiolabeled oligodeoxynucleotides complementary to the MHV68 miR-M1-5, MHV68 tRNA4 and human U6 snRNA sequences (see appendix 1). Filters were washed twice at 30° and 42°, respectively, with SSC 2X, SDS 0.1% and SSC 2X, SDS 0.1%, respectively. Data were recovered by phosphorimager exposition of the filters. Blots were stripped with boiling SDS 0.15% and reprobed several times. Complete stripping of the blot was confirmed by phosphorimaging of the membrane before reprobing.

APPENDIX 1

Primers and probes list

Transcription of the initial tDNA genes pool from genomic DNA

Arg147_for	ATCTCTAATTACGAAGTTAAAGGT
Arg147_rev	ATTCATGAAACAAGTAGGACAAGA
Val148_for	CTTGGAACTTCGCTCCAGGTGGA
Val148_rev	GGAGTGTTAGGACCCATAGTACA
Gly33_for	AGGTCTGAAGGCAAATCAACTGGAC
Gly33_rev	CTGCTTTAGCGGGCCTAGGTGCAT
Trp32_for	ATTATTCCCGAACCCCGTGGTGG
Trp32_rev	AGCTAAGGTGCCGGTCCACTGC
Arg136_for	TGTTTGCTCGGTGTGTCTACTCA
Arg136_rev	GTCCTTCGAAGGAGTGTCCAAGA
Thr163_for	TTCAGCACCTATTCTTATTCTACAG
Thr13_rev	CACAACACTCAATTTATACCATATG

His123_for TATGAAAGAATAGCGCGATCCCCCT
His123_rev ATACAAACTTGGCGACGAGCTGGG

Glu19_for AACCTCTACTTGCACCCATGTG
Glu19_rev TCTACAGGACACGTAGGACATTT

His145_for AGCAAAATATTGGGGACGGG
His145_rev GCTTGGGTAGGGTTGGGAAG

Leu144_for CAGAGTCCTCAGAGTCCCCG
Leu144_rev TACCTTGTGCTCGCAGTGGA

Site specific mutagenesis of Thr163 and Val148 terminators

Val148Tmut_for GCACTTCCACAAGTACATGTCCGTATATCTATCGTTGAAC
Val148Tmut_rev GTTCAACGATAGATATACGGACATGTACTTGTGGAAGTGC

Thr163Tmut_for
CGAATCCCAGCGAGGCCTCTCTATGGCTCCCCCTAAACTTAGGTAAAT
Thr163Tmut_rev
ATTTACCTAAGTTTAGGGGGAGCCATAGAGAGGCCTCGCTGGGATTCG

Val148T4_for*[1] GCACTTCCACAAGTACGATTTTAGTATCTATCGTTGAAC
Val148T4_rev*[1] GTTCAACGATAGATACTAAAATCGTACTTGTGGAAGTGC

His145mut_for GATGTCGCCTCAATTTTTCAACGTACTGAGCAG
His145mut_rev CTGCTCAGTACGTTGAAAAATTGAGGCGACATC

Leu144mut_for CCCACCGCTGCCAGGTTTTTAGCAAAATTAGTAGGGATA
Leu144mut_for TATCCCTACTAATTTTGCTAAAAACCTGGCAGCGGTGGG

*[1] analogous primers containing either the T5, T3 or the non canonical terminators indicated in Figure 5A in place of the underlined sequence were used for each site specific mutagenesis reaction.

T7 RNA polymerase transcriptions

T75'Val148_for TAATACGACTCACTATAGGGGGTGTAGCTCAGTGG
T73'Val148_rev AAAAGACCGGAACTAAGTATTC
T73'ValT5_rev AAAAATCGTACTTGTGGAAGTG
T73'ValT3_rev AAAAACGTTATGTTCAACGATAG

MHV68 7SLUAS-tRNA4-miR-M1-5 expression unit

UAST3_for CCGGGTCCCGACGATTTACAGAGTTGAGATCGG
UAST3_rev CCGATCTCAACTCTGTAAATCGTCGGGACCCGG

Northern probes

MHV68 miR-M1-5	GAGACGACCCGATCTCAACTCT
MHV68 tRNA4	CCGCTCTACCAATTGAGCTACC
U6 snRNA	CGAATTTGCGTGTCATCCTTGC

APPENDIX 2

MHV68 7SLUAS-tRNA4-miR-M1-5 artificial construct

GCGGCCGCAACCGCGCCGATAAGCTTCCGGGGTCCCAACCCTGTAAGGCAGAACGGCTGCCCAGCT
CGGATGACGCCACACTAACGTAGCCTCCAGACCGCCCAGTGTGGGTGTGTCCAAGCTCACGTCGCC
CGGCGTGGCCCCCCCGCTCCCCAATGACGTAACTGCCCTGCAGCTTCTAGTAGCTTTTCGCAGCGT
CTCCGACCGTCGGGGTAGCTCAATTGGTAGAGCGGCAGGCTCATCCCCTGCAGGTTCTCGGTTCAA
TCCCGGGTCCCGACGTTTTTCAGAGTTGAGATCGGGTCGTCTCCCCCTGGCGGAAGGAGGCAAACC
CGAGCTCCTCCTTCTTTTTCTCGAG

Chapter 3

Investigating the role of antisense AluRNAs in post- transcriptional gene silencing

For a long time, Alu elements were thought to be “junk” DNA and dismissed as uninteresting. Nowadays, many examples highlight the impact of transcribed Alu elements on the human genome and it is likely that Alu RNAs may serve as master regulators of gene expression by targeting several different steps. In the present report, we investigated a novel kind of post-transcriptional gene silencing carried out by antisense AluRNAs derived from the 3’-UTR of human mRNAs. We found that one of these AluRNAs, TERF2-AluY, is likely to trigger the silencing of its cognate mRNA by targeting its complementary sequence in the 3’UTR and speculated that the target specificity resides in the unique 3’trailer of the AluRNA.

Introduction

With more than one million copies, Alu elements are the most abundant SINEs in the human genome [96]. Altogether they cover about 10% of the genome mass and occur at higher frequency within non coding regions of gene-rich areas [97]. Despite being scattered through the whole genome, for a long time Alu elements were thought to be "junk" DNA and dismissed as uninteresting. Nowadays, the role of these mobile elements is becoming more and more recognized and several lines of evidence show that they had a great influence on the evolution of the human genome. Due to their high sequence conservation, Alu elements provide abundant nucleation points for homologous recombination events that contribute to a notable number of human diseases [109,110]. Depending on the genomic location, Alu sequences can influence the transcription of neighboring loci by providing new promoter features [111], by changing the methylation status of endogenous promoters [111,112] or by creating alternative splice sites which can inactivate or change the function of a gene product [113]. Very recently, it has been demonstrated that pairs of inverted Alu repeats in the 3'-UTR of human genes can form duplex structures able to repress gene expression [114]. Alu elements carry out many other regulatory functions, but little is known about the post-transcriptional fate of Alu-derived transcripts. Given their origin from the 7SL RNA gene [98], Alu elements inherited an internal promoter closely resembling the type 2 promoter of tRNA genes [99] and can thus be transcribed by RNA polymerase III [100]. Despite this potential, endogenous AluRNAs are expressed at very low levels (10^2-10^3 copies per cell, [115,116]), possibly cause of promoter weakness or CpG methylation at the promoter elements [273-275]. Nonetheless, the observation that various cellular stresses (such as viral infection, heat shock and cycloheximide exposure) transiently increase their expression levels [117], has raised the possibility that AluRNAs may play specific roles in human cells. This hypothesis has been substantiated by two recent studies, in which AluRNAs were shown to influence translation initiation [118] and to block transcription by binding RNA polymerase II during the cellular heat shock response [119]. The recent discovery of Alu elements embedded in the 3'-UTR of known mRNAs poses the fascinating question of whether these genomic repeats can influence gene expression, either at the transcriptional or post-transcriptional level. 3'-UTRs are known to play many crucial roles in the post-transcriptional regulation of gene expression [276] and they are the target of microRNAs which provide the sequence specificity for mRNA destabilization or translational repression through RNA-induced silencing complexes (RISCs) [277]. Abundant antisense transcripts with potential regulatory functions are encoded by eukaryotic genomes [278] and siRNAs derived from genomic repeats have already been described for LINE-1 elements [279]. Likewise, due to their secondary structure similarity with the 7SL RNA [280], antisense AluRNAs could provide double-strand precursors that trigger repression of the target mRNA by interacting with homologous sequences located in its 3'-UTR, either via an RNAi-like pathway or via a novel pathway involving uncharacterized proteins. Moreover, secondary structures created by Alu elements embedded in the 3'-UTR could influence translation in other ways; for example, regulatory factors bound to conserved hairpins within the 3'-UTR are known to modulate mRNA turnover by recruiting the decapping machinery [281] and a similar mechanism may be hypothesized for Alu-derived structures. In the present report we start to elucidate post-transcriptional regulatory mechanisms related to the expression of antisense AluRNAs.

Antisense Alu elements embedded in the 3'UTR of human mRNAs are expressed *in vitro*.

The highly structured UTRs significantly affect gene expression as they play a critical role both in translational control and in the targeting of transcripts to specific subcellular compartments. Like the upstream promoter regions, the UTRs do not encode proteins and they have been free to evolve into battlegrounds for a plethora of RNA-binding proteins and miRNAs, that have recently broaden the range of transacting post-transcriptional regulators. UTRs frequently contain Alu elements that may influence the stability and the rate of translation of mRNAs in various ways, such as through the binding of regulatory factors or modifying enzymes [109,114]. Given the latest progresses in ncRNAs regulation, we considered of potential interest a recent report listing a comprehensive compilation of mRNAs with embedded antisense Alu elements in their 3'-UTRs [282], based on the assumption that antisense AluRNAs could be promising substrates for siRNA biogenesis and guide silencing complexes to their targets. In order to investigate experimentally whether antisense AluRNAs may serve a function in post-transcriptional regulation of gene expression, we selected several antisense elements according to type 2 promoter conservation and occurrence of a $\geq T_4$ RNA polymerase III terminator close to the 3'-end of the coding sequence. Nine of the antisense Alu elements with the best scores (**Table 1**) were PCR-amplified from genomic DNA, cloned in a commercial vector and assayed by *in vitro* transcription in HeLa NE. As shown in **Figure 1**, five elements proved to be relatively active compared to the other Alu templates (EEF2K-AluJo, PDE6B-AluSx, EIF2S3-AluSx, CYCS-AluSp and TERF2-AluY) and gave transcripts of the expected size for termination at the canonical $\geq T_4$ signal. Consistent with the hypothesis that transcription of the Alu templates is supported by RNA polymerase III, synthesis of three of the

Gene name	Ensemble gene	SwissProt-TREMBL annotation	Alu element				
			begin-end	family	A-box score	B-box score	T-distance (from 3'-end)
BCL10	ENSG00000142867	B cell lymphoma/leukemia 10 (B cell CLL/lymphoma 10) (Bcl 10) (CED	1029-714	AluJb	-33.68	-8.75	50 nt
CYCS	ENSG00000172115	Cytochrome c	2858-2559	AluSp	-33.68	-8.75	18 nt
EEF2K	ENSG00000103319	Elongation factor-2 kinase{EI1}.	584-274	AluJo	-30.99	-7.55	59 nt
EIF5A2	ENSG00000163577	Eukaryotic translation initiation factor 5AII (EIF-5A2)	1216-912	AluSx	-34.86	-8.75	134 nt
EIF2S3	ENSG00000130741	Eukaryotic translation initiation factor 2 subunit 3	574-279	AluSx	-32.19	-8.75	74 nt
LAMP3	ENSG00000078081	Lysosomal-associated membrane protein 3{EI1	804-506	AluSq	-33.37	-8.75	111 nt
PDE6B	ENSG00000133256	Rod cGMP-specific 3',5'-cyclic phosphodiesterase beta-subunit	434-133	AluSx	-34.31	-7.55	25 nt
RRP30	ENSG00000148688	Ribonuclease P protein subunit p30 (EC 3.1.26.5) (RNaseP protein p30	1420-1143	AluY	-29.62	-12.39	7 nt
TERF2	ENSG00000132604	Telomeric repeat binding factor 2 (TTAGGG repeat binding factor 2)	377-69	AluY	-34.50	-10.15	89 nt

Table 1. Human mRNAs containing antisense Alu elements in their 3'UTR that were assayed by *in vitro* transcription. Gene name, Ensemble gene name and SwissProt-TREMBL annotation are from a "Compilation of human mRNAs containing Alu elements in their 3'UTR" [282]. For each Alu element are indicated the begin and the end of the coding sequence relative to the 3'UTR of the corresponding mRNA and the distance of a canonical $\geq T_4$ RNA polymerase III terminator from the 3'-end of the Alu coding sequence. Alu repeats were further screened with the Repeat Masker software [283] and classified in their relative subfamily according to [284].

transcriptionally active AluRNAs was completely abolished by a 4 bp substitution in the B-box internal promoter, as measured by a density quantification software (**Figure 2A**).

The RNA polymerase III-derived antisense AluRNAs indicated in figure 2 potentially represent good candidates to test *in vivo* the RNAi-like model presented in the introduction. Nonetheless, the observation of a second reverse oriented AluJo in close proximity to EEF2K-AluJo (**Figure 2B**) gave us a clue to exclude this element from further analysis. Pairs of inverted Alu repeats within the 3'UTR of mRNAs have been recently reported to strongly repress expression of the cognate protein, probably cause of A-to-I RNA editing by the ADAR enzymes [114]. To avoid potential interactions between the two mechanisms, we decided to concentrate our efforts on TERF2-AluY and EIF2S3-AluSx. It shall be noted that a second Alu element in the 3'UTR of EIF2S3 (indicated with AluSx2 in **Figure 2B**) could account for a similar side effect. However, considering that pairs of direct repeats have never been reported to affect gene silencing (probably due to their inability to form an intramolecular dsRNA) and that AluSx2 proved to be transcriptionally inactive by *in vitro* transcription in HeLa NE (data not shown), we considered this second repeat to be trivial for our model.

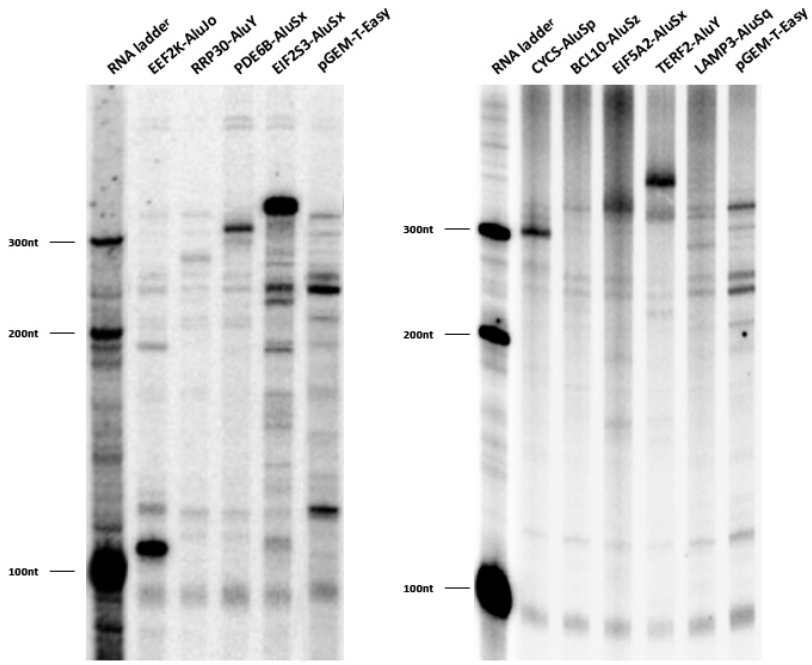


Figure 1. *In vitro* template activity of antisense Alu elements embedded in the 3'-UTR of human mRNAs. The antisense Alu elements indicated in the figure were PCR amplified from buccal cells genomic DNA and cloned in pGEM-T-Easy (*Promega®*) as described in *materials and methods*. *In vitro* transcriptions in HeLa NE were performed under standard conditions using 500ng of each constructs or the empty vector. The size of the transcripts is consistent with RNA polymerase III terminated at the canonical $\geq T_4$. The 0.1-2Kb RNA ladder (*Invitrogen®*) was end-labeled at the 5'-end with T4 polynucleotide kinase (*Promega®*) and [γ-^{32}P]ATP and loaded onto the left-most wells.

