Year : 2019

Genomic arrangements and gene régulation

Yeganeh Seghalaksari Meghdad

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Centre Intégratif de Génomique

Genomic arrangements and gene regulation

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

par

Meghdad Yeganeh Seghalaksari

Master de l'Université de Téhéran

Jury

Prof. Sophie Martin, Président
Prof. Nouria Hernandez, Directeur de thèse
Prof. Bart Deplancke, Expert
Prof. Carol Greider, Expert

Lausanne
(2019)

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole Doctorale
Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Madame	Prof. Sophie	**Martin**
Directeur·rice de thèse	Madame	Prof. Nouria	**Hernandez**
Experts·es	Madame	Prof. Carol	**Greider**
	Monsieur	Prof. Bart	**Deplancke**

le Conseil de Faculté autorise l'impression de la thèse de

Monsieur Meghdad Yeganeh Seghalaksari

Continuous Master Of Science In Biotechnology University of Tehran, Iran

intitulée

Genomic arrangements and gene regulation

Lausanne, le 15 mars 2019

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Sophie Martin

Table of contents

Acknowledgements

First of all, I would like to thank my PhD supervisor Prof. Nouria Hernandez. I really enjoyed working in her lab, and I highly appreciate her support, great supervision, motivation, trust, and kindness. Despite all the tasks she had to do, she was amazingly always available when I needed to discuss something. I learnt a lot from her, both scientific and from her personality.

I am grateful to my thesis jury, Prof. Sophie Martin, Prof. Carol Greider and Prof. Bart Deplancke, for their time and insightful comments.

I would like to thank all my current and former colleagues: Gilles, Sasha, Philippe, Gabriel, François, Simone, Viviane, Pascal, Nicolas and Andrea, for the great atmosphere in the lab, all their helps and fruitful discussions. I thank Nathalie for taking care of all administrative stuff perfectly. Also, I am sincerely thankful to all the people who contributed to the two publications related to my thesis.

Last but not least, I would like to express my sincere gratitude to my friends, my family especially my dear parents, and my beloved Farnaz, for their daily support, encouragements, patience, and for the joy and love they bring to my life; this thesis would not have been possible without them.

Abstract

Gene expression regulation is crucial for every biological process of the cell. In eukaryotes, different RNA polymerases (Pol) transcribe different sets of genes, with Pol II responsible mainly for protein-coding gene transcription, and Pol III transcribing non-coding genes mostly involved in Pol II gene expression. Here we studied two aspects of gene expression regulation with specific genomic arrangements. First, we studied the function of a MIR-SINE located within the first intron of a Pol II-transcribed gene, *Polr3e*, which encodes a subunit of Pol III. This MIR is highly occupied by Pol III, and, strikingly, Pol II accumulates at its 3' region, as revealed by chromatin immunoprecipitation (ChIP) assays. CRISPR/Cas9-mediated removal of the MIR from the genome relieved the accumulation of Pol II at the 3' of the MIR and led to increased levels of *Polr3e* mRNA and POLR3E protein. This suggested that antisense transcription of the MIR by Pol III interfered with Pol II transcription of the *Polr3e* gene, slowing down Pol II elongation and thus leading to Pol II accumulation, a phenomenon we also observed for several other genes.

In the second part, we studied the dynamics of Pol III occupancy at Pol III genes during mouse liver regeneration by performing ChIPs followed by ultra high throughput sequencing (ChIP-seq) with liver samples collected at different times after partial hepatectomy (PH). We observed a general increase in Pol III occupancy at Pol III genes after PH, with genes lowly occupied before PH more affected. This increase was accompanied by an increase in the levels of several pre-tRNAs, as well as of mRNAs encoding several Pol III subunits and transcription factors. We investigated how Pol III genes are regulated differentially during liver regeneration by taking into account the surrounding H3K4me3 and Pol II peaks. We found that Pol II peaks surrounding Pol III peaks generally corresponded to Pol II TSSs. We showed that Pol III occupancy of tRNA genes surrounded by H3K4me3 and Pol II peaks, which have high occupancy levels before PH, tends to change less in liver regeneration than Pol III occupancy of tRNA genes with only a Pol III peak, which were more dynamic. This study revealed two main classes of tRNA genes: those close to Pol II genes, whose Pol III occupancy is relatively stable during liver regeneration, and those devoid of surrounding H3K4me3 and Pol II peaks, whose transcription is dynamic and responds to the increased need of cells for Pol III products during proliferation.

Résumé

La régulation de l'expression des gènes est cruciale pour chaque processus biologique de la cellule. Chez les eucaryotes, différentes ARN polymérases (Pol) transcrivent différents ensembles de gènes; la Pol II étant principalement responsable de la transcription des gènes codant les protéines, et la Pol III transcrivant les gènes non codants principalement impliqués dans l'expression des gènes Pol II. Nous avons étudié deux aspects de la régulation de l'expression des gènes: la régulation d'un gène transcript par la Pol II par le biais d'un gène interne transcript par la Pol III, et la régulation de la transcription par la Pol III durant la régération du foie.

Dans un premier temps, nous avons étudié la fonction d'un MIR-SINE situé dans le premier intron d'un gène transcrit par la Pol II, *Polr3e*, qui code pour une sous-unité de la Pol III. Ce MIR est fortement occupé par la Pol III et, de façon étonnante, plusieurs unités de Pol II s'accumulent dans sa région 3', comme le révèlent les analyses d'immunoprécipitation de chromatine (ChIP). L'élimination de ce MIR du génome par CRISPR/Cas9 a diminué l'accumulation de la Pol II dans la région 3' du MIR et a entraîné une augmentation des taux d'ARNm de *Polr3e* et de la protéine POLR3E. Cela suggère que la transcription antisens du MIR par la Pol III interfère avec la transcription par la Pol II du gène *Polr3e*, ce qui ralentit l'élongation de la Pol II et conduit à l'accumulation d'unités de Pol II. Nous avons ensuite montré que ce genre de phénomène peut être observé pour plusieurs autres gènes.

Dans la deuxième partie, nous avons étudié la dynamique de l'occupation par la Pol III au niveau des gènes Pol III pendant la régénération du foie chez la souris, en effectuant des ChIPs suivies d'un séquençage à très haut débit (ChIP-seq) sur des échantillons de foie prélevés à différents moments après une hépatectomie partielle (HP). Nous avons observé une augmentation générale de l'occupation par la Pol III des gènes de Pol III après HP, les gènes peu occupés avant la HP étant plus affectés. Cette augmentation s'est accompagnée d'une augmentation des niveaux de plusieurs pré-ARNt, ainsi que des ARNm codant spécifiquement pour plusieurs sous-unités de la Pol III et pour des facteurs de transcription utilisés par la Pol III. Nous avons étudié comment différents gènes Pol III sont régulés dans la régénération hépatique en tenant compte des pics H3K4me3 et Pol II environnants. Nous avons constaté que les pics Pol II entourant les pics Pol III correspondaient généralement à des TSSs de la Pol II.

Nous avons montré que l'occupation par la Pol III des gènes d'ARNt entourés de pics H3K4me3 et Pol II, qui ont des niveaux d'occupation élevés avant la HP, tend à être plus stable durant la régénération hépatique que celle des gènes d'ARNt avec seulement un pic Pol III. Ces derniers ont des niveaux d'occupation par la Pol III relativement bas avant la HP, niveaux qui augmentent fortement pendant la régénération hépatique. Cette étude a mis en évidence deux grandes classes de gènes ARNt : ceux proches des gènes Pol II, dont l'occupation par la Pol III est relativement stable pendant la régénération hépatique, et ceux dépourvus des pics H3K4me3 et Pol II environnants, dont la transcription est dynamique et répond au besoin accru des cellules en produits de la Pol III pendant leur prolifération.

Résumé pour le grand public

Arrangements génomiques et régulation des gènes

L'information génétique des cellules est organisée dans leur génome. Cette information génétique, constituée de molécules d'ADN, peut être utilisée pour fabriquer des protéines qui remplissent des fonctions distinctes. La synthèse de différentes protéines à partir de différents gènes comprend la synthèse d'une molécule intermédiaire appelée ARN par un processus appelé transcription, processus qui est effectué par l'enzyme ARN polymérase (Pol) II. Cependant, certains ARN ne codent pas pour les protéines, mais ont des rôles régulateurs divers. La Pol III transcrit de nombreux gènes codant pour ces ARN.

La régulation de l'expression des gènes est cruciale, car sa dérégulation est observée dans de nombreuses maladies. Dans cette thèse, nous nous sommes concentrés sur la régulation de la transcription de la Pol II et de la Pol III. Dans la première partie, nous avons observé que les gènes Pol III qui se trouvent à l'intérieur des gènes Pol II peuvent réguler la transcription des gènes Pol II à l'intérieur desquels ils résident; nous avons ainsi montré que la transcription de la Pol III pouvait interférer avec la transcription de la Pol II. Dans la deuxième partie, nous avons étudié la dynamique de la transcription par la Pol III pendant la régénération du foie chez la souris. Dans cette étude, nous avons effectué une hépatectomie partielle du foie de souris qui a conduit les cellules restantes à entrer en régénération, et étudié les liaisons génomiques par la Pol III et la Pol II à différents stades de cette régénération. Nous avons montré que la transcription par la Pol III augmente pendant la régénération. Nous avons également montré que la transcription des gènes Pol III proches des régions liées à Pol II ne changeait pas beaucoup pendant la régénération, alors que la transcription des gènes Pol III éloignés des gènes Pol II était plus dynamique. En conclusion, dans ces études, nous avons montré certains arrangements génomiques qui sont importants pour la régulation de l'expression des gènes.

Chapter I – General Introduction

Protein-coding gene expression in eukaryotes

In eukaryotes, RNA polymerase II (Pol II) is responsible for transcription of all the mRNA-encoding genes, as well as most snRNA and microRNA genes. The mechanism by which Pol II transcribes protein-coding genes is intensely studied. The recruitment of Pol II to gene transcription start sites (TSSs) is regulated by a number of sequences that can be categorized as regulatory regions and core promoters, the latter including the region encompassing the TSS and immediate downstream sequences. The regulatory regions bind transcription factors that can be activators or repressors and the core promoter regions bind so called general transcription factors (GTFs), which, *in vitro*, are sufficient to direct transcription (Orphanides and Reinberg 2002; Smale and Kadonaga 2003). Pol II transcription entails several steps; pre-initiation complex (PIC) formation on the core promoter, transition of Pol II from the TSS to the region immediately downstream, release of Pol II from this TSS proximal region to efficient elongation, and finally, termination of transcription (Hahn 2004).

As Pol II transcription is highly controlled, gene expression is very dynamic in cells, and it is subject to variation in response to many environmental, signaling, and genetic processes. The control of Pol II transcription may occur at any of its transcription steps (Fuda et al. 2009). Pol II recruitment can be limited by inefficient PIC assembly due, for example, to nucleosomes present in the promoter region, which can be overcome by the action of some activators (Bai and Morozov 2010; Lai and Pugh 2017). Pol II transcription can also be limited at the step during which the polymerase clears the TSS and transitions to the promoter proximal downstream region. This step involves the Pol II C-terminal domain (CTD), a structure at the C-terminus of the Pol II largest subunit consisting of a seven amino acid sequence repeated 52 times in mammals (Egloff and Murphy 2008). The CTD plays critical roles in transcriptional and co-transcriptional events; during Pol II clearing from the TSS, the CTD becomes phosphorylated at Ser5 in the CTD repeats, an event

catalyzed by the general transcription factor TFIIH (Hsin and Manley 2012; Heidemann et al. 2013).

A third important step in regulation of Pol II transcription occurs after transition to the promoter-proximal region, where Pol II pauses. This pause is visualized in ChIP assays as an accumulation of Pol II in the 5' end region of some genes. These genes are in a so-called "poised" state, and they maintain an open promoter structure accessible to GTFs and other factors required for transcription (Gilchrist et al. 2010; Adelman and Lis 2012; Jonkers and Lis 2015). Two main factors are responsible for polymerase pausing, NELF and DSIF. Exit from the pause is regulated by P-TEFb, a kinase that phosphorylates NELF, thus leading to its dissociation from the Pol II complex. P-TEFb also phosphorylates DSIF, thus converting it into a positive elongation factor that remains bound to Pol II during elongation, as well as the Pol II CTD on Ser2 (Figure 1-1). ChIP-seq experiments show that Pol II is more phosphorylated at Ser5 of its CTD near the 5' end of genes, and more at Ser2 of its CTD as it travels towards the 3' end. Ser2 phosphorylation is thus a mark of elongating Pol II (Peterlin and Price 2006; Kwak and Lis 2013).

P-TEFb itself is a highly regulated factor and can be controlled at several steps including transcriptional, post-transcriptional, and post-translational steps. However, the best-studied mechanism of P-TEFb regulation is via the 7SK small nuclear ribonucleoprotein particle (snRNP) (Diribarne and Bensaude 2009). The 7SK snRNP consists of the 7SK RNA bound to two proteins that protect its 5' and 3' ends. When the RNA-binding protein HEXIM joins the complex, it changes its conformation so that the complex can now sequester PTEF-b and keep it in an inactive state. Factors like HIV Tat protein and BRD4 can release PTEF-b from the 7SK snRNP complex, freeing it for action on paused Pol II complexes leading to productive elongation (Peterlin et al. 2012a) (Figure 1-1).

After Pol II starts elongation, several factors contribute to productive elongation in the gene body. TFIIS helps backtracked Pol II make a cut in the nascent RNA 3' end region to realign it with the Pol II catalytic site, and this can restore Pol II elongation. Other factors like TFIIF, a general transcription factor involved in transcription

initiation, and Super Elongation Complex (SEC) consisting of PTEF-b, ELL elongation factors, and several other associated factors, also promote efficient elongation (Shilatifard et al. 2003; Sims et al. 2004; Lin et al. 2010). The polymerase-associated factor complex (PAFc) has an indirect role in elongation as it mediates interaction between Pol II and the SEC (He et al. 2011).

Transcriptional and co-transcriptional events like RNA capping, splicing, and mRNA 3' end processing are highly linked, and the Pol II CTD plays a key role as an integrator of these different processes (Fong and Bentley 2001). DSIF and Ser5 phosphorylation of the Pol II CTD in the early elongation complex are important for nascent mRNA 5' capping, as they cause Pol II to pause, which in turn stimulates the capping enzyme to perform its function. Ser2 phosphorylation of the Pol II CTD and hence the responsible kinase, PTEF-b, are involved in the recruitment of splicing factors and the formation of splicing complexes. Splicing factors themselves can facilitate elongation. mRNA 3' end processing, which includes cleavage and polyadenylation, is stimulated by Pol II CTD Ser2 phosphorylation, by PAFc, and also by the ELL2 component of the SEC (Neugebauer 2002; Zhou et al. 2012).

Pol II elongation is highly dynamic with the rate of elongation varying across a gene. This is required for the co-transcriptional events to be performed efficiently. The rate of Pol II elongation is affected by several factors like presence of nucleosomes, CG rich sequences, and histone marks. For example, in exons, higher CG content, nucleosome occupancy, and also some histone marks lead to decreased rate of Pol II elongation, and this can enhance recruitment of splicing factors as well as co-transcriptional splicing events. Moreover, histone marks and formation of R loops (RNA-DNA hybrids) lead to Pol II pausing after the polyadenylation signal, and this pause leads to efficient termination and 3' end processing of the pre-mRNA, including the 3' cut in the mRNA and subsequent polyadenylation at this newly formed 3' end (Jonkers et al. 2014; Jonkers and Lis 2015).

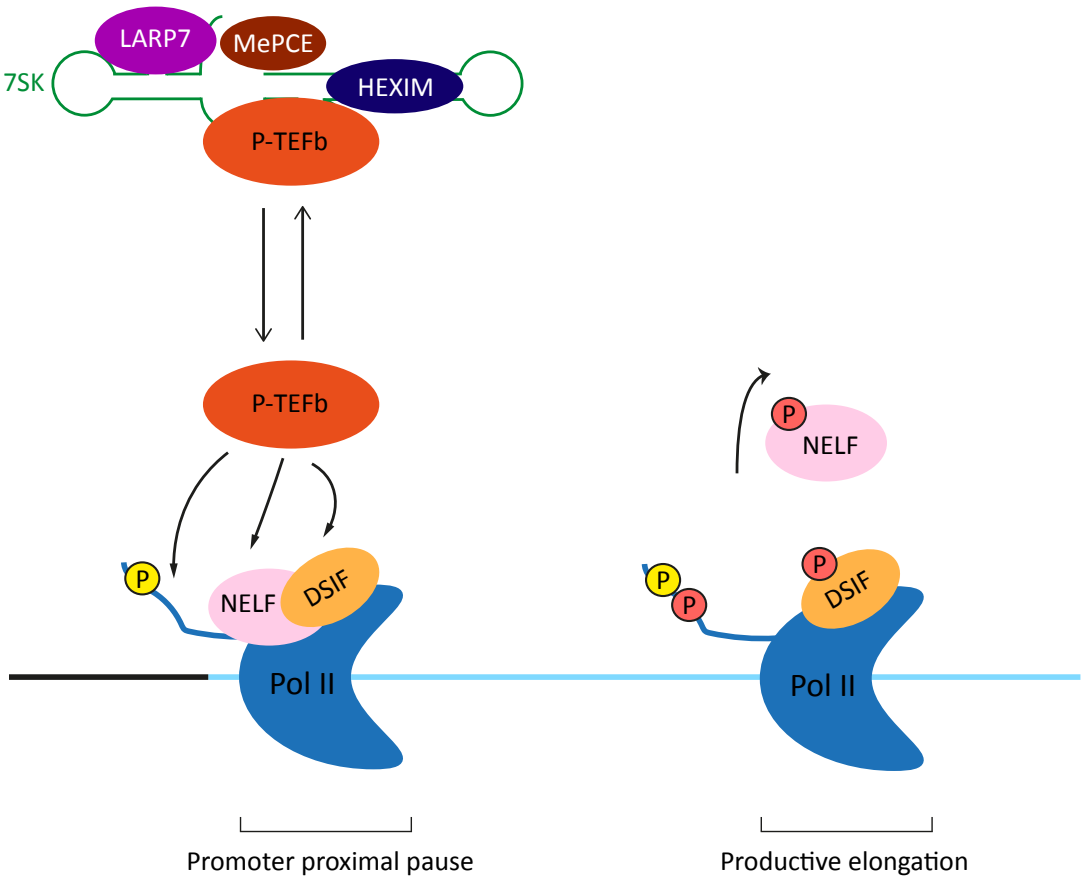


Figure 1-1. Pol II pause and pause release. The promoter proximal pause of Pol II is mediated by DSIF and NELF. Release of PTEF-b from 7SK snRNP complex can lead to productive elongation. Active PTEF-b phosphorylates DSIF, NELF, and Pol II CTD on Ser2.

After the mRNA is synthesized and processed, it can be translated into protein. Unlike bacteria, in which transcription and translation are coupled processes, in eukaryotes, the translation process occurs independently of transcription, in the cytoplasm. As described above, during and right after transcription, the mRNA undergoes several processing events, including m^7G capping at 5' terminus, splicing or alternative splicing, and 3' terminus polyadenylation. The mature mRNA is exported into the cytoplasm, where many factors contribute to its translation, among which ribosomes play a central role. The 80S ribosomes in eukaryotes consist of two subunits: a small, 40S, subunit and a large, 60S, subunit. The 40S subunit consists of 33 proteins and 18S ribosomal RNA (rRNA), whereas the 60S subunit consists of 47 proteins, 28S rRNA, 5.8S rRNA, and 5S rRNA (Melnikov et al. 2012).

Translation occurs in four steps: initiation, elongation, termination, and recycling (Kapp and Lorsch 2004). In brief, 40S ribosomal subunit together with methionyl-tRNAi is recruited to the 5' terminus of capped mRNAs. This complex then scans the 5' UTR of the mRNA until the start codon AUG is recognized. More rarely, an mRNA possesses an internal ribosome entry site (IRES) sequence, which is capable of delivering this complex directly to the start codon, without scanning of the 5' UTR. With the subsequent recruitment of the 60S subunit, the ribosome is assembled, and the next aminoacyl-tRNA is placed in the aminoacyl-tRNA site (A-site) of the ribosome, while the methionyl-tRNAi is located in the peptidyl-tRNA site (P-site). The amino acid on the tRNA in the P-site forms a peptide bond with the one in the A-site, and so the tRNA in P-site becomes uncharged. Then the mRNA is shifted to the next codon, the A-site becomes empty to accept a new aminoacyl-tRNA, and the uncharged tRNA leaves the exit site (E-site) (Hinnebusch and Lorsch 2012). The polypeptide chain is elongated by the formation of peptide bonds between amino acids specified by corresponding codons and the concomitant mRNA shifting codon by codon, and then terminated when a stop codon is placed in A-site of the ribosome. The nascent peptide is then released and the ribosomal subunits are recycled for next rounds of translation (Figure 1-2). In each of these steps, several specific translation factors are involved. eIF2α and eIF4E are two well-known initiation factors, involved in early steps of eukaryotic translation. eIF2α promotes the binding of methionyl-tRNAi to the 40S ribosomal subunit, which then together with some other factors,

forms the pre-initiation complex. eIF4E is involved in recognition of the cap structure of mRNA by the pre-initiation complex. eIF4E activity can be inhibited through interaction with 4E-binding proteins (4E-BPs) (Gebauer and Hentze 2004).

Like transcription, translation is also highly controlled, as it is one of the most energy-consuming processes in the cell. The regulation of translation can occur at any step. However, the main known mechanisms of translational control occur at the initiation step (Sonenberg and Hinnebusch 2009). Translational control can be global, affect a set of mRNAs, or just a single type of mRNA. The regulation of translation initiation can be mediated through effects on the mRNAs directly or through the translational machinery. The main direct effectors on mRNA translation are RNA-binding proteins (RBPs) and microRNAs (miRNAs). RBPs can bind to different regions at the 5' and 3' UTR of mRNAs and play different roles such as translational regulation. This regulation can be mediated through different mechanisms like competing with translation factors binding or interaction with these factors. La-related protein 1 (LARP1) and poly(A)-binding proteins are well-known RPBs involved in translational regulation (Harvey et al. 2018). miRNAs bind the 3' UTR of the mRNA, leading to translation repression (Huntzinger and Izaurralde 2011). As to regulation through the translational machinery, the most-studied mechanism is translational regulation through phosphorylation of the translation initiation factor eIF2α and 4E-BPs, where phosphorylation of the first inhibits, and of the second enhances, translation (Jackson et al. 2010; Hershey et al. 2012).

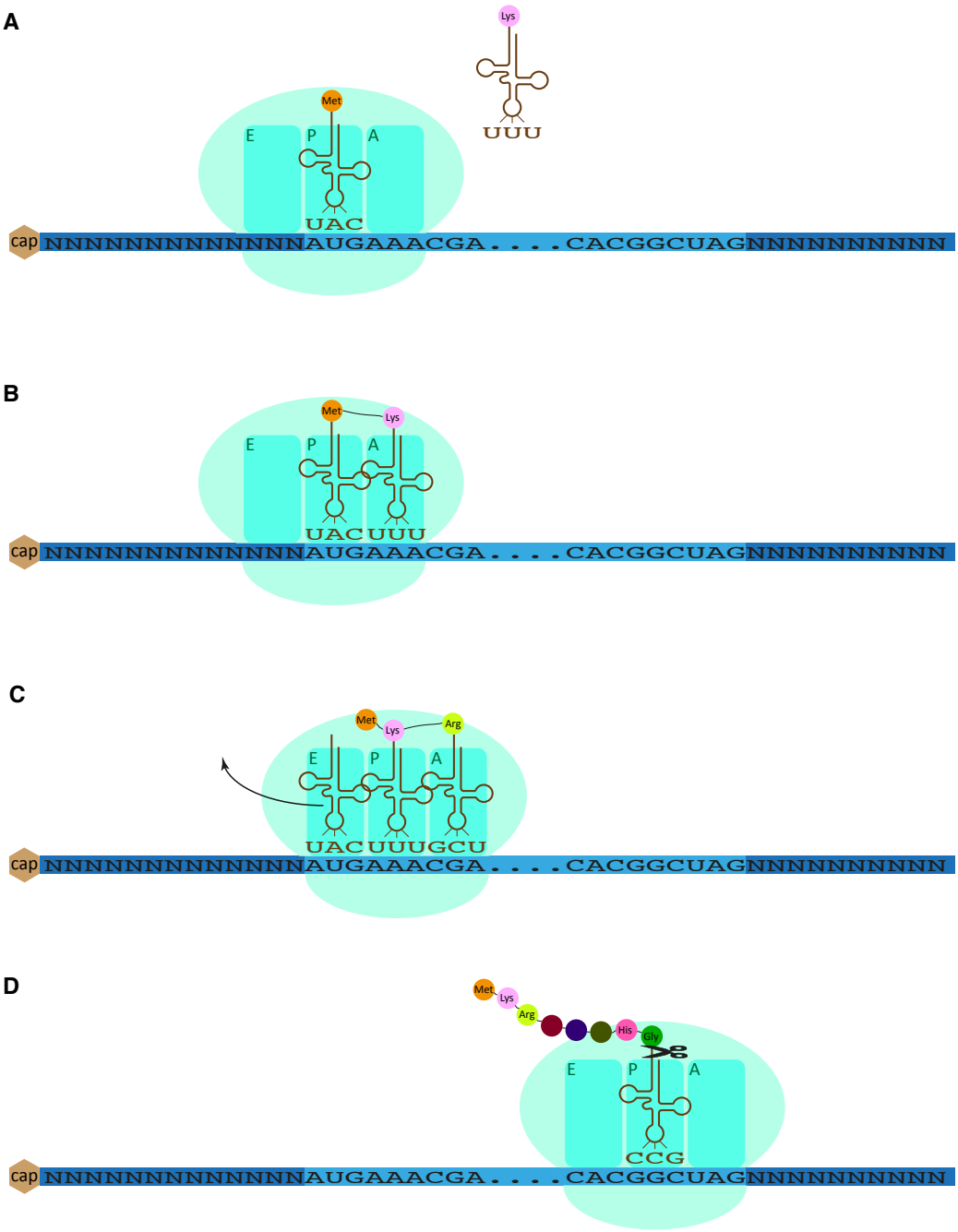


Figure 1-2. Different steps of mRNA translation in eukaryotes. **A)** Small ribosomal subunit together with methionyl-tRNAi is recruited to the cap structure of mRNA, and scans the 5' UTR until the AUG start codon is found. The large ribosomal subunit is recruited, while the methionyl-tRNAi is in P site of the ribosome. **B)** The next aminoacyl-tRNA is placed in the A site of ribosome. A peptide bond is formed between two amino acids. **C)** After shifting of the ribosome to the next codon, the next aminoacyl-tRNA is placed in the A site. Formation of a peptide bond between amino acids in the P and A sites, and release of the uncharged tRNA from the E site, are continued codon by codon. **D)** When a stop codon is localized to the A site, the nascent peptide is released.

RNA polymerase III genes

RNA polymerase III (Pol III) transcribes a wide variety of non-coding RNA genes in eukaryotic cells (Dieci et al. 2013). For example, the tRNAs, which charge different amino acids, and 5S rRNA, which is present in ribosomes, are well-known Pol III RNA products that have essential roles in protein synthesis. tRNA genes are the most abundant Pol III-transcribed genes (Canella et al. 2010; Canella et al. 2012). After transcription, tRNAs go through several steps of maturation, including 5' and 3' end processing, removal of intron (if any), addition of a CCA tail, and chemical modifications at several bases. Finally, the mature tRNAs, folded into a very specific tertiary structure, are able to carry specific amino acids to the translation machinery (Rak et al. 2018). Other genes transcribed by Pol III generate RNAs with a variety of functions. U6 snRNA together with Pol II-transcribed snRNAs play role in splicing of pre-mRNAs (Mroczek and Dziembowski 2013). The 7SL RNA component of the signal recognition particle functions in protein translocation to the endoplasmic reticulum (Walter and Blobel 1982). 7SK RNA negatively regulates Pol II transcription by sequestering PTEF-b in the 7SK snRNP complex (Yang et al. 2001) (see above, Figure 1-1) . RNase P RNA is mainly involved in 5' end processing of pre-tRNAs, while RNase MRP RNA plays role in pre-rRNA processing (Esakova and Krasilnikov 2010; Goldfarb and Cech 2017; Jarrous 2017). Vault RNAs as parts of vault particles are involved in processes like intracellular transport and drug resistance (van Zon et al. 2006; Chen et al. 2018). Y RNAs play role in initiation of chromosomal DNA replication (Krude et al. 2009). Moreover, some short Interspersed Nuclear Elements (SINEs) are still actively transcribed by Pol III and are likely to carry out functions, although these remain to be determined.

SINEs are one of the main types of transposable elements (TEs) in the genome. They constitute about 8 and 15 percent of the mouse and human genomes, respectively (Treangen and Salzberg 2011; Kassiotis 2014). SINEs are retrotransposons, i.e. their transposition occurs through an RNA intermediate. Moreover, they are non-autonomous TEs, as reverse transcription of the RNA intermediate and transposition depend on another type of TEs called Long Interspersed Nuclear Elements (LINEs), which encode in particular the reverse transcriptase required to convert the RNA

intermediate into DNA (Dewannieux et al. 2003; Dewannieux and Heidmann 2005). SINEs are derived from Pol III genes, mostly tRNA and 7SL RNA genes, and also 5S rRNA genes. These genes have gene-internal promoters (see below), which are thus transposed through the RNA intermediate, giving rise to SINEs that may be transcriptionally competent. Indeed, although most SINEs are epigenetically repressed, some are actively transcribed by Pol III (Canella et al. 2010; Canella et al. 2012). The resulting SINE RNAs do not encode proteins and most of them have lengths comprised between 150 to 300 nucleotides. They contain three parts; a 5' part, which resembles the Pol III gene they originate from, a body whose sequence origin is unclear, and, for most of them, an A-rich tail at the 3' end. Some SINEs also have several T residues at the 3' end, which serve as a Pol III transcription termination signal. Depending on both their structure and their origin, SINEs are categorized into different classes. For example B1 SINEs in rodents and Alu elements, which are specific to primates, originate from 7SL RNA, whereas B2 SINEs in rodents originate from tRNAs (Kramerov and Vassetzky 2011; Vassetzky and Kramerov 2013). Mammalian Interspersed Repeats (MIRs) are an ancient family of tRNA-derived SINEs that were amplified mainly before eutherian radiation (Smit and Riggs 1995).

Although SINEs have long been considered junk DNA, they can profoundly impact genome functions. On the one hand, SINEs can mutate or disrupt the sequences in which they are integrated. Although SINE movement in the genome is quite rare, SINEs-mediated disruption in the genome was shown to cause several diseases such as hemophilia and breast cancer (Belancio et al. 2008; Hancks and Kazazian 2012). On the other hand, some SINEs can play gain of function roles such as act as enhancers (Sasaki et al. 2008), introduce a polyadenylation signal inside genes (Lee et al. 2008), or introduce splice sites leading to alternative splicing (Pattanakitsakul et al. 1992). Further, some specific B2 and Alu SINE transcripts are able to bind to Pol II and repress mRNA transcription after heat shock (Espinoza et al. 2004; Mariner et al. 2008).

RNA polymerase III transcription

Pol III consists of seventeen subunits, of which five are shared among Pol III, Pol I, and Pol II, and two between Pol III and Pol I. Five subunits are highly related but not identical among the three polymerases, and correspond to the five subunits of E. coli RNA polymerase. Another five subunits are unique to Pol III (Cramer et al. 2008; Vannini and Cramer 2012). The role of several subunits of Pol III in transcription has been studied. The C53 subunit of Pol III forms with the C37 subunit a heterodimer that resembles the RNA polymerase II transcription factor TFIIF, and plays a role in promoter opening and transcription initiation (Kassavetis et al. 2010; Wu et al. 2011). This heterodimer together with the TFIIS-like subunit C11 also functions in Pol III transcription termination and recycling (Arimbasseri and Maraia 2015; Mishra and Maraia 2018). The Rpc82, Rpc34, and Rpc31 subunits form a TFIIE-related subcomplex, which was shown to play a role in Pol III transcription initiation and elongation (Khoo et al. 2018; Wei and Chen 2018).