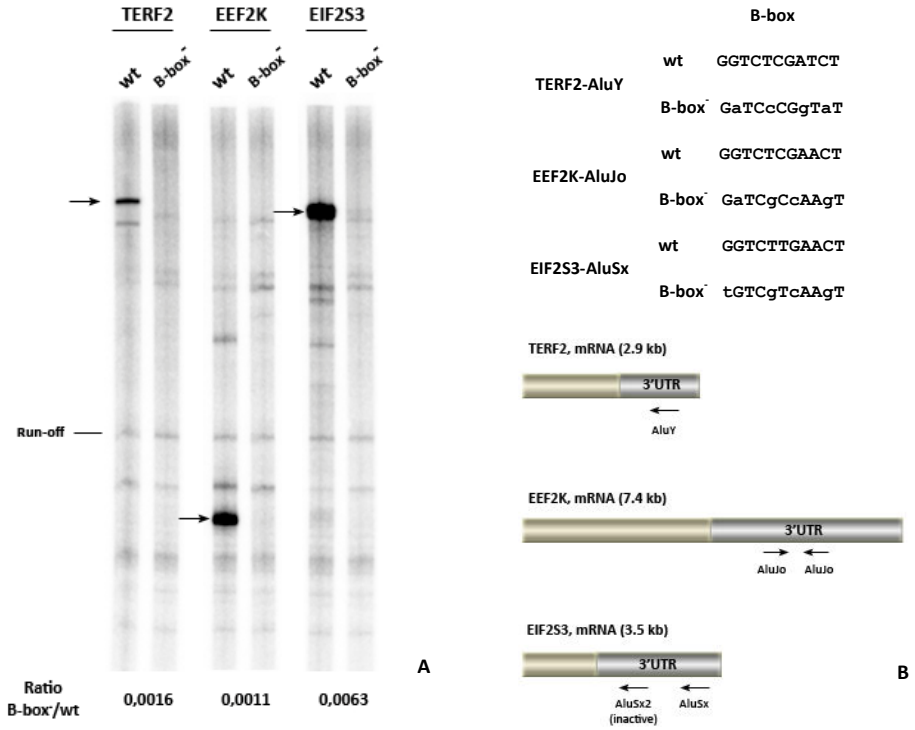


Figure 2. **A**. A 4 bp substitution in the B-box of the internal promoter of the wild-type gene (wt) virtually abolished transcription by the HeLa NE extract, confirming dependence of the selected templates on RNA polymerase III. The amount of each transcript (indicated by the arrows) was calculated by phosphorimager quantification and normalized to a 151 nt internal standard (Run-off). The ratio between the wild-type (wt) and the B-box mutated version (B-box$^-$) of each antisense AluRNAs is indicated at the bottom of the gel. **B.** Schematic representation of TERF2, EEF2K and EIF2S3 mRNAs with Alu elements embedded in their 3'UTRs. The Alu elements in analysis are transcribed in antisense orientation with respect to the sense of transcription of the corresponding mRNA, as indicated by the arrows. A second sense Alu element in the 3'UTR of EEF2K (AluJo) and a second inactive antisense Alu element in the 3'UTR of EIF2S3 (AluSx2) are indicated. Above, are represented the base pairs mutations (in smaller letters) introduced in the B-box of AluY, AluSx and AluJo.

TERF2-AluY and EIF2S3-AluSx are differently expressed in HeLa and HEK293T cells

Despite the high expression levels previously observed *in vitro* for several of the selected sequences, transfected Alu elements with the sole internal promoter usually give hardly detectable transcripts [273,274] and endogenous Alu elements are mostly kept silent in the cell, probably due to promoter weakness or promoter methylation [275,286]. Several studies have demonstrated the importance of upstream flanking sequences for enhanced transcription of endogenous Alu elements [285] and it was reported that the 5'-flanking sequence from several class III genes dramatically stimulates Alu activity both *in vitro* and *in vivo* by transient transfection of human cells. To address whether TERF2-AluY and EIF2S3-AluSx are transcriptionally active *in vivo*, we constructed an Alu expression vector driven either by the 7SL Upstream Activating Sequence (UAS) or by the H1 promoter, the rationale for which is presented in **Figure 3A**. By *in vitro* transcription in

HeLa NE, the 7SL UAS proved to be the most efficient sequence, with the highest expression levels for both TERF2-AluY and EIF2S3-AluSx (**Figure 3B**). Similarly, northern analysis of RNA samples derived from HEK293T and HeLa cells transiently transfected with the same vectors shown that insertion of the Alu elements downstream of the 7SL UAS resulted in a ~2 fold increase in template activity relative to the RNA polymerase III-dependent H1 promoter (**Figure 4A,D**). The data were further confirmed by qRT-PCR, using primers matching either the 3'-trailer (**Figure 4B,E**) or the Alu coding sequence (**Figure 4C**). Notably, a much weaker increase with respect to the empty plasmid was revealed using primers specific for a conserved region of the Alu body: given the high conservation of the coding region among different Alu subfamilies, these data are unsurprising and a stronger background brought about by endogenous Alu transcripts was mostly expected. Worth of mention is the observation that the relative abundance of transfected AluRNAs is much higher in HEK293T than HeLa cells (compare Figure 4B and 4E). As shown by northern blotting, both elements appear to be expressed with comparable efficiency and we tend to exclude that expression of transfected Alu elements in HeLa cells is partially impaired, as previously reported by other authors [273]. Considering that we used probes against the unique 3'-trailer and that basal expression of Alu RNAs in HeLa cells is very poor compared to other cell types [275], we consider unlikely that cross-hybridization of the probes with endogenous transcripts entails for the differences. Among other possible explanations, we speculate that the two Alu elements under examination are endogenously more expressed in HeLa than in HEK293 cells.

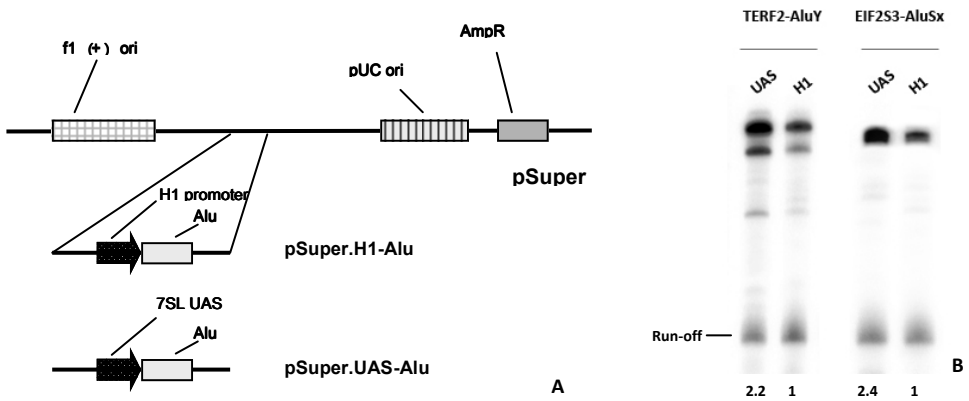


Figure 3. A. To express TERF2-AluY and EIF2S3-AluSx in human cultured cells, the pSuper (*Oligoengine*) vector was modified as follows. In pSuper.H1-Alu, the Alu element (Alu body + 3'trailer + RNA polymerase III termination signal) was cloned downstream of the H1 promoter and transcribed in the sense indicated by the black arrow. In pSuper.UAS-Alu, the H1 promoter was replaced with the 7SL Upstream Activating Sequence [226] and the Alu element was cloned downstream of this promoter region. The 7SL UAS virtually enhances RNA polymerase III recruitment on TFIIIB/TFIIIC-bound type 2 promoters, while the H1 flanking sequence should overcome the need of an internal promoter and allow Alu RNAs expression *in vivo* in case of A-box and B-box weakness. **B.** *In vitro* transcription in HeLa NE reveals that the 7SL UAS support the highest rate of transcription for both Alu elements. The amount of each transcript was calculated by phosphorimager quantification and normalized to a 151 nt internal standard (Run-off). Fold increase relative to H1-driven templates is reported under the gel.

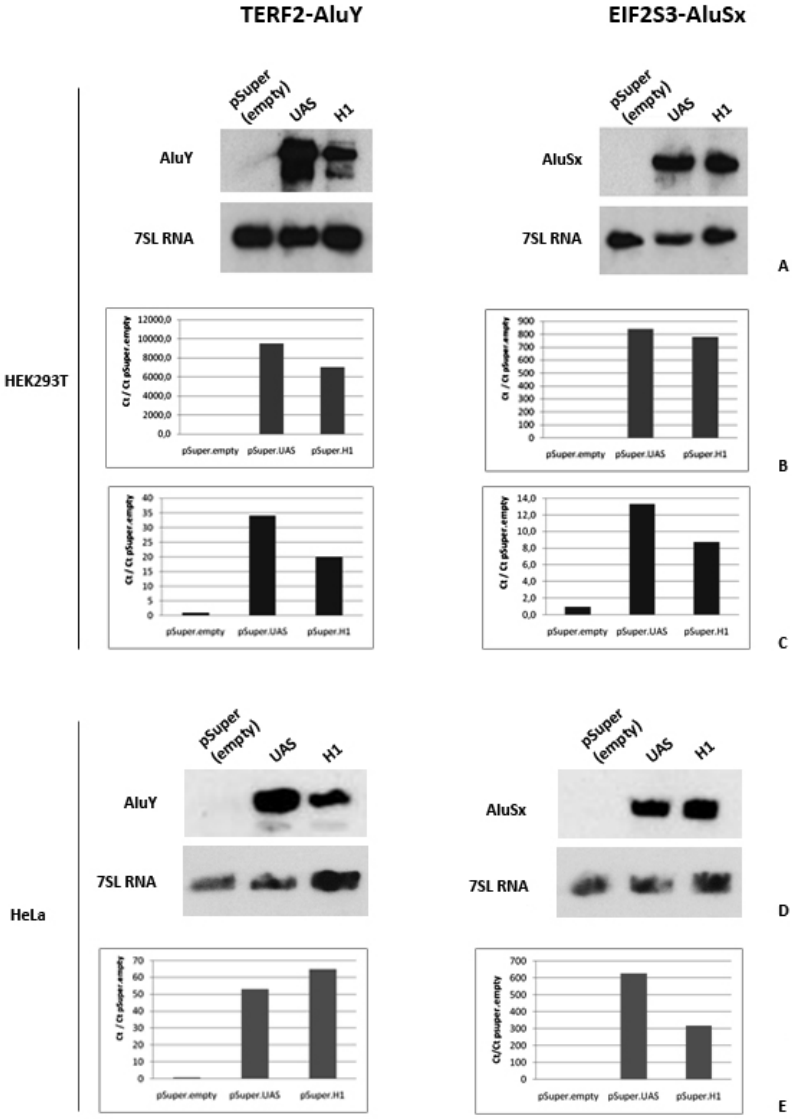


Figure 4. Transfected Alu elements are expressed in HEK293T and HeLa cells. **A.** Northern blot analysis of total RNA (15µg) from HEK293T cells. HEK293T cells were transfected with equal amounts of the AluRNA expressing vectors pSuper.UAS and pSuper.H1, or the empty vector (pSuper.empty). Specific Alu transcripts were detected using end-labeled probes for the unique 3'-trailer. No bands were detected in the mock transfection with the empty vector. As an internal control, the same filters were stripped and reprobed with a probe specific for the S-domain of the 7SLRNA. The 7SLRNA is ancestrally related to the Alu body and is coincidentally 300nt. However, the S-domain is unreleated to any sequence in the Alu body and the Alu RNAs under examination are ~400nt (Alu body + 3'-trailer), thence the bands observed are not due to cross hybridization. **B,C.** Total RNA from HEK293T cells transfected as previously described was analyzed by qReal-Time PCR using primers specific for the 3'-trailer (B) or the coding region (C) of the selected AluRNAs. Expression data for TERF2-AluY and EIF2S3-AluSx are shown as relative to a control sample (RNA from cells transfected with the empty vector) and calculated using the $2^{-\Delta\Delta Ct}$ method [287]. The Ct values of both the control and the samples are normalized for the 7SL RNA. **D**. Similar to panel A, except that total RNA was extracted from transfected HeLa cells (see *materials and methods* for details on the transfection protocol). **E**. qReal-Time PCR anlaysis of total RNA from HeLa cells is as previously described in panel B,C, except that quantitative analysis was not performed with primers for the coding region.

Transfected AluRNAs do not silence gene expression in a luciferase reporter system

To measure possible biological activity from TER2-AluY and EIF2S3-AluSx, we cloned different variants of the target 3'UTRs downstream of a firefly luciferase reporter gene (**Figure 5**). Reporter activity, normalized to renilla luciferase, was initially assessed in HEK293T cells upon co-transfection of pMIR-RL_AluY with increasing amounts of pSuper.UAS-AluY. As shown in **Figure 6A**, expression from pMIR-RL_AluY relative to pMIR-RL control plasmid was basically unchanged, regardless of increasing amounts of co-transfected pSuper.UAS-AluY. Similar results were obtained upon co-transfection of pMIR-RL_AluSx with increasing amounts of pSuper.UAS-AluSx (**Figure 6B**). Taken together, these data suggest a lack of productive interactions between the Alu RNAs and their complementary sequence in the chimeric firefly luciferase mRNA, but does not take into account that additional regions in the 3'UTR, apart from the Alu element, might contribute to gene silencing. Inconsistent with the notion that AluRNAs are generally able to block transcription [119], 48h after transfection of HEK293T cells with a TERF2-AluY RNA expressing vector, the activity of the reporters in analysis was mostly unchanged, if not slightly enhanced (**Figure 6C**), relative to the

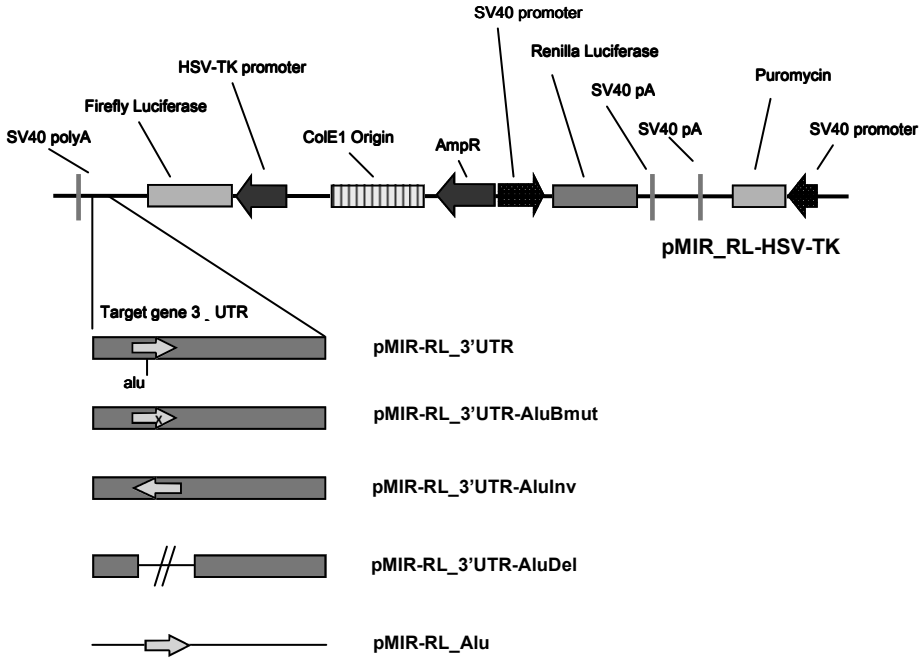
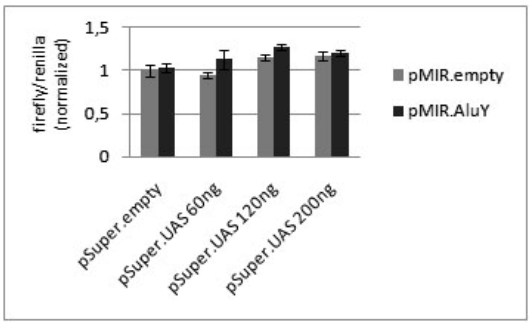


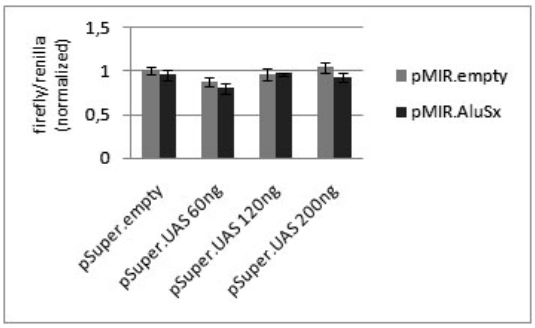
Figure 5. Vectors assayed in the dual luciferase reporter system. To generate reporter plasmids suitable for the dual luciferase reporter assays, the pMIR-REPORT (*Ambion®*) was modified as described in [37]. The vectors were further modified to express the firefly luciferase mRNA with different variants of the TERF2 and EIF2S3 3'UTRs. The pMIR-RL_Alu contains a perfect target site for TERF2-AluY and EIF2S3-AluSx, respectively; the Alu element is in antisense orientation with respect to the sense of transcription and an AluRNA derived from this element would be fully complementary to the corresponding region at the 3'end of the firefly luciferase mRNA. In pMIR-RL_3'UTR, the whole 3'UTR of the potential target mRNAs was cloned behind the firefly luciferase gene; the light grey arrow indicates the sense of transcription of the Alu element. pMIR-RL_3'UTR was then modified as follow to create three additional variants. In pMIR-RL_3'UTR-AluInv the Alu elements was inverted. In pMIR-RL_3'UTR-AluBmut the Alu element was modified by a 4bp mutation at the B-box promoter. In pMIR_3'UTR-AluDel the Alu element was deleted; this last reporter was created for EIF2S3 only, to evaluate possible side effects brought about by the AluSx2 sequence.

mock transfection with pSuper empty vector. Similarly, transfection of HEK293T cells with pSuper.UAS-AluSx gave approximately the same pattern of mock transfection; however, a great variety among pMIR-RL vectors carrying different 3'UTRs was observed in the firefly luciferase readout. Surprisingly, the activity of the reporter with the whole 3'UTR was 3-fold higher than that of the corresponding B-box mutated version (**Figure 6D**). It is not clear why four point mutations at the B-box level affect so markedly luciferase expression. A sequence specific effect (i.e. the creation of a microRNA binding site) is reasonable, while the hypothesis of an RNA polymerase II/RNA polymerase III interplay is unlikely, considering that the activity of the reporter devoid of the Alu element approximate that of the one with the whole 3'UTR.