In addition to known general roles of Pol III subunits in the transcription process (Hoffmann et al. 2015), roles for individual Pol III subunits in different cells have been defined. As a first example, it was shown that the POLR3G subunit of Pol III plays a role in the maintenance of pluripotency in human embryonic stem cells (hESCs), as its siRNA-mediated knockdown in ES cells led to loss of pluripotency and promoted differentiation. Indeed, *POLR3G* knockdown led to a change in transcript levels of several hundred genes, many of which were shown to be bound by pluripotency factors at promoter regions (Lund et al. 2017). The *POLR3G* gene itself was found to be a downstream target of NANOG and OCT4, two pluripotency transcription factors (Wong et al. 2011). Second, mutations in the *POLR3A* and *POLR3B* genes, which encode the largest and second largest subunits of Pol III, were shown to be associated with leukodystrophy in clinical studies (Saitsu et al. 2011). Specific mutations in the *POLR1C* gene, encoding POLR1C subunit of Pol III and Pol I, was also associated with leukodystrophy, and at the molecular level resulted in impaired Pol III assembly and decreased binding of mutant POLR1C to Pol III genes (Thiffault et al. 2015). In zebrafish, a mutation of the *Polr3b* gene resulting in a protein deficient in interaction with the other Pol III subunit Polr3k caused a

deficiency in digestive development (Yee et al. 2007). Yet another study showed that targeted mutation of *Polr3b* in the mouse intestinal epithelium led to impaired cell proliferation and a major mortality of the mice (Kieckhaefer et al. 2016). Finally, mutations in *POLR3E*, *POLR3A*, *POLR3C*, and *POLR3F* could be identified in patients with varicella zoster virus (VZV) infection. These mutations impaired a cytosolic Pol III function required to activate immunity against VZV (Carter-Timofte et al. 2018).

Like the other eukaryotic RNA polymerases, the Pol III enzyme requires promoter elements and transcription factors to initiate transcription at defined positions in the genome. Pol III promoter structures can be classified into three types (Figure 1-3). Two of them are intragenic, whereas the third is extragenic. The type of promoter defines which transcription factors will bind and recruit Pol III. The intragenic type 1 promoters, which are seen in the 5S rRNA genes, consist of an A box, an intermediate element, and a C box. In this type of promoter, TFIIIA, a zinc finger protein, binds first to the intermediate element. This protein-DNA complex can then recruit TFIIIC, followed by Brf1-TFIIIB. Brf1-TFIIIB consists of three subunits, a SANT domain protein known as Bdp1, the TATA box binding protein TBP, and the TFIIB related factor Brf1. These factors together form a transcription pre-initiation complex, which recruits Pol III to start transcription. The intragenic type 2 promoters, which are observed in tRNA genes, differ from type 1 promoters in that they do not need TFIIIA. Instead, TFIIIC binds directly to the promoter elements, called A and B box, followed by TFIIIB binding and Pol III recruitment. The extragenic type 3 promoters, found for example in the U6 and 7SK genes, consist of a distal sequence element (DSE) to which the transcription factors Oct1 and Staf bind. The DSE serves as an enhancer element that activates transcription from the core promoter. The core promoter consists of a proximal sequence element (PSE) to which a five-subunit complex known as SNAPc binds, and a TATA box to which Brf2-TFIIIB (which is identical to Brf1-TFIIIB except that the Brf1 subunit is replaced by another TFIIB-related factor, Brf2) binds. After assembly of these factors, Pol III is recruited to start transcription (Schramm and Hernandez 2002; Dumay-Odelot et al. 2014).

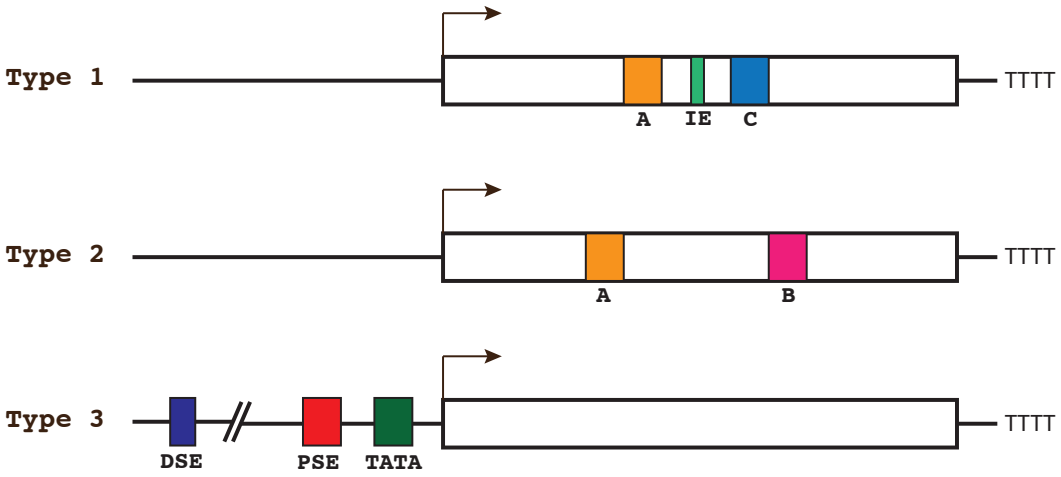


Figure 1-3. Different types of Pol III promoters. Type 1 and type 2 promoters have an intragenic structure, whereas type 3 have an extragenic structure. Pol III transcription is terminated at a short run of T residues.

RNA polymerase III regulation

Like Pol II, Pol III is highly regulated, and its activity level can affect protein translation and cell growth. Pol III regulation has been studied in several model systems and has been documented in processes such as cell differentiation, adaptation to the circadian cycle and response to nutrients, and the cell division cycle. Together, the results point to the involvement of different regulators under different conditions (Figure 1-4).

Several studies have shown that Pol III is regulated in development and cellular differentiation. Regulation of Pol III was observed during development of mouse liver and brain by comparison of different time points of embryonic and adult stages. The most pronounced differences were observed during the embryonic to adult transition (Schmitt et al. 2014). In a study examining Pol III occupancy before and after differentiation of THP-1 monocytes into macrophages, Pol III occupancy at Pol III genes showed a general decrease and highly occupied loci were found to be more affected than that of the lowly occupied ones. In this study, it was shown that active tRNA genes reside in DNA loops, and the regulation of tRNA genes organized in clusters, domains, and interconnected through long-range interactions was found to be somewhat coordinated (Van Bortle et al. 2017). In another study, Pol III was found to bind more Pol III genes in human embryonic stem cells (ES cells) than in differentiated cell types, and Pol III binding regions in ES cells overlapped with regions bound by the pluripotency factors NANOG and OCT4, suggesting that these transcription factors might promote Pol III occupancy in these cells (Alla and Cairns 2014). The detailed mechanisms and factors governing Pol III regulation during the differentiation process remain, however, to be determined.

One known major regulator of Pol III transcription is the protein Maf1. Maf1 was first discovered in yeast, where it is a central player in the repression of Pol III transcription in response to several types of stress including DNA damage, secretory pathway defects, and lack of nutrients (Upadhya et al. 2002). Under stress conditions, yeast Maf1 becomes dephosphorylated, which allows it to bind to, and repress, Pol III. Various yeast kinases have been shown to phosphorylate Maf1 including PKA, Sch9, and TORC1 (Moir et al. 2006; Lee et al. 2009; Wei et al. 2009). Maf1 is

conserved from yeast to human. In mammalian cells, it is directly phosphorylated by mTORC1 (Michels et al. 2010), and like in yeast, it represses Pol III transcription under stress conditions (Reina et al. 2006; Johnson et al. 2007; Bonhoure et al. 2015; Willis and Moir 2018). For example, in IMR90hTert cells, serum starvation led to decreased Pol III occupancy at most of its bound loci, and reduction of endogenous MAF1 levels by RNAi prevented a large part of this effect (Orioli et al. 2016).

Pol III activity is regulated in response to circadian rhythm and nutrient availability. Mice are nocturnal animals and feed during the night. Pol III activity in mouse liver was shown to be higher during the night and lower during daytime (Mange et al. 2017). To decipher whether this effect resulted from circadian regulation or from a response to feeding, Pol III occupancy was measured in mice lacking a central component of the core circadian clock, and thus completely arrhythmic, but fed only at night, and in mice with an intact circadian clock but fed every four hours. The results revealed that Pol III occupancy increases before the night, in other words before the natural feeding time, and that this anticipation effect depends on an intact circadian clock. On the other hand, sustained higher Pol III occupancy during the night and lower occupancy during the day was lost in mice fed every four hours, and the lower occupancy during the day depended on the presence of MAF1. This was consistent with the known inactivation of MAF1 by phosphorylation through the TORC1 pathway, which is itself activated in response to nutrients (Mange et al. 2017).

Several studies have shown Pol III regulation during the cell division cycle. Thus, synthesis of some pre-tRNAs, 5S RNA, and other Pol III products has been found to increase during the transition from a resting G0 state to a proliferating state upon serum stimulation of serum-starved cells (Mauck and Green 1974; Scott et al. 2001). This effect is mediated at least in part by the Ras/Raf/MEK/ERK pathway and involves activated ERK2 phosphorylating BRF1, thereby increasing BRF1 association with TFIIIC and Pol III (Felton-Edkins et al. 2003a).

Pol III is not only modulated upon exit from the G0 state but also during the subsequent phases of the cell division cycle, with high levels of activity in late G1, S,

and G2 phases, followed by repression at mitosis and early G1 (White et al. 1995; Gjidoda and Henry 2013; Herrera et al. 2018). Key players in this cell division cycle regulation include RB and the related p107 and p130 polypeptides as well as the kinases CK2, Plk1, and Cdk1. RB can repress Pol III transcription of genes with type 1 and type 2 promoters through interaction with Brf1-TFIIIB. Phosphorylation of RB mediated by Cyclin D-CDK4 and Cyclin E-CDK2 disrupts this interaction, relieving its repressor activity. In Pol III genes with a type 3 promoter structure, RB could stably associate with SNAPc, TBP, and Pol III on the promoter, thereby inhibiting transcription (Scott et al. 2001; Hirsch et al. 2004; Gjidoda and Henry 2013).

CK2 has a dual role in regulation of Pol III during the cell cycle. Whereas CK2 has a positive effect on Pol III transcription during S phase, as CK2 inhibition at this stage of the cell cycle decreases Pol III transcription, it has a negative effect during mitosis, at which point it phosphorylates Bdp1, leading to its dissociation from Pol III promoters and thus transcription inhibition (Hu et al. 2004). Like CK2, Plk1 can also regulate Pol III transcription positively and negatively. Plk1 regulates transcription of tRNA and 5S rRNA genes by differential phosphorylation of Brf1. Plk1 phosphorylation of Brf1 at serine 450 leads to transcription activation, whereas phosphorylation at threonine 270 during mitosis leads to transcription repression (Fairley et al. 2012). In yeast, yet another kinase, namely Cdk1, was shown to regulate Pol III activity; Cdk1 phosphorylates Bdp1 at tRNA genes during S phase, which boosts their transcription (Herrera et al. 2018).

Pol III deregulation is linked with cancer, with increased Pol III activity in malignant cells (White 2004; White 2005; Johnson et al. 2008; Marshall and White 2008). Such increased activity is manifested for example by genome-wide increased Pol III binding to its targets in a hepatocarcinoma cell line as compared to normal liver cells (Renaud et al. 2014). tRNA overexpression was moreover observed in breast cancer cells (Pavon-Eternod et al. 2009), and upregulation of $tRNA^{Glu}_{UUC}$ and $tRNA^{Arg}_{CCG}$ was shown to be involved in breast cancer metastasis (Goodarzi et al. 2016). Consistent with such observations, the tumor suppressors RB, p53 and PTEN can repress Pol III activity, while the c-MYC oncogene product can activate it. Interaction with TFIIIB was shown to be critical for regulation of Pol III by these factors (Chu et al. 1997;

Cairns and White 1998; Felton-Edkins et al. 2003b; Gomez-Roman et al. 2003; Woiwode et al. 2008).

Along with overexpression of Pol III products, an increase in components of the Pol III transcription machinery has been observed in several transformed cells. Some examples include increased levels of several Pol III transcription factors and subunits in IMR90 fibroblasts following transformation (Durrieu-Gaillard et al. 2018), increased *BRF2* level in Lung Squamous Cell Carcinoma as compared to normal cells (Lockwood et al. 2010), and overexpression of TFIIIC2 in ovarian tumors (Winter et al. 2000).

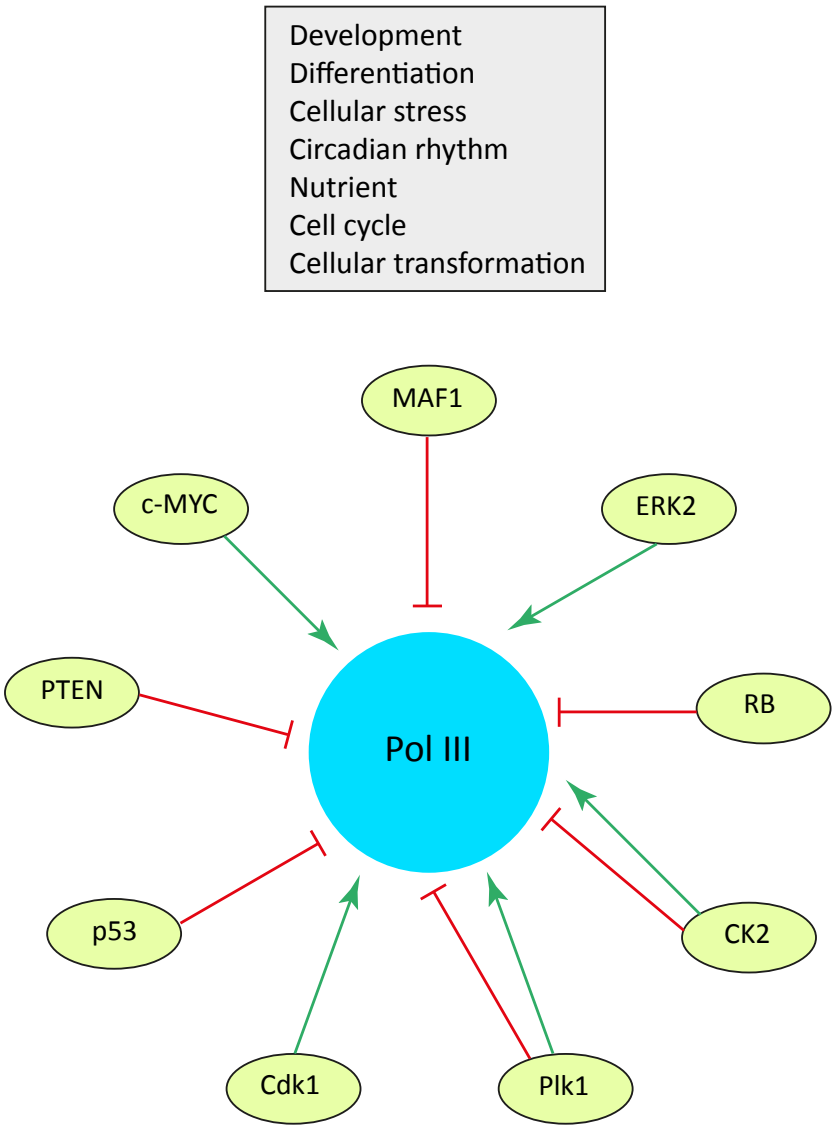


Figure 1-4. Regulators of Pol III transcription. Pol III transcription can be positively or negatively regulated by different regulators in different conditions.

Overlapping genes

Many genes in the genomes of different organisms overlap. This overlap can be partial, or in extreme cases, an entire gene may be located inside another gene, in which case the embedded gene is called a nested gene (Makalowska et al. 2005). In eukaryotes, nested genes located in exons of the host genes are rare (Kumar 2009). Nested genes are mostly located within introns of the host genes, and the introns that include the nested genes are usually larger than the other introns (Yu et al. 2005).

When considering the promoters of overlapping genes, different arrangements are observed: in a first set of two possible arrangements, the promoters themselves are not overlapping and can direct transcription i) in different directions (convergent promoters) or ii) in the same direction (tandem promoters); in a second set of three possible arrangements, the two promoters overlap and can be i) tandem, ii) divergent, or iii) convergent (Figure 1-5). In all these cases, the genes could potentially impact each other's expression through transcriptional interference, which might operate by different mechanisms.

In a "promoter competition" mechanism, binding of RNA polymerase to one promoter inhibits binding of RNA polymerase to the other one. In the "sitting duck" model, an initiation complex sitting on a promoter and slow to fire the polymerase is often dislodged by an elongating RNA polymerase coming from another promoter. In the "occlusion" model, which resembles the "sitting duck" model, the passage of an elongating RNA polymerase through the promoter of another gene transiently inhibits transcription factor and RNA polymerase binding to that promoter. Thus in a study in yeast, transcription of the non-coding RNA *SRG1* from a site upstream of the *SER3* gene was shown to interfere with activators binding to *SER3* gene promoter, leading to its transcription repression (Martens et al. 2004). "Collision" occurs between two elongating RNA polymerases from convergent promoters and can lead to premature termination. For example in yeast, antisense transcription of the non-coding RNA *RME2* interferes with transcription elongation of the overlapping *IME4* gene (Gelfand et al. 2011). Lastly, an immobile DNA-bound factor, including factors tightly bound on a promoter, can act as a "roadblock" for an elongating RNA polymerase (Shearwin et al. 2005).

Transcriptional interference is not the only mechanism by which two overlapping genes can affect each other's expression. This phenomenon can also be mediated at post-transcriptional levels and the mechanisms involved are actively studied in several organisms from bacteria to yeast and higher eukaryotes. Most of these studies focus on antisense long non-coding RNA (lncRNA) genes, which could impact expression of the sense, protein-coding gene at the transcription or post-transcription levels (Katayama et al. 2005; Pelechano and Steinmetz 2013; Huber et al. 2016). Some studies have suggested mechanisms by which transcription or transcript of a gene could affect expression of the other one in an overlapping arrangement. One of the mechanisms intensively studied is regulation through chromatin modifications.

In yeast, transcription of lncRNA *IRT1* recruits histone methyltransferase and deacetylase complexes to the *IME1* gene promoter, which results in a silent chromatin state and inhibition of transcription activators binding (van Werven et al. 2012). In another example, antisense transcription leads to repression of *PHO84* gene through histone deacetylation (Castelnuovo et al. 2013). At the *PHO5* gene promoter, antisense transcription leads to histone eviction and transcription activation (Uhler et al. 2007). And at the *CDC28* gene, osmotic stress activates transcription of an antisense lncRNA from 3' region of the gene, which by a gene looping mechanism leads to chromatin remodeling at *CDC28* gene promoter and its transcription activation (Nadal-Ribelles et al. 2014).

In human cells, *HAS2* antisense RNA was shown to be sufficient for chromatin opening around *HAS2* gene promoter and its increased transcription (Vigetti et al. 2014). The lncRNA *ANRIL* was found to interact with polycomb repressive complex 1 (PRC1) and to repress *INK4A* gene transcription through H3K27me3 maintenance (Yap et al. 2010). siRNA-mediated knockdown of the *BDNF* antisense transcript led to a decrease of the repressive mark H3K27me3 at the *BDNF* locus and elevated *BDNF* transcription in different mouse and human cell lines (Modarresi et al. 2012). In hepatocellular carcinoma cells, *AChE* gene repression due to histone methylation at the promoter mediated by an antisense RNA was observed (Xi et al. 2014). In another study in hepatocellular carcinoma cells, *GPC3* antisense transcript recruited histone acetyltransferase P300/CBP-associated factor to *GPC3* gene body, which in this case

enhanced the histone marks H3K4ac, H3K27ac, and H4ac, leading to *GPC3* transcription upregulation (Zhu et al. 2016).

Another mechanism by antisense genes that can affect sense gene expression is through DNA methylation. Transcription of an antisense RNA leads to mouse *Igf2r* gene promoter methylation and its repression on the paternal allele (Wutz et al. 1997). Similarly, transcription of antisense RNA was shown to be involved in *HBA2* gene silencing through its CpG island methylation (Tufarelli et al. 2003).

A third mechanism involves post-transcriptional regulation of the sense gene by the antisense RNAs, which requires formation of double strand RNAs. This type of regulation can be mediated by increased mRNA stability. Thus, an antisense transcript was shown to increase *BACE1* mRNA stability by masking a miRNA-binding site (Faghihi et al. 2010). Moreover, two mRNAs, *WDR83* and *DHPS,* were shown to from a duplex at their overlapping regions leading to their increased mutual stability (Su et al. 2012). Yet in another example, an antisense RNA increased stability of *PDCD4* mRNA by regulating its association with RNA-binding proteins (Jadaliha et al. 2018).

And fourth, antisense RNA can be involved in translation regulation. *Nos1* antisense RNA was shown to decrease NOS1 protein levels, and not *Nos1* mRNA levels (Korneev et al. 2015). Further, an antisense lncRNA, which does not change *Uchl1* mRNA levels, leads to an increase in UCHL1 protein levels in mouse cells. This antisense transcript contains an embedded SINEB2 domain that is essential for its function (Carrieri et al. 2012).

The aforementioned studies in eukaryotes are all focused on overlapping Pol II transcribed genes. There are, however, also a few known examples of Pol II-Pol I and Pol II-Pol III gene overlaps. In *Saccharomyces cerevisiae,* a mitochondrial protein gene named TAR1 is transcribed antisense of 25S rRNA gene. The nested TAR1 gene is transcribed by Pol II, whereas the 25S rDNA is transcribed by Pol I. However, whether transcription of one of these two transcription units might impact on the other has not been examined (Coelho et al. 2002). In a study in yeast, TFIIIB bound to a

tRNA gene located between two Pol II transcription units could block Pol II transcription, inhibiting read-through transcription of the upstream Pol II gene. Inactivation of TFIIIB binding led to production of an abnormal transcript extending over both Pol II genes (Korde et al. 2014). In another study, performed in *Arabidopsis thaliana*, it was shown that antisense Pol III transcription of a set of nine proline tRNA genes nested in the AtNUDT22 gene negatively correlated with Pol II sense transcription of the AtNUDT22 gene (Lukoszek et al. 2013).

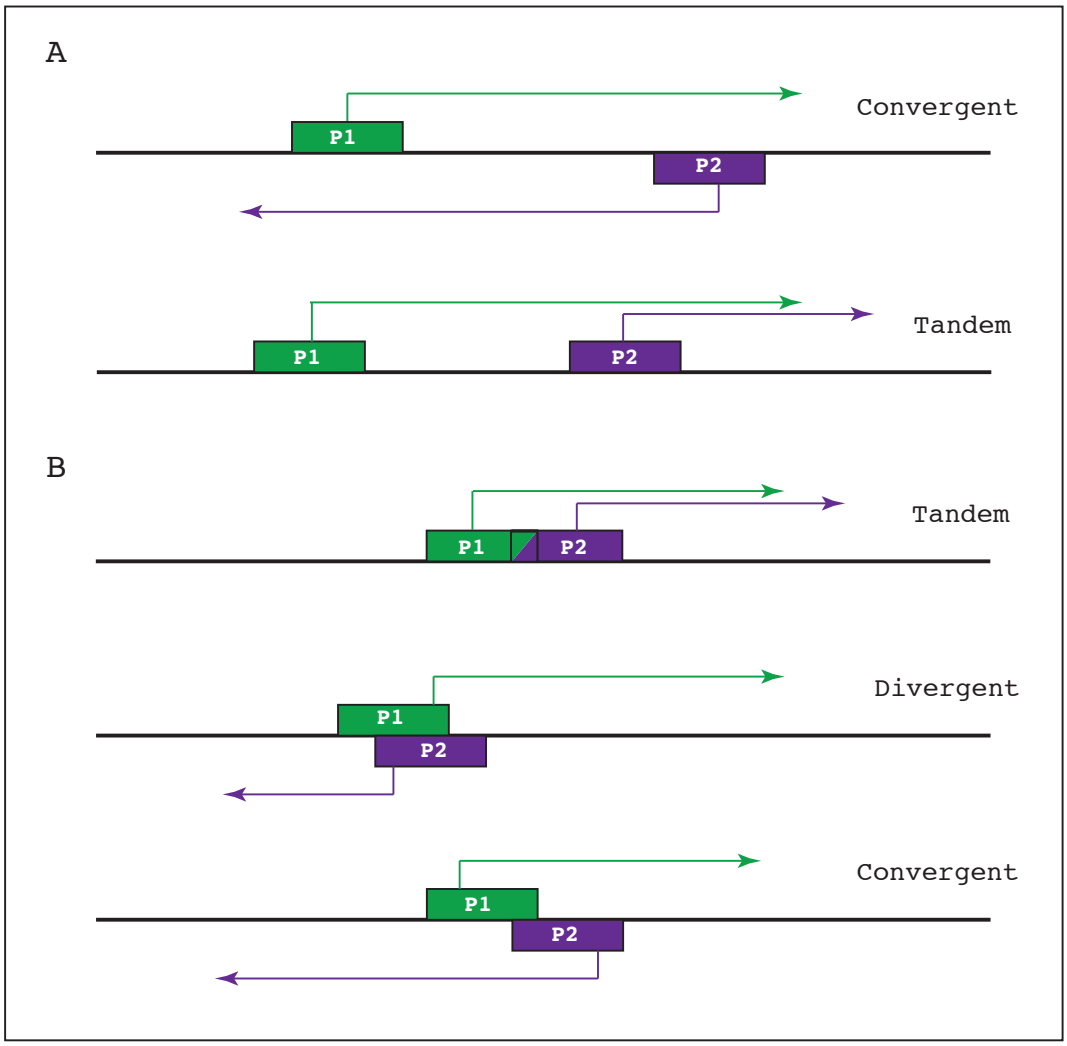


Figure 1-5. Schematic view of different promoter arrangements in overlapping genes. **A)** Non-overlapping promoters, in which transcription of the overlapping genes can be convergent or tandem. **B)** Overlapping promoters, in which transcription of overlapping genes can be tandem, divergent, or convergent.

RNA polymerase III effects on protein-coding genes expression

Although different RNA polymerases transcribe different sets of genes in eukaryotes, there is interplay between the RNA polymerases. In particular, there are several mechanisms by which Pol III, directly or through its products, can affect expression of protein-coding genes, at the transcriptional or post-transcriptional levels (Table 1). As mentioned above, physical overlaps between Pol III and Pol II genes can influence Pol II transcription through transcriptional interference, with at least the two documented cases just mentioned, i.e. the roadblock to Pol II transcription constituted by TFIIIB binding to a tRNA gene (Korde et al. 2014), and the anticorrelation observed between Pol III transcription of proline tRNA genes and Pol II transcription of the gene in which the tRNA genes are nested (Lukoszek et al. 2013).

Physical location of Pol III relative to Pol II genes seems to play a role even when the genes do not overlap. Several studies suggest that active Pol III genes are often located close to Pol II peaks and Pol II TSS, and that the transcription by Pol III and nearby Pol II might be co-regulated (Barski et al. 2010; Moqtaderi et al. 2010; Oler et al. 2010; Canella et al. 2012; Van Bortle et al. 2017). It was also suggested that during differentiation, some Pol III genes change in concert with distant Pol II genes through long-range interactions (Van Bortle et al. 2017). Although the mechanisms involved have not been explored in these genome-wide studies, there are a few specific examples for which more information is available. For example, a B2 SINE located upstream of the growth hormone locus was shown to act as a boundary element. Bidirectional transcription of the SINE by Pol III and Pol II changes the chromatin structure, which leads to transcription of the growth hormone gene during development (Lunyak et al. 2007; Ponicsan et al. 2010). Moreover in neuronal cells, a class of SINEs with enhancer activity was identified (eSINEs). In response to neuronal depolarization, these eSINEs are transcribed by Pol III, which activates transcription of their nearby Pol II genes (Policarpi et al. 2017).

Two cases in which a Pol III product directly affects Pol II transcription have been described. The first is Pol III synthesis of 7SK small nuclear RNA (Yang et al. 2001), which a described above can repress Pol II transcription by sequestering the positive

transcription elongation factor b (P-TEFb) into the P-TEFb/HEXIM/7SK complex. Upon some stimulation such as signaling pathways, stress, UV irradiation and transcription inhibitors, the active form of P-TEFb is released, which leads to productive transcription elongation (Diribarne and Bensaude 2009; Peterlin et al. 2012b). The second is Pol III synthesis of some classes of SINE RNAs, specifically B2 RNAs in the mouse and Alu RNAs in human cells, in response to cellular stress including heat shock. These SINE RNAs can bind directly to Pol II, leading to transcription repression (Espinoza et al. 2004; Mariner et al. 2008).

Pol III is indirectly involved in translation of mRNAs into proteins and is thus involved in another layer of regulation. Indeed, Pol III transcribes 5S rRNA, which has significant structural roles in the ribosome (Ciganda and Williams 2011; Gongadze 2011). Furthermore, Pol III transcribes all tRNAs, which deliver the different amino acids to the translation machinery, so changes in level of tRNAs can be a potential regulator of translation. tRNA genes have been shown to be differentially expressed in different human tissues (Dittmar et al. 2006), and altered tRNA levels have been associated with several diseases (Kirchner and Ignatova 2015). Some studies suggest that the differences in tRNA expression under different conditions coincide with differential codon usage and therefore translation of specific mRNAs (Plotkin et al. 2004; Gingold et al. 2014).

tRNAs are heavily modified at different nucleosides. Although some of these modifications are important for tRNA structure, those in or close to the anticodon loop of the tRNA are critical for translation efficiency and accuracy (Duechler et al. 2016). These modifications are highly dynamic and can vary in response to environmental changes such as cellular stress and nutrient, leading to change in translation, either general translation or translation of some specific mRNAs (Gu et al. 2014; Liu et al. 2016), or to an unfolded protein stress response (Nedialkova and Leidel 2015). Consistent with the importance of tRNA modifications, changes in these modifications are observed in several diseases including cancer (Kirchner and Ignatova 2015; Dong et al. 2016).

tRNAs can be cleaved by different nucleases leading to generation of tRNA halves and short tRNA-derived fragments (tRFs) (Sobala and Hutvagner 2011). tRNA halves usually result from cleavage within the anticodon loop of mature tRNAs in stress conditions, and can play roles in translation repression (Kumar et al. 2016). As to short tRFs, they are often the products of cleavage at the 5' or 3' ends of mature or precursor tRNAs. Some of these tRFs were found to associate with Argonaute proteins, displaying a miRNA-like translation repression, whereas others associate with polysomes to repress translation in a global manner (Keam and Hutvagner 2015).

SINE RNAs can also impact gene expression at the translation level. Some Alus were shown to be located upstream of miRNA genes in the human genome, leading to transcription of these miRNAs by Pol III, which are then processed and can, like other miRNAs, play a role in translational regulation (Borchert et al. 2006). Alu RNA itself was moreover shown to stimulate selective mRNAs translation without affecting global protein synthesis (Rubin et al. 2002). Further, 7SL was found to interact with the 3' UTR of p53 mRNA and suppress its translation (Abdelmohsen et al. 2014), and small RNAs generated from vault RNAs (svRNAs) were shown to act like miRNAs to regulate gene expression (Persson et al. 2009). Finally, BC1 RNA, which is expressed in a subset of neuronal cells, can repress translation through interaction with eIF4A and poly(A) binding protein (Wang et al. 2005).

Effect mediator	Effect level
Genomic arrangements	Transcription
7SK	Transcription
SINE RNA	Transcription
5S rRNA	Translation
tRNA and tRNA modifications	Translation
tRNA-derived fragments	Translation
SINE RNA	Translation
7SL	Translation
Small vault RNA (svRNA)	Translation
BC1 RNA	Translation

Table 1. Pol III effects on Pol II gene expression. Different effect mediators and their effect level are indicated.

Chapter II – Effect of a MIR-SINE on expression of the *Polr3e* gene

Summary

Genomic overlap of genes is a common phenomenon in different organisms. In eukaryotes, especially in yeast, these overlaps can mediate gene expression regulation. Most of the studies have focused on Pol II-transcribed long non-coding RNA, stemming from sequences partially overlapping, or nested in, a protein-coding gene, and often in an antisense orientation. However, overlaps between genes transcribed by Pol II and Pol III, and the possible effects of such overlaps, are poorly studied.

A MIR-SINE located in the first intron of the *Polr3e* gene, which encodes for a subunit of Pol III, is an example of a Pol III gene nested in a Pol II gene in an antisense orientation. ChIP-seq experiments have shown that this MIR is highly occupied by Pol III in mouse liver, and that there is an unexpected accumulation of Pol II at the region corresponding to its 3' end. So we asked whether this MIR is playing a role in the regulation of *Polr3e* expression (Figure 2-1).

We showed that the sequence, and the genomic arrangement of the MIR, i.e. in the first intron of the *Polr3e* gene, are conserved in different mammalian species. We further used available datasets to show that this MIR is highly occupied by Pol III not only in liver, but also in several other cell types. Furthermore, the absence of NELF in the region showing accumulation of Pol II at the 3' of the MIR argues against the presence of a Pol II TSS in this region, and thus against the possibility that the accumulation of Pol II reflects the accumulation usually seen at TSSs of actively transcribed genes.

To examine whether the MIR regulates Pol II transcription through the *Polr3e* gene and thus POLR3E expression, we performed a number of different experiments. We first placed the MIR, either wild-type or with mutations debilitating its promoter, in the intron of an EGPF gene, and observed that active antisense transcription of the

MIR decreased EGFP expression. Second, we used CRISPR/Cas9 to remove the MIR sequence from the genome in cultured cells, and showed that removal of the MIR increased *Polr3e* mRNA and protein levels. Overexpression of the MIR from distant loci in the genome did not change *Polr3e* levels, suggesting a role for MIR transcription, rather than for the MIR transcript, in *Polr3e* regulation. ChIP experiments revealed the absence of Pol II accumulation at 3' region of the MIR in MIR-deleted cell line, further suggesting a transcriptional interference mechanism between Pol II transcribing *Polr3e* and Pol III transcribing the MIR.

We then used available ChIP-seq datasets for Pol II and Pol III and showed that in several cases where a Pol III gene is located inside a Pol II gene, and where both Pol II and Pol III genes are occupied by their respective RNA polymerases, there is a peak of Pol II coinciding with the Pol III gene. These findings suggest that down-regulation of Pol II gene expression by Pol III transcription of a gene nested inside the Pol II gene constitutes a so far unsuspected layer of gene expression regulation.

In this work, I did all the experiments, wrote the first version of the manuscript, and then contributed to revised versions (Yeganeh et al. 2017).

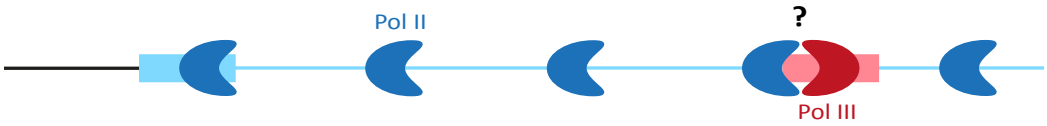


Figure 2-1. Schematic view of the project summary. In this project, the aim was to study whether a Pol III gene nested in a Pol II gene plays a role in its expression regulation, and whether transcriptional interference is involved.

Publication

Transcriptional interference by RNA polymerase III affects expression of the *Polr3e* gene

Meghdad Yeganeh,[1] Viviane Praz,[1,2] Pascal Cousin,[1] and Nouria Hernandez[1]

[1]Center for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne, 1015 Lausanne, Switzerland; [2]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

Overlapping gene arrangements can potentially contribute to gene expression regulation. A mammalian interspersed repeat (MIR) nested in antisense orientation within the first intron of the *Polr3e* gene, encoding an RNA polymerase III (Pol III) subunit, is conserved in mammals and highly occupied by Pol III. Using a fluorescence assay, CRISPR/Cas9-mediated deletion of the MIR in mouse embryonic stem cells, and chromatin immunoprecipitation assays, we show that the MIR affects *Polr3e* expression through transcriptional interference. Our study reveals a mechanism by which a Pol II gene can be regulated at the transcription elongation level by transcription of an embedded antisense Pol III gene.

[*Keywords*: RNA polymerase; gene expression regulation; *Polr3e* gene; antisense transcription; transcriptional interference; SINE]

Supplemental material is available for this article.

Received November 15, 2016; revised version accepted February 15, 2017.

In eukaryotes, RNA polymerase II (Pol II) is responsible for transcription of all of the mRNA-encoding genes as well as most genes encoding small nuclear RNA (snRNA) and microRNAs. Pol II-dependent transcription occurs in several steps, each of which can be subjected to regulation in response to environmental and genetic signaling processes (Fuda et al. 2009). One of the highly regulated steps is the transition from initiation to productive elongation, which is controlled by several positive and negative regulatory factors (Zhou et al. 2012). In particular, the 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF) cause the polymerase to pause just downstream from the transcription start site (TSS). Entering productive elongation involves DSIF and NELF phosphorylation and loss of NELF from the transcription complex. Once the polymerase has entered productive elongation, its elongation rate is highly dynamic and varies across a gene, in particular to allow cotranscriptional events to be performed efficiently. The largest Pol II accumulations detected by chromatin immunoprecipitation (ChIP) assays typically correspond to promoter-proximal pausing before productive elongation and slowing down near the gene 3′ end for cotranscriptional polyadenylation of the transcript (Jonkers and Lis 2015). The distribution of DSIF follows a similar pattern, whereas NELF is typically found only at the promoter-proximal pause region (Zhou et al. 2012).

Genes often lie in overlapping arrangements on either the same strand or opposite strands. The overlap can be partial, or an entire gene may be located or "nested" inside another gene, usually within an intron (Kumar 2009). A frequent arrangement found in both yeast and mammals but studied mostly in yeast is a long noncoding RNA (lncRNA; natural antisense transcript) gene present in antisense orientation relative to a protein-coding gene (Katayama et al. 2005; Huber et al. 2016). The different arrangements of overlapping genes can contribute to the regulation of gene expression by a number of mechanisms involving, in general, the natural antisense transcripts and/or, in some cases, the process of overlapping transcription (Pelechano and Steinmetz 2013). As recent examples, in yeast, a *CDC28* antisense lncRNA induced upon osmotic stress mediates gene looping and the transfer of Hog1 and associated factors from the 3′ untranslated region (UTR) to the *CDC28* TSS region, resulting in *CDC28* transcription activation (Nadal-Ribelles et al. 2014). In mammalian cells, the lncRNA Wrap53, an antisense transcript originating from the *p53* locus, binds CTCF and contributes to p53 regulation (Saldana-Meyer et al. 2014), and as a third example, a lncRNA antisense

Corresponding author: nouria.hernandez@unil.ch
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.293324.116. Freely available online through the *Genes & Development* Open Access option.

to the *Nos1* locus down-regulates Nos1 protein but not *Nos1* mRNA, suggesting a post-transcriptional effect (Korneev et al. 2015).

Most cases of gene regulation by an overlapping ncRNA gene in eukaryotes concern arrangements in which both genes are transcribed by Pol II, with very few studied examples of Pol II–Pol I and Pol II–Pol III gene overlaps. Among these, the Pol II TAR1 gene is nested within the Pol I 25S rRNA gene in *Saccharomyces cerevisiae,* but the influence, if any, of this layout on gene expression has not been examined directly (Coelho et al. 2002). In *Arabidopsis thaliana*, a set of nine Pol III proline transfer RNA (tRNA) genes is located antisense of the Pol II AtNUDT22 gene, and the two genes display negatively correlated expression levels, consistent with the possibility that they influence each other's expression (Lukoszek et al. 2013).

In addition to genes of known function, Pol III transcribes some short interspersed nuclear elements (SINEs). SINEs originated by retrotransposition mostly from tRNA and *Rn7sl* genes, often resulting in the transposed element carrying the gene-internal promoter of the source gene (Dieci et al. 2013). Although most SINEs are epigenetically repressed, some are actively transcribed by Pol III as independent transcription units (Roberts et al. 2003; Barski et al. 2010; Canella et al. 2010, 2012; Moqtaderi et al. 2010; Oler et al. 2010; Raha et al. 2010; Renaud et al. 2014). SINEs have long been considered as junk DNA, but it is now clear that they can profoundly impact genome functions both in *cis* (for example, by constituting new enhancers or splice sites) and in *trans* (for example, by producing RNAs that affect Pol II transcription) (Kramerov and Vassetzky 2011). Here we examined the role of a member of the mammalian interspersed repeat (MIR) family, an ancient family of tRNA-derived SINEs that were amplified before the mammalian radiation (Smit and Riggs 1995). This MIR is nested in antisense orientation within the first intron of the *Polr3e* gene, which codes for one of the Pol III subunits. We show that this arrangement is conserved in different mammalian species and that it directly impacts on Pol II transcription elongation through the *Polr3e* gene. Thus, the Pol III transcribed MIR can contribute to regulation of a Pol III subunit-encoding gene.

Results

A MIR in the first intron of the Polr3e *gene is conserved among mammalian species and highly occupied by both Pol III and Pol II*

In both the mouse and human genomes, the first intron of the *Polr3e* gene contains an antisense MIR SINE (Canella et al. 2012). ChIP-seq (ChIP combined with high-throughput sequencing) data obtained from mouse livers reveal that this MIR is as highly occupied by Pol III as a tRNA Leu gene located upstream of the *Polr3e* TSS (Fig. 1A). Indeed, this MIR was found to be highly occupied by Pol III as compared with the mean occupancy scores of either all Pol III-occupied loci or just SINEs in not only mouse livers but also a mouse hepatocarcinoma cell line and human IMR90 and IMR90Tert cell lines (Fig. 1B; Renaud et al. 2014; Orioli et al. 2016), consistent with its high occupancy also in HeLa cells (Oler et al. 2010).

The high Pol III occupancy of this particular MIR is in contrast to the low occupancy of most SINEs and prompted us to search for its presence in other species. We found MIR-related sequences located antisense in the first intron of the *Polr3e* genes of all examined mammalian species, including the monotreme platypus (*Ornithorhynchus anatinus*), as illustrated by the sequence similarity tree in Figure 1C. The sequence alignment in Supplemental Figure S1 shows that all of these MIRs have potentially functional type 2 Pol III promoters; i.e., gene-internal A and B boxes separated by 25–26 base pairs (bp). This conservation is consistent with MIRs having amplified before the mammalian radiation (Smit and Riggs 1995) and suggests that the MIR in the first intron of the *Polr3e* gene might have a function.

When examining the Pol II and Pol III occupancy patterns in Figure 1A, we noticed a striking accumulation of Pol II not only at the TSS, as expected from pausing before escape into productive elongation, but also just before the antisense Pol III MIR (Fig. 1A; see also Canella et al. 2012). ChIP-seq data from HeLa cells (Liu et al. 2014) show DSIF accumulation near both the TSS and the MIR but NELF accumulation only near the TSS, arguing against Pol II accumulation at the MIR resulting from a second, unannotated TSS in this region (Fig. 2). A possible interpretation is that the MIR contributes to a Pol II accumulation at its 3′ end through either transcription interference or a *trans*-acting mechanism involving the MIR RNA.

Active transcription of the MIR in antisense, but not sense, orientation within an EGFP-expressing construct leads to decreased fluorescence intensity

To examine the effect of the MIR on expression of an overlapping Pol II gene, we placed the MIR (either wild type or with mutated A and B boxes in either sense or antisense orientation) within an intron inserted into the EGFP-coding sequence (Fig. 3A,B; Santillan et al. 2014). In vitro transcription assays with these constructs revealed robust and intact A-box- and B-box-dependent expression of both the sense and antisense MIR (Fig. 3C). We thus used these constructs to create stable inducible cell lines by cotransfection into Flp-In T-REx 293 cells along with a plasmid expressing Flp-recombinase and selection of the transfected cells with hygromycine. Northern blotting revealed weak but clearly detectable A-box- and B-box-dependent expression from both the sense and antisense MIRs (Fig. 3D). The relatively weak signal, which is in contrast to robust MIR expression in vitro, suggests rapid degradation of the MIR transcript in the cell.

We measured EGFP expression by FACS in either noninduced cells or cells induced for EGFP expression by doxycycline. When the MIR was antisense relative to EGFP transcription, fluorescence intensity was decreased slightly (20%–30%) but reproducibly in cells containing the

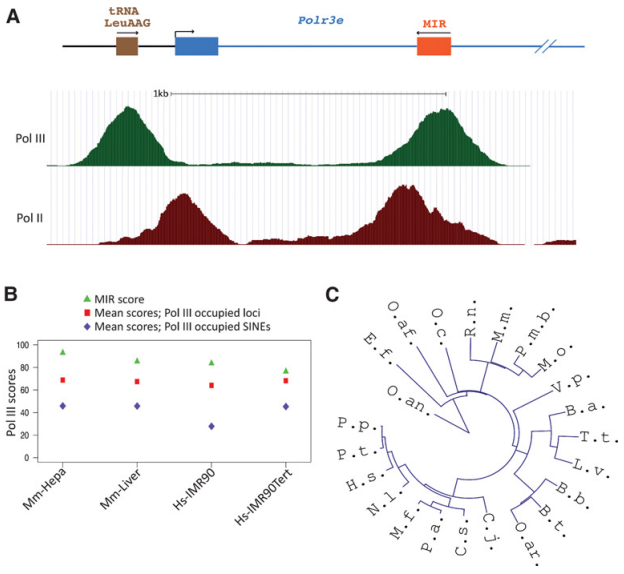


Figure 1. A MIR in the first intron of the *Polr3e* gene is conserved among mammals and highly occupied by Pol III and Pol II. (*A*) The genomic arrangement of the beginning of the mouse *Polr3e* gene is shown at the *top*. A tRNA Leu gene is located upstream of the *Polr3e* TSS. The MIR in the first intron is transcribed in the antisense direction. The *bottom* part shows a University of California at Santa Cruz (UCSC) genome browser view of Pol III and Pol II ChIP-seq profiles in mouse livers. (*B*) MIR, average Pol III-occupied locus, and average Pol III-occupied SINE Pol III occupancy scores [$\log_2$(immunoprecipitation/input)] shown as percentages in mouse hepa1–6 cells, mouse livers, IMR90 cells, and IMR90tert cells. (*C*) Alignment tree of MIR sequences in the first intron of *Polr3e* in different mammalian species, including *Balaenoptera acutorostrata* (B.a.), *Bos taurus* (B.t.), *Bubalus bubalis* (B.b.), *Callithrix jacchus* (C.j.), *Chlorocebus sabaeus* (C.s.), *Eptesicus fuscus* (E. f.), *Homo sapiens* (H.s.), *Lipotes vexillifer* (L.v.), *Macaca fascicularis* (M.f.), *Microtus ochrogaster* (M.o.), *Mus musculus* (M.m), *Nomascus leucogenys* (N.l.), *Ornithorhynchus anatinus* (O.an.), *Orycteropus afer* (O.af.), *Oryctolagus cuniculus* (O.c.), *Ovis aries* (O.ar.), *Pan paniscus* (P.p.), *Pan troglodytes* (P.t.), *Papio anubis* (P.a.), *Peromyscus maniculatus bairdii* (P.m.b), *Rattus norvegicus* (R.n.), *Tursiops truncates* (T.t), and *Vicugna pacos* (V.p.). The tree is based on the alignment in Supplemental Figure S1 and was generated from Jalview with the neighbor joining based on percent identity; the file was processed at the Interactive Tree of Life (iTOL) Web site (http://itol.embl.de) for circular graphic representation.

wild-type MIR as compared with cells containing the mutant MIR construct (Fig. 3E). This was true at both the low leaky EGFP expression levels in the absence of doxycycline and at high doxycycline-induced EGFP expression levels. In contrast, no measurable effect was observed in this assay when the MIR was in the sense orientation (Fig. 3F). Thus, a Pol III transcribed MIR can reduce expression of a Pol II gene in which it is embedded in antisense orientation.

CRISPR/Cas9-mediated deletion of the MIR leads to increased expression of Polr3e

To determine the effect of the MIR in its natural genomic context, we used the CRISPR/Cas9 system to delete MIR genomic sequences in mouse embryonic stem (ES) cells. A schematic view of deletions obtained in different ES cell clones is shown in Figure 4A. We engineered deletions that left intact the 3′ region of the MIR where our ChIP-seq data had revealed accumulation of Pol II. As expected, MIR RNA was absent from these cell lines as determined by RT-qPCR (Fig. 4B), and at least the one cell line (KO11) that we tested still expressed the three pluripotency transcription factor-encoding genes *Oct4*, *Sox2*, and *Nanog* despite a number of cell passages imposed by genome engineering and single-cell cloning (Fig. 4C).

We then measured *Polr3e* mRNA expression levels in the various MIR knockout cells and wild-type cells by RT-qPCR with qPCR primers located inside a single exon toward the 3′ end of the *Polr3e* transcription unit (i.e., measuring both processed and unprocessed transcripts) and observed increased *Polr3e* total mRNA levels in all MIR knockout cell lines as compared with wild-type cells (Fig. 4D).

We focused on the MIR-deleted cell line KO11. To determine whether the increase in total mRNA level of *Polr3e* reflected increased transcription or any post-transcriptional effect such as increased stability, we measured levels of *Polr3e* pre-mRNA and mature mRNA in wild-

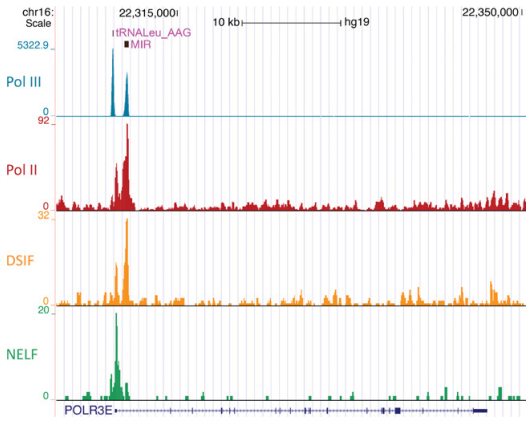


Figure 2. UCSC genome browser views of Pol III, Pol II, DSIF, and NELF ChIP-seq profiles in HeLa cells in the *POLR3E* gene region. Tracks are from ENCODE (Pol III subunit RPC1) and Liu et al. (2014) (Pol II, DSIF, and NELF).

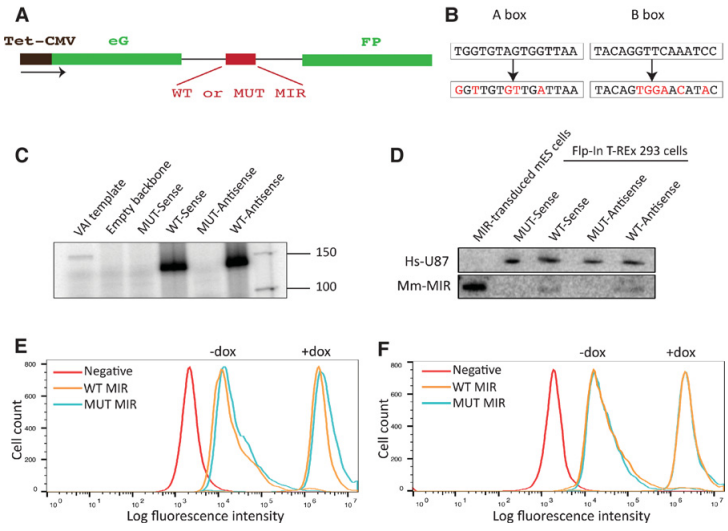


Figure 3. Active transcription of the MIR in antisense, but not sense, orientation within an EGFP-expressing construct leads to decreased fluorescence intensity. (*A*) Schematic view of the EGFP construct. EGFP-coding sequence under a tetracycline-inducible promoter is interrupted by an intron in which wild-type (WT) or mutant (MUT) MIR in sense or antisense orientation is inserted. (*B*) Mutations introduced in the A and B boxes of the MIR. Changes are indicated in red. (*C*) In vitro transcription performed with the templates indicated at the *top*: mutant MIR or wild-type MIR inserted in the EGFP intron in sense or antisense orientation. The Adenovirus 2 VAI gene was used as a positive control. (*D*) Northern blot analysis of mouse MIR expression in human Flp-In T-REx 293 cells transfected with the constructs indicated at the *top*. Total RNA extracted from mouse embryonic stem (ES) cells transduced with a construct containing the same MIR insert as in the EGFP constructs served as a positive control for MIR size. As an internal control for loading, the blot was probed with an oligonucleotide complementary to human U87. (*E*) FACS analysis of cells transfected with EGFP constructs containing wild-type and mutant MIR in antisense orientation represented as cell count (*Y*-axis) versus log fluorescence intensity (*X*-axis) either noninduced or induced with 2 µg/mL doxycycline and collected 72 h after induction. Non-transfected Flp-In T-REx 293 cells were used as a negative control. The experiment was repeated four times. $P\text{-value}_{-\text{dox}} = 0.000023$; $P\text{-value}_{+\text{dox}} = 0.000016$, calculated using Student's *t*-test. (*F*) As in *E* but with cells transfected with wild-type or mutant sense constructs. $P\text{-value}_{-\text{dox}} = 0.131$; $P\text{-value}_{+\text{dox}} = 0.167$.

type and KO11 cells by RT-qPCR. We designed (1) qPCR primers inside a single intron and (2) a primer at an exon–exon junction paired with another within an exon to amplify specifically precursor and mature mRNAs, respectively. We observed increased levels of both unspliced (Fig. 4E) and spliced (Fig. 4F) *Polr3e* mRNA in KO11 cells relative to wild-type cells, and this increase in turn resulted in an almost 1.5-fold increase in protein levels as determined by Western blot (Fig. 4G). However, these increased POLR3E levels did not lead to increased levels of 5S rRNA, pre-tRNA Ile, and U6 snRNA (Supplemental Fig. S2), suggesting that this Pol III subunit is not limiting for Pol III activity under the conditions tested.

The regulation of Pol II gene expression by a nested Pol III gene shown here suggests the existence of another layer of gene expression regulation by interplay between RNA polymerases. To determine how general this phenomenon might be, we examined all of the Pol III (RPC1) ENCODE peaks from HeLa cells and extracted those embedded in a Pol II transcription unit. We found 984 ENCODE Pol III peaks within Pol II genes, but examination of several of those RPC1 peaks in Pol II-occupied genes (for example, peaks located within the *CDA*, *KPNA6*, *YARS*, *STK40*, *CTH*, *USP33*, *DNAJB4*, and *SOAT1* Pol II transcription units on chromosome 1) did not reveal corresponding accumulations of Pol II. However, in the human genome, the analysis of anti-Pol III ChIP-seq experiments is complicated by the presence of a large number of repetitive sequences derived from Pol III transcription units; as a result, a large proportion of the sequence tags obtained in such experiments matches several locations (sometimes several hundred locations) in the genome. We therefore remapped the ENCODE sequence tags and recalculated peak scores using our previously described method (Canella et al. 2012), which assigns different weights to tags according to the number of times they were sequenced and the number of corresponding matches in the genome. The results confirmed 22 of the 984 peaks, of which 15 coincided with RPC4 peaks observed in IMR90Tert cells (see Supplemental Table S1; Orioli et al. 2016). These 15 Pol III peaks were located in nine different Pol II transcription units, five of which were clearly occupied by Pol II (Liu et al. 2014) and, importantly, clearly displayed Pol II accumulations at locations of Pol III peaks, as shown in Figure 5 for the *VAC14*, *SHF*, *CTC1*, and *HES7* genes. Such accumulations occurred when the Pol III transcription unit was orientated sense or antisense relative to the Pol II gene (Fig. 5). The results show that at least in HeLa and IMR90Tert cells, there are very few cases of Pol III-occupied transcription units leading to intragenic Pol II accumulations; they further suggest that both sense and antisense Pol III transcription units can lead to Pol II roadblocks, although the final effect on gene expression is likely to depend on both Pol II and Pol III transcription levels.

The MIR effect on Polr3e *is mediated by transcriptional interference*

To determine whether MIR RNA is sufficient to regulate *Polr3e* expression, we overexpressed the MIR using a

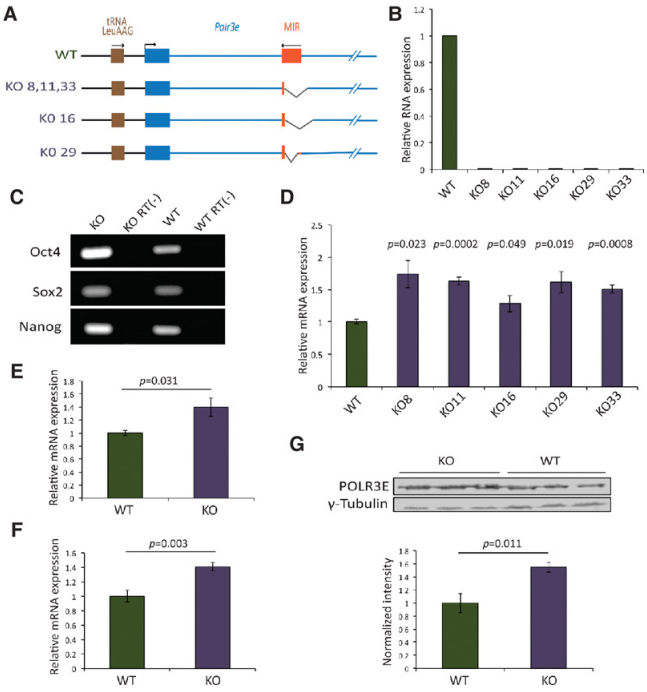


Figure 4. CRISPR/Cas9-mediated deletion of the MIR leads to increased expression of *Polr3e*. (*A*) Schematic view of CRISPR/Cas9-mediated deletions in different clones (#8, #11, #16, #29, and #33). All deletions maintain intact the region corresponding to the 3′ end of the MIR. (*B*) RT-qPCR performed with MIR-specific primers for both reverse transcription and qPCR with RNA extracted from the indicated cell lines. The qPCR results were normalized to *Gapdh* expression. (*C*) RT–PCR performed with total RNA from KO11 and wild-type ES cells to monitor expression of the pluripotency genes *Oct4*, *Sox2*, and *Nanog*. In the control RT(−) lanes, no reverse transcriptase was included in the reactions. (*D*) Results of RT-qPCR performed with primers detecting both pre-mRNA and mature *Polr3e* mRNA in the indicated ES knockout cell lines relative to wild-type cells. The qPCR results were normalized to *Actb* mRNA. Error bars represent ± SEM. $n = 3$. The *P*-values were calculated using Student's *t*-test and are relative to the wild type. (*E*) Results of RT-qPCR performed with primers detecting *Polr3e* pre-mRNA in the knockout ES cells relative to wild-type cells. The results were normalized to *Actb* expression. Error bars and *P*-values are as in *D*. (*F*) As in *E* but with primers detecting only mature *Polr3e* mRNA. (*G*, *top*) Western blot performed with anti-POLR3E and anti-γ-tubulin antibodies with protein extracts from knockout and wild-type ES cells. (*Bottom*) The POLR3E band intensities were quantified and normalized to their corresponding γ-tubulin band intensities. Error bars and *P*-values are as in *D*.

lentiviral vector coexpressing GFP. We did so not in only wild-type but also KO11 cells, in case the endogenous MIR levels in wild-type cells were already saturating. Although we could confirm MIR overexpression in both wild-type and KO11 ES cells by RT-qPCR (Fig. 6A, B, left panels), we did not observe any significant change in the level of total *Polr3e* mRNA (Fig. 6A,B, right panels) or POLR3E protein (Fig. 6C,D). This suggests that MIR RNA overexpressed from other loci in the genome does not affect *Polr3e* transcription or *Polr3e* mRNA translation.

The results above (Fig. 1A) revealed Pol II accumulation toward the 3′ end of the MIR, consistent with the MIR creating a roadblock to elongation. We performed ChIP-qPCR assays in mouse ES cells using antibodies against the RPC4 (POLR3D) subunit of Pol III and the RPB2 (POLR2B) subunit of Pol II. The positions of the primers used for qPCR are shown in Figure 6E and Supplemental Figure S3A. As expected, we detected a strong Pol III signal on the tRNA Leu gene upstream of the *Polr3e* TSS in both wild-type and KO11 cells, and the Pol III signal in the MIR region was detectable only in the wild-type cells (Fig. 6F; Supplemental Fig. S3B). With the anti-RPB2 antibody, we observed similar Pol II enrichment on the *Polr3e* TSS in both wild-type and MIR KO11 cells, suggesting that the presence or absence of the MIR did not affect Pol II recruitment (Fig. 6G; Supplemental Fig. S3C). Similarly, Pol II recruitment was unaffected at several locations along the *Polr3e* gene with one notable exception: The Pol II accumulation in wild-type cells just before the MIR was absent in MIR KO11 cells (Fig. 6G; Supplemental Fig. S3C). These results indicate that removal of the MIR and thus suppression of antisense Pol III transcription within the first *Polr3e* intron relieve Pol II pausing, leading to the increased expression of *Polr3e* observed above.

Discussion

We describe a genomic arrangement, conserved in mammalian cells, in which a MIR antisense Pol III transcription unit creates a roadblock within the *Polr3e* Pol II transcription unit, leading to decreased expression of the *Polr3e* gene at both the level of mRNA and protein accumulation. The effect on protein expression was relatively modest (1.5-fold) but in line with the conclusions of a recent systematic study in yeast, where Pol II antisense transcription of ncRNAs was shown to lead, on average, to a less than twofold reduction in the protein levels encoded by the sense genes (Huber et al. 2016). However, the effect of antisense transcription was different in different conditions; in this respect, it is possible that the MIR affects *Polr3e* levels differently in different cells and tissues and under different conditions.

Insertion of the MIR into an intron placed within the EGFP-coding sequence leads to decreased EGFP expression under conditions of both low (noninduced) and high (doxycycline-induced) EGFP transcription. This is in

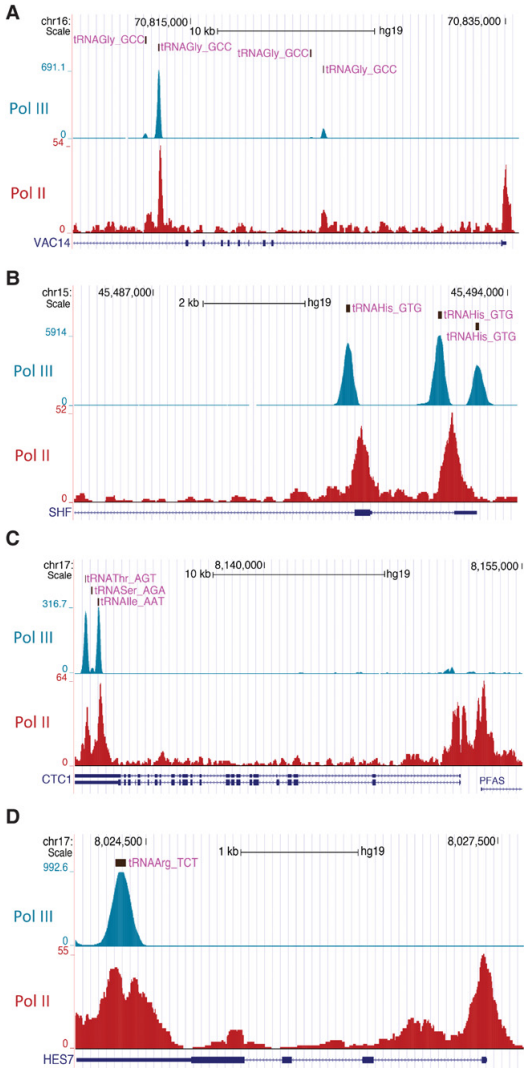


Figure 5. UCSC genome browser views showing examples of Pol II accumulations overlapping with Pol III peaks within the *VAC14* (*A*), *SHF* (*B*), *CTC1* (*C*), and *HES7* (*D*) genes. Tracks are from ENCODE (RPC1) and Liu et al. (2014) (Pol II).

contrast to ncRNA Pol II transcription in yeast, which seems to suppress specifically low levels of gene expression. Thus, in a genome-wide study, antisense ncRNA transcription was found to switch off corresponding sense genes under conditions of low, but not high, expression (Xu et al. 2011). Similarly, a ncRNA transcribed antisense of the GAL10 gene suppressed leaky expression of GAL10 and GAL1 in glucose-containing repressing medium but not in galactose-containing inducing medium (Lenstra et al. 2015). The inhibitory effect of the MIR under conditions of both low and high Pol II transcription may be a specificity of an antisense Pol III transcription unit as opposed to an antisense Pol II ncRNA transcription unit, or the range of the EGFP assay may not cover Pol II expression levels that might be differentially affected by MIR expression.