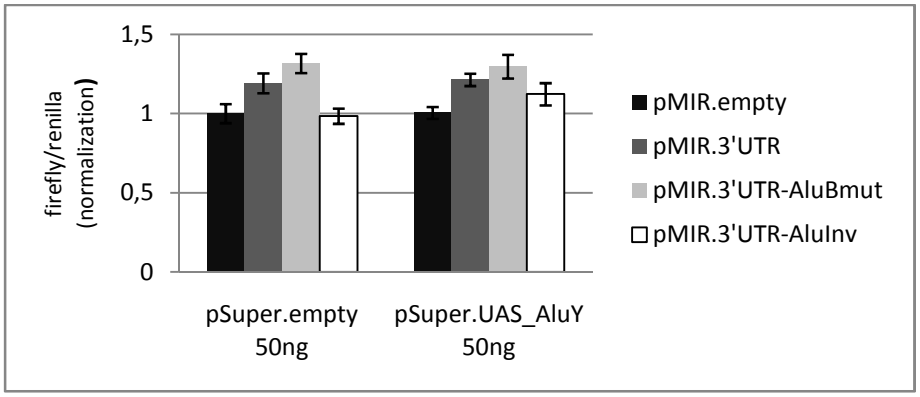
In summary, the data from the luciferase assays appear to be against the proposed RNAi-like model and suggest that the TERF2 and EIF2S3 3'UTRs are not targeted by the corresponding antisense AluRNAs. Nonetheless, we cannot exclude that the artificial expression of a reporter gene is biased in several ways. In the titration assays, expression from the strong HSV-TK promoter could lead to sub-saturating amounts of the firefly luciferase mRNA, making it insensitive to co-expressed AluRNAs. Moreover, comparing luciferase activity from pMIR-RL vectors carrying different 3'UTRs does not take into account that the variability in the downstream sequences could affect the stability of the corresponding mRNA. Lastly, a silencing pathway that entails co-transcriptional instead of post-transcriptional gene targeting would rely on genomic features, such as nucleosome assembly or chromatin modifications, that the transfected plasmids probably lack. For these and other reasons, we decided to investigate the effect of AluRNAs expression on the endogenous cellular targets TERF2 and EIF2S3.



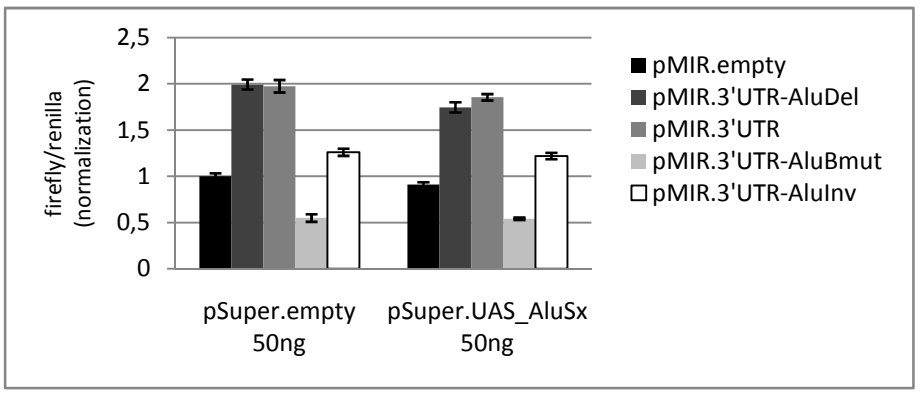
A



B



C



D

Figure 6. Effects of TERF2-AluY and EIF2S3-AluSx RNA on the target 3'UTRs in a luciferase reporter system. **A.** A reporter construct containing a perfect complementary site for TERF2-AluY RNA behind the firefly luciferase gene (pMIR-RL_AluY) or the empty vector (pMIR.empty), were cotransfected with increasing amount of pSuper.UAS-AluY as indicated in the figure. 48hrs after transfection, firefly luciferase activity was normalized to renilla activity. Error bars are derived from eight independent transfections and the same experiment was repeated twice. **B.** Similar to panel A, with the difference that the titration assay was carried out for EIF2S3-AluSx and the relative target site in pMIR-RL_AluSx. **C.** Several reporter constructs containing different versions of the TERF2 3'UTR (for a detailed description, see Figure 5) were cotransfected in HEK293T cells together with pSuper.UAS_AluY or pSuper.empty (mock). 48hrs after transfection, firefly luciferase activity was normalized to renilla activity. Error bars are derived from five independent transfections and the same experiment was repeated twice. **D**. Similar to panel C, with the difference that reporter activity was measured upon co-transfection of pSuper.UAS_AluSx with pMIR-RL constructs engineered with different variant of the EIF2S3 3'-UTR. With respect to the previous analysis, an additional vector containing the 3'UTR devoid of AluSx was made. 48hrs after transfection, firefly luciferase activity was normalized to renilla activity. Error bars are derived from five independent transfections and the same experiment was repeated twice.

Over-expression of TERF2-AluY and EIF2S3-AlSx differently affects target mRNA levels in transfected HEK293T, but not HeLa cells

The data provided so far tend to exclude the RNAi-like model proposed in this project, but do not show any evidence that the same scenario applies to the AluRNA regulation of endogenous targets. Based on the previous observation that the 7SL UAS supported higher expression levels for

both Alu elements and that H1 driven RNAs are generally more toxic for the cell [288,289], the relative amount of TERF2 and EIF2S3 mRNA was analyzed via qReal-Time PCR upon transfection of HEK293T cells with pSuper.UAS-AluY or pSuper.UAS-AluSx. Considering that AluRNAs are able to bind RNA polymerase II and are found at the promoter of certain mRNA genes [119,290], data were normalized for two RNA polymerase III transcripts commonly used in qRT-PCR, namely U6 snRNA and 5S rRNA. As shown in **Figure 7**, the levels of four nuclear transcripts surprisingly increased after over-expression of AluSx, including those of the potential target EIF2S3 and the housekeeping genes GAPDH and B2M. Similarly, over-expression of AluY led to an enhancement, although less marked, in B2M, GAPDH and EIF2S3 mRNA levels, but, strikingly, determined a sharp reduction in the TERF2 mRNA expression levels relative to the mock transfection. This result strongly argue for a specific role of TERF2-AluY RNA in the silencing of its cognate mRNA, which is likely to rely on the targeting of its complementary sequence in the 3'UTR. On the other hand, EIF2S3Sx appears to globally stimulate RNA polymerase II transcription, in contrast with a previous paper reporting that increase expression of AluRNAs upon heat-shock has the effect to physically

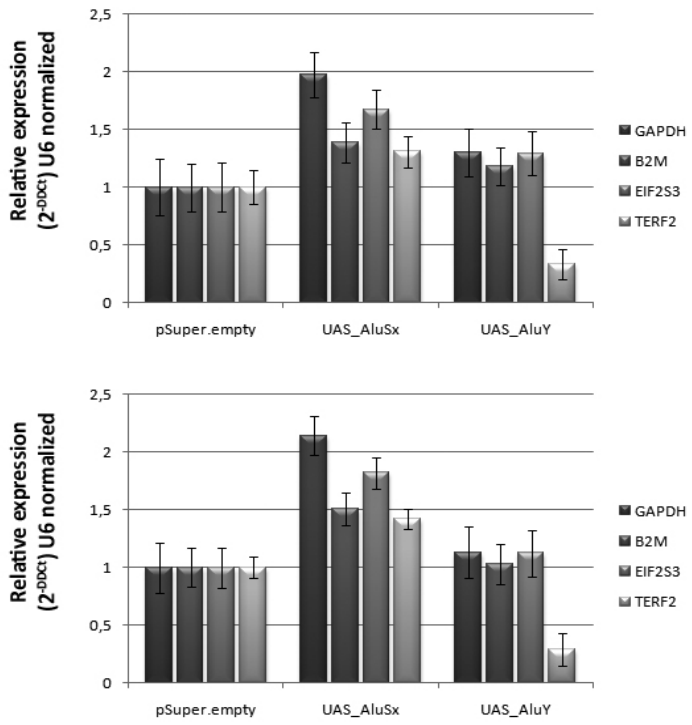


Figure 7. Expression of AluSx and AluY differently affects mRNA levels of the endogenous targets and two housekeeping genes. Total RNA from HEK293T cells transfected with pSuper.empty, pSuper.UAS-AluSx or pSuper.UAS-AluY was analyzed by qReal-Time PCR to assess the expression levels of four messenger RNAs: TERF2 (telomeric repeat binding factor2), EIF2S3 (eukaryotic translation initiation factor 2 subunit 3 gamma), GAPDH (glyceraldehyde 3-phosphate dehydrogenase) and B2M (Beta 2 microglobulin). Expression data are shown as relative to a control sample (RNA from cells transfected with the empty vector) and calculated using the $2^{-\Delta\Delta Ct}$ method [287]. The Ct values of both the control and the samples are normalized to 5S rRNA or U6 snRNA as indicated in the figure. Each assay was done in triplicate and each bar is the average of two independent RNA samples.

repress RNA polymerase II [290]. Similar assay were performed in HeLa cells, but the expression levels of the four mRNAs were mostly unchanged upon transfection with either TERF2-AluY or EIF2S3-AluSx (data not shown).

Investigating the role of the 3'trailer of TERF2-AluY in target gene silencing

Given the high similarity between the coding sequences of EIF2S3-AluSx and TERF2-AluY, we consider unlikely that the underlined specificity towards the target mRNAs resides in this region, otherwise a similar decrease in TERF2 mRNA levels would have been observed after over-expression of EIF2S3-AluSx. Alu elements are thought to diffuse and integrate in the human genome via a mechanism known as retroposition [104]. This model implies that the integration occurs between RNA and DNA near a nicking site generated by an endonucleolytic enzyme, preferably at the consensus sequence TTAAAA. Alu elements probably undergo retroposition at this site cause of their A-rich 3'-tail, which may specifically interact via Watson-Crick base pairing with chromosomal targets prior to integration. As a result of retroposition, the 3'trailer of Alu elements bears unique sequence features that are not shared among the members of this family. Along with this observation we hypothesized that the sequence specificity of TERF2-AluY for its cognate target mRNA could eventually reside in the unique 3'-trailer. To validate this hypothesis, we deleted the 3'trailer from TERF2-AluY and cloned the corresponding PCR fragment in pSuper.UAS to generate pSuper.UAS-AluY3'less. As expected, the plasmid gave rise to a shorter AluRNA by *in vitro* transcription in HeLa NE, which was expressed with efficiency comparable to the wild-type Alu element, plus a longer read-through transcript most likely due to transcription termination at a 20bp downstream T-rich site inside the vector (**Figure 8**). Notably, TERF2-AluY3'less co-migrated with a smaller band derived from TERF2-AluY transcription, that we had also previously found by northern blotting (**Figure 4A**), but considered aspecific since the filter had been hybridized with a probe for the unique 3'-trailer. Given that this transcript appears in all the transcription assay with the wild-type Alu element and that it co-migrate with the TERF2-AluY RNA lacking the 3'trailer, we decided to investigate more in depth the mechanism leading to its biogenesis. To further clarify this point, *in vitro* transcription in HeLa NE of TERF2-AluY was carried out under standard conditions and subsequent chased with UTP 2mM for the indicated times (**Figure 9**). The experiment shown a general, aspecific degradation of the AluRNA, but failed to identify any enrichment in the smaller sized products, thus excluding that it comes from the processing of the longer one and suggesting that it rather arise from premature termination. To better understand the nature of the shorter transcript, we performed a cold *in vitro* transcription

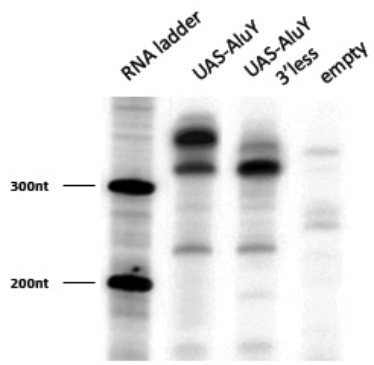


Figure 8. *In vitro* transcriptions in HeLa NE were performed under standard conditions using 500ng of pSuper (empty), pSuper.UAS-AluY or a deletion mutant were the 3'trailer of TERF2-AluY was removed (a canonical RNA polymerase III terminator was introduced at the end of the poly-A tail to efficiently terminate transcription). The size of the transcripts is consistent with RNA polymerase III terminated at the canonical $\geq T_4$. The 0.1-2Kb RNA ladder (*Invitrogen®*) was end-labeled at the 5'-end with T4 polynucleotide kinase (*Promega®*) and [γ-^{32}P]ATP and loaded onto the left-most well.

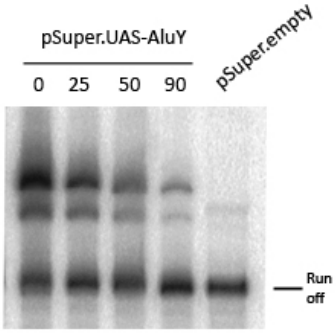


Figure 9. Pulse-chase reaction. *In vitro* transcription in HeLa NE of TERF2-AluY was carried out under standard conditions and subsequent chased with UTP 2mM for the indicated times. A 220nt labeled run-off transcripts was loaded into each well as an internal standard.

assay with pSuper.UAS-AluY and pSuper.UAS-AluY3'less, followed by hybridization of the unlabeled transcripts with probes specific for different regions of the TERF2-AluY RNA, as depicted in **Figure 10**. Consistent with the hypothesis that the shorter transcript originates from premature termination, the probe against the coding sequence of TERF2-AluY (YNBtra2) resolved into a doublet, while the probe against the unique 3'-trailer (YNBtra2) gave a single signal (**Figure 11**). Similar results were obtained for the transcription assay with pSuper.UAS-AluY3'less, with the exception that YNBtra2 found no transcripts, as expected for the Alu element lacking the 3'-trailer. Remarkably, the probe matching the sequence immediately downstream the poly(A)-tail (YNBtraLong, the same probe used in the *in vivo* northern experiment) proved to target the shorter transcripts in both lanes, probably due to cross-hybridization of its 3'end with the Alu elements. We speculate that this cross-hybridization is caused by premature termination at the non-canonical element TTATT (underlined in **Figure 10**), which was identified in the previous chapter. *In vitro* and *in vivo* assay to further investigate the role of this non-canonical terminator in the synthesis of the downstream 3'-trailer are currently on the way, as well as experiment to further verify the role of the unique 3'-trailer in the targeting of the TERF2 mRNA.