Our results show that the levels of the POLR3E subunit of Pol III can be regulated by the MIR. Like its yeast ortholog, Rpc37, with Rpc53, POLR3E (RPC5) forms a dimer with POLR3D (RPC4) that resembles TFIIF (Hu et al. 2002; Cramer et al. 2008). In yeast, this TFIIF-like dimer contributes to promoter opening and transcription initiation (Kassavetis et al. 2010) and is necessary, together with Rpc11, for formation of the pretermination complex and transcription termination (Arimbasseri and Maraia 2015). Mammalian POLR3E is also essential for Pol III transcription, as immunodepletion of this subunit from the Pol III complex debilitated Pol III transcription in vitro (Hu et al. 2002). Thus, the POLR3E subunit plays essential roles in the Pol III transcription process, and regulation of its levels may be critical. Although we did not observe higher levels of several Pol III products in MIR KO11 cells as compared with wild-type cells, suggesting that POLR3E was not limiting under our experimental conditions, it is likely that in other cell types or conditions in which the *Polr3e* gene is less transcribed, inhibition by the MIR, which, as mentioned above, reduces expression of lowly expressed EGFP, leads to less Pol III activity. Under such a condition, inhibition by the MIR might constitute a negative feedback loop, where overactivated Pol III would lead to increased MIR transcription and thus decreased expression of POLR3E, leading in turn to decreased Pol III activity.

How frequent is inhibition of Pol II transcription by an embedded Pol III transcription unit? A stringent analysis of Pol III occupancy in HeLa and IMR90Tert cells revealed only a handful of Pol III-occupied loci embedded within Pol II-occupied genes, but, in all of these cases, the Pol III peaks coincided with accumulations of Pol II. Thus, in these particular cultured cells, there are few potential cases. However, there is a very large number of unoccupied SINEs within Pol II transcription units: The observed tissue-specific expression of SINEs and tRNA genes (Dittmar et al. 2006; Faulkner et al. 2009) raises the possibility that some of the embedded Pol III transcription units create roadblocks for expression of their host Pol II genes in a tissue-specific manner. Moreover, there might be mechanisms other than elongation block for regulation of Pol II genes by nested Pol III transcription units.

The ineffectuality of MIR RNA overexpression to impact on *Polr3e* expression even in MIR KO11 cells, the inhibitory effect on EGFP expression of an actively transcribed MIR embedded antisense within the EGFP transcription unit, and, perhaps most telling, the accumulation of Pol II in the first *Polr3e* intron when—and only when—the MIR is present all argue for a mechanism of inhibition entailing a transcriptional interference mechanism. Transcriptional interference in overlapping genes can be modeled in several ways (Shearwin et al. 2005), one of which is disruption of transcription factor binding

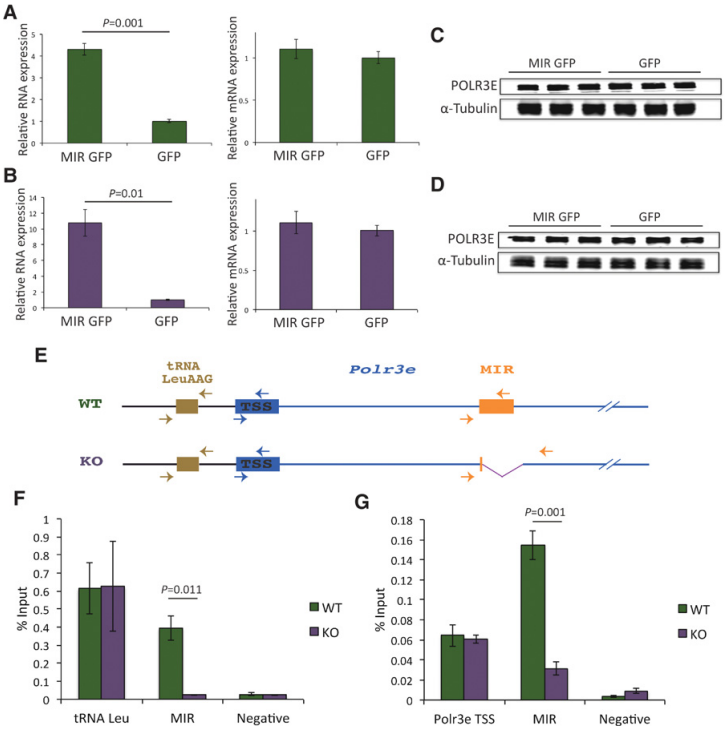


Figure 6. The MIR effect on *Polr3e* is mediated in *cis*. (*A*) RT-qPCR detecting the MIR (*left* panel) and total *Polr3e* mRNA (*right* panel) in wild-type ES cells transduced with a GFP lentiviral construct either containing the MIR (MIR GFP) or without the MIR (GFP), as indicated on the *X*-axis. The results were normalized to *Actb* mRNA. Error bars represent ±SEM. $n = 3$. The *P*-values were calculated using Student's *t*-test. (*B*) As in *A* but in KO11 cells. (*C*) Western blot performed with anti-POLR3E and anti-α-tubulin antibodies with protein extracts from wild-type ES cells transduced with MIR GFP or GFP lentiviral constructs. (*D*) As in *C* but with KO11 ES cell extracts. (*E*) Positions of primers used for qPCR after ChIP in wild-type and KO11 cells. (*F*) ChIP-qPCR performed with anti-RPC4 antibody. The tRNA Leu gene upstream of the *Polr3e* TSS served as a positive control. A Pol II-occupied locus (*Mycbp*) served as a negative control. The qPCR signals were normalized to input. Error bars and *P*-values are as in *A*. (*G*) ChIP-qPCR performed with anti-RPB2 antibody. The *Polr3e* TSS served as a positive control, and a Pol III-occupied locus (a tRNA Leu gene on chromosome 13) served as a negative control. The qPCR signals were normalized to the input. Error bars and *P*-values are as in *A*.

by the traveling polymerase. In our case, the MIR transcription unit does not overlap with the *Polr3e* promoter and must thus be interfering with Pol II elongation within the *Polr3e* gene. Interference might result from one of the Pol III transcription factors forming a roadblock to Pol II elongation. TFIIIC, which binds directly to the A and B boxes and recruits TFIIIB, has a low enrichment relative to TFIIIB subunits in ChIP-seq analyses, consistent with it detaching from the template after establishment of the Pol III transcription initiation complex (Roberts et al. 2003, 2006; Moqtaderi and Struhl 2004; Soragni and Kassavetis 2008). TFIIIB, on the other hand, has very high occupancy and is responsible, in yeast, for the inhibition by a Pol III tRNA gene of readthrough Pol II transcription from an upstream lncRNA gene, an effect that was independent on orientation (Korde et al. 2014). A TFIIIB-mediated effect might account for some of the cases that we observed in the human genome, which involve sense as well as antisense Pol III transcription units (Fig. 5), but, in the case of the MIR, the results of the EGFP assay suggest that a major roadblock function occurs only when the MIR is antisense. Thus, a more likely possibility is that, in this case, inhibition results from head-to-head collision of Pol II and Pol III. The Pol II encountering Pol III might either just slow down, pause, or have to backtrack to allow cleavage of the nascent RNA 3′ end and realignment into the catalytic site. In yeast, the collision of two Pol II machineries transcribing from convergent promoters was shown to block transcription and lead to Pol II polyubiquitylation and degradation (Hobson et al. 2012). Thus, an interesting question is whether full-length *Polr3e* mRNAs can be generated only when the MIR happens to not be transcribed by Pol III or whether Pol II, perhaps in collaboration with some bypass factors, can, at least on occasion, transcribe through the roadblock.

Materials and methods

Cell culture, transfection, and lentiviral transduction

V6.5 mouse ES cells were maintained on 0.1% gelatin in DMEM/F-12 GultaMAX (Gibco) supplemented with 15% ES cell-qualified fetal bovine serum (Gibco), 100 U/mL penicillin, 100 µg/mL streptomycin, 0.1 mM nonessential amino acids, 0.1 mM 2-mercaptoethanol, and 1000 U/mL LIF (Merck Millipore). Flp-In T-REx 293 cells were cultured in DMEM containing 10% tetracycline-free fetal calf serum (Bioconcept) and penicillin/streptomycin. Mouse ES cells and 293 cells were transfected with 1:4 and 1:3 (microgram:microliter) ratios of DNA to FuGENE HD transfection reagent (Promega), respectively. For production of lentiviral particles, 293FT cells were cotransfected with psPAX2, pMD2.G, and pRRLSIN.cPPT.PGK-GFP.WPRE plasmid (Addgene) containing the MIR and ~140 bp of 5′ and 3′ flanking genomic sequence (the same sequence used for EGFP assay). The supernatant of transfected cells was collected 48 and 72 h after transfection, and the lentiviral particles were concentrated by ultracentrifugation. Transduced ES cells were selected by FACS.

CRISPR/Cas9 genome engineering

CRISPR/Cas9 genome engineering was performed as described (Ran et al. 2013). To delete the MIR genomic sequence, preserving its 3′ end region, we used two short guide RNAs (sgRNAs) (see Supplemental Fig. S4A). After cotransfection of mouse ES cells with pSpCas9(BB)-2A-GFP vectors containing sgRNA1 and sgRNA2 and selection by FACS, we screened for MIR deletions by PCR on genomic DNA of single-cell-derived ES cells with a pair of primers flanking the MIR (Supplemental Fig. S4B) and by sequencing of some of the PCR products (Supplemental Fig. S4C). We obtained different deletion lengths (Supplemental Fig. S4D).

RNA extraction and RT-qPCR

Total RNA was extracted and DNase I-treated with the miRNeasy minikit (Qiagen). One microgram of RNA was reverse-transcribed with M-MLV reverse transcriptase (Promega) with either gene-specific primers or random hexamers. The sequences of the qPCR primers are listed in Supplemental Table S2.

Western blot, Northern blot, and in vitro transcription

The primary antibodies for Western blots (anti-POLR3E [RPC5], CS1542 [Hu et al. 2002], and anti-γ- and α-Tubulin [Santa Cruz Biotechnology]) were used at 1:1000 dilutions. For Northern blots, RNA was extracted with TRIzol reagent (Ambion) according to the manufacturer's instructions, and 20 µg (Fig. 3D) or 10 µg (Supplemental Fig. S2) of total RNA was used. The oligonucleotide probe sequences are listed in Supplemental Table S2. In vitro transcription was performed according to Lobo et al. (1992).

ChIPs

ChIPs were performed as described in Orioli et al. (2016). Chromatin was sheared with a Bioruptor sonicator (Diagenode). Sonicated chromatin from 5 million cells was used for each ChIP. The antibodies for immunoprecipitation were anti-RPB2 (POLR2B) (Santa Cruz Biotechnology, H-201) and anti-RPC4 (POLR3D) (Canella et al. 2012). The sequences of qPCR primers used after ChIP are listed in Supplemental Table S2.

Acknowledgments

We thank Vincent Dion for discussion and help with the EGFP assay, Bart Deplancke and Carol W. Greider for discussion, and Philippe Lhote for help with cell culture. This work was funded by the University of Lausanne and Swiss National Science Foundation (SNSF) grants 31003A_132958 and 31003A_169233.

References

Arimbasseri AG, Maraia RJ. 2015. Mechanism of transcription termination by RNA polymerase III utilizes a non-template strand sequence-specific signal element. *Mol Cell* **58:** 1124–1132.

Barski A, Chepelev I, Liko D, Cuddapah S, Fleming AB, Birch J, Cui K, White RJ, Zhao K. 2010. Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat Struct Mol Biol* **17:** 629–634.

Canella D, Praz V, Reina JH, Cousin P, Hernandez N. 2010. Defining the RNA polymerase III transcriptome: genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res* **20:** 710–721.

Canella D, Bernasconi D, Gilardi F, LeMartelot G, Migliavacca E, Praz V, Cousin P, Delorenzi M, Hernandez N, Cycli XC. 2012. A multiplicity of factors contributes to selective RNA polymerase III occupancy of a subset of RNA polymerase III genes in mouse liver. *Genome Res* **22:** 666–680.

Coelho PS, Bryan AC, Kumar A, Shadel GS, Snyder M. 2002. A novel mitochondrial protein, Tar1p, is encoded on the antisense strand of the nuclear 25S rDNA. *Genes Dev* **16:** 2755–2760.

Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, et al. 2008. Structure of eukaryotic RNA polymerases. *Annu Rev Biophys* **37:** 337–352.

Dieci G, Conti A, Pagano A, Carnevali D. 2013. Identification of RNA polymerase III-transcribed genes in eukaryotic genomes. *Biochim Biophys Acta* **1829:** 296–305.

Dittmar KA, Goodenbour JM, Pan T. 2006. Tissue-specific differences in human transfer RNA expression. *PLoS Genet* **2:** e221.

Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, Schroder K, Cloonan N, Steptoe AL, Lassmann T, et al. 2009. The regulated retrotransposon transcriptome of mammalian cells. *Nat Genet* **41:** 563–571.

Fuda NJ, Ardehali MB, Lis JT. 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461:** 186–192.

Hobson DJ, Wei W, Steinmetz LM, Svejstrup JQ. 2012. RNA polymerase II collision interrupts convergent transcription. *Mol Cell* **48:** 365–374.

Hu P, Wu S, Sun Y, Yuan CC, Kobayashi R, Myers MP, Hernandez N. 2002. Characterization of human RNA polymerase III identifies orthologues for *Saccharomyces cerevisiae* RNA polymerase III subunits. *Mol Cell Biol* **22:** 8044–8055.

Huber F, Bunina D, Gupta I, Khmelinskii A, Meurer M, Theer P, Steinmetz LM, Knop M. 2016. Protein abundance control by non-coding antisense transcription. *Cell Rep* **15:** 2625–2636.

Jonkers I, Lis JT. 2015. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16:** 167–177.

Kassavetis GA, Prakash P, Shim E. 2010. The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. *J Biol Chem* **285:** 2695–2706.

Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, et al. 2005. Antisense transcription in the mammalian transcriptome. *Science* **309:** 1564–1566.

Korde A, Rosselot JM, Donze D. 2014. Intergenic transcriptional interference is blocked by RNA polymerase III transcription factor TFIIIB in *Saccharomyces cerevisiae*. *Genetics* **196:** 427–438.

Korneev SA, Maconochie M, Naskar S, Korneeva EI, Richardson GP, O'Shea M. 2015. A novel long non-coding natural antisense RNA is a negative regulator of Nos1 gene expression. *Sci Rep* **5:** 11815.

Kramerov DA, Vassetzky NS. 2011. SINEs. *Wiley Interdiscip Rev RNA* **2:** 772–786.

Kumar A. 2009. An overview of nested genes in eukaryotic genomes. *Eukaryot Cell* **8:** 1321–1329.

Lenstra TL, Coulon A, Chow CC, Larson DR. 2015. Single-molecule imaging reveals a switch between spurious and functional ncRNA transcription. *Mol Cell* **60:** 597–610.

Liu P, Xiang Y, Fujinaga K, Bartholomeeusen K, Nilson KA, Price DH, Peterlin BM. 2014. Release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear

ribonucleoprotein (snRNP) activates hexamethylene bisacetamide-inducible protein (HEXIM1) transcription. *J Biol Chem* **289:** 9918–9925.

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

Lukoszek R, Mueller-Roeber B, Ignatova Z. 2013. Interplay between polymerase II- and polymerase III-assisted expression of overlapping genes. *FEBS Lett* **587:** 3692–3695.

Moqtaderi Z, Struhl K. 2004. Genome-wide occupancy profile of the RNA polymerase III machinery in *Saccharomyces cerevisiae* reveals loci with incomplete transcription complexes. *Mol Cell Biol* **24:** 4118–4127.

Moqtaderi Z, Wang J, Raha D, White RJ, Snyder M, Weng Z, Struhl K. 2010. Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat Struct Mol Biol* **17:** 635–640.

Nadal-Ribelles M, Sole C, Xu Z, Steinmetz LM, de Nadal E, Posas F. 2014. Control of Cdc28 CDK1 by a stress-induced lncRNA. *Mol Cell* **53:** 549–561.

Oler AJ, Alla RK, Roberts DN, Wong A, Hollenhorst PC, Chandler KJ, Cassiday PA, Nelson CA, Hagedorn CH, Graves BJ, et al. 2010. Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat Struct Mol Biol* **17:** 620–628.

Orioli A, Praz V, Lhote P, Hernandez N. 2016. Human MAF1 targets and represses active RNA polymerase III genes by preventing recruitment rather than inducing long-term transcriptional arrest. *Genome Res* **26:** 624–635.

Pelechano V, Steinmetz LM. 2013. Gene regulation by antisense transcription. *Nat Rev Genet* **14:** 880–893.

Raha D, Wang Z, Moqtaderi Z, Wu L, Zhong G, Gerstein M, Struhl K, Snyder M. 2010. Close association of RNA polymerase II and many transcription factors with Pol III genes. *Proc Natl Acad Sci* **107:** 3639–3644.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR–Cas9 system. *Nat Protoc* **8:** 2281–2308.

Renaud M, Praz V, Vieu E, Florens L, Washburn MP, l'Hote P, Hernandez N. 2014. Gene duplication and neofunctionalization: POLR3G and POLR3GL. *Genome Res* **24:** 37–51.

Roberts DN, Stewart AJ, Huff JT, Cairns BR. 2003. The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proc Natl Acad Sci* **100:** 14695–14700.

Roberts DN, Wilson B, Huff JT, Stewart AJ, Cairns BR. 2006. Dephosphorylation and genome-wide association of Maf1 with Pol III-transcribed genes during repression. *Mol Cell* **22:** 633–644.

Saldana-Meyer R, Gonzalez-Buendia E, Guerrero G, Narendra V, Bonasio R, Recillas-Targa F, Reinberg D. 2014. CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev* **28:** 723–734.

Santillan BA, Moye C, Mittelman D, Wilson JH. 2014. GFP-based fluorescence assay for CAG repeat instability in cultured human cells. *PLoS One* **9:** e113952.

Shearwin KE, Callen BP, Egan JB. 2005. Transcriptional interference–a crash course. *Trends Genet* **21:** 339–345.

Smit AF, Riggs AD. 1995. MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. *Nucleic Acids Res* **23:** 98–102.

Soragni E, Kassavetis GA. 2008. Absolute gene occupancies by RNA polymerase III, TFIIIB, and TFIIIC in *Saccharomyces cerevisiae*. *J Biol Chem* **283:** 26568–26576.

Xu Z, Wei W, Gagneur J, Clauder-Munster S, Smolik M, Huber W, Steinmetz LM. 2011. Antisense expression increases gene expression variability and locus interdependency. *Mol Syst Biol* **7:** 468.

Zhou Q, Li T, Price DH. 2012. RNA polymerase II elongation control. *Annu Rev Biochem* **81:** 119–143.

Supplementary data

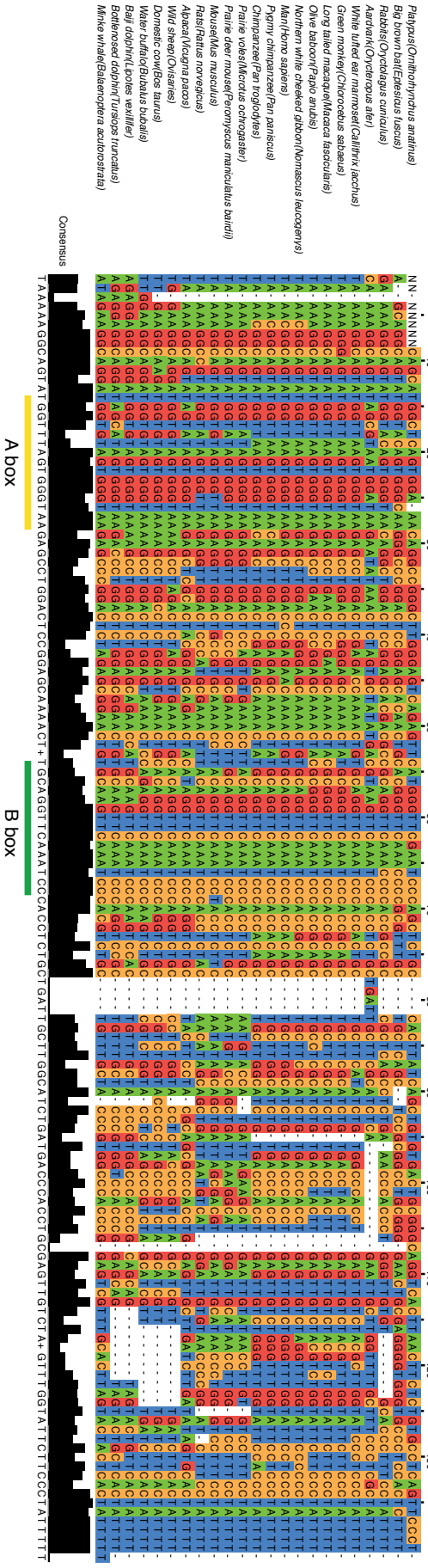


Figure S1. Alignment of the MIR-like sequence in the first intron of the *Polr3e* gene in different mammalian species. The alignment was performed with ClustalW with the default settings for nucleotide sequences; the graphic representation was generated with Jalview. The A and B boxes are indicated.

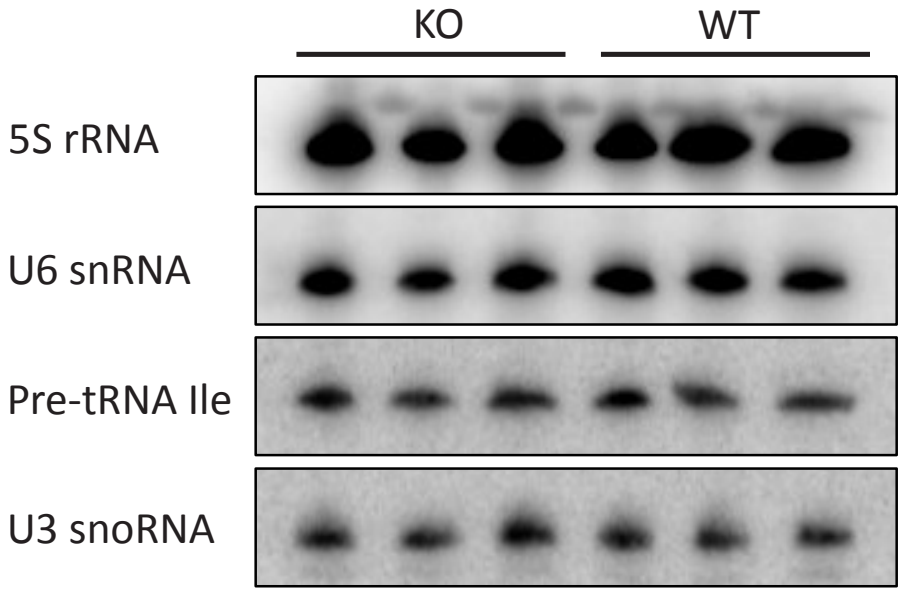


Figure S2. Northern blot for Pol III transcribed 5S rRNA, U6 snRNA, and pre-tRNA Ile in MIR KO and WT ES cells. The Pol II-transcribed U3 snoRNA was used as an internal control for RNA loading.

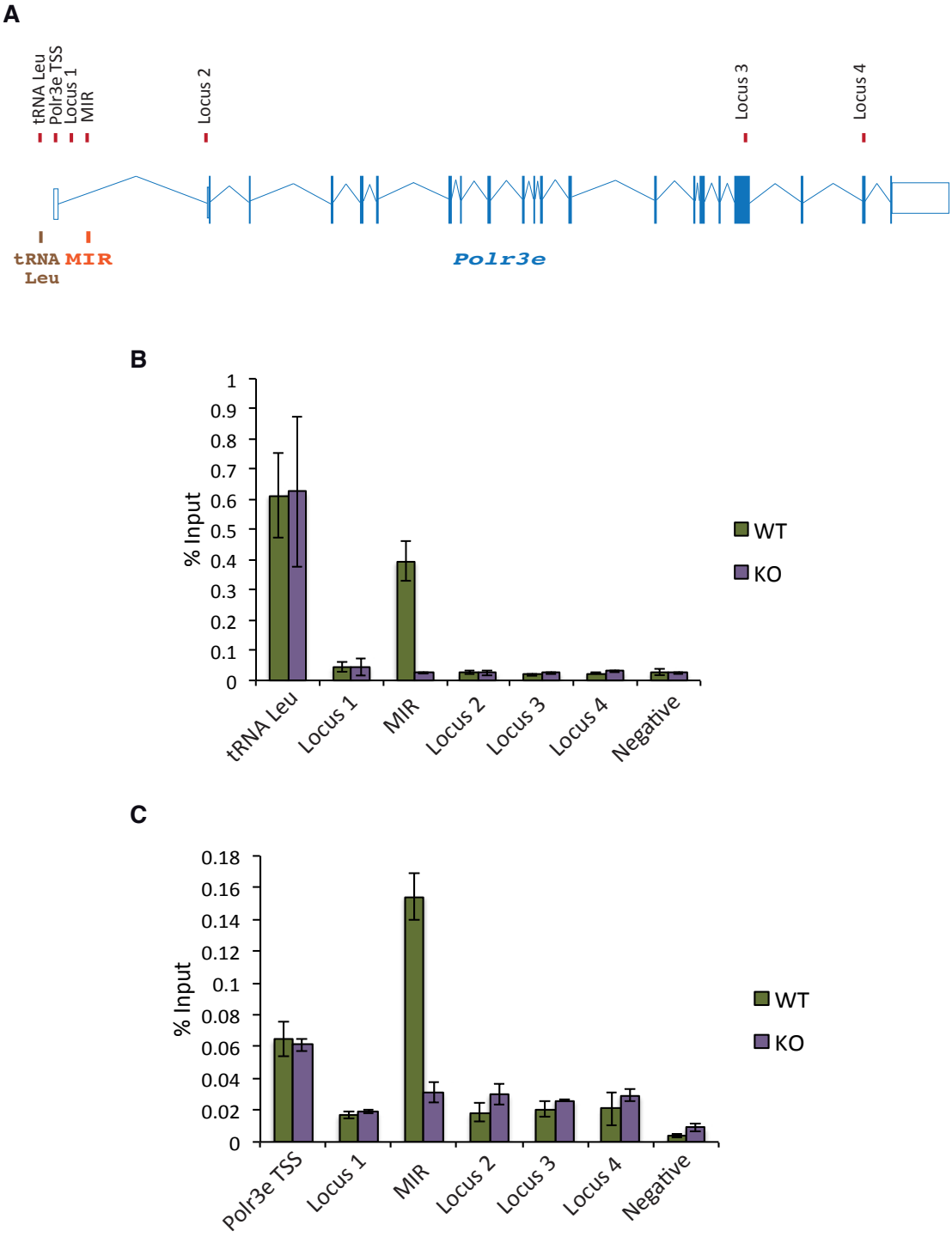
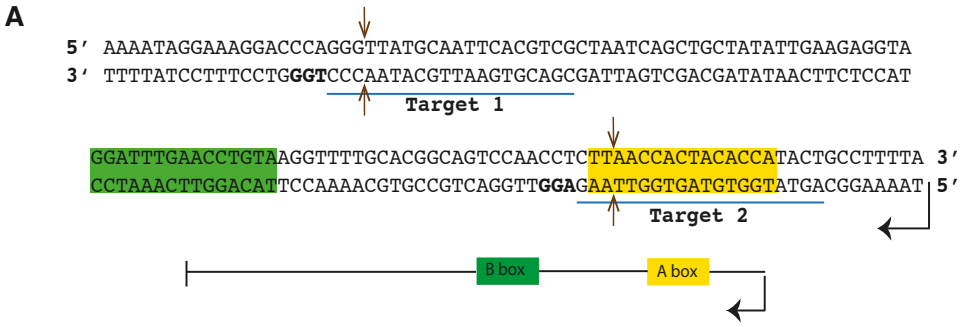
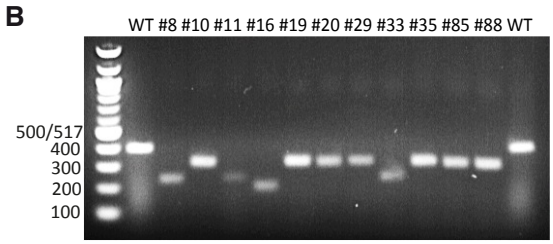


Figure S3. Pol III and Pol II occupancy sampling over the *Polr3e* gene. **A)** Position of primers used for qPCR after ChIP. The exon-intron structure of the *Polr3e* gene is redrawn from the Ensembl genome browser. **B)** ChIP-qPCR performed with an anti-RPC4 antibody. **C)** ChIP-qPCR performed with an anti-RPB2 antibody.





C

```
KO 8,11,33 CTAGGTTCGCCTTATCCCTTTGAAGCCGGTGTGTGGTTTAAAATAGGAAAGGACC----------
WT         CTAGGTTCGCCTTATCCCTTTGAAGCCGGTGTGTGGTTTAAAATAGGAAAGGACCCAGGGTTATGCAATTCACGTCGCTAATCAGCTGCTATATTGAAGA

KO 8,11,33 ----------
WT         GGTAGGATTTGAACCTGTAAGGTTTTGCACGGCAGTCCAACCTCTTAACCACTACACCATACTGCCTTTTACCAAAGAGATGTGCAGCATGTTAGTTAAG

KO 8,11,33 ----------TTGGCATCAGCTAGATTTGGATTTCTTTACCTGTTTTCTTGGGGGGCTGTGACCTTGGGTAAACCAATTCGAAGTT
WT         AATAAAACCAAGGCTCTTGGCATCAGCTAGATTTGGATTTCTTTACCTGTTTTCTTGGGGGGCTGTGACCTTGGGTAAACCAATTCGAAGTT
```

```
KO 16      CTAGGTTCGCCTTATCCCTTTGAAGCCGGTGTGTGGTTTAAAATAGGAAAGGACCCAGGG----------
WT         CTAGGTTCGCCTTATCCCTTTGAAGCCGGTGTGTGGTTTAAAATAGGAAAGGACCCAGGGTTATGCAATTCACGTCGCTAATCAGCTGCTATATTGAAGA

KO 16      ----------
WT         GGTAGGATTTGAACCTGTAAGGTTTTGCACGGCAGTCCAACCTCTTAACCACTACACCATACTGCCTTTTACCAAAGAGATGTGCAGCATGTTAGTTAAG

KO 16      ----------GCTGTGACCTTGGGTAAACCAATTCGAAGTT
WT         AATAAAACCAAGGCTCTTGGCATCAGCTAGATTTGGATTTCTTTACCTGTTTTCTTGGGGGGCTGTGACCTTGGGTAAACCAATTCGAAGTT
```

```
KO 29      CTAGGTTCGCCTTATCCCTTTGAAGCCGGTGTGTGGTTTAAAATAGGAAAGGACCCAGGG----------
WT         CTAGGTTCGCCTTATCCCTTTGAAGCCGGTGTGTGGTTTAAAATAGGAAAGGACCCAGGGTTATGCAATTCACGTCGCTAATCAGCTGCTATATTGAAGA

KO 29      ----------AACCACTACACCATACTGCCTTTTACCAAAGAGATGTGCAGCATGTTAGTTAAG
WT         GGTAGGATTTGAACCTGTAAGGTTTTGCACGGCAGTCCAACCTCTTAACCACTACACCATACTGCCTTTTACCAAAGAGATGTGCAGCATGTTAGTTAAG

KO 29      AATAAAACCAAGGCTCTTGGCATCAGCTAGATTTGGATTTCTTTACCTGTTTTCTTGGGGGGCTGTGACCTTGGGTAAACCAATTCGAAGTT
WT         AATAAAACCAAGGCTCTTGGCATCAGCTAGATTTGGATTTCTTTACCTGTTTTCTTGGGGGGCTGTGACCTTGGGTAAACCAATTCGAAGTT
```

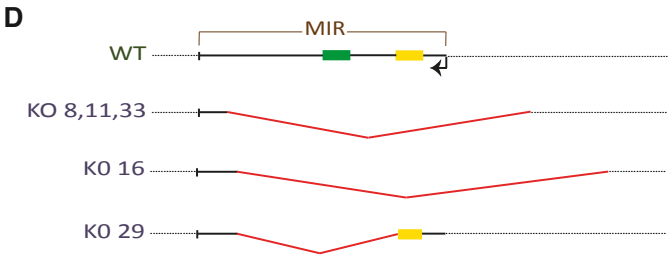


Figure S4. CRISPR/Cas9-mediated deletion of the MIR in mouse ES cells. **A)** sgRNA design for deletion of the MIR. The MIR sequence, the A and B boxes, and the target sequences are shown. The PAM sequences are shown in bold, and the predicted double strand break sites are indicated by double arrows. **B)** Products of PCRs performed on genomic DNA from several clones with a pair of primers flanking the MIR, separated on an agarose gel. **C)** Characterization of the deletions in different MIR KO clones. The sequencing results of the PCR products obtained from several clones are shown aligned to the WT sequence. The A (yellow) and B (green) boxes are highlighted; the MIR sequence is shown in bold. **D)** Schematic view of the deletions in the different clones. The MIR A (yellow) and B (green) boxes, TSS (arrow), and termination site (end of black line) are shown.

Table S1. List of ENCODE Pol III peaks located inside human Pol II genes. The table lists all ENCODE RPC1 peaks located inside Pol II transcription units. Column A gives links to a UCSC genome browser session showing RPC4, RPC1, ENCODE RPC1, Pol II, DSIF, and NELF tracks. The pol II, DSIF, and NELF tracks are from Liu et al., 2014. Column D indicates the direction of the Pol III transcription unit relative to the genome. Column G refers to scores (columns R, S, and T) calculated as in Canella et al., 2012, for the regions indicated (columns P and Q), for anti-RPC4 ChIP-seq data obtained from IMR90Tert cells under serum replete conditions (Orioli et al. 2016). Column H refers to scores (columns U, V, W) calculated as in Canella et al., 2012, for the regions indicated (columns P and Q), for the ENCODE anti-RPC1 ChIP-seq data obtained from HeLa cells. Column I refers to scores (column X), calculated as in ENCODE, for the regions indicated (columns P and Q). Y stands for yes, N for no. Column L indicates the direction of the Pol II gene relative to the genome.

"Table S1 is available at http://genesdev.cshlp.org/content/31/4/413/suppl/DC1"

RT-qPCR	
Oligonucleotide name	Sequence (5’ to 3’)
MIR- Fwd	GACTGCCGTGCAAAACCTTA
MIR- Rev	TTATGCAATTCACGTCGCTAA
Total *Polr3e*- Fwd	AGAGACAGTTTGTGCTCACG
Total *Polr3e*- Rev	TGAGACGCCACTGAAGAGTA
Pre-*Polr3e*-Fwd	CTGTTGGCTGCTACTGAACA
Pre-*Polr3e*-Rev	CTCTTCGCAGTGCTTTGACT
Mature *Polr3e*- Fwd	CACATATTCCTCAAAGCTGATGG
Mature *Polr3e*- Rev	GATGCCATGTAAAGGTGTCAGG
Actb- Fwd	CTAAGGCCAACCGTGAAAAGAT
Actb- Rev	CACAGCCTGGATGGCTACGT
Gapdh- Fwd	AGGTCGGTGTGAACGGATTTG
Gapdh- Rev	TGTAGACCATGTAGTTGAGGTCA

Northern blot	
Oligonucleotide name	Sequence (5’ to 3’)
MIR	TTATGCAATTCACGTCGCTAA
U87 snoRNA	TCACACCCATGACTGCCACT
5S rRNA	TTAGCTTCCGAGATCAGACG
U6 snRNA	CACGAATTTGCGTGTCATCC
Pre-tRNA Ile	ATCGCTTACGCCTAGCACTG
U3 snoRNA	GGAGGGAAGAACGATCATCA

ChIP-qPCR	
Oligonucleotide name	Sequence (5’ to 3’)
MIR- Fwd	TTCGCCTTATCCCTTTGAAG
MIR- Rev (WT)	CAGTATGGTGTAGTGGTTAAG
MIR- Rev (KO11)	TACTCTTCTCATTAGCTGTGC
chr7-tRNA Leu- Fwd	TTAGAAAACGACGTCAACAGC
chr7-tRNA Leu- Rev	GACAAAAGAAAAAGCCTGCCT
Polr3e TSS- Fwd	CATTGTGGGTAAGAGGAAGC
Polr3e TSS- Rev	TATCAGGCAGCGGCCATGTTC
Mycbp- Fwd	ACTCGAAGCGCGAGCAGT
Mycbp- Rev	CTCACCTTTCGTCAGCGTGT
chr13- tRNA Leu- Fwd	AGGTTCACGGAAGGTCTG
chr13- tRNA Leu- Rev	CTATGGCTTCCTCGCTCTG
Locus 1- Fwd	GCTTTCGGAAGAGTGGGAAG
Locus 1- Rev	AGAGTTGACCAGGTTCAACG
Locus 2- Fwd	CACAGACACAGAAAGGAGACC
Locus 2- Rev	ACCTCACAGTCCTCAACTCG
Locus 3- Fwd	AGAGACAGTTTGTGCTCACG
Locus 3- Rev	TGAGACGCCACTGAAGAGTA
Locus 4- Fwd	CATCGGCAGGTTTTGCTTG
Locus 4- Rev	CTTTAGCACTTTATCCACCTCC

Table S2. Sequences of oligonucleotides used in RT-qPCR, ChIP-qPCR, and northern blots.

Regulation of the MIR in differentiation

We showed that the MIR located in the first intron of the *Polr3e* gene is highly occupied by Pol III in different cells and tissues, and that this reduces Pol II transcription of the *Polr3e* gene. An interesting question is whether differential occupation of the Pol III gene under different conditions might result in differential expression of the *Polr3e* gene. To address this question, we first tried to increase Pol III transcription of the MIR by knocking down the Pol III repressor Maf1 in ES cells, in order to see whether this would reduce *Polr3e* mRNA level. However, Maf1 knockdown did not increase MIR RNA levels in ES cells, suggesting that MIR Pol III transcription is already maximal (data not shown). We then took advantage of the ES cells capability for differentiation to examine MIR, and *Polr3e,* regulation. To induce exit from the stem cell state and differentiation, we kept the cells in a medium lacking leukemia inhibitory factor (LIF). LIF is a cytokine whose usage in mouse ES cells medium stimulates their self-renewal and inhibits their differentiation. In the differentiated cells as compared to ES cells, we observed an up-regulation of the MIR, while the *Polr3e* level was slightly decreased (Figure 2-2).

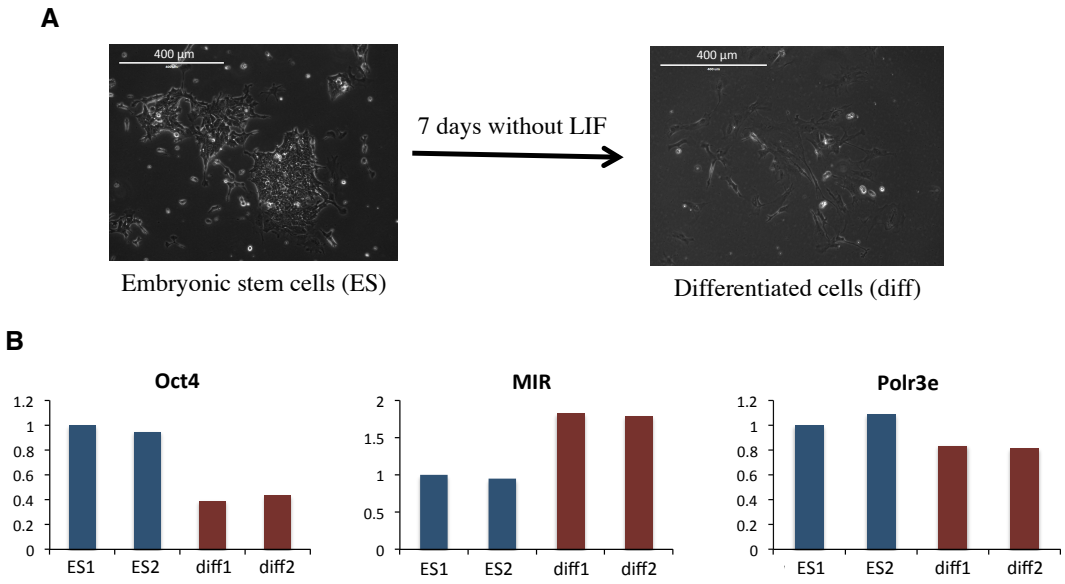


Figure 2-2. Regulation of the MIR in differentiation. **A)** Light microscopy of mouse ES cells, and differentiated cells. **B)** qPCR results for pluripotency factor *Oct4*, MIR, and *Polr3e*, indicated as relative to ES1. The results were normalized to *Gapdh* mRNA level.

Given these results, we decided to induce a more precise, lineage-specific differentiation of the ES cells. We induced neural differentiation by producing embryoid bodies in a suspension medium, followed by seeding the cells in a neural differentiation-inducing medium. Neural differentiation was confirmed by microscopy, immunostaining, and qPCR of mRNAs coding for pluripotency and neural markers. We showed that there is an increase in the level of MIR RNA in differentiated cells versus ES cells, accompanied by a decrease in *Polr3e* mRNA levels (Figure 2-3). These results indicate that transcription of the MIR can be regulated under different conditions, and suggest that this impacts directly on Pol II transcription of the *Polr3e* gene.

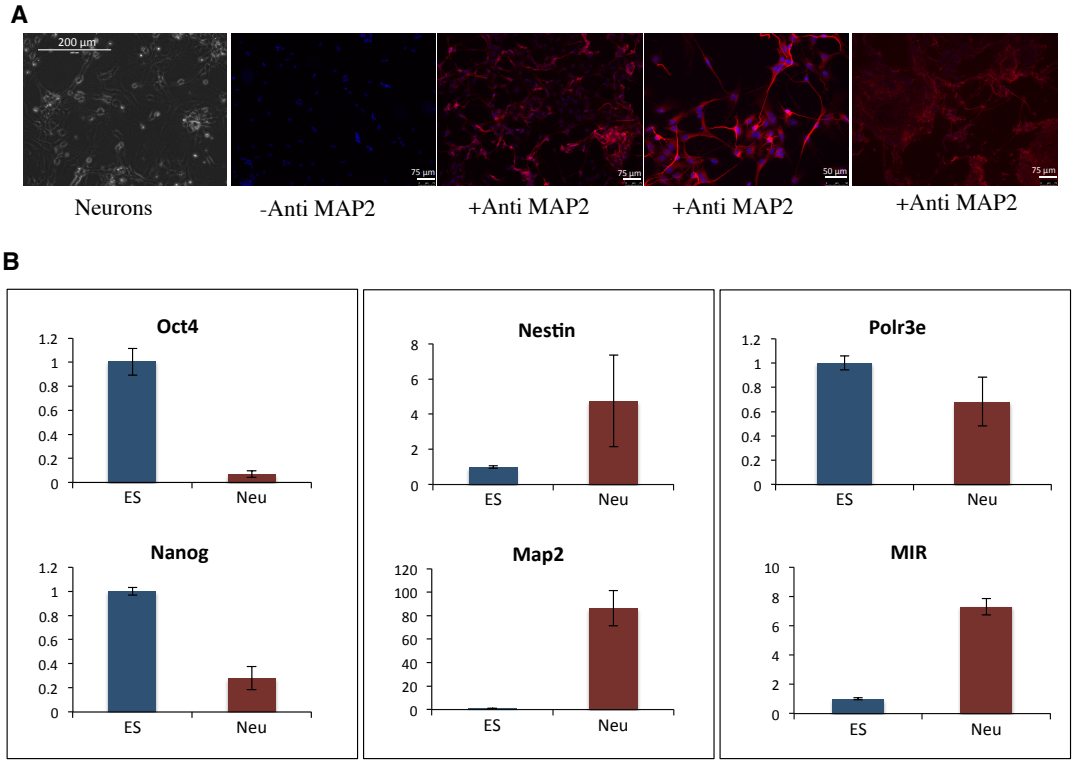


Figure 2-3. Regulation of the MIR in neural differentiation. **A)** Light microscopy and anti-MAP2 staining of the differentiated neural cells. **B)** qPCR results for pluripotency factors *Oct4* and *Nanog*, neural markers *Nestin* and *Map2*, *Polr3e* and MIR, indicated as relative to ES. The results were normalized to *Gapdh* mRNA level. Error bars represent ± SD and n=3.

Chapter III – RNA polymerase III regulation in liver regeneration

Summary

Pol III is highly regulated during development, differentiation, and proliferation, probably to adapt the cell to any changes requiring either quantitative or qualitative changes in its translational program. Dynamic changes of Pol III activity have been studied in several cell lines, but very rarely in normal tissues. Furthermore, although it has been shown that many active Pol III genes reside close to Pol II genes and H3K4me3 chromatin mark, it was not shown whether the dynamics of Pol III transcription is coordinated with that of Pol II transcription and H3K4me3 occupancy. In this study, we have used a partial hepatectomy (PH) system followed by Pol III, H3K4me3, and Pol II ChIP-seq in mouse liver to study the dynamics of Pol III occupancy when the liver enters regeneration and hepatocytes start dividing, and to determine whether Pol III occupancy changes are correlated with changes in surrounding H3K4me3 and Pol II occupancy (Figure 3-1).

First, we showed that before PH, tRNA genes that have surrounding H3K4me3 and Pol II peaks have a generally higher Pol III occupancy score as compared with those with Pol III peak only. For SINEs however, these scores were not much different. For both tRNA genes and SINEs, the surrounding Pol II peaks generally corresponded to the TSS of a Pol II gene.

Next, we examined the dynamics of Pol III after PH. We showed that Pol III occupancy is generally higher in regenerating liver cells 36 hours after PH (TP36) versus in resting cells (TP0), an effect that was most pronounced for lowly occupied loci. We identified some genes not occupied by Pol III at TP0 and becoming occupied at TP36. This increase in Pol III occupancy was accompanied by an increase in levels of several pre-tRNAs and mRNAs coding for Pol III subunits and transcription factors. The increase in Pol III occupancy on tRNA genes classified according to isoacceptors and isotypes was very general, with no clear specific effect on one or the other isoacceptor or isotype class.

Last, we investigated how different Pol III genes respond differentially to liver regeneration. We showed that tRNA genes without surrounding Pol II and H3K4me3 peaks, which tend to be lowly occupied at TP0, are those that changed the most with respect to Pol III occupancy after PH. On the other hand, tRNA genes with surrounding H3K4me3 and Pol II peaks generally changed less with respect to Pol III occupancy. Consistent with this observation, only a very low percentage of tRNA genes that changed significantly from TP0 to TP36 had surrounding Pol II peaks. When we looked at Pol III occupancy changes by chromosome, we observed a chromosomal bias with chromosome 13 having a disproportionately large number of changing genes. We showed that there are a few tRNA gene clusters on chromosome 13 and other chromosomes with the genes in the cluster regulated more similarly to each other than to genes outside the clusters on the same chromosome.

In this work, I analyzed and interpreted data, did the qPCR experiments, wrote a manuscript draft and contributed to revised versions (Yeganeh et al. 2018).

Figure 3-1. Schematic view of the project summary. In this project, the main aim was to compare Pol III occupancy levels in Pol III genes close to, or far from, Pol II peaks, and to determine whether these categories of genes were regulated differentially during liver regeneration.

Nucleic Acids Research, 2018 **1**
doi: 10.1093/nar/gky1282

Differential regulation of RNA polymerase III genes during liver regeneration

Meghdad Yeganeh[1,†], Viviane Praz[1,2,†], Cristian Carmeli[1,3,†], Dominic Villeneuve[1], Leonor Rib[1], Nicolas Guex[2], Winship Herr[1], Mauro Delorenzi[3,4], Nouria Hernandez[1,*] and the CycliX consortium

[1]Center for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne, 1015 Lausanne, Switzerland, [2]SIB Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland, [3]Bioinformatics Core Facility, SIB Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland and [4]Department of Fundamental Oncology and the Ludwig Center for Cancer research, Faculty of Biology and Medicine, University of Lausanne, 1015 Lausanne, Switzerland

Received September 19, 2018; Revised November 22, 2018; Editorial Decision December 12, 2018; Accepted December 14, 2018

ABSTRACT

Mouse liver regeneration after partial hepatectomy involves cells in the remaining tissue synchronously entering the cell division cycle. We have used this system and H3K4me3, Pol II and Pol III profiling to characterize adaptations in Pol III transcription. Our results broadly define a class of genes close to H3K4me3 and Pol II peaks, whose Pol III occupancy is high and stable, and another class, distant from Pol II peaks, whose Pol III occupancy strongly increases after partial hepatectomy. Pol III regulation in the liver thus entails both highly expressed housekeeping genes and genes whose expression can adapt to increased demand.

INTRODUCTION

Compared to RNA polymerase (Pol) II promoters, Pol III promoters are quite simple with just three main types of structures. Type 1 promoters occur only in the 5S genes, type 2, by far the most abundant, are present in tRNA genes, most SINEs, and some other genes, and type 3 are present in less than fifteen annotated genes in both the human and mouse genomes (1,2). Despite this relative uniformity, different annotated Pol III genes have very different levels of Pol III occupancy, which correlate with different levels of transcriptional activity (3–5). In fact, one of the surprises of early genomic studies was the discovery that some 40–50% of annotated Pol III genes are not occupied by Pol III and transcriptionally silent, an observation that is only partially explained by poor promoter sequences (2–4,6–9).

Expressed Pol III loci differ from silent ones by the nearby presence of histone marks such as H3K4me3 (2,6–8,10) typical of chromatin regions that are or have been actively transcribed by Pol II ((11) and references therein). Moreover, active Pol III loci tend to reside close to Pol II TSSs and to peaks of Pol II occupancy, which suggests that transcription of nearby Pol II and Pol III genes is somehow co-regulated (2,5–8,10). Pol III transcription is indeed tightly regulated, allowing the cell to adapt to changing needs in biosynthetic capacity resulting from, for example, cell growth and proliferation. Furthermore, overexpression of Pol III genes is observed in many transformed cells (12–17). Similarly, genome-wide Pol III occupancy comparisons of mouse hepatocarcinoma cells with normal mouse liver cells (18), or of precursors with induced pluripotent cells and human H1 ES cells (10), all point to higher Pol III transcription in dividing as compared to differentiated cells.

Genome-wide Pol III occupancy and transcription have been studied in only a few dynamic systems, and very rarely in a normal tissue. Studies comparing Pol III occupancy in human serum-starved versus serum-replete IMR90Tert cells (4), in the mouse liver at different times during the circadian cycle (19), in mouse liver and brain at different stages of development (20), or in THP-1 cells and THP-1-derived macrophages by PMA treatment (5) have all emphasized that different Pol III genes respond differently to changing cellular conditions. In the last case, concerted down-

*To whom correspondence should be addressed. Tel: +41216922010; Fax: +41216923905; Email: nouria.hernandez@unil.ch
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.
Present addresses:
Cristian Carmeli, Institute of Social and Preventive Medicine, University Hospital of Lausanne (CHUV), 1015 Lausanne, Switzerland.
Leonor Rib, Bioinformatics Center, Department of Biology & Biotech Research and Innovation Center University of Copenhagen, Copenhagen, Denmark.

regulation of certain tRNA genes in clusters and contact domains was observed.

Here, we have taken advantage of the synchronous hepatocyte proliferation occurring after partial hepatectomy (PH) to examine, in a normal tissue, the dynamics of Pol III occupancy upon transition from a resting G0 state to a proliferating state. We find two classes of active Pol III genes, one class with high and relatively static Pol III occupancy, often characterized by proximity to Pol II TSSs and Pol II peaks, and a second class with much lower Pol III occupancy, devoid of nearby Pol II peaks, but highly dynamic. The resulting picture is one where a network of Pol III genes, often located close to Pol II TSSs, ensures steady production of essential Pol III RNA products in the differentiated tissue, whereas another, expressed at low levels in the differentiated tissue, ensures the increased synthesis of Pol III products needed in preparation for cell division.

MATERIALS AND METHODS

Animals, partial hepatectomies, and chromatin immunoprecipitations (ChIP)

C57/BL6 12–14-week-old male mice were housed under a 12 h light/12 h dark cycle regimen for two weeks with food available during the night. Two-third partial hepatectomies were performed as described (21–23). Three pools of three mice were processed in one batch between ZT01.5 and ZT02.5, with three mice operated every 20 minutes. The livers of the three mice were pooled for each timepoint. ChIPs were performed as described (24). The following antibodies were used: anti-RPC4 (CS681) (2), anti-H3K4me3 (Abcam, ab8580) and anti-RPB2 (Santa Cruz Biotechnology, sc-673-18). It should be noted that the anti-H3K4me3 antibody used scored as the best ENCODE-validated anti-H3K4me3 antibody but is 60–66% specific for H3K4me3, with crossreaction to H3K4me2 and very weak crossreaction with H3K4me1 (25).

Ultra-high-throughput sequencing and tag alignment

Ten nanogram of immunoprecipitated chromatin was used to prepare sequencing libraries with the Diagenode MicroPlex Library Preparation kit (catalog no C05010011) as specified by the manufacturer, with a total of 14 amplification cycles. One or several bar-coded sequencing libraries were then loaded onto one lane of a HiSeq 2000 flow cell and paired-end sequenced at 50 or 100 cycles. For each condition, we sequenced both input chromatin and the corresponding ChIP samples.

The first fifty nucleotides of each sequence were mapped onto the UCSC mouse genome version NCBI37/mm9 via the eland_extended mode of ELAND v2e in the Illumina CASSAVA pipeline v1.8.2. We first retrieved fragments with unique matches at both ends. The tags with multiple matches on the mouse genome were then aligned via fetchGWI with an allowed maximum of 500 matches per tag. When tag alignment revealed multiple possible fragments for a tag pair, all possible fragments were kept and given a weight corresponding to the normalized size probability as determined from the size probability distribution of the unique fragments. Only fragments sequenced once (non-redundant fragments) were considered in the analysis.

Score computation

We first normalized the number of fragments aligned onto the genome in each sample relative to the median number of fragments aligned onto the genome across all samples (all time points). Scores were computed as the $\log_2$-ratio between counts in paired ChIPs and Input samples to which a pseudocount value of 16 was added. For RPC4 ChIP scores, the counts were assigned to a previously defined list of mouse annotated Pol III genes and Pol III occupied loci (see (18)), with each annotated RNA-coding region extended by 150 bp on each side. As shown in Supplementary Figure S1, the correlation between the RPC4 scores calculated from paired-end sequencing (this work) and single-end sequencing (18) was very high ($R = 0.962$). For H3K4me3 and RPB2, the regions considered extended 1000 bp upstream and downstream of the Pol III loci TSSs. One fragment was worth one count, and fractional counts were attributed in case of a partial overlap between a fragment and an extended locus.

To establish whether the score value at a particular locus was significantly higher than in other regions of the genome, the score was compared to the distribution of scores computed in non-overlapping bins (400 bp for RPC4 and 2000 bp for H3K4me3 and RPB2) over the entire genome. The threshold value was obtained applying a Bonferroni correction for the multiplicity of loci. The calculated scores for all samples are listed in Supplementary Table S1.

Data quality

To monitor sample quality, we produced mean-difference scatter plots of mouse genome bin counts and Pol III loci (see (26)). Samples displaying well-differentiated RPC4 occupancy scores on Pol III loci compared to scores on the rest of the genome were considered of good quality. 134 tRNA genes out of 433 (31%) had scores below the cutoff in all replicates and all time points. Supplementary Figures S2A and B show the score reproducibility for biological replicates obtained at TP0, TP36, TP48, and TP60, for RPC4 and RPB2 occupancy, respectively. The scores from replicates are, in some cases, slightly shifted relative to the X = Y line, but the correlation coefficients are always larger than 0.9.

Statistical modeling

To investigate dynamical changes of RPC4 scores we considered a general linear model with four levels (corresponding to the four time points, i.e. TP0, TP36, TP48 and TP60) and two replicates per level and per ChIP-seq. The model was estimated through empirical Bayes method implemented in the Limma package (27). That method allows a robust estimation of the variance ensuing from the variance across genes. For each contrast of interest (e.g. TP36 versus TP0) a *P*-value and a T statistic were computed for each locus. To take into account the multiplicity of null hypotheses, false discovery rates (FDR) were estimated from

the *P*-values (28). In particular, we estimated the proportion of null *P*-values through a smoothing spline approach (29) to increase power. Statistical significance was called for genes showing a FDR <0.05.

Gene expression analysis

Total RNA was extracted using miRNeasy mini kit (Qiagen) with on-column DNase I treatment. One microgram of RNA was used for cDNA synthesis with random hexamers and M-MLV Reverse Transcriptase (Promega). qPCRs were performed with appropriate primers (see Supplementary Table S2) and SensiFAST SYBR kit (Bioline) under the following conditions: an initial denaturing step (95°C 10′) followed by 40 cycles of incubations at 95°C for 15″, 58°C for 20″, and 72°C for 20″. qPCR reactions were performed in triplicate, with reactions lacking reverse transcriptase serving as negative controls. Melting curve analyses revealed single peaks for all reactions, and calculation of primers efficiencies using standard curves indicated comparably high efficiencies (from 86 to 102%) for the different primer sets. The relative expressions were analyzed with the $2^{\Delta\Delta Ct}$ method.

RESULTS AND DISCUSSION

Association of RPC4 occupancy with H3K4me3 and RPB2 occupancy at TP0

In our experiments, mice were kept in 12 h light/12 h dark cycle, with food available only during the night. PH was performed always at the very beginning of the day (ZT01.5-ZT02.5). We profiled H3K4me3, Pol II (RPB2 subunit), and Pol III (RPC4 subunit) occupancy in parallel at different time points after PH by chromatin immunoprecipitation followed by high throughput sequencing (ChIP-seq). The Pol II and H3K4me3 occupancy data have been used to study occupancy dynamics at Pol II genes (11). Here, we used the same data as well as the Pol III data to determine occupancy scores at Pol III genes, as described before (18) (Supplementary Table S1).

We first explored the relationship between RPC4 occupancy and presence of H3K4me3 and RPB2 peaks at TP0, before PH. To avoid confusing signals from overlapping peaks at closely spaced loci, we considered only 'isolated' loci, i.e., Pol III loci separated by at least 1.5 kb, unless otherwise mentioned (437 out of 646 loci, Supplementary Table S3, column D). We classified these loci into eight groups, i.e., loci with a peak for each of the three factors tested, with peaks for two of them, with a peak for only one of them, and with no peaks (Table 1). The Rn5S and the Rn4.5S genes, which are each tandemly repeated in the genome, displayed only RPC4 peaks, and both RPC4 and H3K4me3 peaks, respectively (Table 1 and Supplementary Table S3). We found SINEs in all groups, whereas nearly all 'other' Pol III genes, corresponding mostly to genes with type 3 promoters (see Supplementary Table S3, genes labeled 'other', for a list), had peaks for RPC4+H3K4me3+RPB2 (Table 1).

For tRNA genes, the largest group was the one with no peaks (Table 1). Considering just Pol III occupancy, ∼56% of isolated tRNA genes had no RPC4 peaks against ∼45% of all tRNA genes (see Supplementary Table S3), indicating that the proportion of silent tRNA genes is higher among isolated genes than among all tRNA genes. This is in line with the finding that in human cells, the median expression level for tRNA genes increases with the number of neighboring tRNA genes (5). The second and third largest groups of isolated tRNA genes were the ones with both RPC4+H3K4me3 peaks, and with RPC4 + H3K4me3 + RPB2 peaks, respectively. The group with RPC4 peaks only comprised about 10% of isolated tRNA genes, and the latter proportion was similar for all tRNA genes (Table 1 and Supplementary Table S3).

We then investigated whether the RPC4 occupancy scores differed across groups and types of genes. We focused on the groups with RPC4 peaks either i) alone or ii) with H3K4me3 peaks or iii) with both H3K4me3 and RPB2 peaks. The average RPC4 tag density profile and the individual gene RPC4 occupancy scores were lowest in the first group and much higher in the groups with H3K4me3 peaks or H3K4me3 + RPB2 peaks for isolated tRNA genes (Figures 1A and B), an effect not seen for isolated SINEs (Supplementary Figures S3A and B). RPC4 scores of tRNA genes with RPC4 + H3K4me3 peaks, or RPC4 + H3K4me3 + RPB2 peaks, were similar to each other but clearly higher than scores of tRNA genes with only RPC4 peaks (Figure 1B). The SINEs in the three groups had very similar scores (Supplementary Figure S3B). Comparing tRNA gene and SINE scores in each of the three groups revealed no significant difference for loci with only RPC4 (*P*-value = 0.1162 and 0.0126 for replicates 1 and 2) but much higher scores for tRNA genes with RPC4 + H3K4me3 peaks, or RPC4 + H3K4me3 + RPB2 peaks (*P*-value < 2e–4).

The average tag density RPC4 profiles revealed low peaks of RPC4 in RPC4 alone group (Figure 1A and Supplementary Figure S3A, see also Supplementary Figure S4A). H3K4me3 profiles for genes with RPC4+H3K4me3 peaks revealed a prominent H3K4me3 peak upstream of the TSS, and for tRNA genes a smaller one downstream, reflecting the aggregate of individual tRNA genes with only an upstream peak, only a downstream peak, or both (see Supplementary Figure S4B). In the case of tRNA genes and SINEs in the third group, the profiles of accumulated tags were very similar, with major upstream and minor downstream peaks (Figure 1A, Supplementary Figures S3A and S4C). This directionality, with the Pol II peak generally upstream of the Pol III peak, is consistent with previous observations in cultured human cells (6,7,10,30).

CpG islands (within 1 kb around the TSS) (red dots in Figure 1B and Supplementary Figure S3B, Supplementary Table S3, column F) were strongly associated with the presence of RPB2 peaks and depleted in tRNA genes or SINEs with RPC4 peaks only. Thus, at TP0, isolated SINEs tend to have low Pol III occupancy regardless of the presence of H3K4me3 or RPB2 peaks, whereas isolated tRNA genes can have high Pol III occupancy scores, which are generally associated with the presence of nearby H3K4me3 and RPB2 peaks.

The general proximity of active Pol III transcription units to H3K4me3 and Pol II peaks, and to Pol II TSSs, has been noted before (2,5–8,10,30), but the relationship among these Pol II peaks, H3K4me3 peaks, and Pol II TSSs is not clear. Separation of isolated Pol III genes into two groups,

Table 1. Pol III loci grouped according to presence or absence of RPC4, H3K4me3 and RPB2 peaks

		Pol III	Pol III	Pol III	——-	Pol III	——-	——-	——-
		Pol II	Pol II	——-	Pol II	——-	Pol II	——-	——-
		H3K4me3	——-	H3K4me3	H3K4me3	——-	——-	H3K4me3	——-
tRNA genes (233)	No.	37/29	0/0	41/49	2/2	24/22	3/1	1/1	125/129
	%	15.9/12.4	0/0	17.6/21	0.9/0.9	10.3/9.4	1.3/0.4	0.4/0.4	53.6/55.4
SINEs (123)	No.	18/9	9/2	10/5	7/14	42/23	2/6	5/12	30/52
	%	14.6/7.3	7.3/1.6	8.1/4.1	5.7/11.4	34.1/18.7	1.6/4.9	4.1/9.8	24.4/42.3
Rn5s (51)	No.	0/0	0/0	0/0	0/0	40/40	0/0	0/0	11/11
	%	0/0	0/0	0/0	0/0	78.4/78.4	0/0	0/0	21.6/21.6
Rn4.5s (17)	No.	0/0	0/0	17/17	0/0	0/0	0/0	0/0	0/0
	%	0/0	0/0	100/100	0/0	0/0	0/0	0/0	0/0
Other Pol III genes (13)	No.	11/11	0/0	1/1	1/1	0/0	0/0	0/0	0/0
	%	84.6/84.6	0/0	7.7/7.7	7.7/7.7	0/0	0/0	0/0	0/0

The columns correspond to isolated loci grouped by presence or absence (as defined by a score above or below the cut-off) of Pol III, H3K4me3 or Pol II, at TP0. The lanes correspond to different types of Pol III loci. The first number in each group corresponds to RPC4 and RPB2 replicates 1, and H3K4me3 replicate 2, the second to RPC4, RPB2 and H3K4me3 replicates 2. For each group, the percentage of total number of genes in the corresponding gene type (shown in parenthesis in the first column) is indicated.