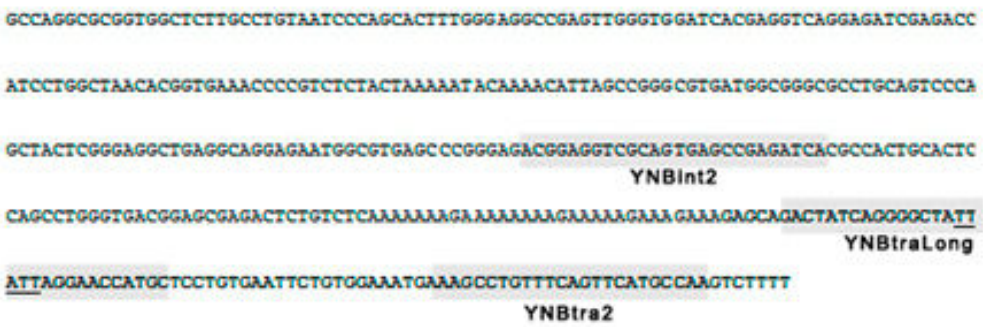


Figure 10. Sequence of TERF2-AluY with indicated probes complementary to the coding sequence (YNBint2) or the 3'trailer (YNBtraLong and YNBtra2).

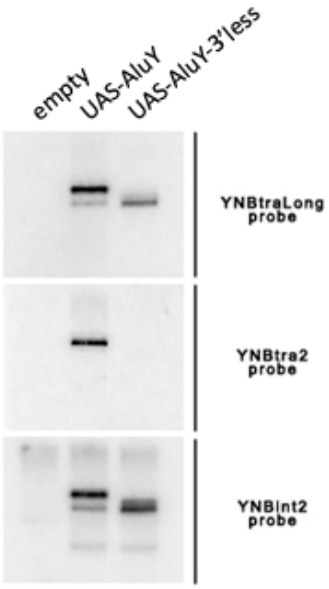


Figure 11. *In vitro* transcription in HeLa nuclear extract of the three vectors pSuper.UAS-AluY, pSuper.UAS-AluY3'less and pSuper.empty was performed under standard condition, with the exception that the reaction was run without [α-^{32}P]UTP. Unlabeled RNAs were purified as described in materials and methods, loaded onto a 6% polyacrylamide gel and fixed to a nylon membrane. Filters were probed either with YNBtraLong, YNBtra2 or YNBint2 and analyzed at the phosphorimager.

Discussion

At their origin, Alu elements were probably not more than selfish DNA, but the genome clearly adapted to their presence by assigning them some important functions, which most likely participated in primate evolution and helped in their divergence from other mammals [291]. Alu elements should rather be considered as a huge reservoir of potential regulatory functions and many examples highlighted their impact either at the post-transcriptional level. Whereas insertion of Alu elements may have beneficial effects, the same process could also be detrimental for the organism. By inserting into essential genes, Alu elements can indeed annihilate important physiological functions; moreover, because of their high level of sequence homology, they can also mediate important disease-causing genomic rearrangements [109,110,292]. The duality of the effect of Alu element insertion demonstrates that Alu amplification has not been a neutral process; on the contrary, it shows that Alu elements actively participated in genome evolution. Alu elements and more specifically RNAs, have been recently reported to affect gene expression in several ways, either they are transcribed thanks to their own RNA polymerase III promoter or as part of longer mRNAs (embedded Alu RNAs). A-to-I editing of Alu sequences has been shown to modulate gene expression of inverted repeat Alu containing mRNAs by titrating the amount of mRNA that is allowed to reach the cytoplasm [114]. Alu RNAs have been found to stimulate the translation of reporter mRNAs in a cell free translation system, while Alu RNPs act as general inhibitors of protein translation [118]. Earlier this year, B2 and Alu RNAs were shown to prevent RNA polymerase II from properly engaging the DNA during closed complex formation, resulting in complexes with an altered conformation that are transcriptionally inert.

In the present report we focused on the effect of transcribed antisense Alu elements on post-transcriptional gene silencing. Human 3'UTRs contain 953 embedded Alu elements in antisense orientation with respect to the sense of transcription of the corresponding RNA [282]. Considering that 3'UTRs are preferred targets of small antisense RNAs in a process known as RNA interference, we speculated that antisense Alu RNAs on their own, or small antisense RNAs derived from the latter, could repress gene expression in a similar way as previously reported for miRNAs. Once selected several of these elements based on A-box / B-box conservation and terminator distance from the 3'-end of the coding sequence, we found most antisense Alu elements to be transcribed in a cell-free system. Three of these elements (two full-length and one left-monomer Alu RNAs) gave particularly abundant RNAs of the expected size and were further assayed to validate RNA polymerase III transcription from their own internal promoter by simple B-box mutagenesis; for reasons explained in the text, we focused our attention on EIF2S3-AluSx and TERF2-AluY. Both Alu RNAs were efficiently expressed *in vivo* under the control of the 7SL UAS or the H1 promoter, but none of them efficiently repressed its antisense target sequence in a dual luciferase reporter assay. However, in contrast with this first observation, TERF2-AluY, but not EIF2S3-AluSx, proved to able to specifically repress its cognate target TERF2 mRNA, upon transfection of HEK293T cells with a 7SLUAS-driven TERF2-AluY expressing vector. Based on the high sequence conservation among different Alu elements and the retroposition mechanism by which Alu elements spread through the genome, we hypothesized that the target specificity of TERF2-AluY would reside in its 3'trailer. By generating a vector that express a TERF2-AluY version lacking the 3'trailer, we found that this unique sequence is likely to be transcribed as result of partial RNA polymerase III read-through at a non-canonical terminator downstream of the poly(A)-tail. In the immediate future, TERF2-AluY3'less will be assayed *in vivo* by transfection of HEK293T cells and the data will be further confirmed by Western Blotting to investigate a hopeful decrease in the TERF2 protein levels. If a 3'trailer mediated effect on the target mRNA will be found in the future, TERF2-AluY could represent an example of biogenesis of a regulatory RNA thanks to RNA polymerase III read-through at a non canonical site.

MATERIAL AND METHODS

Construction of plasmids

Using oligonucleotides defined in appendix 1, nine constructs containing the different Alu sequences indicated in Table 1 were generated by PCR on 100ng of buccal cells genomic DNA with recombinant Taq DNA polymerase (*Fermentas*) as described by the manufacturer. The resulting PCR products were purified with Nucleospin Extract II columns (*Fermentas*) and TA-cloned into pGEM-T-Easy® (*Promega*) as described by the manufacturer, with the exception that the cloning reaction was carried out with 40ng of plasmid DNA and 8U of T4 DNA ligase. Plasmids were transfected and propagated in XL1-B cells, recovered with the *QIAGEN* MIDI plasmid kit and verified by base sequencing. The final sequences were the same available in the NCBI database, with the exception of mutations T→C at position 21 of PDE6B-AluSx, T→C at position 105 of EIF2S3-AluSx, G→A at position 57 and 66 of CYCS-AluSp and T→C at position 150 of TERF2-AluY, which are likely to be cloning mutations. The 4-bp substitution in the B-box of TERF2-AluY, EIF2S3-AluSx and EEF2K-AluJo is as indicated in Figure 2B and was introduced by site specific mutagenesis with Pfu DNA polymerase (Promega) and the primers indicated in appendix 1. To prepare the plasmids employed in the *in vivo* transcriptional assay, TERF2AluY and EIF2S3AluSx were cloned downstream of the H1 and 7SL RNA genes upstream flanking region. To generate pSuper.H1-TERF2AluY and pSuper.H1-EIF2S3AluSx we cloned a BglII/XhoI TERF2AluY or EIF2S3AluSx-containing fragment (Alu RNA coding region + 3'trailer + RNA polymerase III terminator) in the MCS of an Amp^R → Kan^R version of pSuper (*Oligoengine*), downstream of the H1 promoter. Restriction sites were introduced in the primers used for the PCR amplification of the two Alu elements from 50ng of the corresponding pGEM-T-Easy vectors indicated above. To generate pSuper.UAS-TERF2AluY and pSuper.UAS-EIF2S3AluSx, we replaced the XhoI/NotI H1 promoter-containing fragment of pSuper with the -206/+1 upstream flanking region of the human 7SL RNA gene (Accession No. NC_000014.8, also known as RN7SL2 or 7SL1c) fused by PCR to EIF2S3-AluSx and TERF2-AluY, respectively; the fragments were fused in such a way to locate the TSS of each Alu element immediately downstream of the 7SLUAS, resembling the endogenous genomic organization of the 7SLRNA gene. To create the construct pSuper.UAS-TERF2AluY3'less we first PCR amplified from 50ng of pSuper.UAS-TERF2AluY a TERF2AluY fragment lacking part of the 3'trailer: the reactions were carried out with the primers indicated in appendix 1 which also introduced a XhoI site 3' to the Alu coding sequence. The TERF2AluY3'less fragment was then PCR combined with the 7SLUAS as previously described and XhoI/NotI cloned in pSuper.

To generate AluRNAs reporter plasmids, the pMIR-RL vector described in [37] was modified by inserting different target sequences downstream of the firefly luciferase gene; unless otherwise reported, cloning strategy was the same for generation of both TERF2AluY and EIF2S3AluSx pMIR-RL reporter plasmids. pMIR-RL_Alu plasmids were constructed by inserting either EIF2S3AluSx or TERF2AluY into the PmeI site (this and other restriction sites described herein for generation of luciferase reporter plasmids are located into the MCS between the stop codon of the firefly luciferase gene and the SV40 polyA tail sequence). The two Alu elements were PCR amplified from 1 ng of pGEM-T-Easy®-EIF2S3AluSx or pGEM-T-Easy®-TERF2AluY, respectively, with Phusion DNA polymerase (*Finnzymes*) under conditions suggested by the manufacturer; PCR fragments were cloned blunt into pMIR-RL with the one-step method described in [296]. pMIR-RL_3'UTR plasmids were generated by inserting the whole 3'UTR of TERF2 and EIF2S3 mRNAs into the SpeI site of pMIR-RL. The 3'UTRs were amplified via touch-down PCR from 100ng of genomic DNA using

Phusion DNA polymerase (*Finnzymes*) and tailed primers to introduce a SpeI restriction site at both ends. By site specific mutagenesis of these vectors with the same primers previously employed for B-box mutation of EIF2S3AluSx or TERF2AluY, we generated the variants pMIR-RL_3'UTR-AluBmut. pMIR-RL_3'UTR-AluInv plasmids were generated by inverting the sequence +982/+1840 of the TERF2 3'UTR or the sequence +256/+1588 of the EIF2S3 3'UTR; the same inverted region was deleted in pMIR-RL_3'UTR-EIF2S3AluSxDel.

***In vitro* transcription**

In vitro transcriptions in HeLa nuclear extract and the pulse chase reaction were performed as described in the previous chapter. In several reactions, 2500-5000 cpm of a [α-^{32}P]UTP labeled run-off RNA were added as a loading control prior to phenol:chloroform extraction. Labeled run-off RNAs were generated by T7 RNA polymerase transcription on pGEM-T-Easy vectors digested with restriction enzymes leaving 5'-overhanged or blunt ends.

Cell culture and transfection

HEK293T cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat inactivated fetal bovine serum, 100 unit/ml penicillin and 100 µg/ml streptomycin at 37°C in 5% CO2 atmosphere. Approximately 5h prior to transfection, cells were plated at 30% confluency per 10 cm dish. 10 µg of plasmid DNA diluted in 33 mM CaCl2 and 1× HEPES buffered saline (137 mM NaCl, 0.75 mM Na2HPO4, 27.3 mM HEPES-KOH pH 7.1) was added drop-wise under gentle agitation. HeLa cells were grown in monolayers in Dulbecco-modified Eagle's medium supplemented with 10% heat inactivated fetal bovine serum, 100 unit/ml penicillin and 100 µg/ml streptomycin at 37°C in 5% CO2 atmosphere. Approximately 12h prior to transfection cells were plated at 30% confluency per 10 cm dish. 25 µg of plasmid DNA was diluted in half the amounts of Lipofectamine 2000 (*Invitrogen*) and OptiMEM (*Invitrogen*) indicated by the manufacturer and added dropwise to the culture medium. For several reactions, transfections were performed in 6-wells plates, scaling down the reagents volume according to the plate surface. All transfections for luciferase assays were performed in quintuplicate in a 96-well plate containing 40-50% confluent cells grown without antibiotics. Both HEK293T and HeLa cells were transfected with 50ng of pMIR-RL variants and the amount of pSuper vectors indicated in the text, using Lipofectamine 2000 and OptiMEM (*Invitrogen*) as described in manufacturer's instructions. Transfected cells were then incubated at 37°C. 5% CO2 until harvested for luciferase assays or total RNA isolation.

Northern and Real-Time PCR Analysis

30 µg of total RNA isolated from either HEK293 or HeLa cells using Trifast (*Peqlab*) was separated by 6% denaturing RNA PAGE (7M urea, MOPS 1x) and transferred to a nylon membrane (*GE Healthcare*) by semidry electroblotting. Membranes were crosslinked by 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) chemical crosslink incubating for 1hr 20' at 60°C as described in [295]. Membranes were pre-hybridizated for 1h and hybridizated overnight at 30°C with Denhardt reagent 5X, SSC 5X, SDS 0.5%, salmon sperm DNA 0.1 mg/ml and ^{32}P-radiolabeled oligodeoxynucleotides complementary to the Alu RNAs 3'trailer (see appendix 1). After hybridization, membranes were washed once 10 min with 5X SSC, SDS 1%, once 10min with 1X SSC, SDS 1%, and once 15 min with SSC 1X, SDS 0.1%. Radioactive signals were detected by

exposure of BioMax MS film (*Kodak*) using an intensifying screen (*GE Healthcare*). To normalize RNA quantification, filters were stripped twice with boiling SDS 0.15% and reprobed with a ^{32}P-radiolabeled oligodeoxynucleotide complementary to the 7SLRNA S-domain (see appendix 1). Hybridization was carried out overnight at 42°C and membranes were washed twice with 5X SSC, SDS 1% and 1X SSC, SDS 1%, respectively, before exposure.

For northern analysis on RNAs synthesized *in vitro* by transcription in HeLa nuclear extracts, reaction conditions were as described in the previous chapter, with the difference that RNAs were resolved onto a 6% polyacrylamide gel and cross-linked to the membrane by UV-treatment and heat incubation (1hr at 60°C).

For Real-Time PCR analysis, 3µg of total RNA were treated with 3U of DNaseI (*Fermentas*) and reverse transcribed with random hexamers using the cDNA synthesis kit (*Fermentas*) as described by the manufacturer. 1/40th of the resulting cDNA was amplified in triplicate using Mesa Green qPCR MasterMix Plus (Eurogentec) and primers indicated in appendix 1, including a minus RT (Reverse Transcriptase) control reaction to ensure that genomic DNA had not contaminated the RNA samples. Reactions were run for 40 cycles and the mean Ct (only data with standard deviation lower than 0.3 were considered significative) was calculated from each triplicate sample. Expression data were calculated using the $2^{-\Delta\Delta Ct}$ method [287].

Dual luciferase assays

All dual luciferase assays were conducted in quintuplicate in 96-well plates using the Dual-Luciferase Reporter Assay System (Promega) following the manufacturer's protocol. 48 hrs post transfection, medium was removed and 25 µl of *Promega* Passive Lysis buffer 1X rocked at room temperature for 15 min were added to each well. 2.5µl (HEK293T) or 10µl (HeLa) of lysed cells were transferred into a 96-well plate for the luciferase reading which was performed as described by the manufacturer with Renilla and Firefly Sample Buffers (*Promega*).