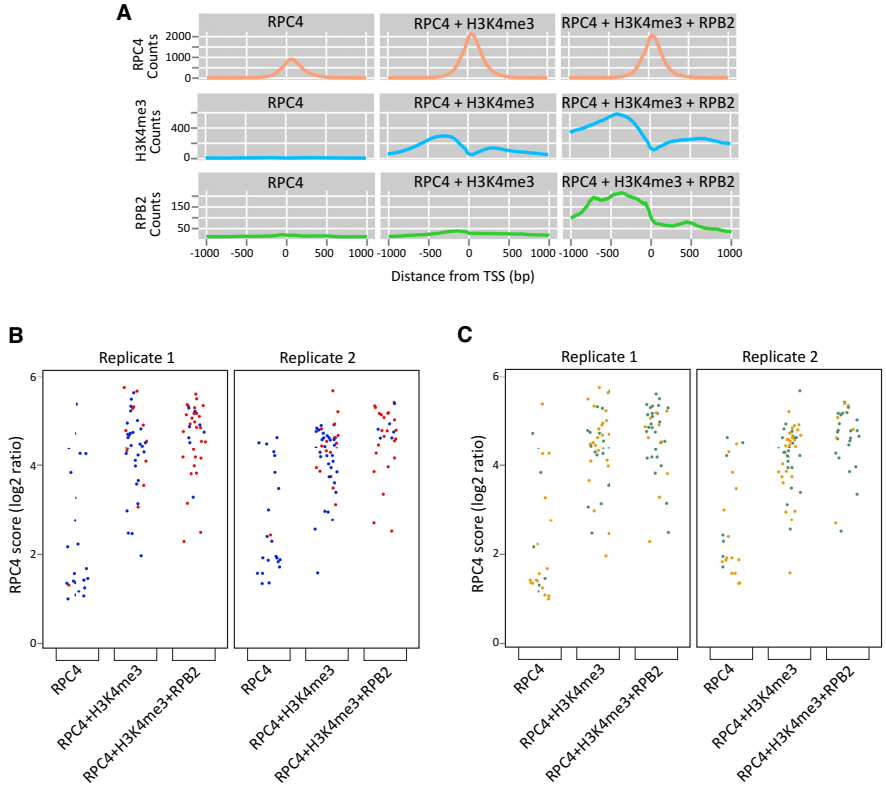


Figure 1. (**A**) Average tag density profiles for the factors indicated in the left, for the isolated tRNA gene groups (columns). The profiles were computed from replicate 2. (**B**) Scatterplots of RPC4 scores for isolated tRNA genes in replicate 1 and 2. The grouping into tRNA gene groups was done for each replicate independently (for H3K4me3, only replicate 2 was used). *P*-value < 2e–4 for the two comparisons in both replicates (see text), permutation based t test with 10000 permutations. Red and blue dots, genes with and without associated CpG islands, respectively. (**C**) As in B, but the orange and green dots indicate genes >2.65 kb and <2.65 kb, respectively, from TSS or poly A site of Pol II genes.

those within 2.65 kb of, and those removed by more than 2.65 kb from, a Pol II gene (TSS or poly A signal) (Supplementary Table S3, column D) revealed that the large majority (about 76% for tRNA genes and 89% for SINEs) of the loci with a RPB2 (and H3K4me3) peak were close to known Pol II genes (Figure 1C and Supplementary Figure S3C).

Visual examination of each of these <2.65 kb loci showed that in 26 out if 27 cases for tRNA genes, and in 13 out of 16 cases for SINEs, the RPB2 and H3K4me3 peaks overlapped a Pol II TSS (See Supplementary Figure S5A for examples), and the three remaining SINEs were located within Pol II transcription units. Pol II TSSs and tRNA genes were often divergent, consistent with previous findings (8) and explaining the major Pol II peak upstream of tRNA genes observed above (Figure 1A). Further, of the 10 isolated tRNA loci with an RPB2 peak >2.65 kb away from a Pol II TSS, half contained a CpG island (see Supplementary Figure S5B for examples with and without a CpG island). Thus, for most isolated tRNA genes and most SINEs with RPC4 + H3K4me3 + RPB2 peaks, the H3K4me3 and RPB2 peaks are explained by Pol II occupancy of a nearby Pol II transcription unit TSS or, for a few tRNA genes, of a nearby CpG island devoid of annotated TSSs. In HeLa cells, Oler *et al*. (8) observed a much more general presence of Pol II peaks upstream of Pol III-occupied tRNA genes, independent of the presence of Pol II TSSs. This may reflect a human-mouse cell difference or, more likely, a more spurious Pol II occupancy in the highly transformed HeLa cells as opposed to normal mouse liver cells.

In the RPC4 + H3K4me3 tRNA gene group close to Pol II loci (Figure 1C, green dots), the H3K4me3 peak was on a nearby Pol II TSS in only half the cases; in the other half, the tRNA gene was often surrounded by two H3K4me3 peaks with no associated annotated feature (see Supplementary Figure S5C for examples). In RPC4 + H3K4me3 tRNA gene group far from Pol II loci, there was generally a single major H3K4me3 peak on one side of the tRNA gene, in rare cases associated with a CpG island (Supplementary Figure S5D). These findings confirm the occurrence of H3K4me3 peaks close to active Pol III genes independent of the presence of Pol II peaks. Indeed, the 4.5S genes are a striking example of this configuration (See Table 1). This suggests that H3K4 trimethylation can occur at Pol III transcription units independently of the nearby presence of Pol II and Pol II transcription units. Such trimethylation might be brought about through transcription factors such as MYC, which can associate both with active Pol III genes (31), and with the WDR5 subunit of H3K4 methylase complexes (32).

Thus, at tRNA genes, high RPC4 occupancy scores are associated with proximity to either just H3K4me3 or H3K4me3 + RPB2 peaks, and in the latter case the H3K4me3 and RPB2 peaks are generally associated with Pol II TSSs. At SINEs, the RPC4 scores are similar in all groups, but where RPB2 and H3K4me3 peaks are present, these peaks are again mostly associated with Pol II TSSs or lie within Pol II transcription units.

Changes in Pol III genome association after PH

The results above describe Pol III gene occupancy in the liver, an organ composed largely of differentiated cells in a G0 state. After PH, most of the remaining hepatocytes exit from the G0 state and progress though the cell division cycle. Under the conditions used here for PH, S phase was reached after about 36 h (time point (TP) 36), G2/M phase at 44–48 h (TP48), followed by a second, less synchronous round of division (23). We examined RPC4 occupancy scores after PH. In general, RPC4 occupancy at all Pol III loci (isolated or not) increased sharply from TP0 to TP36, and then stayed quite elevated at TP48 and TP60 (Figure 2A). We built a one-way ANOVA statistical model with samples from TP0, TP36, TP48 and TP60 and ran an estimation of the contrasts of interests with the empirical Bayes method (see Materials and Methods). We did not find statistical differences (score changes with an FDR < 0.05) between the samples at TP36, TP48 and TP60. There was, however, a significant difference between TP0 and the later time points. We focused on the TP0 to TP36 transition, corresponding to liver cells in G0 (TP0) and S phase (TP36).

To illustrate the RPC4 score changes from TP0 to TP36, we plotted the T statistic as a function of the average score at TP0 (Figure 2B; see Supplementary Table S4 for FDR and T statistic values). The RPC4 scores of most loci increased, but 143 loci showed significant changes (red and blue dots; FDR < 0.05) and corresponded to those with low to medium RPC4 scores at TP0; in fact, 44 loci with significant changes from TP0 to TP36 had an RPC4 score below cutoff at TP0 (blue dots). Thus, out of 433 tRNA genes, 295 were occupied by Pol III in at least one sample at TP0 or TP36. By comparison, 300 tRNA genes were found to be active in at least one stage of mouse liver development (20), among them the 295 found active here. This indicates that as suggested for human tRNA genes (4), a population of mouse tRNA genes is in a lasting repressed state.

In contrast to loci with low scores at TP0, most loci with high scores at TP0 did not change significantly. Interestingly, out of the 143 changing genes with significant changes, the most numerous ones (n=107), and the ones with the largest changes (representing 100% of the genes in the upper tertile) were tRNA genes. The remaining changing loci were SINEs (n=31), Rn5s (n=3), and other Pol III genes, namely Rny1 and Rpph1.

To determine whether increased Pol III occupancy reflected increased transcription, we measured levels of intron-containing tRNA precursors, which are unstable and, therefore, better reflect transcription activity than mature tRNAs. RT-qPCR revealed increases in precursor tRNA levels derived from the 985-tRNAArg-TCT, 85-tRNALeu-CAA, and 960-tRNATyr-GTA loci (Figure 2C), consistent with increased Pol III occupancy reflecting increased ongoing transcription (4,5).

The increase in Pol III transcription was accompanied by increased levels of mRNAs coding for several transcription factors and Pol III subunits including TBP, BDP1, RPC8 (*Polr3h* mRNA), RPC4 (*Polr3d* mRNA), and RPC5 (*Polr3e* mRNA) (Figure 2D), in line with the observation that upon transformation of cultured cells, the levels of some Pol III transcription factors and subunits increase (17). We also observed an increase in the ratio of *Polr3g* to *Polr3gl* mRNA levels (Figure 2E), consistent with the observation that the *Polr3gl* gene provides a basic protein level and the *Polr3g*

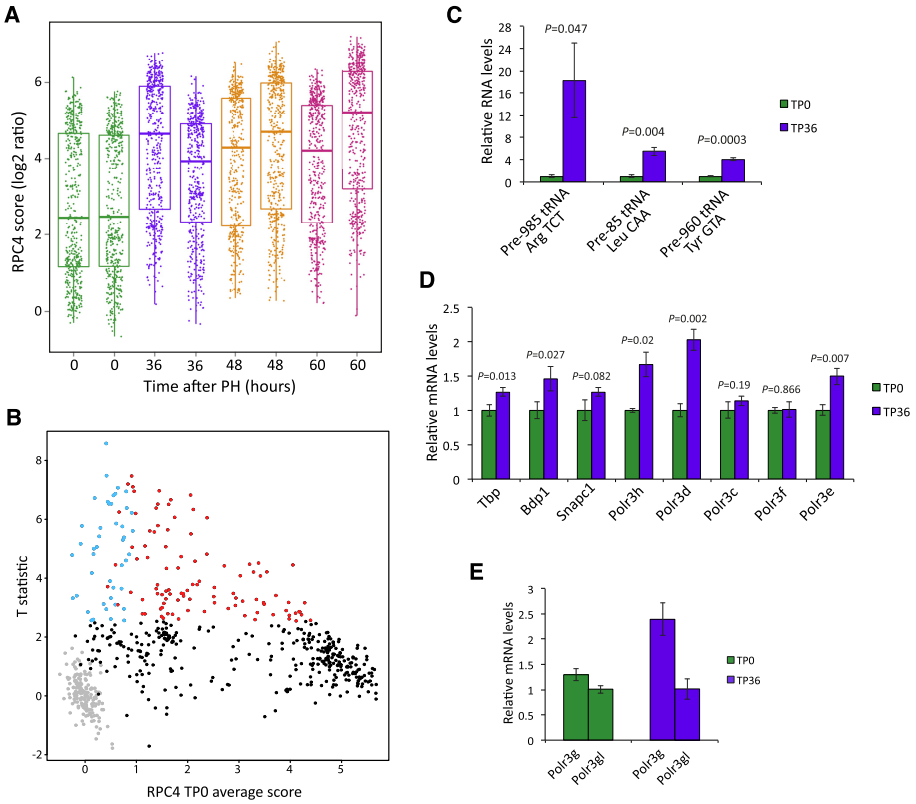


Figure 2. (**A**) Distribution of RPC4 scores for Pol III genes across time and replicates. Dots represent Pol III genes, and their distribution is summarized with their median (thicker horizontal bar), lower/upper quartile (box) and ±1.5 times the interquartile range (end of whiskers). Genes with an RPC4 occupancy score below the cutoff in all time points and replicates are not shown. (**B**) Plot of T statistics estimated for the TP0 to TP36 transition as a function of the average RPC4 score across replicates at TP0. A positive value of the T statistic indicates a score increase from TP0 to TP36. Red and blue dots, loci with an associated FDR<0.05, with blue dots indicating loci with an RPC4 score below cutoff at TP0. Black dots, loci with an associated FDR >0.05. Gray dots, loci with scores below the cutoff at both TP0 and TP36. (**C**) RT-qPCR quantification (normalized to Actb mRNA level) of indicated precursor tRNAs, relative to the values at TP0. Error bars represent ± SD. $n = 3$. The *P*-values were calculated using Student's *t*-test. (**D**) RT-qPCR quantification (normalized to Actb mRNA level) of the indicated mRNAs, relative to the values at TP0. Error bars and *P*-values as in C. (**E**) RT-qPCR quantification of *Polr3g* relative to *Polr3gl* mRNA. Error bars as in C.

gene allows adaptation to conditions when more protein is needed ((18) and references therein)

The observation that only a subset of tRNA genes displayed significantly increased RPC4 occupancy at TP36 as compared to TP0 prompted us to investigate how this increased occupancy affected tRNAs by isotypes and isoacceptors. As shown in Supplementary Figures S6A and S6B, the increase was general among tRNA isoacceptors and isotypes, suggesting a translation globally more active in dividing cells as compared to resting ones. We further wondered whether tRNA genes differentially occupied by Pol III at TP0 and TP36 were the same as the ones changing during development, and thus compared them with those changing between embryonic stage E15.5 and adult mouse liver (33). Of the 107 tRNA genes with changing scores from TP0 to TP36 and the 97 genes with changing scores from E15.5 to adult liver, 51 were common, clearly more than expected by chance (Supplementary Figure S7). Thus, there seems to be a pool of tRNA genes that are especially susceptible to differential regulation.

Changes in Pol III occupancy after PH correlate weakly with changes in H3K4me3 but not with changes in RPB2 occupancy

As shown above and consistent with previous observations in cultured cell lines, hepatocytes respond to the need to divide by differential increased Pol III transcription at individual Pol III loci. Do these loci also display changes in nearby H3K4me3 or Pol II occupancy? Both H3K4me3 and RPB2 scores around Pol III loci were overall quite stable (Figures 3A and B). We observed a weak correlation (Pearson correlation coefficient = 0.51) between changes in

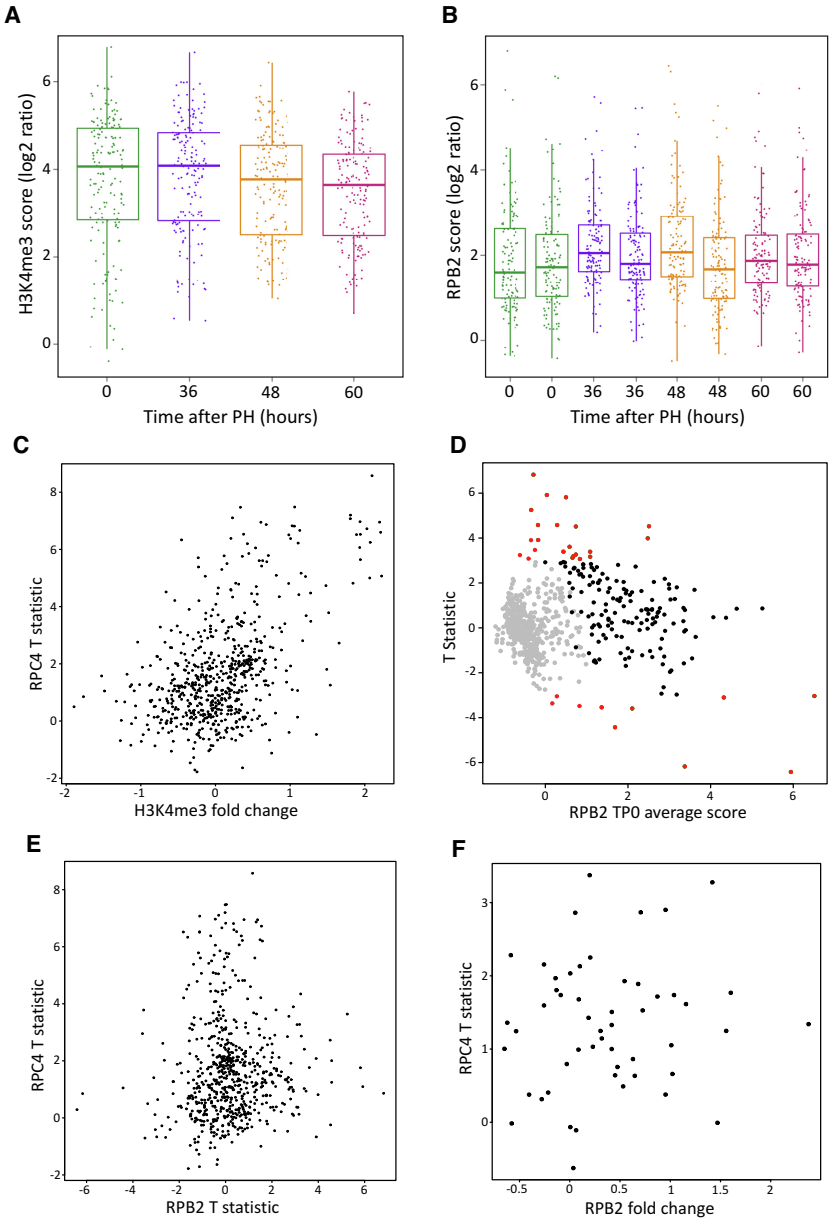


Figure 3. (**A**) Distribution of H3K4me3 scores around Pol III genes across time. Dots represent isolated Pol III loci whose H3K4me3 scores are above the cutoff. (**B**) Distribution of RPB2 scores around Pol III genes across time and replicates. Dots represent isolated Pol III loci whose RPB2 scores are above the cutoff. (**C**) Scatterplot of RPC4 scores T statistic and H3K4me3 fold change from TP0 to TP36. All Pol III loci are shown ($n = 646$). (**D**) Plot of RPB2 scores T statistic estimated for the transition TP0 to TP36 as a function of the RPB2 average score, across replicates, at TP0. RPB2 scores were computed in a region of ±1 kb around Pol III gene TSS. A positive value of the T statistic indicates an increase of score from TP0 to TP36. Red dots and black dots, genes with an associated FDR< and >0.05, respectively; grey dots, genes with RPB2 scores below cut-off. (**E**) Scatterplot of RPC4 and RPB2 scores T statistic, from TP0 to TP36. All Pol III loci are shown ($n = 646$). (**F**) Scatterplot of RPC4 T statistic and fold change of RPB2 occupancy at nearby Pol II TSS (±250 bp of the Pol II TSS). The Pol III loci considered (both isolated and not isolated) had RPC4 scores above the cutoff and were within 2.65 kb of a Pol II TSS with a corresponding RPB2 peak ($n = 52$).

Pol III and surrounding H3K4me3 peaks at all Pol III loci (Figure 3C, see also Supplementary Table S5), which was slightly higher for Pol III loci with RPC4 scores above cutoff in at least one sample at TP0 or TP36 (Pearson correlation coefficient = 0.56, data not shown). In contrast, whereas most Pol III occupancy changes were positive from TP0 to TP36, Pol II occupancy changes went in both directions (Figure 3D and Supplementary Table S4), and the correlation was very low (Figure 3E and Supplementary Table S5, Pearson correlation coefficient = 0.04). On the other hand, Pol III occupancy changes and Pol II occupancy changes occurring specifically at nearby Pol II TSSs (±250 bp) were generally positive for both Pol II and Pol III, consistent with the observations of Van Bortle *et al*. (5), who observed tRNA genes in domains to be down-regulated similarly to nearby protein-coding genes during differentiation (Supplementary Table S5). However, the Pearson correlation was again very low (0.12) (Figure 3F). Thus, locally (±1 kb), on a gene per gene basis, the dynamical changes are not quantitatively associated. This argues against a simple model in which Pol III genes are submitted to the same transcription regulation as nearby Pol II transcription units.

Pol III occupancy dynamics after PH in different groups

We next wondered whether there might be differences in how Pol III occupancy scores change in Pol III genes with different associated peaks. Whereas for SINEs, there were no significant differences between the groups, for tRNA genes, T statistic scores were highest in the RPC4 peak only group (two-sided KS test *P*-value = 0.001439 for RPC4 > RPC4 + H3K4me3, and 3.073e–11 for RPC4 > RPC4 + H3K4me3 + RPB2), and slightly higher in the RPC4 + H3K4me3 peak group than in the RPC4 + H3K4me3 + RPB2 group (*P*-value = 0.0004566) (Figures 4A and B, see also Supplementary Table S6). We focused on the differences between the RPC4 peak only and the RPC4 + H3K4me3 + RPB2 groups, as these differences were the largest. Out of 49 isolated tRNA genes with significantly different scores between TP0 and TP36, ∼53% had RPC4 peaks alone and 6% had RPC4 + H3K4me3 + RPB2 peaks, whereas among 82 non-changing Pol III-occupied isolated tRNA genes, these numbers were about 11% for genes with RPC4 peaks alone, and 52% for genes with RPC4 + H3K4me3 + RPB2 peaks (Supplementary Figure S8A). Thus, the majority of genes with changing scores, and those with the largest score differences between TP0 and TP36, were tRNA genes with RPC4 peaks alone, which tended to have lower scores at TP0 than genes in other groups (Figure 1). This was not a peculiarity of isolated tRNA genes, as a similar tendency was observed when all RPC4-occupied stable and changing tRNA genes were considered (Supplementary Figure S8B).

We used MEME to search for de novo motifs and CentriMo to search for known motifs within tRNA genes and their flanking sequences (±250 bp), or just within flanking sequences. We compared tRNA genes grouped according to surrounding peaks (tRNA genes in Figure 4A) as well as tRNA genes with or without significant RPC4 occupancy changes from TP0 to TP36 (tRNA genes in Figure 2B). We could not, however, identify any motifs differentially enriched in a specific group of tRNA genes, or in tRNA genes with changing versus stable RPC4 occupancy (data not shown), suggesting that differential expression after hepatectomy is not directly linked to the presence of specific DNA elements within tRNA genes and their immediate flanking sequences. Consistent with this observation, with the exception of some differences in the A and B boxes, the known promoter elements of tRNA genes, previous studies failed to identify *cis*-regulatory sequences correlating with differential tRNA gene expression in the adult liver (2), during development (20), or in proliferating versus diffenrentiated cells (34).

Last, we examined the chromosomal location of all of the 143 genes with significant increased RPC4 occupancy from TP0 to TP36. As shown in Figure 4C, the median of T statistics was highest for genes on chromosome 13, which, together with chromosomes 1 and 3, also seemed to harbor a disproportionate number of changing genes. We performed a chi-square independence test testing the null hypothesis that the distribution of changing versus stable Pol III genes on each chromosome was uniform. We obtained a *P*-value of 4.21e−12, indicating a strong association between chromosomal location and changes in Pol III occupancy. As shown in the mosaic plot in Supplementary Figure S9, chromosome 13 was indeed strongly enriched in changing genes, and depleted in stable genes, relative to expected values. Chromosomes 1 and 3 harbored a relatively high number of changing genes but contributed little to the total chi-square score.

tRNA genes in clusters and contact domains have been observed to share similar transcriptional dynamics during differentiation, as determined by comparing transcription changes at tRNA genes with the median fold change across tRNA gene clusters and domains to which they belong (5). Chromosome 13 contains a number of clustered tRNA genes. To explore how these and other tRNA gene clusters behaved, we first defined tRNA clusters as regions containing three or more tRNA genes, each spaced less than 5 kb from the next one, which gave a total of 42 clusters (listed in Supplementary Table S7). For each chromosome, we performed Kolmogorov-Smirnov tests to compare the distribution of RPC4 fold change among genes outside and within clusters, testing the null hypothesis that the two distributions are similar. With this stringent criterion, some clusters did indeed display a significantly different distribution of fold changes than non-clustered genes on the same chromosome, as shown by their *P*-values (Figure 4D (red bars indicate *P*-values < 0.05), Supplementary Table S7), but many did not. Thus, in our system, and as defined with stringent criteria, the fold changes for genes in clusters are not systematically differently distributed than for non-clustered genes on the same chromosome.

Supplementary Figure S10 shows examples of clusters with significantly different RPC4 fold change distribution from non-clustered genes on chromosomes 13 and 1 (see Figure 4D). The clusters contain very high proportions of changing Pol III loci and are strikingly devoid of nearby RPB2 peaks and, in many cases, of nearby H3K4me3 peaks, consistent with the results above indicating that the most changing tRNA genes are those devoid of RPB2 peaks (Figure 4A, Supplementary Figures S8A and B).

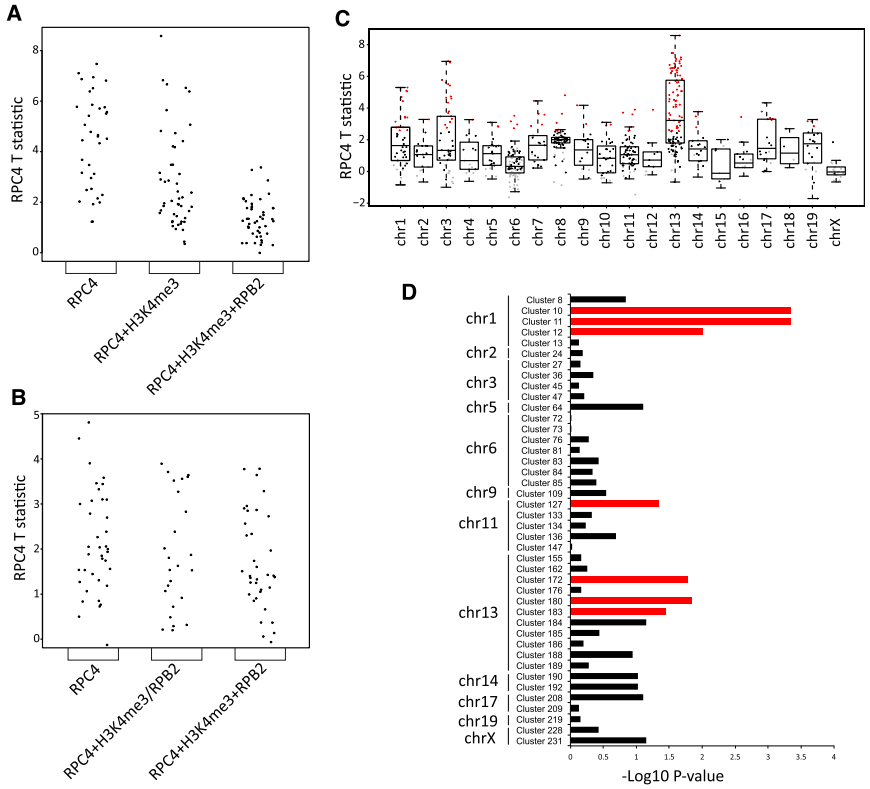


Figure 4. (**A**) Scatterplot of RPC4 T statistic (between TP0 and TP36) for the indicated groups of isolated tRNA genes (RPC4 peak only, $n = 35$; RPC4 + H3K4me3 peaks, $n = 48$); RPC4 + H3K4me3+RPB2 peaks ($n = 46$). There are significant differences between groups (see text for *P*-values). (**B**) Scatterplot of T statistic (between TP0 and TP36) for the indicated groups of isolated SINEs (RPC4 peak only, $n = 41$; RPC4+H3K4me3 peaks, $n = 11$; RPC4 + RPB2 peaks, $n = 14$; RPC4 + H3K4me3 + RPB2 peaks, $n = 38$). There was no significant difference between various groups (RPC4 only versus RPC4+H3K4me3/RPB2, *P*-value = 0.2417; RPC4 versus RPC4 + H3K4me3 + RPB2, *P*-value = 0.1018; RPC4 + H3K4me3/RPB2 versus RPC4 + H3K4me3 + RPB2, *P*-value = 0.9975). (**C**) RPC4 T statistic for TP0 to TP36 transition across different chromosomes. Red dots, loci with significant changes, grey dots, loci with RPC4 scores below cutoff. (**D**) Comparison of RPC4 occupancy fold change distribution between tRNA genes in clusters and outside clusters on the same chromosome. Red bars indicate clusters with KS test *P*-value <0.05.

The Pol III occupancy changes described here contrast with the observation that during differentiation in human THP-1 cells, the most occupied genes were the ones with the biggest occupancy changes after differentiation (5). They are, however, consistent with previous results comparing Pol III occupancy in serum-replete versus serum-starved IMR90Tert human cells, which identified highly and relatively stably occupied Pol III genes, and Pol III genes with lower but more dynamic Pol III occupancy (4). It is furthermore striking that many of the genes with scores changing between TP0 and TP36 also changed scores between embryonic E15.5 and adult mouse liver (33). In the PH system, we further observe that stable, but not changing, genes, are close to H3K4me3 peaks and Pol II-occupied TSSs. It thus appears that cells have a network of Pol III genes with relatively stable expression, and genes outside this network to adapt Pol III transcription quickly and reversibly to the increasing need of dividing cells for Pol III products.

DATA AVAILABILITY

The ChIP-seq data of this study have been submitted to he NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE114650. UCSC Genome Browser session.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Keith Harshman, Director of the Lausanne Genomic Technologies Facility, where all the ultra-high-throughput sequencing was performed, and Ioannis Xenarios, Director of the Vital-IT (http://www.vital-it.ch) Center for High Performance Computing of the Swiss Institute of Bioinformatics.

CYCLIX CONSORTIUM

Nouria Hernandez[1], Mauro Delorenzi[2,3,4], Bart Deplancke[5], Béatrice Desvergne[1], Nicolas Guex[6], Winship Herr[1], Felix Naef[5], Jacques Rougemont[7], Ueli Schibler[8], Teemu Andersin[8], Pascal Cousin[1], Federica Gilardi[1], Pascal Gos[8], Fabienne Lammers[1], Maykel Lopes[1], François Mange[1], Shilpi Minocha[1], Sunil Raghav[5], Dominic Villeneuve[1], Roberto Fabbretti[6], Volker Vlegel[6], Ioannis Xenarios[1,2,6], Eugenia Migliavacca[1,6], Viviane Praz[1,2], Fabrice David[2,7], Yohan Jarosz[2,7], Dmitry Kuznetsov[6], Robin Liechti[6], Olivier Martin[6], Julien Delafontaine[2,7], Julia Cajan[5], Cristian Carmeli[1,3], Kyle Gustafson[1], Irina Krier[5], Marion Leleu[2,7], Nacho Molina[5], Aurélien Naldi[7], Leonor Rib[1], Jonathan Sobel[5], Laura Symul[5], Gergana Bounova[1,2], Philippe Jacquet[2,7]

[1]Center for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne, 1015 Lausanne, Switzerland

[2]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

[3]Bioinformatics Core Facility, Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

[4]Department of Oncology and Ludwig Center for Cancer Research, Faculty of Biology and Medicine, University of Lausanne, 1011 Lausanne, Switzerland

[5]Interfaculty Institute of Bioengineering, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

[6]Vital IT, Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

[7]Bioinformatics and Biostatistics Core Facility, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

[8]Department of Molecular Biology, Faculty of Sciences, University of Geneva, 1211 Geneva, Switzerland

FUNDING

University of Lausanne; originally by CycliX, a grant from the Swiss SystemsX.ch initiative evaluated by the Swiss National Science Foundation (SNSF); SNSF Sinergia grant [CRSII3_160798]. Funding for open access charge: Swiss National Science Foundation and University of Lausanne.
Conflict of interest statement. None declared.

REFERENCES

1. James Faresse,N., Canella,D., Praz,V., Michaud,J., Romascano,D. and Hernandez,N. (2012) Genomic study of RNA polymerase II and III SNAPc-bound promoters reveals a gene transcribed by both enzymes and a broad use of common activators. *PLos Genet.*, **8**, e1003028.
2. Canella,D., Bernasconi,D., Gilardi,F., LeMartelot,G., Migliavacca,E., Praz,V., Cousin,P., Delorenzi,M., Hernandez,N. and Cycli,X.C. (2012) A multiplicity of factors contributes to selective RNA polymerase III occupancy of a subset of RNA polymerase III genes in mouse liver. *Genome Res.*, **22**, 666–680.
3. Canella,D., Praz,V., Reina,J.H., Cousin,P. and Hernandez,N. (2010) Defining the RNA polymerase III transcriptome: genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res.*, **20**, 710–721.
4. Orioli,A., Praz,V., Lhote,P. and Hernandez,N. (2016) Human MAF1 targets and represses active RNA polymerase III genes by preventing recruitment rather than inducing long-term transcriptional arrest. *Genome Res.*, **26**, 624–635.
5. Van Bortle,K., Phanstiel,D.H. and Snyder,M.P. (2017) Topological organization and dynamic regulation of human tRNA genes during macrophage differentiation. *Genome Biol.*, **18**, 180.
6. Barski,A., Chepelev,I., Liko,D., Cuddapah,S., Fleming,A.B., Birch,J., Cui,K., White,R.J. and Zhao,K. (2010) Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat. Struct. Mol. Biol.*, **17**, 629–634.
7. Moqtaderi,Z., Wang,J., Raha,D., White,R.J., Snyder,M., Weng,Z. and Struhl,K. (2010) Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat. Struct. Mol. Biol.*, **17**, 635–640.
8. Oler,A.J., Alla,R.K., Roberts,D.N., Wong,A., Hollenhorst,P.C., Chandler,K.J., Cassiday,P.A., Nelson,C.A., Hagedorn,C.H., Graves,B.J. *et al.* (2010) Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat. Struct. Mol. Biol.*, **17**, 620–628.
9. Raha,D., Wang,Z., Moqtaderi,Z., Wu,L., Zhong,G., Gerstein,M., Struhl,K. and Snyder,M. (2010) Close association of RNA polymerase II and many transcription factors with Pol III genes. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 3639–3644.
10. Alla,R.K. and Cairns,B.R. (2014) RNA polymerase III transcriptomes in human embryonic stem cells and induced pluripotent stem cells, and relationships with pluripotency transcription factors. *PLoS One*, **9**, e85648.
11. Rib,L., Villeneuve,D., Minocha,S., Praz,V., Hernandez,N., Guex,N., Herr,W. and Cycli,X.C. (2018) Cycles of gene expression and genome response during mammalian tissue regeneration. *Epigenet. Chromatin*, **11**, 52.
12. White,R.J. (2005) RNA polymerases I and III, growth control and cancer. *Nat. Rev. Mol. Cell Biol.*, **6**, 69–78.
13. White,R.J. (2004) RNA polymerase III transcription and cancer. *Oncogene*, **23**, 3208–3216.
14. Johnson,S.A., Dubeau,L. and Johnson,D.L. (2008) Enhanced RNA polymerase III-dependent transcription is required for oncogenic transformation. *J. Biol. Chem.*, **283**, 19184–19191.
15. Pavon-Eternod,M., Gomes,S., Geslain,R., Dai,Q., Rosner,M.R. and Pan,T. (2009) tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.*, **37**, 7268–7280.
16. Goodarzi,H., Nguyen,H.C.B., Zhang,S., Dill,B.D., Molina,H. and Tavazoie,S.F. (2016) Modulated expression of specific tRNAs drives gene expression and cancer progression. *Cell*, **165**, 1416–1427.
17. Durrieu-Gaillard,S., Dumay-Odelot,H., Boldina,G., Tourasse,N.J., Allard,D., Andre,F., Macari,F., Choquet,A., Lagarde,P., Drutel,G. *et al.* (2018) Regulation of RNA polymerase III transcription during transformation of human IMR90 fibroblasts with defined genetic elements. *Cell Cycle*, **17**, 605–615.
18. Renaud,M., Praz,V., Vieu,E., Florens,L., Washburn,M.P., l'Hote,P. and Hernandez,N. (2014) Gene duplication and neofunctionalization: POLR3G and POLR3GL. *Genome Res.*, **24**, 37–51.
19. Mange,F., Praz,V., Migliavacca,E., Willis,I.M., Schutz,F., Hernandez,N. and Cycli,X.C. (2017) Diurnal regulation of RNA polymerase III transcription is under the control of both the feeding-fasting response and the circadian clock. *Genome Res.*, **27**, 973–984.
20. Schmitt,B.M., Rudolph,K.L., Karagianni,P., Fonseca,N.A., White,R.J., Talianidis,I., Odom,D.T., Marioni,J.C. and Kutter,C. (2014) High-resolution mapping of transcriptional dynamics across tissue development reveals a stable mRNA-tRNA interface. *Genome Res.*, **24**, 1797–1807.
21. Mitchell,C. and Willenbring,H. (2008) A reproducible and well-tolerated method for 2/3 partial hepatectomy in mice. *Nat. Protoc.*, **3**, 1167–1170.
22. Mitchell,C. and Willenbring,H. (2014) Addendum: a reproducible and well-tolerated method for 2/3 partial hepatectomy in mice. *Nat. Protoc.*, **9**, 1532.
23. Minocha,S., Villeneuve,D., Rib,L., Moret,C., Guex,N. and Herr,W. (2017) Segregated hepatocyte proliferation and metabolic states within the regenerating mouse liver. *Hepatol. Commun.*, **1**, 871–885.
24. Le Martelot,G., Canella,D., Symul,L., Migliavacca,E., Gilardi,F., Liechti,R., Martin,O., Harshman,K., Delorenzi,M., Desvergne,B. *et al.* (2012) Genome-wide RNA polymerase II profiles and RNA

accumulation reveal kinetics of transcription and associated epigenetic changes during diurnal cycles. *PLoS Biol.*, **10**, e1001442.
25. Shah,R.N., Grzybowski,A.T., Cornett,E.M., Johnstone,A.L., Dickson,B.M., Boone,B.A., Cheek,M.A., Cowles,M.W., Maryanski,D., Meiners,M.J. *et al.* (2018) Examining the roles of H3K4 methylation states with systematically characterized antibodies. *Mol. Cell*, **72**, 162–177.
26. Bonhoure,N., Bounova,G., Bernasconi,D., Praz,V., Lammers,F., Canella,D., Willis,I.M., Herr,W., Hernandez,N., Delorenzi,M. *et al.* (2014) Quantifying ChIP-seq data: a spiking method providing an internal reference for sample-to-sample normalization. *Genome Res.*, **24**, 1157–1168.
27. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, doi:10.2202/1544-6115.1027.
28. Strimmer,K. (2008) A unified approach to false discovery rate estimation. *BMC Bioinformatics*, **9**, 303.
29. Storey,J.D. and Tibshirani,R. (2003) Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 9440–9445.
30. Listerman,I., Bledau,A.S., Grishina,I. and Neugebauer,K.M. (2007) Extragenic accumulation of RNA polymerase II enhances transcription by RNA polymerase III. *PLos Genet.*, **3**, e212.
31. Gomez-Roman,N., Grandori,C., Eisenman,R.N. and White,R.J. (2003) Direct activation of RNA polymerase III transcription by c-Myc. *Nature*, **421**, 290–294.
32. Thomas,L.R., Wang,Q., Grieb,B.C., Phan,J., Foshage,A.M., Sun,Q., Olejniczak,E.T., Clark,T., Dey,S., Lorey,S. *et al.* (2015) Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC. *Mol. Cell*, **58**, 440–452.
33. Rudolph,K.L., Schmitt,B.M., Villar,D., White,R.J., Marioni,J.C., Kutter,C. and Odom,D.T. (2016) Codon-Driven translational efficiency is stable across diverse mammalian cell states. *PLos Genet.*, **12**, e1006024.
34. Gingold,H., Tehler,D., Christoffersen,N.R., Nielsen,M.M., Asmar,F., Kooistra,S.M., Christophersen,N.S., Christensen,L.L., Borre,M., Sorensen,K.D. *et al.* (2014) A dual program for translation regulation in cellular proliferation and differentiation. *Cell*, **158**, 1281–1292.

Supplementary data

Differential regulation of RNA polymerase III genes during liver regeneration

Meghdad Yeganeh[†], Viviane Praz[†], Cristian Carmeli[†], Dominic Villeneuve, Leonor Rib, Nicolas Guex, Winship Herr, Mauro Delorenzi, Nouria Hernandez[*], and the CycliX consortium

† First co-authors

*Corresponding author

Supplemental Inventory

Supplementary Figures

Supplementary Tables (Separate files)

"Supplementary Tables are available at https://academic.oup.com/nar/article/47/4/1786/5266946#supplementary-data

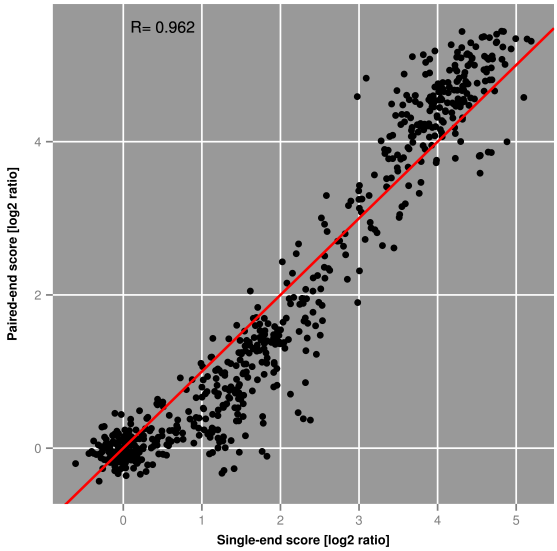


Figure S1. Comparison of scores obtained with paired-end sequencing (current study) and single-end sequencing (Renaud et al., Genome Res, 2014. 24(1): p. 37-51). The scores are log2 ratios with a pseudocount value of 30. The scores are the average of two biological replicates (mouse liver) for the single-end sequencing, and the average of two replicates at TP0 for the paired-end sequencing. Single-end and paired scores were normalized to the median number of counts of the four samples. The Pearson correlation coefficient of the two sets of scores is 0.962. Each dot represents a gene, the red line is the x=y line.

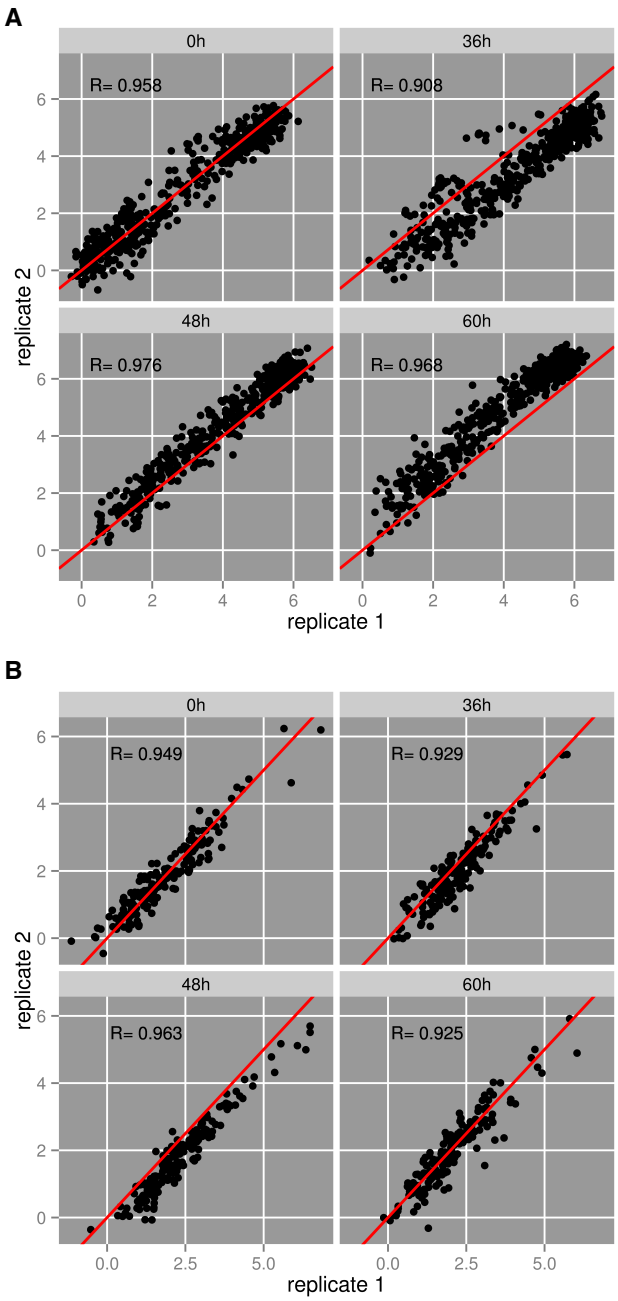


Figure S2. Scatterplot of RPC4 and RPB2 scores between the two replicates for each time point. The x=y line is in red. R is the Pearson correlation coefficient. **A)** RPC4 scores for those loci whose score is higher than the corresponding cutoff value for at least one time point. **B)** RPB2 scores for loci whose score is higher than the corresponding cutoff for at least one time point.

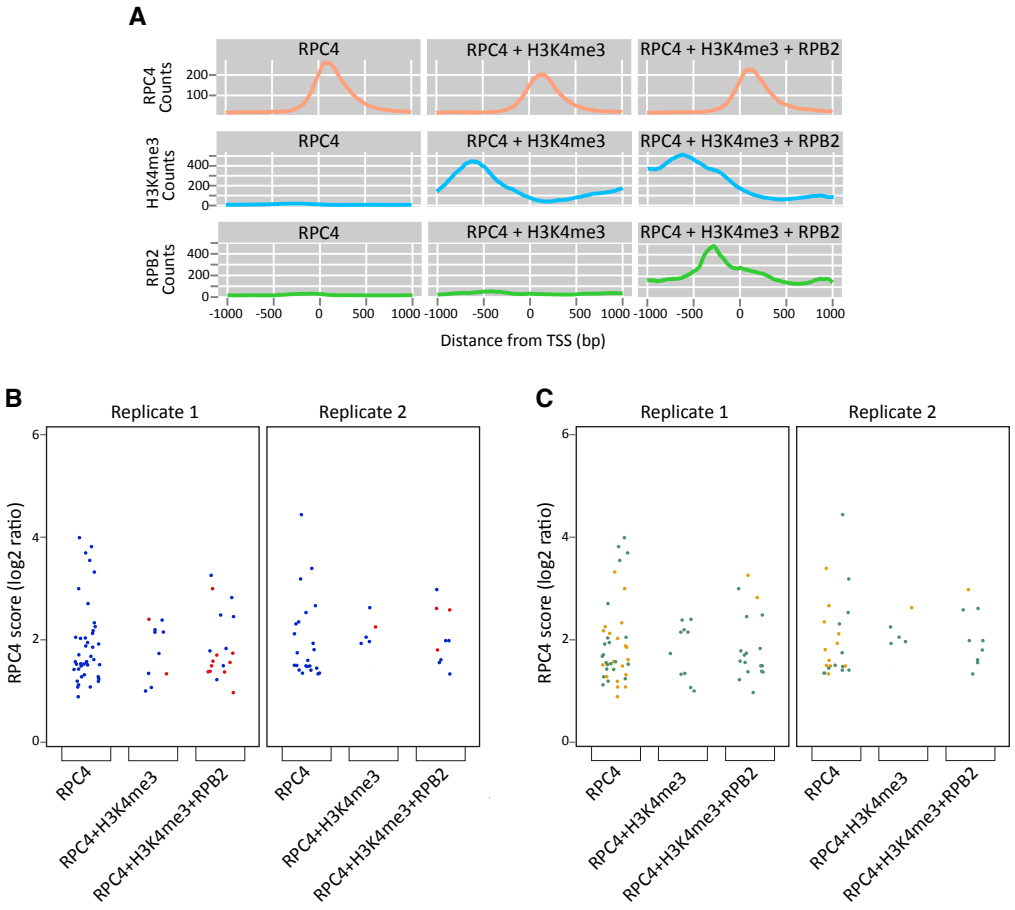


Figure S3. **A)** Average tag density profiles for RPC4, H3K4me3, and RPB2, as indicated on the left, for isolated SINEs with RPC4 peaks only (left column), RPC4 and H3K4me3 peaks (middle column), and RPC4, H3K4me3, and RPB2 peaks (right column). The profiles were computed from replicate 2 in a region of +/- 1Kbp around the annotated TSSs of isolated SINEs. **B)** Scatterplots of RPC4 scores for isolated SINEs in replicates 1 and 2 as indicated. The SINE groups are as in A, and the grouping into groups is done for each replicate independently (for H3K4me3, only replicate 2 was used). Scores in different groups are similar, permutation based t test P-value > 0.29) Red and blue dots, SINEs with and without associated CpG islands, respectively. **C)** As in B, but the orange and green dots indicate SINEs >2.65 Kbp and <2.65 Kbp, respectively, from TSS or poly A sites of Pol II genes.

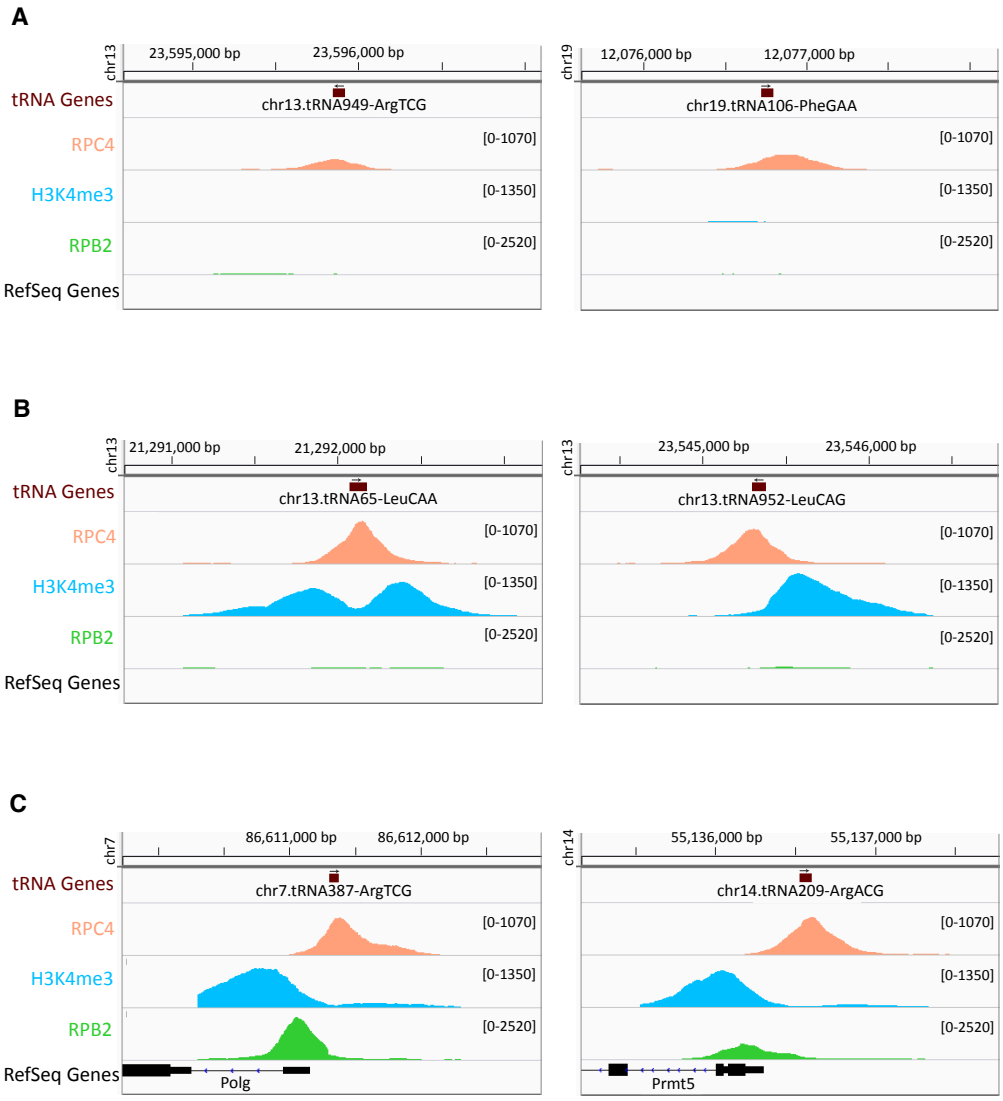


Figure S4. Integrative Genomics Viewer (IGV) views of genome regions surrounding isolated tRNA genes with: **A)** RPC4 peaks only. **B)** RPC4 and H3K4me3 peaks. **C)** RPC4, H3K4me3, and RPB2 peaks. The ChIP-seq tracks are from replicate 2.

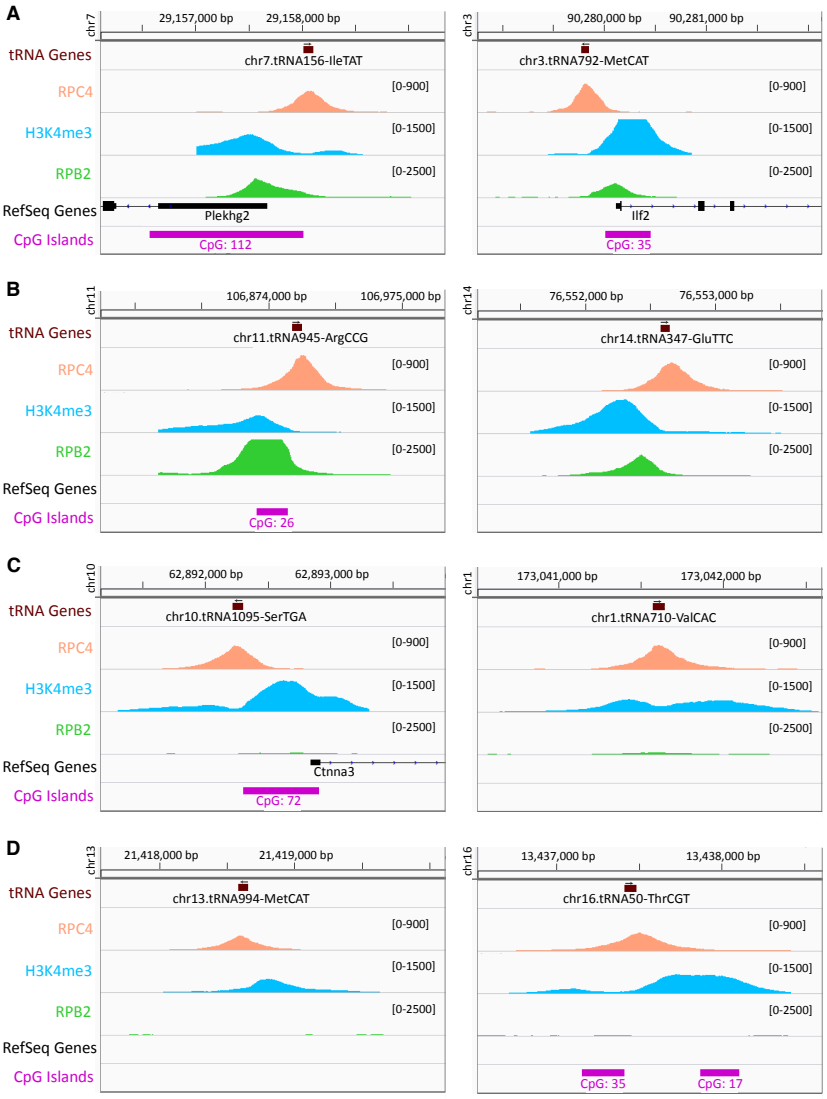


Figure S5. IGV views of genome regions surrounding isolated tRNA genes with different features: **A)** with RPC4, H3K4me3, and RPB2 peaks, located <2.65 Kbp from a Pol II gene, and with the RPB2 peak located on the Pol II gene TSS. **B)** with RPC4, H3K4me3, and RPB2 peaks, located >2.65 Kbp from a Pol II gene, and either containing (left), or not containing (right), a CpG island. **C)** with RPC4 and H3K4me3 peaks, and no RPB2 peak, located <2.65 Kbp from a Pol II gene, and with the main H3K4me3 peak either coinciding (left), or not coinciding (right), with a Pol II TSS and CpG island. **D)** with RPC4 and H3K4me3 peaks, and no RPB2 peak, located >2.65 Kbp from a Pol II gene. The ChIP-seq tracks are from replicate 2.

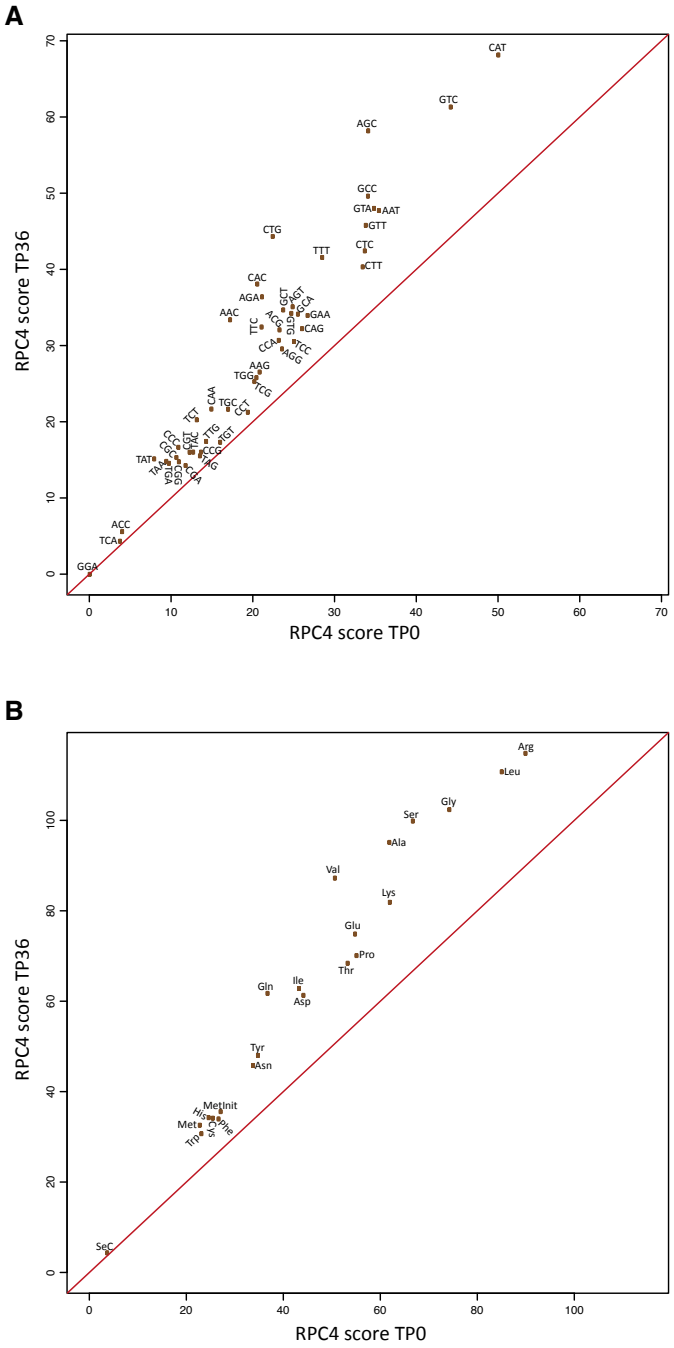


Figure S6. Scatter plots of RPC4 scores at TP0 and TP36 for all tRNA genes by **A)** Isoacceptors **B)** Isotypes. The scores were calculated as the mean of the two replicates for each gene, and then summed by isoacceptors or isotypes.

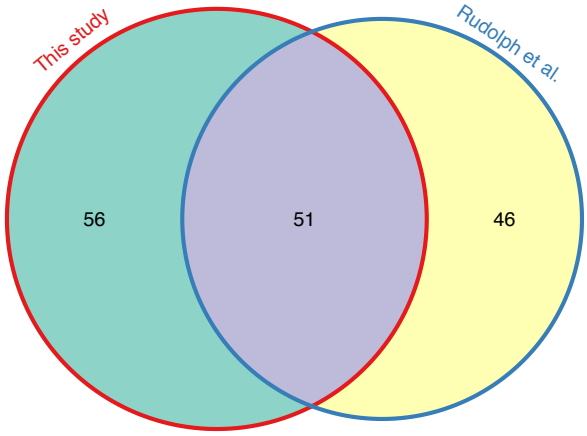


Figure S7. Venn diagram comparing tRNA genes differentially occupied by Pol III at TP0 and TP36 in this study with those different between embryonic and adult mouse liver (Rudolph et al., PLoS Genet, 2016. 12(5): p. e1006024). The number of active tRNA genes was 295 in our study and 311 in Rudolph et al.

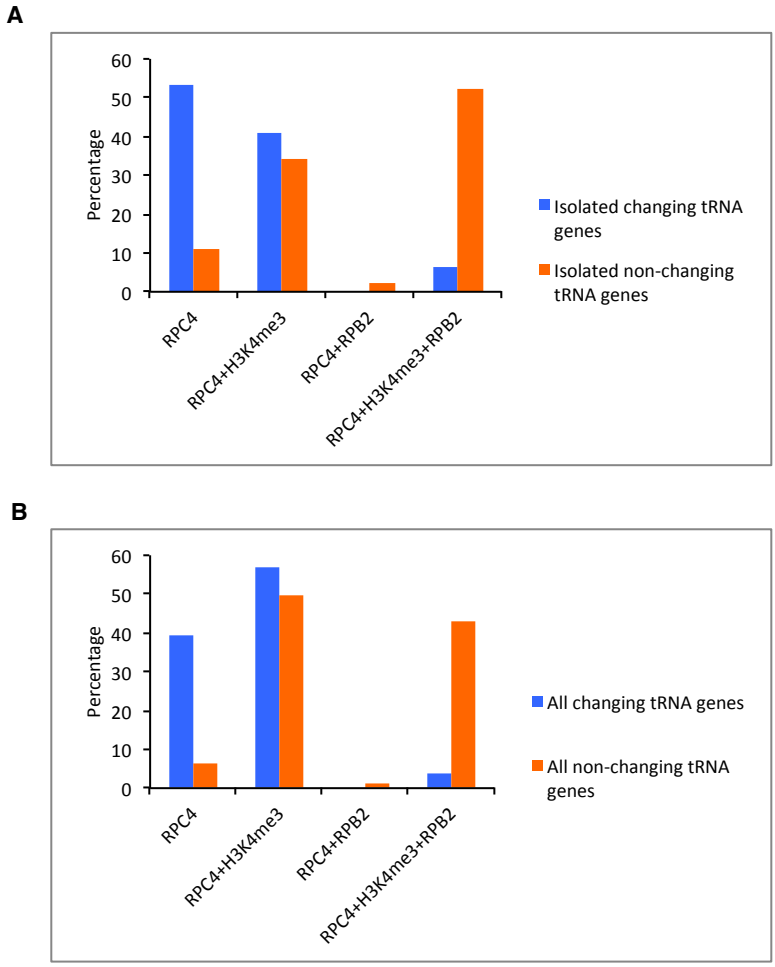


Figure S8. Percentage of tRNA genes in each group according to presence of different peaks, in changing and non-changing tRNA genes (with RPC4 score above cutoff) for: **A)** Isolated tRNA genes **B)** All (isolated and non-isolated) tRNA genes.