APPENDIX 1

Primers and probes list

Transcription of the initial Alu elements pool from genomic DNA*[1]

EEF2KAluJo_for	TTGATCAACCCTAAATGCAAAAGC
EEF2KAluJo_rev	GCAGTCAACACTGAAGTCTTGGC
RPP30_for	ATGCAATATTAATGTAAGGGCTCTA
RPP30_rev	CATTGGTGAGGTTATAAAGCTTTC
PDE6B_for	CCTATGGCTCCCTCAATCTTCAC
PDE6B_rev	GGCTACAGTTTATTCACAGTTGGC
EIF2S3_for	GTGACATAGATTATACTACTACTAAT
EIF2S3_rev	CAACGTCAGCAAGAGTTGCAAAT

CYCS_for CAGATGAGTAACTGCCATGTTCTAG
CYCS_rev CACACTGGACAGCAAAAGTCTACAA

BCL10_for TGATACTGGAGGAACACTTGATGG
BCL10_rev AAAGACAAATCTGGCTAATGGC

EIF5A2_for ACAGTGGCAAAGAGGGTTGATTAA
EIF5A2_rev TCTACTTAGAAATGGTTGGAAAGG

TERF2_for AGCGAGTGGGTCAAGGCTGGAGTGC
TERF2_rev TGCTTCTTAGCTCCATGATCTCCAAC

LAMP3_for ATTGTAACTAATACTACTGTGTGTGC
LAMP3_rev TCTTAGTCTAGCGATTCTGGAGCAG

B-box site specific mutagenesis of EIF2S3AluSx, TERF2AluY and EEF2KAluJo

TERF2Bmut_for CGTGTCAGCCAGGATGATCCCGGTATCCTGACCTCGTGATC
TERF2Bmut_rev GATCACGAGGTCAGGATACCGGGATCATCCTGGCTGACACG

EIF2S3Bmut_for CATGTTGCCCAGGCTTGTCGTCAAGTCCTAACCTCAGGTG
EIF2S3Bmut_rev CACCTGAGGTTAGGACTTGACGACAAGCCTGGGCAACATG

EEF2KBmut_for CTGTGTTGCCCAGGCTGATCGCCAAGTCCTGGACTTAAGTGAGC
EEF2KBmut_rev GCTCACTTAAGTCCAGGACTTGGCGATCAGCCTGGGCAACACAG

Amplification of Alu elements cloned downstream of the pSuper (*Oligoengine*) H1 promoter.

EIF2S3SxBgl_for CGGAGATCTGCCGGGCAGAGTGGCTCA
EIF2S3SxXho_rev GGCACTCGAGCAAAAGGTCAAGAAGTAAAAGATG

TERF2AluYBgl_for CGGAGATCTGCCAGGCGCGGTGGCTC
TERF2AluYXho_rev GGCACTCGAGGCTGTGATGATTAAGGATCGC

Amplification of Alu elements cloned downstream of the 7SLUAS

TERF2YUAS_for GTCTCCGACCGCCAGGCGCGGTGGCTC
TERF2AluYXho_rev (as above)

EIF2S3SxUAS_for GTCTCCGACCGCCGGGCAGAGTGGCTCA
EIF2S3SxXho_rev (as above)

TERF2_3lessXhoI_for[*2] GATCTCGAGCAAAAGCCCCTGATAGTCTGCTCT

Amplification of the whole 3'UTR of EIF2S3 and TERF2 mRNAs

EIF2S3whole_for2 GCGACTAGTGAAGTGGATGTAAGATTTATG
EIF2S3whole_rev2SpeI CCGACTAGTATATGCCCATTACACTCAGTA

TERF2whole_for GCGACTAGTGCGAGTGGGTCAAGGCTGG
TERF2whole_rev CCGACTAGTAATGAAGAATGCTACAAGTATGGA

Northern probes*[3]

7SLSd	CTCCTTAGGCAACCTGGTGGTC
SxNBtraLong	CTTATTCCCCATGTCCCTATACTTCGTGTG
YNBint2	TGATCTCGGCTCACTGCGACCTCCGT
YNBtra2	TGGCATGAACTGAAACAGGCTT
YNBtraLong	GCATGGTTCCTAATAATAGCCCCTGATAGTC

Real-Time PCR primers

7SLRT_rev	CTGCTCCGTTTCCGACCT
7SLRT_for	GCACTAAGTTCGGCATCAATA
B2MRT_for	AGGCTATCCAGCGTACTCCA
B2MRT_rev	TCAATGTCGGATGGATGAAA
EIF2S3RT_for	CTGTCCTGGCCACGATATTT
EIF2S3RT_rev	AGGTGTTCCGATGTCTGAGG
SxRTint_for	GTTAGGAGTTCAAGACCAGCC
SxRTtra_for	ACTCTAGCCTAGGTGACGGA
GAPDHRT_for	AGCTCATTTCCTGGTATGACAACG
GAPDHRT_rev	CTTCCTCTTGTGCTCTTGCTGG
TERF2RT_for	AAGTTCAGGCAGCACCAGAT
TERF2RT_rev	TTTAGAAATGGCAGCCCAGT

*[1] Primers to amplify the Alu elements indicated in the text (Table 1) are named according to the mRNA in which the Alu element is embedded.

*[2] to create the 3'less version of TERF2AluY together with TERF2YUAS_for.

*[3] 7SL indicates a probe for the S-domain of the 7SL RNA; "Y" and "Sx" refers to TERF2AluY and EIF2S3AluSx, respectively; "int" and "tra" refers to the Alu coding region and the 3'trailer, respectively.

REFERENCES

1 **White**, R.J. RNA polymerases I and III, growth control and cancer. *Nat. Rev. Mol. Cell Biol.* (2005) 6:69–78

2 **Moss**, T., Langlois, F., Gagnon-Kugler, T. and Stefanovsky, V. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis, *Cell. Mol. Life Sci.* (2007) 64:29–49.

3 **Woychik,** N.A., Hampsey, M. The RNA polymerase II machinery: Structure illuminates function. *Cell (2002) 108:453–463*

4 **Lee**, T.I. and Young, R.A. Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* (2000) 34:77–137

5 **Panning**, B. and Taatjes, D.J. Transcriptional Regulation: It Takes a Village. *Mol.Cell* (2008) 31:622-629

6 **Goodrich,** J.A. and Kugel, J.F. Opinion: Non-coding-RNA regulators of RNA polymerase II transcription. *Nature Reviews Molecular Cell Biology* (2006) 7:612-616

7 **Herr**, A.J. *et al.*, RNA polymerase IV directs silencing of endogenous DNA. *Science* (2005) 308:118–120.

8 **Schramm** L., Hernandez N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev* (2002) 16:2593–2620.

9 **Chédin**, S., Ferri, M.L., Peyroche, G., Andrau, J.C., Jourdain S. and Lefebvre, O. The yeast RNA polymerase III transcription machinery: a paradigm for eukaryotic gene activation, *Cold Spring Harb. Symp. Quant. Biol* (1998)63:381–389.

10 **Flores**, A., J.-F. Briand, O. Gadal, J.-C. Andrau, L. Rubbi, V. Van Mullem, C. Boschiero, M. Goussot, C. Marck, C. Carles, P. Thuriaux, A. Sentenac and M. Werner. A protein-protein interaction map of yeast RNA polymerase III. *Proc. Natl Acad. Sci. USA* (1999) 96:7815–7820.

11 **Hu**, P., Wu, S., Sun, Y., Yuan, C.-C., Kobayashi, R., Myers, M.P., and Hernandez, N. The characterization of human RNA polymerase III identifies orthologues for yeast RNA polymerase III subunits. *Mol Cell* Biol (2002).

12 **Chedin** S., Riva M., Schultz P., Sentenac A., Carles C. The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. *Genes* Dev (1998) 12:3857–3871.

13 **Jasiak**, A.J., Armache, K.J., Martens, B., Jansen, R.P. and Cramer, P. Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. *Mol Cell* (2006) 23:71–81.

14 **Fernandez-Tornero**, C., Bottcher, B., Riva, M., Carles, C. and Steuerwald, U. Insights into transcription initiation and termination from the electron microscopy structure of yeast RNA polymerase III. *Mol Cell* (2007) 25:813–23.

15 **Cramer**, P. Recent structural studies of RNA polymerases II and III. *Biochemical Society Transactions* (2006) 34:1058–1061.

16 **Wang**, Z and, Roeder, R.G. Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. *Genes Dev* (1997) 11:1315–26.

17 **Landrieux**, E., Alic, N., Ducrot, C., Acker, J., Riva, M., and Carles, C. A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. *EMBO J* (2006) 25: 118–128.

18 **Ferri**, M.L., Peyroche, G., Siaut, M., Lefebvre, O., Carles, C., Conesa, C. and Sentenac, A. A novel subunit of yeast RNA polymerase III interacts with the TFIIB-related domain of TFIIIB70. *Mol. Cell Biol*. (2000) 20:488−495.

19 **Soll**, D. and Raj Bhandary ,T.L. tRNA: Structure, Biosynthesis, and Function. *American Society for Microbiology Press* (1995).

20 **Crick**, F.H.C. Codon – anticodon pairing: the wobble hypothesis. *J. Mol. Biol.* (1966) 19:548–555.

21 **Cusack,** S. Aminoacyl-tRNA synthetases. *Curr Opin Struct Biol.* (1997) 7:881-889.

22 **Ibba**, M. and Soll, D. Aminoacyl-tRNA synthesis. *Annu Rev Biochem* (2000) 69:617–650.

23 **Eriani**, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* (1990) 347:203–206.

24 **Kutay**, U. *et al.* Identification of a tRNA-specific nuclear export receptor. *Mol Cell* (1998) 1:359–369.

25 **Arts**, G. J., Fornerod, M. and Mattaj, I. W. Identification of a nuclear export receptor for tRNA. *Curr Biol* (1998) 8:305–314.

26 **Simos**, G. and Hurt, E. Transfer RNA biogenesis: a visa to leave the nucleus. *Curr Biol (1999)* 9:R238–R241.

27 **Lipowsky**, G. *et al.* Coordination of tRNA nuclearexport with processing of tRNA. *RNA Publ RNA Soc* (1999) 5:539–549.

28 **Arts**, G. J., Kuersten, S., Romby, P., Ehresmann, B. and Mattaj, I. W. The role of exportin-t in selective nuclear export of mature tRNAs. *EMBO J.* (1998) 17:7430–7441.

29 **Yoshihisa**, T., Yunoki-Esaki, K., Ohshima, C., Tanaka, N. and Endo, T. Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. *Mol. Biol. Cell* (2003) 14:3266–3279.

30 **Takano**, A., Endo, T. and Yoshihisa, T. tRNA actively shuttles between the nucleus and cytosol in yeast. *Science* (2005) 309:140–142.

31 **Cate**, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N. and Noller, H.F. X-ray Crystal Structures of 70*S* Ribosome Functional Complexes. *Science* (1999) Volume 285, 5436:2095–2104.

32 **Spahn,** C.M., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G. and Frank, J. Structure of the 80S ribosome from Saccharomyces cerevisiae--tRNA-ribosome and subunit-subunit interactions. *Cell* (2001) 107(3):373-86.

33 **Passmore**, L.A., Schmeing, T., Maag, D., Applefield, D., Acker, M., Algire, M., Lorsch, J. and Ramakrisnan, V. The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell* (2007) 26:41–50**.**

34 **Siridechadilok**, B., Fraser, C.S., Hall, R.J., Doudna, J.A. and Nogales, E. Structural roles for human translation factor eIF3 in initiation of protein synthesis. *Science* (2005) 310:1513–1515.

35 **Rheinberger**, H.J., Sternbach, H., and Nierhaus, K.H. Three tRNA binding sites on Escherichia coli ribosomes. *PNAS September* (1981) volume 78, 9:5310-5314.

36. **Myasnikov**, A.G., Simonetti, A., Marzi, S. and Klaholz, B.P. Structure-function insights into prokaryotic and eukaryotic translation initiation. *Curr Opin Struct Biol* (2009) 19(3):300-9.

37 **Sprinzl**, M. and Vassilenko, K.S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* (2005) 33:139-40.

38 **Frank**, D.N. and Pace, N.R. RIBONUCLEASE P: Unity and Diversity in a tRNA Processing Ribozyme. *Annu Rev Biochem* (1998) 67:153–180.

39 **Wolin,** S.L. and Matera, A.G. The trials and travels of tRNA. *Genes Dev.* (1999) 13:1-10.

40 **Srisawat**, C., Houser-Scott, F., Bertrand, E., Xiao, S., Singer, R.H. and Engelke, D.R. An active precursor in assembly of yeast nuclear ribonuclease P. *RNA* (2002) 8:1348-1360.

41 **Jarrous**, N. Human ribonuclease P: Subunits, function, and intranuclear localization. *RNA* (2002) 8: 1-7.

42 **Lee**, Y., Kindelberger, D.W., Lee, J.Y., McClennen, S., Chamberlain, J. and Engelke, D.R. Nuclear pre-tRNA terminal structure and RNase P recognition. *RNA* (1997) 3:175-185.

43 **Bertrand**, E., Houser-Scott, F., Kendall, A., Singer, R.H. and Engelke, D.R. Nucleolar localization of early tRNA processing. *Genes Dev* (1998) 12:2463–2468.

44 **Jacobson**, M.R., Cao, L.G., Taneja, K., Singer, R.H., Wang, Y.L. and Pederson, T. Nuclear domains of the RNA subunit of RNase P. *J Cell Sci* (1997) 110:829–837.

45 **Reiner**, R., Ben-asouli, Y., Krilovetzky, I. and Jarrous, N. A role for the catalytic ribonucleoprotein RNase P in RNA polymerase III transcription. *Genes Dev* (2006) 20:1621–1635.

46 **Li**, Z. and Deutscher, M.P. Maturation pathways for *E. coli* tRNA precursors: A random multienzyme process in vivo. *Cell* (1996) 86:503–512.

47 **Li**, Z. and Deutscher, M.P. RNase E plays an essential role in the maturation of *Escherichia coli* tRNA precursors. *RNA* (2002) 8:97–109.

48 **Redko**, Y., Li de la Sierra-Gallay, I. and Condon, C. When all's zed and done: the structure and function of RNaseZ in prokaryotes. *Nat Rev Microbiol* (2007) 5:278 – 286.

49 **Van Horn**, D. J., Yoo, C. J., Xue, D., Shi, H. and Wolin, S. L. The La protein in *Schizosaccharomyces pombe*: a conserved yet dispensable phosphoprotein that functions in tRNA maturation. *RNA* (1997) 3:1434-1443.

50 **Spath**, B., Canino, G. and Marchfelder, A. tRNase Z: the end is not in sight. *Cell Mol Life Sci* (2007) 64, 2404-2412.

51 **Schiffer**, S., Rçsch, S. and Marchfelder, A. Assigning a function to a conserved group of proteins: the tRNA 3'-processing enzymes. *EMBO J* (2002) 21:2769 – 2777.

52 **Zhao**, Z., Su, W., Yuan, S. and Huang, Y. Functional conservation of tRNase ZL among Saccharomyces cerevisiae, Schizosaccharomyces pombe and humans. *Biochemical Journal* (2009) 422:483-492.

53 **Deutscher**, M.P. tRNA nucleotidyltransferase. *The Enzymes* (1982) 15:183−215.

54 **Xiong**, Y. and Steitz, T. A. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* (2004) 430:640−645.

55 **Tomita**, K., Fukai, S., Ishitani, R., Ueda, T., Takeuchi, N., *et al.* Structure of a class II CCA-adding enzyme complexed with primer tRNA and an incoming ATP analog. Nature (2004) 430:700−704.

56 **Yoshihisa**, T. *et al.*, Cytoplasmic splicing of tRNA in *Saccharomyces cerevisiae*, *Genes Cells* (2007), 12:285–297.

57 **Paushkin**, S.V. *et al.* Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3′ end formation. *Cell* (2004) 117:311–321.

58 **Agris**, P.F. Wobble position modified nucleosides evolved to select transfer RNA codon recognition: a modified-wobble hypothesis. *Biochimie* (1991) 73:1345–1349.

59 **Rodnina**, M.V. and Wintermeyer, W. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem* (2001) 70**:**415–435.

60 **Cochella**, L. and Green, R. Fidelity in protein synthesis. *Curr Biol* (2005) 15:536–540.

61 **Burge**, C.B., Tuschl, T.H. and Sharp, P.A. Splicing of precursors to mRNAs by the spliceosomes. In RNA World II. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,USA.* (1999) 525–560.

62 **Nilsen**, T.W. RNA–RNA interactions in nuclear pre-mRNA splicing. In RNA Structure and Function. *Cold Spring HarborLaboratory Press, Cold Spring Harbor, NY, USA* (1998) 279–307.

63 **Reddy**, R., Henning, D., Das,G., Harless,M. and Wright, D. The capped U6 small nuclear RNA is transcribed by RNA polymerase III. *J Biol Chem* (1987) 262:75–81.

64 **Kunkel**, G.R., Maser, R.L., Clavet, J.P. and Pederson, T. U6 small nuclear RNA is transcribed by RNA polymerase III. *Proc Natl Acad Sci USA* (1986) 83:8575–8579.

65 **Fairley**, J.A., Kantidakis, T., Kenneth, N.S., Intine, R.V., Maraia, R.J., and White, R.J. Human La is found at RNA polymerase III-transcribed genes *in vivo*. *PNAS* (2005) volume 102, 51:18350-18355.

66 **Moenne**, A., Camier, S., Anderson, G., Margottin, F., Beggs, J. and Sentenac, A. The U6 gene of Saccharomyces cerevisiae is transcribed by RNA polymerase C (III) in vivo and in vitro. *EMBO J* 9(1): 271–277.

67 **Singh**, R. and Reddy, R. γ-Monomethyl-phosphate: a cap structure in spliceosomal U6 snRNA. Proc *Natl Acad Sci USA* (1989) 86:8280–8283.

68 **Booth**, B.L. and Pugh, B.F. Identification and characterization of a nuclease specific for the 3' end of the U6 small nuclear RNA. *J Biol Chem* (1997) 272:984–991.

69 **Trippe**, R., Sandrock, B. and Benecke, B.J. A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. *Nucleic Acids Res* (1998) 26:3119–3126.

70 **Bartkiewicz**, M., Gold, H. and Altman, S. Identification and characterization of an RNA molecule that copurifies with RNase P activity from HeLa cells. *Genes Dev* (1989) 3:488–99.

71 **Kikovska**, E., Svärd, S.G. and Kirsebom, L.A. Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc Natl Acad Sci U S A* (2007) 104(7): 2062–2067.

72 **Dahmus**, M.E. Phosphorylation of the C-terminal domain of RNA polymerase II. *Biochim Biophys Acta* (1995) 1261:171-182.

73 **Price**, D.H. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol Cell Biol* (2000) 20(8):2629-34.

74 **Michels**, A.A., Fraldi, A., Li, Q., Adamson, T.E., Bonnet, F., Nguyen, V.T., Sedore, S.C., Price, J.P., Price, D.H., Lania, L., and Bensaude O. Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J* (2004) 23:2608-2619.

75 **Nguyen**, V.T., Kiss, T., Michels, A.A., and Bensaude O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* (2001) 414:322-325.

76 **Murphy**, S., Di Liegro, C., and Melli, M. The in vitro transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. *Cell* (1987) 51:81-87.

77 **Kruger**, W., and Benecke, B. J. Structural and functional analysis of a human 7 S K RNA gene. *J Mol Biol* (1987) 195:31-41.

78 **Egloff**, S., Van Herreweghe, E. and Kiss, T. Regulation of polymerase II transcription by 7SK snRNA: two Distinct RNA elements direct P-TEFb and HEXIM1 binding. *Molecular and Cellular Biology* (2006) volume 26, 2:630-642.

79 **Smirnov**, A.V., Entelis, N.S., I.A., Krasheninnikov, R.M. and Tarassov, I.A. Specific Features of 5S rRNA Structure – Its Interactions with Macromolecules and Possible Functions. *Biochemistry (Moscow*) (2008) Volume 73, 13:1418-1437.

80 **Lin**, E., Lin, S.W. and Lin, A. The participation of 5 S rRNA in the cotranslational formation of eukaryotic 5 S ribonucleoprotein complex. *Nucleic Acids Res* (2001) 29:2510-2516.

81 **Pittman**, R.H., Andrews, M.T. and Setzer, D.R. A feedback loop coupling 5 S rRNA synthesis to accumulation of a ribosomal protein. *J Biol Chem* (1999) *274*:33198-33201.

82 **Ban**, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science* (2000) 289:905-920.

83 **Rosa**, M.D., Gottlieb, E., Lerner, M.R. and Steitz, J.A. Striking similarities are exhibited by two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated ribonucleic acids VAI and VAII. *Mol Cell Biol* (1981) 1:785–796.

84 **Lerner**, M.R., Andrews, N.C., Miller, G. and Steitz, J.A. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus *Proc Natl Acad Sci USA* (1981) 78:805–809.

85 **Chiu**, Y.H., MacMillan, J.B. and Chen, Z.J. RNA Polymerase III Detects Cytosolic DNA and Induces Type I Interferons through the RIG-I Pathway. *Cell* Volume 138 3: 576-591.

86 **Kedersha**, N.L. and Rome, L.H. Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA. *J Cell Biol* (1986) 103: 699– 709.

87 **Kedersha**, N.L., Miquel, M.C., Bittner, D. and Rome, L.H. Vaults. II. Ribonucleoprotein structures are highly conserved among higher and lower eukaryotes. *J Cell Biol* (1990) 110: 895–901.

88 **Kickhoefer**, V.A., Vasu, S.K. and Rome L.H. Vaults are the answer, what is the question? *Trends Cell Biol* (1996) 6:174–178.

89 **Izquierdo**, M.A., Scheffer, G.L., Flens, M.J., Giaccone, G., Broxterman, H.J., Meijer, C.J. *et al*. Broad distribution of the multidrug resistance-related vault lung resistance protein in normal human tissues and tumors. *Am J Pathol* (1996) 148:877–887.

90 **van Zon**, A., Mossink, M.H., Schoester, M., Scheffer, G.L., Scheper, R.J., Sonneveld, P. *et al*. Multiple human vault RNAs. Expression and association with the vault complex. *J Biol Chem* (2001) 276:37715–37721.

91 **Kickhoefer**, V.A., Searles, R.P., Kedersha, N.L., Garber, M.E., Johnson, D.L. and Rome, L.H. Vault ribonucleoprotein particles from rat and bullfrog contain a related small RNA that is transcribed by RNA polymerase III. *J Biol Chem* (1993) 268: 7868–7873.

92 **Pfeffer**, S., Zavolan, M., Grässer, F.A., Chien, M., Russo, J.J., Ju, J., John, B., Enright, A.J., Marks, D., Sander, C. and Tuschl, T. Identification of Virus-Encoded MicroRNAs. *Science* (2004) volume 304, 5671:734-736.

93 **Pfeffer**, S., Sewer, A., Lagos-Quintana, M., Sheridan, R., Sander, C., Grässer, F.A., van Dyk, L.F., Ho C.K., Shuman, S., Chien, M., Russo, J.J., Ju, J., Randall, G., Lindenbach, B.D., Rice, C.M., Simon, V., Ho, D.D., Zavolan, M. and Tuschl, T. Identification of microRNAs of the herpesvirus family. *Nature Meth* (2005) 2:269-276.

94 **Bowden**, R.J., Simas, J.P., Davis, A.J. and Efstathiou, S. 1997. Murine gammaherpesvirus 68 encodes tRNA-like sequences which are expressed during latency. *J Gen Virol* (1997)78:1675.

95 **Borchert**, G.M., Lanier, W., Davidson, B.L. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* (2006) volume 13, 12:1097-1101.

96 **International Human Genome Sequencing Consortium**. Initial sequencing and analysis of the human genome. *Nature* (2001) 409:860-921.

97 **Bennett**, P. Demystified ... microsatellites. *Mol Pathol* (2000) volume 53, 4:177-183.

98 **Singer**, M.F. Sines and lines: highly repeated short and long interspersed sequences in mammalian genomes. *Cell* (1982) 28:433–444.

99 **Singer**, M.F. Highly repeated sequences in mammalian genomes. *Int Rev Cytol* (1982) 76:67–112.

100 **Kramerov**, D.A., Vassetzky, N.S. Short retroposons in eukaryotic genomes. *Int Rev Cytol* (2005) 247:165-221.

101 **Umylny**, B., Presting, G., Efirid, J.T., Klimovitsky, B.I. and Ward, S. Most human Alu and murine B1 repeats are unique. *Journal of Cellular Biochemistry* (2007) 102:110-121.

102 **Quentin**,Y. Origin of the Alu family: a family of Alu-likemonomers gave birth to the left and the right arms of the Alu elements. *Nucleic Acids Res* (1992) 20:3397–3401.

103 **Mighell**, A.J., Markham, A.F., Robinson, P.A. Alu sequences. *FEBS Lett* (1997) volume 417, 1:1-5.

104 **Jurka**, J. and Klonowski, P. Integration of retroposable elements in mammals: selection of target sites. *J Mol Evol* (1996) 43:685-689.

105 **Shen**, M.R., Brosius, J. and Deininger, P.L. *BC1* RNA, the transcript from a master gene for ID element amplification, is able to prime its own reverse transcription. *Nucleic Acids Res* (1997) 25:1641–1648.

106 **Perez-Stable**, C., Ayres, T.M. and Shen, C.M. Distinctive sequence organization and functional programming of an *Alu* repeat promoter. *Proc Natl Acad Sci USA* (1984) 81:5291–5295.

107 **Fuhrman**, S.A., Deininger, P.L., LaPorte, P., Friedmann, T. and Geiduschek, E.P. Analysis of transcription of the human *Alu* family ubiquitous repeating element by eukaryotic RNA polymerase III. *Nucleic Acids Res* (1981) 9:6439–6456.

108 **Dewannieux**, M., Esnault, C. and Heidmann, T. LINE mediated retrotransposition of marked Alu sequences. *Nat Genet* (2003) 35:41–48.

109 **Deininger**, P.L. and Batzer, M.A. Alu repeats and human disease. *Mol Genet Metab* (1999) 67:183-193.

110 **Jurka**, J. Evolutionary impact of human Alu repetitive elements. *Curr Opin Genet Dev* (2004) 14:603-608.

111 **Tomilin**, N.V. Regulation of mammalian gene expression by retroelements and non-coding tandem repeats. *Bioessays* (2008) volume 30, 4:338-48.

112 **Cox**, G.S., Gutkin, D.W., Haas, M.J. and Cosgrove, D.E. Isolation of an *Alu* repetitive DNA binding protein and effect of CpG methylation on binding to its recognition sequence. *Biochimica et Biophysica Acta-Gene Structure and Expression* (1998) volume 1396, 1:67–87.

113 **Sorek**, R., Ast, G. and Graur, D. Alu-containing exons are alternatively spliced. *Genome Res* (2002) 12:1060–1067.

114 **Chen**, L.L., Decerbo, J.N. and Carmichael, G.G. Alu element-mediated gene silencing. *EMBO J* (2008) volume 27, 12:1694-705

115 **Liu**, W-M., Maraia, R.J., Rubin, C.M. and Schmid, C.W. *Nucleic Acids Res* (1994) 22:1087–1095.

116 **Umylny**, B., Presting, G. and Ward, S. Evidence of Alu and B1 expression in dbEST. *Arch Androl* (2007) volume 53, 4:207-218.

117 **Liu**, W.M., Chu, W.M., Choudary, P.V. and Schmid, C.W. Cell stress and translational inhibitors transiently increase the abundance of mammalian SINE transcripts. *Nucleic Acids Res* (1995) 23:1758–1765.

118 **Hasler**, J. and Strub,K. AluRNP and AluRNA regulate translation initiation in vitro. *Nucleic Acids Res* (2006) 34:2374–2385.

119 **Mariner**, P.D., Walters, R.D., Espinoza, C.A., Drullinger, L.F., Wagner, S.D., Kugel, J.F. and Goodrich, J.A. Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Molecular Cell* (2008) volume 29, 4:499-509.

120 **Keenan**, R.J., Freymann, D.M., Stroud, R.M. and Walter, P. The signal recognition particle. *Annu Rev Biochem* (2001) 70:755–775

121 **Siegel**, V. and Walter, P. Removal of the Alu structural domain from signal recognition particle leaves its protein translocation activity intact. *Nature* (1986) volume 320, 6057:81-84.

122 **Sauer-Eriksson**, A.E. and Hainzl, T. S-domain assembly of the signal recognition particle. *Curr Opin Struct Biol* (2003) volume 13, 1:64-70.

123 **Bovia**, F., Wolff, N., Ryser, S. and Strub, K. The SRP9/14 subunit of the human signal recognition particle binds to a variety of Alu-like RNAs and with higher affinity than its mouse homolog. *Nucleic Acids Res* (1997) 25:318–326.

124 **Chang**, D.Y., Hsu, K. and Maraia, R.J. Monomeric scAlu and nascent dimeric Alu RNAs induced by adenovirus are assembled into SRP9/14-containing RNPs in HeLa cells. *Nucleic Acids Res* (1996) 24:4165–4170.

125 **Bovia**, F., Fornallaz, M., Leffers, H. and Strub, K. The SRP9/14 subunit of the signal recognition particle (SRP) is present in more than 20-fold excess over SRP in primate cells and exists primarily free but also in complex with small cytoplasmic Alu RNAs. *Mol Biol Cell* (1995) 6:471–484.

126 **Engelke**, D.R., Ng S.Y., Shastry, B.S. *et al*. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell* (1980) 19:717–728.

127 **Miller**, J., McLachlan, A.D. and Klug, A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J* (1985) 4:1609–1614.

128 **Moorefield**, B. and Roeder, R.G. Purification and characterization of human transcription factor IIIA. *J Biol Chem* (1994) 269:20857–20865.

129 **Hanas**, J.S., Hazuda, D.J., Bogenhagen, D.F. *et al*. *Xenopus* transcription factor A requires zinc for binding to the 5S RNA gene. *J Biol Chem* (1983) 258:14120–14125.

130 **Lim**, V.I. Analysis of action of the wobble adenine on codon reading within the ribosome. J Mol Biol (1995) 252:277–282.

131 **Lim**, V.I. Analysis of interactions between the codon-anticodon duplexes with the ribosome: their role in translation. *J Mol Biol* (1997) 266:877–890.

132 **Björk**, G.R., Durand, J.M., Hagervall, T.G., Leipuviene, R., Lundgren, H.K., Nilsson, K. *et al*. Transfer RNA modification: influence on translational frameshifting and metabolism. *FEBS Letters* (1999) 452:47–51.

133 **Agris**, P.F., The importance of being modified: roles of modified nucleosides and Mg^{2+} in RNA structure and function, *Prog Nucl Acids Res Mol Biol* (1996) 53:79–129.

134 **Lobo**, S.M., Tanaka, M., Sullivan, M.L. and Hernandez, N. A TBP complex essential for transcription from TATA-less but not TATA-containing RNA polymerase III promoters is part of the TFIIIB fraction. *Cell* (1992) 71:1029-1040.

135 **Taggart**, A.K., Fisher, T.S. and Pugh, B.F. The TATA-binding protein and associated factors are components of pol III transcription factor TFIIIB. *Cell* (1992) 71:1015-1028.

136 **Teichmann**, M. and Seifart, K.H. Physical separation of two different forms of human TFIIIB active in the transcription of the U6 or the VAI gene in vitro. *Embo J* (1995) 14:5974-5983.

137 **Wang**, Z. and Roeder,R.G. Structure and function of a human transcription factor TFIIIB subunit that is evolutionarily conserved and contains both TFIIB- and highmobility-group protein 2-related domains. *Proc Natl Acad Sci U S A* (1995) 92:7026-7030.

138 **Henry**, R.W., Ford, E., Mital, R., Mittal, V. and Hernandez, N. Crossing the line between RNA polymerases: transcription of human snRNA genes by RNA polymerases II and III. *Cold Spring Harb Symp Quant Biol* (1998) 63:111-120.

139 **Willis**, I.M. A universal nomenclature for subunits of the RNA polymerase III transcription initiation factor TFIIIB. *Genes Dev* (2002) 16:1337-1338.

140 **Teichmann**, M., Wang, Z. and Roeder, R.G. A stable complex of a novel transcription factor IIB related factor, human TFIIIB50, and associated proteins mediate selective transcription by RNA polymerase III of genes with upstream promoter elements. *Proc Natl Acad Sci U S A* (2000) 97:14200-14205.

141 **Schramm**, L., Pendergrast, P.S., Sun, Y. and Hernandez, N. Different human TFIIIB activities direct RNA polymerase III transcription from TATA-containing and TATA-less promoters. *Genes Dev* (2000) 14:2650-2663.

142 **Zhu**, W., Zeng, Q., Colangelo, C.M., Lewis, M., Summers, M.F., and Scott, R.A. The N-terminal domain of TFIIB from *Pyrococcus furiosus* forms a zinc ribbon. *Nat Struct Biol* (1996) 3:122–124.

143 **Mital**, R., Kobayashi, R. and Hernandez, N. RNA polymerase III transcription from the human U6 and adenovirus type 2 VAI promoters has different requirements for human BRF, a subunit of human TFIIIB. *Mol Cell Biol* (1996) 16:7031-7042.

144 **Kassavetis**, G.A., Joazeiro, C.A., Pisano, M., Geiduschek, E.P., Colbert, T., Hahn, S. and Blanco, J.A. The role of the TATA-binding protein in the assembly and function of the multisubunit yeast RNA polymerase III transcription factor, TFIIIB. *Cell* (1992) 71:1055-1064.

145 **Aasland**, R., Stewart, A.F., Gibson, T. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIIB. *Trends Biochem Sci* (1996) volume 21, 3:87-88.

146 **Yoshinaga**, S.K., Boulanger, P.A. and Berk, A.J. Resolution of human transcription factor TFIIIC into two functional components. *Proc Natl Acad Sci USA* (1987) 84:3585-3589.

147 **Dean**,N. and Berk,A.J. Separation of TFIIIC into two functional components by sequence specific DNA affinity chromatography. *Nucleic Acids Res* (1987) 15:9895-9907.

148 **Oettel**, S., Hartel, F., Kober, I., Iben, S., and Seifart, K.H. Human transcription factors $IIIC_2$, $IIIC_1$ and a novel component $IIIC_0$ fulfil different aspects of DNA binding to various pol III genes. *Nucleic Acids Res* (1997) 25:2440-2447.

149 **Schultz**, M.C., Reeder, R.H. and Hahn, S. Variants of the TATA-binding protein can distinguish subsets of RNA polymerase I, II, and III promoters. *Cell* (1992) 69: 697–702.

150 **Wang**, Z., and Roeder, R.G. TFIIIC1 acts through a downstream region to stabilize TFIIIC2 binding to RNA polymerase III promoters. *Mol Cell Biol* (1996) 16*:*6841-6850.

151 **Weser**, S., Gruber, C., Haffner, H.M., Teichmann, M., Roeder, R.G., Seifart, K.H. and Meiβner W. The 250 kDa form of HsBdp1 (TFNR) is an essential component of human TFIIIC1 activity. *J Biol Chem* (2004) 279:27022-27029.

152 **Wang**, Z., and Roeder, R.G. DNA topoisomerase I and PC4 can interact with human TFIIIC to promote both accurate termination and transcription reinitiation by RNA polymerase III. *Mol Cell* (1998) 1:749-757.

153 **Kovelman**, R. and Roeder, R.G. Purification and characterization of two forms of human transcription factor IIIC. *J Biol Chem* (1992) 267:24446–24456.

154 **Dumay-Odelot**, H., Marck, C., Durrieu-Gaillard, S., Lefebvre, O., Jourdain, S., Prochazkova, M., Pflieger, A. and Teichmann, M. Identification, molecular cloning and characterization of the sixth subunit of human transcription factor TFIIIC. *J Biol Chem* (2007) 282:17179-17189.

155 **Lagna**, G., Kovelman, R., Sukegawa, J. and Roeder, R.G. Cloning and characterization of an evolutionary divergent DNA-binding subunit of mammalian TFIIIC. *Mol Cell Biol* (1994) 14: 3053–3064.

156 **Hernandez**, N. snRNA genes: A model system to study fundamental mechanisms of transcription. *J Biol Chem* (2001) 276: 26733–26736.

157 **Rincon**, J.C., Engler, S.K., Hargrove, B.W. and Kunkel, G.R. Molecular cloning of a cDNA encoding human SPH-binding factor, a conserved protein that binds to the enhancer-like region of the U6 small nuclear RNA gene promoter. *Nucl Acids Res* (1998) 26:4846-4852.

158 **Myslinski**, E., Krol, A. and Carbon, P. ZNF76 and ZNF143 are two human homologs of the transcriptional activator Staf. *J Biol Chem* (1998) 273:21998–22006.

159 **Herr**, W. and Cleary, M.A. The POU domain: versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. *Genes Dev* (1995) volume 9, 14:1679-1693.

160 **Henry**, R.W., Mittal, V., Ma, B., Kobayashi, R., and Hernandez, N. SNAP19 mediates the assembly of a functional core promoter complex (SNAPc) shared by RNA polymerases II and III. *Genes Dev* (1998) 12: 2664–2672.

161 **Maraia**, R.J. and Intine, R.V. Recognition of nascent RNA by the human La antigen: Conserved and diverged features of structure and function. *Mol Cell Biol* (2001) 21:367–379.

162 **Teplova**, M., Yuan, Y.R., Phan, A.T., Malinina, L., Ilin, S., Teplov, A. and Patel, D.J. Structural basis for recognition and sequestration of UUU-OH 3'-termini of nascent RNA pol III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* (2006) 21:75–85.

163 **Mathews**, M.B. and Francoeur, A.M. La antigen recognizes and binds to the 39-oligouridylate tail of a small RNA. *Mol Cell Biol* (1984) 4:1134–1140.

164 **Stefano**, J.E. Purified lupus antigen La recognizes an oligouridylate stretch common to the 39 termini of RNA polymerase III transcripts. *Cell* (1984) 36**:**145–154.

165 **Rinke**, J. and Steitz, J.A. Precursor molecules of both human 5S ribosomal RNA and transfer RNAs are bound by a cellular protein reactive with anti-La lupus antibodies. *Cell* (1982) 29:149–159.

166 **Wolin**, S.L. and Cedervall, T. The La protein. *Annu Rev Biochem* (2002) 71:375–403.

167 **Fan**, H., Goodier, J.L., Chamberlain, J.R., Engelke, D.R. and Maraia, R.J. 5’ processing of tRNA precursors can be modulated by the human La antigen phosphoprotein. *Mol Cell Biol* (1998) 18:3201–3211.

168 **Bayfield**, M.A. and Maraia R.J. Precursor-product discrimination by La protein during tRNA *Nat Struct Mol Biol*. (2009) volume 16, 4:430–437.

169 **Gottlieb**, E. and Steitz, J.A. Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III. *EMBO J* (1989) 8:851–861.

170 **Gottlieb**, E. and Steitz, J.A. The RNA binding protein La influences both the accuracy and the efficiency of RNA polymerase III transcription *in vitro*. *EMBO J* (1989) 8:841–850.

171 **Goodier**, J.L. and Maraia, R.J. Terminator-specific recycling of a B1-Alu transcription complex by RNA polymerase III is mediated by the RNA terminus-binding protein La. *J Biol Chem* (1998) 273:26110–26116.

172 **Maraia**, R.J. Transcription termination factor La is also an initiation factor for RNA polymerase III. *Proc Natl Acad Sci USA* (1996) 93:3383–3387.

173 **Weser**, S., Bachmann, M., Seifart, K.H. and Meißner, W. Transcription efficiency of human polymerase III genes in vitro does not depend on the RNP-forming autoantigen La. *Nucleic Acids Res* (2000) 28:3935–3942.

174 **Fairley**, J.A., Kantidakis, T., Kenneth, N.S., Intine, R.V., Maraia, R.J. and White, R.J. Human La is Found at RNA Polymerase III-Transcribed Genes In Vivo. *Proc Nat Acad Sci USA* (2005) 102:18350–18355.

175 **Wang**, Z., Luo, T. and Roeder, R.G. Identification of an autonomously initiating RNA polymerase III holoenzyme containing a novel factor that is selectively inactivated during protein synthesis inhibition. *Genes Dev* (1997) 11:2371–2382.

176 **Wang**, Z., Bai, L., Hsieh, Y.J. and Roeder, R.G. Nuclear factor 1 (NF1) affects accurate termination and multiple-round transcription by human RNA polymerase III. *EMBO J* (2000) 19:6823–6832.

177 **Intine**, R.V., Sakulich, A.L., Koduru, S.B., Huang, Y., Pierstorff, E., Goodier, J.L., Phan, L. and Maraia, R.J. Transfer RNA maturation is controlled by phosphorylation of the human La antigen on serine 366. *Mol Cell* (2000) 6:339–348.

178 **Theunissen**, O., Rudt, F. and Pieler, T. Structural determinants in 5 S RNA and TFIIIA for 7 S RNP formation. *Eur J Biochem* (1998) 258:758-767.

179 **Szymanski**, M., Barciszewska, M.Z., Erdmann, V.A. and Barciszewski, J. 5 S rRNA: structure and interactions. *Biochem J* (2003) 371:641-651.

180 **Bogenhagen**, D.F. The intragenic control region of the *Xenopus* 5 S RNA gene contains two factor A binding domains that must be aligned properly for efficient transcription initiation. *J Biol Chem* (1985) 260:6466–6471.

181 **Pieler**, T., Hamm, J. and Roeder, R.G. The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. *Cell* (1987) 48:91–100.

182 **Challice**, J.M. and Segall, J. Transcription of the 5 S rRNA gene of *Saccharomyces cerevisiae* requires a promoter element at +1 and a 14-base pair internal control region. *J Biol Chem* (1989) 264:20060–20067.

183 **Nolte**, R.T., Conlin, R.M., Harrison, S.C. and Brown, R.S. Differing role for zinc fingers in DNA recognition : structure of a six-finger transcription factor IIIA complex. *Proc Natl Acad Sci USA* (1998) 95:2938–2943.

184 **Engelke**, D.R., Ng, S.Y., Shastry, B.S. and Roeder, R.G. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell* (1980) 19:717–728.

185 **Keller**, H.J., Romaniuk, P.J. and Gottesfeld, J.M. Interaction of Xenopus TFIIIC with the TFIIIA.5 S RNA gene complex. *J Biol Chem* (1992) 267:18190–18198.

186 **Fradkin**, L.G., Yoshinaga, S.K., Berk, A.J. and Dasgupta, A. Human transcription factor TFIIIC2 specifically interacts with a unique sequence in the Xenopus laevis 5S rRNA gene. *Mol Cell Biol* (1989) 9:4941–4950.

187 **Kassavetis**, G.A., Braun, B.R., Nguyen, L.H. and Geiduschek, E.P. S.cerevisiae TFIIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIIA and TFIIIC are assembly factors. *Cell* (1990) 60:235-245.

188 **Weser**, S., Riemann, J., Seifart, K.H. and Meissner, W. Assembly and isolation of intermediate steps of transcription complexes formed on the human 5S rRNA gene. *Nucleic Acids Research* (2003) Volume 31, 9:2408-2416.

189 **Lassar**, A.B., Martin, P.L. and Roeder, R.G. Transcription of class III genes: formation of preinitiation complexes. *Science* (1983) 222:740-748.

190 **Bieker**, J.J., Martin, P.L. and Roeder, R.G. Formation of a rate-limiting intermediate in 5S RNA gene transcription. *Cell* (1985) 40:119-127.

191 **Galli**, G., Hofstetter, H. and Birnstiel, M.L. Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. *Nature* (1981) 294:626–631.

192 **Hofstetter**, H., Kressman, A. and Birnstiel, M.L. A split promoter for a eucaryotic tRNA gene. *Cell* (1981) 24:573–585.

193 **Sharp**, S., DeFranco, D., Dingermann, T., Farrell, P. and Soll, D. Internal control regions for transcription of eukaryotic tRNA genes. *Proc Natl Acad Sci* (1981) 78:6657–6661.

194 **Yoshinaga**, S.K., L'Etoile, N.D. and Berk, A.J. Purification and characterization of transcription factor IIIC2. *J Biol Chem* (1989) 264:10726–10731.

195 **Hsieh**, Y.J., Kundu, T.K., Wang, Z., Kovelman, R. and Roeder, R.G. The TFIIIC90 subunit of TFIIIC interacts with multiple components of the RNA polymerase III machinery and contains a histone-specific acetyltransferase activity. *Mol Cell Biol* (1999) 19:7697–7704.

196 **Hsieh**, Y.J., Wang, Z., Kovelman, R. and Roeder, R.G. Cloning and characterization of two evolutionarily conserved subunits (TFIIIC102 and TFIIIC63) of human TFIIIC and their involvement in functional interactions with TFIIIB and RNA polymerase III. *Mol Cell Biol* (1999) 19:4944–4952.

197 **Arnold**, G.J. and Gross, H.J. Unrelated leader sequences can efficiently promote human tRNA gene transcription. *Gene* (1987) 51:237-246.

198 **Arnold**, G.J., Schmutzler, C. and Gross, H.J. Functional dissection of 5' and 3' extragenic control regions of human $tRNA^{Val}$ genes reveals two different regulatory effects. *DNA* (1988) 7:87-97.

199 **Capone**, J.P. Modulation of the phenotypic expression of a human serine tRNA gene by 5'-flanking sequences. *DNA* (1988) 7:459-468.

200 **Gonos**, E.S. and Goddard, J.P. The role of the 5'-flanking sequence of a human $tRNA^{Glu}$ gene in modulation of its transcriptional activity in vitro. *Biochem J* (1990) 272:797-803.

201 **Palida**, F., Hale, C. and Sprague, K.U. Transcription of a silkworm tRNAAlaC gene is directed by two AT-rich upstream sequence elements. *Nucleic Acids Res* (1993) 21:5875-5881.

202 **Kober**, I., Teichmann, M. and Seifart, K.H. hTFIIIBß stably binds to pol II promoters and recruits RNA polymerase III in a hTFIIIC1 dependent way. *J Mol Biol* (1998) 284:7-20.

203 **Pavesi**, A., Conterio, F., Bolchi, A., Dieci, G. and Ottonello, S. Identification of new eukaryotic tRNA genes in genomic DNA databases by a multistep weight matrix analysis of transcriptional control regions. *Nucl Acids Res* (1994) 22:1247–1256.

204 **Tapping**, R.I., Syroid, D.E. and Capone, J.P. Upstream interactions of functional mammalian tRNA gene transcription complexes probed using a heterologous DNA-binding protein. *J Biol Chem* (1994) 269:21812–21819.

205 **Kunkel**, G.R. and Pederson, T. Upstream elements required for efficient transcription of a human U6 RNA gene resemble those of U1 and U2genes even though a different polymerase is used. *Genes Dev* (1988) 2:196-204.

206 **Goomer**, R.S. and Kunkel, G.R. The transcriptional start site for a human U6 small nuclear RNA gene is dictated by a compound promoter element consisting of the PSE and the TATA box. *Nucleic Acids Res* (1992) volume 20, 18**:**4903–4912.

207 **Kunkel**, G.R. and Pederson, T. Transcription of a human U6 small nuclear RNA gene in vivo withstands deletion of intragenic sequences but not of an upstream TATATA box. *Nucleic Acids Res* (1989) 17:7371–7379.

209 **Danzeiser**, D.A., Urso, O. and Kunkel, G.R. Functional characterization of elements in a human U6 small nuclear RNA gene distal control region. *Mol Cell Biol* (1993) 13:4670–4678.

210 **Myslinski**, E., Krol, A. and Carbon P. Optimal tRNA((Ser)Sec) gene activity requires an upstream SPH motif. *Nucleic Acids Res* (1992) 20:203–209.

211 **Kunkel**, G.R., Cheung, T.C., Miyake, J.H., Urso, O., McNamara-Schroeder, K.J. and Stumph, W.E. Identification of a SPH element in the distal region of a human U6 small nuclear RNA gene promoter and characterization of the SPH binding factor in HeLa cell extracts. *Gene Expr* (1996) 6:59–72.

212 **Kunkel**, G.R. and Hixson, J.D. The distal elements, OCT and SPH, stimulate the formation of preinitiation complexes on a human U6 snRNA gene promoter in vitro. *Nucleic Acids Res* (1998) 26:1536–1543.

213 **Mittal**, V., Ma, B. and Hernandez, N. SNAP(c): a core promoter factor with a built-in DNA-binding damper that is deactivated by the Oct-1 POU domain. *Genes Dev* (1999) 13:1807-1821.

214 **Stunkel**, W., Kober, I. and Seifart, K.H. A nucleosome positioned in the distal promoter region activates transcription of the human U6 gene. *Mol Cell Biol* (1997) 17:4397–4405.

215 **Mittal**, V. and Hernandez, N. Role for the amino-terminal region of human TBP in U6 snRNA transcription. *Science* (1997) 275:1136–1140.

216 **Tichelaar**, J.W., Wieben, E.D., Reddy, R., Vrabel, A. and Camacho, P. In vivo expression of a variant human U6 RNA from a unique, internal promoter. *Biochemistry* (1998) 37:12943-12951.

217 **Yoon**, J.B., Murphy, S., Bai, L., Wang, Z. and Roeder, R.G. 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* (1995) volume 15, 4:2019–2027.

218 **Paule**, M.R. and White, R.J. Transcription by RNA polymerases I and III. *Nucleic Acids Res* (2000) 28:1283–1298.

219 **Kassavetis**, G.A., Blanco, J.A., Johnson, T.E. and Geiduschek, E.P. Formation of open and elongating transcription complexes by RNA polymerase III. *J Mol Biol* (1992) 226:47-58.

220 **Kenneth**, N.S., Marshall, L. and White, R.J. Recruitment of RNA polymerase III in vivo. *Nucleic Acids Res* (2008) volume 36, 11:3757-64.

221 **Kassavetis**, G.A., Prakash, P. and Shim, E. The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. *J Biol Chem* (2009) 285:2695-2706.

222 **Carbon**, P. and Krol, A. Transcription of the *Xenopus laevis* selenocysteine tRNASer(Sec) gene: a system that combines an internal B box and upstream elements also found in U6 snRNA genes. *EMBO J* (1991) volume 10, 3:599–606.

223 **Howe**, J.G. and Shu, M.D. Epstein–Barr virus small RNA (EBER) genes: unique transcription units that combine RNA polymerase II and III promoter elements. *Cell* (1989) volume 57, 5:825–834.

224 **Kickhoefer**, V.A., Emre, N., Stephen, A.G., Poderycki, M.J. and Rome, L.H. Identification of conserved vault RNA expression elements and a non-expressed mouse vault RNA gene. *Gene* (2003) volume 309, 2:65–70.

225 **Ullu**, E. and Weiner, A.M. Human genes and pseudogenes for the 7SL RNA component of signal recognition particle. *EMBO J* (1984) volume 3, 13:3303–3310.

226 **Englert,** M., Felis, M., Junker, V. and Beier, H. Novel upstream and intragenic control elements for the RNA polymerase III-dependent transcription of human 7SL RNA genes. *Biochimie* (2004) volume 86, 12:867–874.

227 **Müller**, J. and Benecke, B.J. Analysis of transcription factors binding to the human 7SL RNA gene promoter. *Biochem Cell Biol* (1999) volume 77, 5:431–438.

228 **Bredow**, S., Sürig, D., Müller, J., Kleinert, H. and Benecke, B.J. Activating-transcription factor (ATF) regulates human 7S L RNA transcription by RNApolymerase III in vivo and in vitro. *Nucleic Acids Res* (1990) volume 18, 23:6779–6784.

229 **Lee**, K.A., Hai, T.Y., SivaRaman, L., Thimmappaya, B., Hurst, H.C., Jones, N.C. and Green, M.R. A cellular protein, activating transcription factor, activates transcription of multiple E1A-inducible adenovirus early promoters. *Proc Natl Acad Sci USA* (1987) volume 84, 23:8355–8359.

230 **Hai**, T. and Hartman, M.G. The molecular biology and nomenclature of the activating transcription factor/cAMP responsive element binding family of transcription factors: activating transcription factor proteins and homeostasis. *Gene* (2001) volume 273, 1:1–11.

231 **Bogenhagen**, D.F. and Brown, D.D. Nucleotide sequences in Xenopus 5S DNA required for transcription termination. *Cell* (1981) volume 24, 1:261–270.

232 **Akusjärvi**, G., Mathews, M.B., Andersson, P., Vennström, B. and Pettersson, U. Structure of genes for virus-associated RNAI and RNAII of adenovirus type 2. *Proc Natl Acad Sci USA* (1980) volume 77, 5:2424–2428.

233 **Weinmann**, R., Brendler, T.G., Raskas, H.J. and Roeder, R.G. Low molecular weight viral RNAs transcribed by RNA polymerase III during adenovirus 2 infection. *Cell* (1976) volume 7, 4:557-566.

234 **Baer**, M., Nilsen, T.W., Costigan, C. and Altman, S. Structure and transcription of a

human gene for H1 RNA, the RNA component of human RNase P. *Nucleic Acids Res* (1990) volume 18, 1:97-103.

235 **Murphy**, S., Di Liegro, C. and Melli, M. The In Vitro Transcription of the 7SK RNA Gene by RNA Polymerase Ill Is Dependent Only on the Presence of an Upstream Promoter. *Cell* (1987) volume 51, 1:81-87.

236 **Chang**, Y.N., Pirtle, I.L. and Pirtle, R.M. Nucleotide sequence and transcription of a human tRNA gene cluster with four genes. *Gene* (1986) volume 48, 1:165-174.

237 **Vnencak-Jones**, C.L., Wahab, S.Z., Zehner, Z.E. and Holmes, W.M. A human tRNA(iMet) gene produces multiple transcripts. *Mol Cell Biol* (1987) volume 7, 11:4134-4138.

238 **Hess**, J., Perez-Stable, C., Wu, G.J., Weir, B., Tinoco, I. Jr and Shen, C.K. End-to-end transcription of an Alu family repeat. A new type of polymerase-III-dependent terminator and its evolutionary implication. *J Mol Biol* (1985) volume 184, 1:7–21.

239 **Thomann**, H.U., Schmutzler, C., Hüdepohl, U., Blow, M. and Gross, H.J. Genes, variant genes and pseudogenes of the human tRNA(Val) gene family. Expression and pre-tRNA maturation in vitro. *J Mol Biol* (1989) volume 209, 4:505-523.

240 **Emerson**, B.M. and Roeder, R.G. DNA sequences and transcription factor interactions of active and inactive forms of mammalian 5 S RNA genes. *J Biol Chem* (1984) volume 259, 12:7926–7935.

241 **Braglia**, P., Percudani, R. and Dieci, G. Sequence Context Effects on Oligo(dT) Termination Signal Recognition by *Saccharomyces cerevisiae* RNA Polymerase III. *JBC* (2005) 280:19551–19562.

242 **Gunnery**, S., Ma, Y. and Mathews, M.B. Termination sequence requirements vary among genes transcribed by RNA polymerase III. *JMB* (1999) 286:745-757.

243 **Cozzarelli**, N.R., Gerrard, S.P., Schlissel, M., Brown, D.D. and Bogenhagen, D.F. Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. *Cell* (1983) 34:829–835.

244 **Campbell**, F.E. Jr, Setzer, D.R. Transcription termination by RNA polymerase III: Uncoupling of polymerase release from termination signal recognition. *Mol Cell Biol* (1992) 12:2260–2272.

245 **Dieci**, G. and Sentenac, A. Facilitated Recycling Pathway for RNA Polymerase III. *Cell* (1996) volume 84, 2:245–252

246 **Ferrari**, R., Rivetti, C., Acker, J. and Dieci, G. Distinct roles of transcription factors TFIIIB and TFIIIC in RNA polymerase III transcription reinitiation. *Proc Natl Acad Sci U S A* (2004) volume 101, 37:13442–13447.

247 **Ferrari**, R. and Dieci, G. The transcription reinitiation properties of RNA polymerase III in the absence of transcription factors. *Cell Mol Biol Lett* (2008) volume 13, 1:*112*–118.

248 **Cabart**, P., Lee, J. and Willis, I.M. Facilitated recycling protects human RNA polymerase III from repression by Maf1 in vitro. *JBC* (2008) volume 283, 52:36108-36117.

249 **Johnson**, S.S., Zhang, C., Fromm, J., Willis, I.M. and Johnson, D.L. Mammalian Maf1 is a negative regulator of transcription by all three nuclear RNA polymerases. *Mol Cell* (2007) volume 26, 3:367–379.

250 **Reina**, J.H., Azzouz, T.N. and Hernandez, N. Maf1, a new player in the regulation of human RNA polymerase III transcription. *PLoS ONE* (2006) volume 1, 1:e134.

251 **Proudfoot**, N.J., Furger, A. and Dye, M.J. Integrating mRNA processing with transcription. *Cell* (2002) volume 108, 4:501-512.

252 **Koski**, R.A., Clarkson, S.G., Kurjan, J., Hall, B.D. and Smith, M. Mutations of the yeast SUP4 tRNATyr locus: transcription of the mutant genes in vitro. *Cell* (1980) 22:415-425.

253 **Kuhn**, A., Normann, A., Bartsch, I. and Grummt, I. The mouse ribosomal gene terminator consists of three functionally separable sequence elements. *EMBO J* (1988) 7:1497–1502.

254 **Evers**, R. and Grummt, I. Molecular coevolution of mammalian ribosomal gene terminator sequences and the transcription termination factor TTF-I. *Proc Natl Acad Sci* (1995) 92:5827–5831.

255 **Kuhn**, A. and Grummt, I. 3′-End formation of mouse pre-rRNA involves both transcription termination and a specific processing reaction. *Genes Dev* (1989) 3:224–231.

256 **Jansa**, P., Burek, C., Sander, E.E. and Grummt, I. The transcript release factor PTRF augments ribosomal gene transcription by facilitating reinitiation of RNA polymerase I. *Nucleic Acids Res* (2001) volume 29, 2:423-429.

257 **Allison** D.S. and Hall, B.D. Effects of alterations in the 3′ flanking sequence on in vivo and in vitro expression of the yeast SUP4-o tRNATyr gene. *EMBO J* (1985) 4:2657–2664.

258 **Reyes**, V.M., Newman, A. and Abelson, J. Mutational Analysis of the Coordinate Expression of the Yeast tRNAArg_tRNAAs Gene Tandem. Mol Cell Biol (1986) volume 6, 7:2436-2442.

259 **Percudani**, R., Pavesi, A. and Ottonello, S. Transfer RNA gene redundancy and translational selection in *Saccharomyces cerevisiae*. *J Mol Biol* (1997) 268:322–330.

260 **Hani**, J. and Feldmann, H. tRNA genes and retroelements in the yeast genome *Nucl Acids Res* (1998) 26:689–696.

261 **Kim**, V.N. MicroRNA biogenesis: coordinated cropping and dicing NATURE *Nat Rev Mol Cell Biol* (2005) volume 6, 5:376-85.

262 **Lowe**, T.M. and Eddy, S.R. tRNAscan-SE: a program for improved detection of transferRNA genes in genomic sequence. *Nucleic Acids Res* (1997) 25: 955-964.

263 **Kohany**, O., Gentles, A.J., Hankus, L. and Jurka, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics* (2006) 7:474.

264 **Traut**, T.W. Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem* (1994) volume 140, 1:1–22.

265 **Lee**, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H. and Kim, V.N.. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* (2004) volume 23, 20:4051-4060.

266 **Diebel**, K.W., Smith, A.L. van Dyk, L.F. Mature and functional viral miRNAs transcribed from novel RNA polymerase III promoters. *RNA* (2010) volume 16, 1:170-185.

267 **Bogerd**, H.P., Karnowski, H.W., Cai, X., Shin, J., Pohlers, M. and Cullen, B.R. A mammalian herpesvirus uses non-canonical expression and processing mechanisms to generate viral microRNAs. *Molecular Cell* (ahead of print).

268 **Lau**, N.C., Lim, L.P., Weinstein, E.G. and Bartel, D.P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* (2001) 294:858–862.

269 **Lagos-Quintana**, M., Rauhut, R., Lendeckel, W. and Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science* (2001) 294:853–858.

270 **Ozsolak**, F., Poling, L.L., Wang, Z., Liu, H., Liu, X.S., Roeder, R.G., Zhang, X., Song, J.S. and Fisher, D.E. Chromatin structure analyses identify miRNA promoters. *Genes Dev* (2008) 22:3172–3183.

271 **Borchert**, G.M., Lanier, W. and Davidson, B.L. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* (2006) 13:1097–1101.

272 **Adeniyi-Jones**, S., Romeo, P.H. and Zasloff, M. Generation of long read-through transcripts in vivo and in vitro by deletion of 3' termination and processing sequences in the human tRNAmet gene. *Nucleic Acids Res* (1984) volume 12, 2:1101-1115.

273 **Chu** W.M., Liu, W.M. and Schmid, C.W. RNA polymerase III promoter and terminator elements affect Alu RNA expression. *Nucleic Acids Res* (1995) 23:1750–1757.

274 **Roy**, A.M., West, N.C., Rao, A., Adhikari, P., Alemán, C., Barnes, A.P. and Deininger, P.L. Upstream flanking sequences and transcription of SINEs. *J Mol Biol* (2000) volume 302, 1:17-25.

275 **Li**, T.H., Kim, C., Rubin, C.M. and Schmid, C.W. K562 cells implicate increased chromatin accessibility in Alu transcriptional activation. *Nucleic Acids Res* (2000) volume 28, 16:3031-3039.

276 **Mignone**, F., Gissi, C., Liuni, S. and Pesole, G. Untranslated regions of mRNAs. *Genome Biol* (2002) 3:reviews0004.1–0004.10.

277 **Rana**, T.M. Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* (2007) volume 8, 1:23-36.

278 **Katayama**, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi,

T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engström, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C. RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. Antisense transcription in the mammalian transcriptome. *Science* (2005) volume 309, 5740:1564-1566.

279 **Yang**, N. and Kazazian, H.H. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol* (2006) 9:763-71.

280 **Sinnett**, D., Richer, C., Deragon, J.M. and Labuda, D. Alu RNA secondary structure consists of two independent 7 SL RNA-like folding units. *J Biol Chem* (1991) 266:8675–8678.

281 **Badis**, G., Saveanu, C., Fromont-Racine, M. and Jacquier, A. Targeted mRNA degradation by deadenylation-independent decapping. *Molecular Cell* (2004) volume 15, 1:5-15.

282 **Häsler**, J., Samuelsson, T. and Strub, K. *Alu* RNAs embedded in 5' and 3' UTRs of human mRNAs. Online at *http://www.unige.ch/sciences/biologie/bicel/Strub/researchAlu.html* (2007).

283 **Smit,** A.F.A., Hubley, R. and Green, P. RepeatMasker Open-3.0. *http://www.repeatmasker.org* (2006).

284 **Batzer**, M.A., Deininger, P.L., Hellmann-Blumberg, U., Jurka, J., Labuda, D., Rubin, C.M., Schmid, C.W., Zietkiewicz, E. and Zuckerkandl, E. Standardized Nomenclature for Alu Repeats. *J Mol Evol* (1996) 42:3–6

285 **Chesnokov**, I. and Schmid, C.W. Flanking sequences of an Alu source stimulate transcription in vitro by interacting with sequence-specic transcription factors. *J Mol Evol* (1996) 42:30-36.

286 **Liu**, W.M. and Schmid,C.W. Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. *Nucleic Acids Res* (1993) 21:1351–1359.

287 **Livak**, K.J. and Schmittgen, D.T. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method. *Methods* (2001) 25:402-408.

288 **Mäkinen**, P.I., Koponen, J.K., Kärkkäinen, A.M., Malm, T.M., Pulkkinen, K.H., Koistinaho, J., Turunen, M.P. and Ylä-Herttuala, S. Stable RNA interference: comparison of U6 and H1 promoters in endothelial cells and in mouse brain. *J Gene Med* (2006) volume 8, 4:433-441.

289 **Gasior**, S.L., Palmisano, M. and Deininger, P.L. Alu-linked hairpins efficiently mediate RNA interference with less toxicity than do H1-expressed short hairpin RNAs. *Anal Biochem* (2006) volume 349, 1:41-48.

290 **Yakovchuk**, P., Goodrich, J.A. and Kugel, J.F. B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. *Proc Natl Acad Sci U S A* (2009) volume 106, 14:5569-5574.

291 **Brosius**, J. and Gould, S.J. On 'genomenclature': a comprehensive (and respectful) taxonomy for pseudogenes and other 'junk DNA'. *Proc Natl Acad Sci USA* (1992) 89:10706–10710.

292 **Kolomietz**, E., Meyn, M.S., Pandita, A. and Squire, J.A. The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. *Genes Chromosomes Cancer* (2002) 35:97–112.

293 **Studier**, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* (1990) 185:60-89.

294 **Dignam**, J.D., Lebovitz, R.M. and Roeder, R.G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Research* (1983) volume 11, 5:1475-1489.

295 **Pall**, G.S. and Hamilton, A.J. Improved northern blot method for enhanced detection of small RNA. *Nature protocols* (2008) volume 3, 6:1077-1084.

296 **Bolchi,** A., Ottonello, S. and Petrucco, S. A general one-step method for the cloning of PCR products. *Biotechnology and Applied Biochemistry* (2005) 42:205–209.