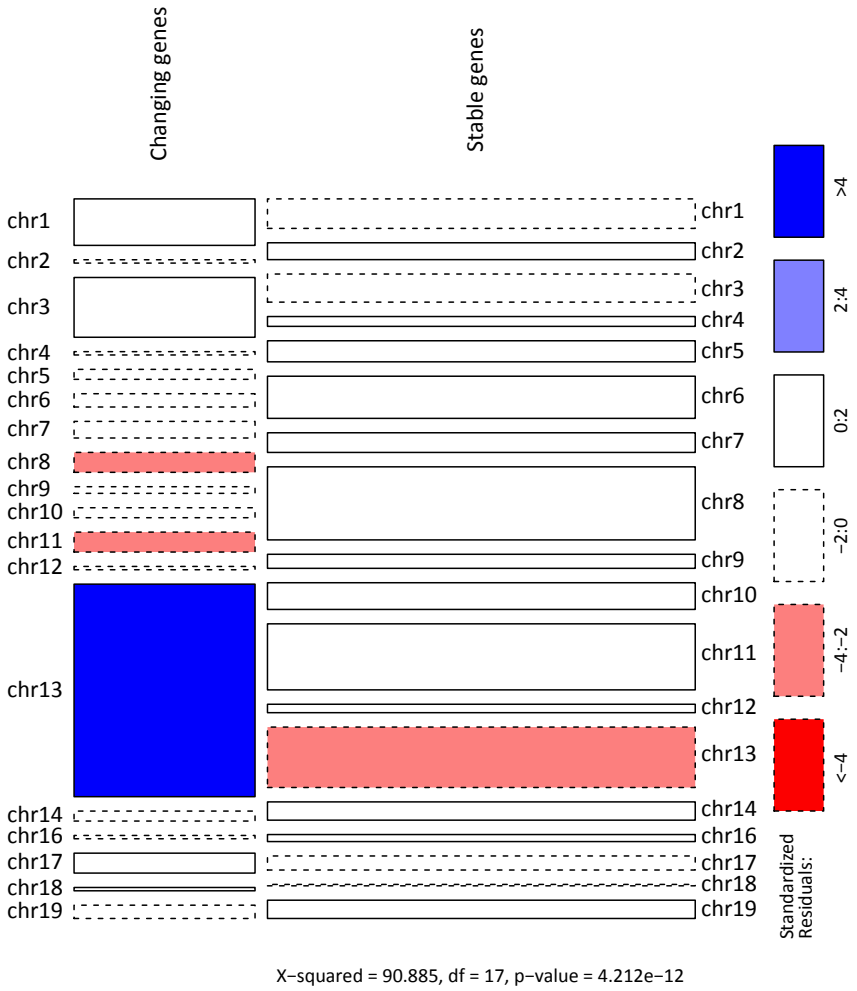
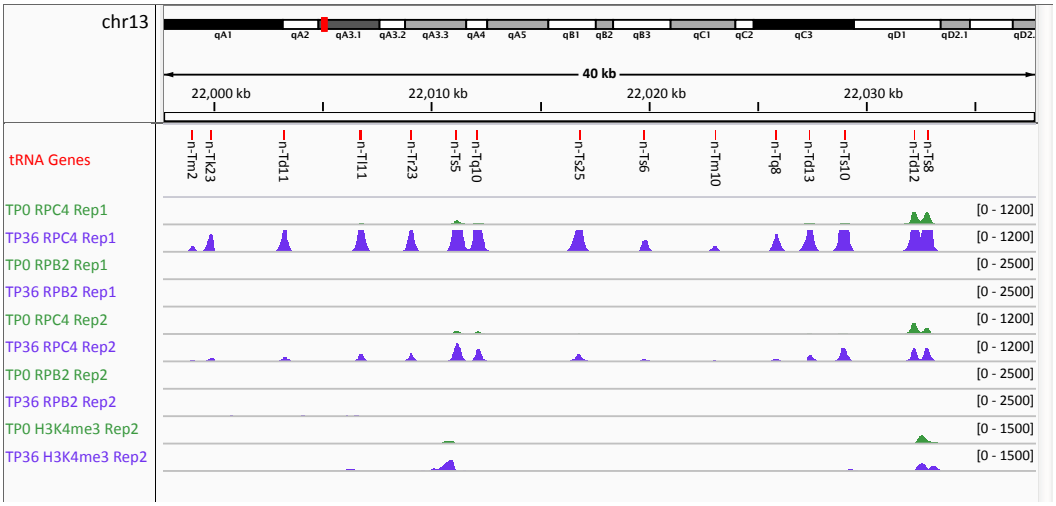


Figure S9. Chromosome 13 is enriched in Pol III genes with changes in occupancy between TP00 and TP36. Mosaic plot where the width of the two columns is related to the number of genes with significant (left column) or non-significant (right column) changes in Pol III occupancy scores between TP00 and TP36 relative to the total number of Pol III genes. The height of each rectangle in the first and second column is related to the number of changing or stable genes in each chromosome relative to the total number of changing or stable genes. For each cell, full or dashed lines indicate whether the relative contribution of each chromosome to the total Chi-square score is positive or negative, respectively, with blue and red colors indicating that the observed values are higher and lower, respectively, than expected (random) values.

A



B

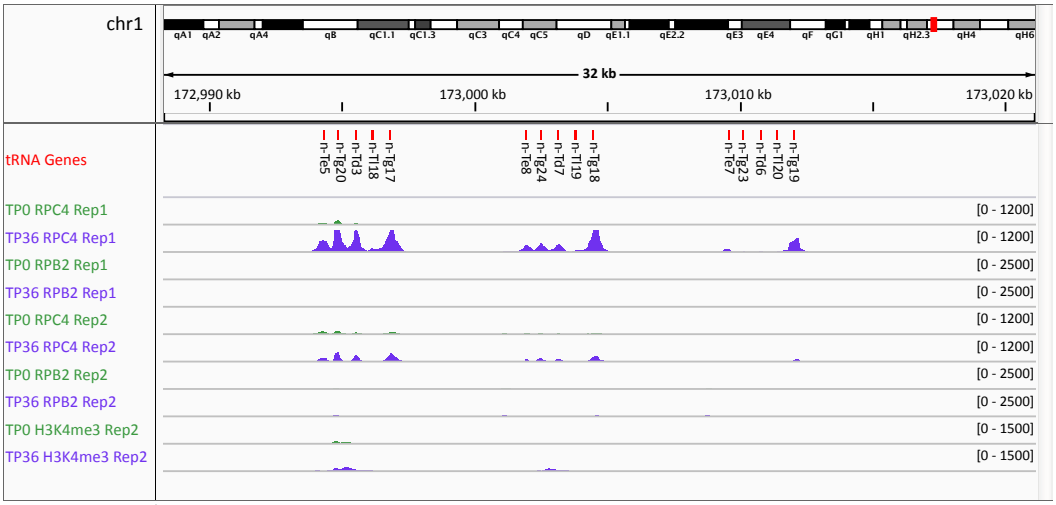


Figure S10. IGV views of genome regions corresponding to some clusters of tRNA genes on chromosomes 13 (cluster 172) and 1 (clusters 10,11,12). The tracks are as indicated on the left.

Chapter IV – Discussion and Perspectives

In eukaryotes, different RNA polymerases transcribe different sets of genes. Nevertheless, there is an important interplay between different polymerases. For example, studies have shown that several factors like c-Myc, p53, and RB can co-regulate Pol I, Pol II, and Pol III (Hannan et al. 2000; Zhai and Comai 2000; Felton-Edkins et al. 2003b; Gomez-Roman et al. 2006; White 2008; Beckerman and Prives 2010). In this thesis project, we studied the interplay between Pol II and Pol III at the genomic level and showed how some specific genomic arrangements can be involved in the regulation of transcription by Pol II and Pol III.

In the first part of this thesis project, we investigated how a Pol III transcription unit nested in a Pol II gene can regulate its transcription. We showed that antisense transcription of a MIR-SINE by Pol III interferes with Pol II transcription of the *Polr3e* gene.

By inserting the MIR within the intron of the EGFP gene, we showed that it could reduce EGFP expression only when it is in antisense, and not in sense, orientation relative to EGFP gene. This was a surprising observation because in a study in yeast, it was shown that TFIIIB, the key factor binding to Pol III promoter and recruiting RNA polymerase III, could interfere with Pol II transcription regardless of its binding orientation on DNA (Korde et al. 2014). Furthermore, in several other cases of gene arrangements with a Pol III gene within a Pol II gene, we observed a Pol II peak, suggestive of Pol II slow-down, on the Pol III gene whether the Pol III gene was sense or antisense relative to the Pol II gene. This further suggests that Pol III transcription can interfere with Pol II transcription regardless of the relative orientation of the two genes. It thus remained to be determined why, in the EGFP model system, interference was dependent on an antisense orientation of the MIR. We can probably exclude differences in levels of Pol III transcription of the MIR, since equivalent levels of MIR RNAs were produced from both the sense and antisense constructs. It seems possible, however, that interference by Pol III is dependent not only on the level of Pol III transcription, but also on that of Pol II transcription of the host gene, and that the exact levels of Pol II and Pol III transcription affect the ability of sense or

antisense Pol III transcription to interfere with Pol II transcription. It would be interesting to confirm the effect observed with the MIR-*Polr3e* arrangement in some of the other candidate genes where we observed Pol II accumulation on a nested Pol III gene by removing the Pol III gene with CRISPR/Cas9, as we did for the MIR, and confirm increased Pol II transcription of the Pol II host gene. It would be interesting to perform this experiment in a case where the Pol III gene is oriented sense relative to the Pol II gene, as this would confirm that the effect can be independent of relative orientation of the two transcription units.

An interesting aspect about the MIR and the *Polr3e* gene has been that *Polr3e* encodes a subunit of Pol III, and there might be a negative feedback loop, where increased Pol III activity leads to increased MIR transcription, then decreased *Polr3e* transcription, and ultimately decreased Pol III activity. To test this hypothesis, we first tried to increase Pol III transcription from the MIR. However, when we knocked down MAF1, the Pol III repressor, in ES cells, we did not observe any increase in the level of the MIR. The MIR is one of the most highly occupied Pol III targets in several cell types. It is thus possible that the MIR is simply transcribed at maximum levels even in the presence of MAF1, and thus removal of MAF1 has no effect. It would therefore be worthwhile to test more conditions leading to increase or, probably more important, decrease in Pol III activity. A possibility would be to use RNAi to decrease the levels of one of the Pol III subunits, for example the largest one, and determine whether this has an effect on *Polr3e* transcription and Pol II occupancy.

Another requirement to bring support to the negative feedback loop hypothesis is that changes in levels of POLR3E should affect Pol III transcription. Thus, studies should be performed to test the effect of up- or down-regulation of POLR3E on Pol III transcription. In the MIR-KO cells, where we observed an increase in POLR3E, we did not see any difference in the levels of some selected Pol III products as assessed by northern blot. However, given that Pol III genes can be selectively regulated (Orioli et al. 2016; Mange et al. 2017; Van Bortle et al. 2017), a more comprehensive method to assess Pol III transcription levels, such as genome-wide examination of Pol III occupancy by ChIP-seq, would need to be performed.

Yet another requirement in support of a regulatory role of the MIR for differential *Polr3e* expression is that the MIR itself be subject to differential regulation. To examine this possibility, we sought conditions under which one might expect down-regulation of general Pol III activity, and thus down-regulation of the *Polr3e* gene. One such condition is the passage from a cellular proliferation state to a differentiated state, in which one might expect demand for Pol III products to be lower. We thus examined the situation during differentiation of ES cells. We observed that levels of MIR RNA increased, with a slight decrease in *Polr3e* mRNA level, consistent with our hypothesis. It is, however, well known that during differentiation, the expression profile of many genes changes. So, to determine whether this change in *Polr3e* expression level is indeed mediated by the MIR rather than reflecting a decrease in transcription initiation due to other mechanisms of regulation, we would need to differentiate MIR-KO ES cells, probably from several different clones, with efficiencies comparable to WT cells, and determine whether *Polr3e* expression is regulated differently in the absence of the MIR.

SINEs have long been considered "junk DNA", until several studies suggested that they might have evolved functions depending on where they got integrated in the genome. For example there are SINEs that act as polyadenylation sites, splicing acceptors, or enhancers. Here, we have revealed another potential role of a SINE mediated by its transcription. We have also shown that there are more SINEs and tRNA genes that have the same arrangements as the MIR, that could potentially play the same roles.

Many of the previous studies on overlapping genes focused on Pol II long non-coding RNA (lncRNA) genes arranged in antisense orientation of protein-coding genes (Pelechano and Steinmetz 2013; Huber et al. 2016). This arrangement can regulate the protein-coding gene at the transcription or post-transcriptional levels through distinct mechanisms. In the case of the MIR located within the *Polr3e* gene, and in several other cases of Pol III genes nested within a Pol II gene, the mechanism by which the Pol III gene impacts on expression of Pol II gene seems to be through transcriptional interference, as evidenced by the peaks of Pol II occupancy occurring next to the peak of Pol III occupancy located on the Pol III gene. An interesting possibility is that in

addition to transcription interference, nested Pol III genes regulate the overlapping Pol II gene through their transcripts, which in many cases are complementary to the mRNA. However, our observation that overexpression of the MIR from a plasmid has no impact on *Polr3e* expression suggests that any RNA-mediated effect would be dependent on the RNA being synthesized in close proximity to the genomic location of the Pol II gene. In any case, given that many Pol II and Pol III genes are expressed differentially under different conditions, or in a tissue-specific manner (Dittmar et al. 2006; Sonawane et al. 2017), the interplay between Pol II and Pol III transcription may constitute a so far unsuspected additional layer of gene expression regulation.

In the second part of the thesis, we studied the dynamics of Pol III gene occupancy during mouse liver regeneration, and compared them with those of H3K4me3 and Pol II occupancy in the vicinity of Pol III genes. We used a two-third partial hepatectomy (PH) of the liver, which causes the remaining hepatocytes to enter the cell division cycle in a synchronous manner, combined with Pol III, Pol II, and H3K4me3 ChIP-seq.

First, we looked at Pol III occupancy at Pol III genes before PH. As shown before in mouse liver as well as in human cells (Canella et al. 2010; Canella et al. 2012), many tRNA genes were not occupied by Pol III. Further, because several studies suggested that active Pol III genes lie close to Pol II and H3K4me3 peaks (Barski et al. 2010; Moqtaderi et al. 2010; Oler et al. 2010; Raha et al. 2010; Canella et al. 2012; Alla and Cairns 2014; Van Bortle et al. 2017), we looked at H3K4me3 and Pol II peaks around tRNA genes and SINEs. We found that the presence of Pol II peaks is generally associated with the presence of CpG islands and that these peaks correspond to a Pol II gene TSS. In some cases for SINEs, the surrounding Pol II occupancy corresponded to an arrangement where the Pol III gene is located within a Pol II gene, a phenomenon we had observed before (Yeganeh et al. 2017). We also found Pol III genes with surrounding H3K4me3, and no Pol II peak nor Pol II gene. Interestingly, we found that Pol III occupancy at tRNA genes is generally higher when H3K4me3 and Pol II peaks are present.

We then examined the dynamics of occupancy from 0 (TP0) to 36 hours (TP36) after PH. We observed that there is a general increase in Pol III occupancy at Pol III genes at TP36 as compared to TP0. This increase was more pronounced for loci lowly occupied by Pol III at TP0, consistent with a previous study comparing Pol III occupancy in serum-replete versus serum-starved human IMR90Tert cells (Orioli et al. 2016). We showed that there was a corresponding increase in the level of several pre-tRNAs as assessed by qPCR. In a recent study of a mouse lacking the Pol III repressor MAF1, it was shown that increased Pol III activity led to increased levels of some tRNA precursors but not mature-length tRNAs (Bonhoure et al. 2015). It would thus be worth to assess the levels of some mature tRNAs after PH by techniques such as northern blots, as mature tRNA levels are not efficiently assessed by qPCR due to the presence of many modifications that can affect the primer elongation step.

To get a first idea of whether increased Pol III activity was accompanied by increased production of Pol III subunits and Pol III transcription factors, we assessed the levels of mRNAs encoding several Pol III subunits by qPCR, and indeed found an increase. Although suggestive of increased levels of Pol III and transcription factors, these results would need to be extended by western blot analyses to allow any firm conclusion.

The observation that different tRNA genes are regulated differentially during liver regeneration raises several questions, in particular, why, and how? To address the first question, we checked whether differential increase in tRNA occupancy might lead to differential increase in some tRNA isoacceptor classes, which in theory could enhance translation of mRNAs enriched in specific codons, as has been observed before (Gingold et al. 2014; Goodarzi et al. 2016). However, when we compared Pol III occupancy of tRNA genes at TP0 and TP36 by isotypes and isoacceptors, we observed a general increase at TP36, suggesting the translation being generally higher in proliferating cells as compared to resting ones. This is consistent with a study in which despite differential expression of mRNA-coding genes and individual tRNA genes in different stages of mouse liver and brain development, the codon usage of mRNAs and the relative levels of tRNAs by isoacceptor and isotype classes were stable (Schmitt et al. 2014).

How can differential dynamics of Pol III genes be explained? Several studies have explored whether differences in promoter elements or any motifs flanking tRNA genes could explain differential expression under various conditions. In all cases, including one comparing highly occupied and non-occupied tRNA genes in mouse liver (Canella et al. 2012), another one comparing active tRNA genes in differentiated and proliferating cells (Gingold et al. 2014), and yet another one comparing tRNA genes differentially expressed in embryonic and adult mouse liver (Schmitt et al. 2014), no DNA motifs different in one group compared to the other were found, with the exception of some small differences in the A and B boxes of the tRNA genes. Nevertheless, we searched for motifs within and around changing versus non-changing tRNA genes, and tRNA genes close to Pol II peaks versus those with only Pol III peaks. Consistent with previous findings, we did not find any motifs differentially enriched in any of these groups. On the other hand, we did find that tRNA genes that are close to Pol II peaks, and thus close to Pol II gene TSSs, were not changing much in Pol III occupancy during liver regeneration, whereas those without surrounding H3K4me3 and Pol II were the most dynamic.

Several studies suggest that active Pol III genes are associated with H3K27ac histone marks (Oler et al. 2010; Alla and Cairns 2014; Van Bortle et al. 2017). It was also shown that the upstream region of tRNA genes active in proliferating cells is more enriched with this modification than those of genes active in differentiated cells (Gingold et al. 2014). It would thus be very interesting to investigate the dynamics of H3K27ac at tRNA genes during liver regeneration, and to determine whether this histone modification is different in changing tRNA genes versus non-changing ones. Furthermore, in studies where the dynamics of Pol III transcription were examined either during serum starvation of cultured cells or during differentiation of THP-1 monocytes into macrophages, enhanced binding of MAF1 at down-regulated tRNA genes was shown (Orioli et al. 2016; Van Bortle et al. 2017). To investigate whether the MAF1 repressor plays a role in differential regulation of tRNA genes during liver regeneration, PHs could be performed in *Maf1* KO mice or, better, in liver-specific *Maf1* KO mice, to determine whether regulation of Pol III occupancy would be altered.

When we looked at changes in Pol III occupancy by chromosome, we observed a chromosomal bias, with changing genes being over-represented on chromosome 13. For a few tRNA gene clusters on this chromosome and for some on other chromosomes, we observed that genes within the clusters were regulated more similarly to each other than to genes outside the clusters on the same chromosomes. Some of these clusters contained highly changing tRNA genes. However, we could not conclude a systematic co-regulation of genes in clusters. Co-regulation of tRNA genes in clusters was suggested in THP-1 monocytes differentiation into macrophages. In this study, tRNA gene clusters were defined as regions with neighboring genes less than 20 kb apart, and transcription dynamics of individual tRNA genes (down-regulated, non-differential, up-regulated) was compared with the median fold change of the clusters to which they belonged (Van Bortle et al. 2017). However, we used more stringent criteria. We defined clusters as regions with at least three tRNA genes, with distance between two neighbor genes less than 5 kb, and we compared the Pol III occupancy fold change distribution between tRNA genes in clusters and tRNA genes outside clusters on the same chromosome.

It was shown that many active tRNA genes are located in genomic clusters (Schmitt et al. 2014), DNA loops (Van Bortle et al. 2017), and overlapping with CTCF binding sites (Alla and Cairns 2014). Moreover, examination of Pol III occupancy dynamics during THP-1 monocytes differentiation into macrophages suggested that tRNA genes in clusters changed coordinately with distant Pol II genes connected through DNA loops (Van Bortle et al. 2017). It would thus be interesting to determine whether the clusters examined in our study are located within DNA loops, by characterization of CTCF binding sites and mapping the interactions using Hi-C, and whether these DNA loops are dynamic during liver regeneration. The study of possible distant interactions with Pol II genes could be combined with that of the dynamics of Pol II genes during PH (Rib et al. 2018), to determine whether the dynamics of Pol III and Pol II gene expression might be coordinated among distant, but connected, genes.

In summary, in this thesis, we showed how genomic arrangements could impact gene expression regulation. In the first part, we revealed a mechanism of Pol II transcription regulation by nested Pol III genes, and in the second part, we showed

that Pol III genes distant from, or in the vicinity of, Pol II peaks on Pol II TSSs are more or less prone to changes in Pol III occupancy during the transition from a resting to an actively dividing cellular state.

References

Abdelmohsen K, Panda AC, Kang MJ, Guo R, Kim J, Grammatikakis I, Yoon JH, Dudekula DB, Noh JH, Yang X et al. 2014. 7SL RNA represses p53 translation by competing with HuR. *Nucleic Acids Res* **42**: 10099-10111.

Adelman K, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* **13**: 720-731.

Alla RK, Cairns BR. 2014. RNA polymerase III transcriptomes in human embryonic stem cells and induced pluripotent stem cells, and relationships with pluripotency transcription factors. *PLoS One* **9**: e85648.

Arimbasseri AG, Maraia RJ. 2015. Mechanism of Transcription Termination by RNA Polymerase III Utilizes a Non-template Strand Sequence-Specific Signal Element. *Mol Cell* **58**: 1124-1132.

Bai L, Morozov AV. 2010. Gene regulation by nucleosome positioning. *Trends in genetics : TIG* **26**: 476-483.

Barski A, Chepelev I, Liko D, Cuddapah S, Fleming AB, Birch J, Cui K, White RJ, Zhao K. 2010. Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. *Nat Struct Mol Biol* **17**: 629-634.

Beckerman R, Prives C. 2010. Transcriptional regulation by p53. *Cold Spring Harbor perspectives in biology* **2**: a000935.

Belancio VP, Hedges DJ, Deininger P. 2008. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome Res* **18**: 343-358.

Bonhoure N, Byrnes A, Moir RD, Hodroj W, Preitner F, Praz V, Marcelin G, Chua SC, Jr., Martinez-Lopez N, Singh R et al. 2015. Loss of the RNA polymerase III repressor MAF1 confers obesity resistance. *Genes Dev* **29**: 934-947.

Borchert GM, Lanier W, Davidson BL. 2006. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* **13**: 1097-1101.

Cairns CA, White RJ. 1998. p53 is a general repressor of RNA polymerase III transcription. *EMBO J* **17**: 3112-3123.

Canella D, Bernasconi D, Gilardi F, LeMartelot G, Migliavacca E, Praz V, Cousin P, Delorenzi M, Hernandez N, Cycli XC. 2012. A multiplicity of factors contributes to selective RNA polymerase III occupancy of a subset of RNA polymerase III genes in mouse liver. *Genome Res* **22**: 666-680.

Canella D, Praz V, Reina JH, Cousin P, Hernandez N. 2010. Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res* **20**: 710-721.

Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, Pesce E, Ferrer I, Collavin L, Santoro C et al. 2012. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491**: 454-457.

Carter-Timofte ME, Paludan SR, Mogensen TH. 2018. RNA Polymerase III as a Gatekeeper to Prevent Severe VZV Infections. *Trends Mol Med* **24**: 904-915.

Castelnuovo M, Rahman S, Guffanti E, Infantino V, Stutz F, Zenklusen D. 2013. Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nat Struct Mol Biol* **20**: 851-858.

Chen J, OuYang H, An X, Liu S. 2018. Vault RNAs partially induces drug resistance of human tumor cells MCF-7 by binding to the RNA/DNA-binding protein PSF and inducing oncogene GAGE6. *PLoS One* **13**: e0191325.

Chu WM, Wang Z, Roeder RG, Schmid CW. 1997. RNA polymerase III transcription repressed by Rb through its interactions with TFIIIB and TFIIIC2. *J Biol Chem* **272**: 14755-14761.

Ciganda M, Williams N. 2011. Eukaryotic 5S rRNA biogenesis. *Wiley Interdiscip Rev RNA* **2**: 523-533.

Coelho PS, Bryan AC, Kumar A, Shadel GS, Snyder M. 2002. A novel mitochondrial protein, Tar1p, is encoded on the antisense strand of the nuclear 25S rDNA. *Genes Dev* **16**: 2755-2760.

Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ et al. 2008. Structure of eukaryotic RNA polymerases. *Annu Rev Biophys* **37**: 337-352.

Dewannieux M, Esnault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet* **35**: 41-48.

Dewannieux M, Heidmann T. 2005. L1-mediated retrotransposition of murine B1 and B2 SINEs recapitulated in cultured cells. *J Mol Biol* **349**: 241-247.

Dieci G, Conti A, Pagano A, Carnevali D. 2013. Identification of RNA polymerase III-transcribed genes in eukaryotic genomes. *Biochim Biophys Acta* **1829**: 296-305.

Diribarne G, Bensaude O. 2009. 7SK RNA, a non-coding RNA regulating P-TEFb, a general transcription factor. *RNA Biol* **6**: 122-128.

Dittmar KA, Goodenbour JM, Pan T. 2006. Tissue-specific differences in human transfer RNA expression. *PLoS Genet* **2**: e221.

Dong C, Niu L, Song W, Xiong X, Zhang X, Zhang Z, Yang Y, Yi F, Zhan J, Zhang H et al. 2016. tRNA modification profiles of the fast-proliferating cancer cells. *Biochem Biophys Res Commun* **476**: 340-345.

Duechler M, Leszczynska G, Sochacka E, Nawrot B. 2016. Nucleoside modifications in the regulation of gene expression: focus on tRNA. *Cell Mol Life Sci* **73**: 3075-3095.

Dumay-Odelot H, Durrieu-Gaillard S, El Ayoubi L, Parrot C, Teichmann M. 2014. Contributions of in vitro transcription to the understanding of human RNA polymerase III transcription. *Transcription* **5**.

Durrieu-Gaillard S, Dumay-Odelot H, Boldina G, Tourasse NJ, Allard D, Andre F, Macari F, Choquet A, Lagarde P, Drutel G et al. 2018. Regulation of RNA polymerase III transcription during transformation of human IMR90 fibroblasts with defined genetic elements. *Cell Cycle*: 1-11.

Egloff S, Murphy S. 2008. Cracking the RNA polymerase II CTD code. *Trends Genet* **24**: 280-288.

Esakova O, Krasilnikov AS. 2010. Of proteins and RNA: the RNase P/MRP family. *RNA* **16**: 1725-1747.

Espinoza CA, Allen TA, Hieb AR, Kugel JF, Goodrich JA. 2004. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat Struct Mol Biol* **11**: 822-829.

Faghihi MA, Zhang M, Huang J, Modarresi F, Van der Brug MP, Nalls MA, Cookson MR, St-Laurent G, 3rd, Wahlestedt C. 2010. Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol* **11**: R56.

Fairley JA, Mitchell LE, Berg T, Kenneth NS, von Schubert C, Sillje HH, Medema RH, Nigg EA, White RJ. 2012. Direct regulation of tRNA and 5S rRNA gene transcription by Polo-like kinase 1. *Mol Cell* **45**: 541-552.

Felton-Edkins ZA, Fairley JA, Graham EL, Johnston IM, White RJ, Scott PH. 2003a. The mitogen-activated protein (MAP) kinase ERK induces tRNA synthesis by phosphorylating TFIIIB. *EMBO J* **22**: 2422-2432.

Felton-Edkins ZA, Kenneth NS, Brown TR, Daly NL, Gomez-Roman N, Grandori C, Eisenman RN, White RJ. 2003b. Direct regulation of RNA polymerase III transcription by RB, p53 and c-Myc. *Cell Cycle* **2**: 181-184.

Fong N, Bentley DL. 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes & development* **15**: 1783-1795.

Fuda NJ, Ardehali MB, Lis JT. 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461**: 186-192.

Gebauer F, Hentze MW. 2004. Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* **5**: 827-835.

Gelfand B, Mead J, Bruning A, Apostolopoulos N, Tadigotla V, Nagaraj V, Sengupta AM, Vershon AK. 2011. Regulated antisense transcription controls expression of cell-type-specific genes in yeast. *Mol Cell Biol* **31**: 1701-1709.

Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, Adelman K. 2010. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143**: 540-551.

Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, Christophersen NS, Christensen LL, Borre M, Sorensen KD et al. 2014. A dual program for translation regulation in cellular proliferation and differentiation. *Cell* **158**: 1281-1292.

Gjidoda A, Henry RW. 2013. RNA polymerase III repression by the retinoblastoma tumor suppressor protein. *Biochim Biophys Acta* **1829**: 385-392.

Goldfarb KC, Cech TR. 2017. Targeted CRISPR disruption reveals a role for RNase MRP RNA in human preribosomal RNA processing. *Genes Dev* **31**: 59-71.

Gomez-Roman N, Felton-Edkins ZA, Kenneth NS, Goodfellow SJ, Athineos D, Zhang J, Ramsbottom BA, Innes F, Kantidakis T, Kerr ER et al. 2006. Activation by c-Myc of transcription by RNA polymerases I, II and III. *Biochem Soc Symp*: 141-154.

Gomez-Roman N, Grandori C, Eisenman RN, White RJ. 2003. Direct activation of RNA polymerase III transcription by c-Myc. *Nature* **421**: 290-294.

Gongadze GM. 2011. 5S rRNA and ribosome. *Biochemistry (Mosc)* **76**: 1450-1464.

Goodarzi H, Nguyen HCB, Zhang S, Dill BD, Molina H, Tavazoie SF. 2016. Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* **165**: 1416-1427.

Gu C, Begley TJ, Dedon PC. 2014. tRNA modifications regulate translation during cellular stress. *FEBS Lett* **588**: 4287-4296.

Hahn S. 2004. Structure and mechanism of the RNA polymerase II transcription machinery. *Nature structural & molecular biology* **11**: 394-403.

Hancks DC, Kazazian HH, Jr. 2012. Active human retrotransposons: variation and disease. *Curr Opin Genet Dev* **22**: 191-203.

Hannan KM, Hannan RD, Smith SD, Jefferson LS, Lun M, Rothblum LI. 2000. Rb and p130 regulate RNA polymerase I transcription: Rb disrupts the interaction between UBF and SL-1. *Oncogene* **19**: 4988-4999.

Harvey RF, Smith TS, Mulroney T, Queiroz RML, Pizzinga M, Dezi V, Villenueva E, Ramakrishna M, Lilley KS, Willis AE. 2018. Trans-acting translational regulatory RNA binding proteins. *Wiley Interdiscip Rev RNA* **9**: e1465.

He N, Chan CK, Sobhian B, Chou S, Xue Y, Liu M, Alber T, Benkirane M, Zhou Q. 2011. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **108**: E636-645.

Heidemann M, Hintermair C, Voss K, Eick D. 2013. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys Acta* **1829**: 55-62.

Herrera MC, Chymkowitch P, Robertson JM, Eriksson J, Boe SO, Alseth I, Enserink JM. 2018. Cdk1 gates cell cycle-dependent tRNA synthesis by regulating RNA polymerase III activity. *Nucleic acids research*.

Hershey JW, Sonenberg N, Mathews MB. 2012. Principles of translational control: an overview. *Cold Spring Harb Perspect Biol* **4**.

Hinnebusch AG, Lorsch JR. 2012. The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol* **4**.

Hirsch HA, Jawdekar GW, Lee KA, Gu L, Henry RW. 2004. Distinct mechanisms for repression of RNA polymerase III transcription by the retinoblastoma tumor suppressor protein. *Mol Cell Biol* **24**: 5989-5999.

Hoffmann NA, Jakobi AJ, Moreno-Morcillo M, Glatt S, Kosinski J, Hagen WJ, Sachse C, Muller CW. 2015. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528**: 231-236.

Hsin JP, Manley JL. 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**: 2119-2137.

Hu P, Samudre K, Wu S, Sun Y, Hernandez N. 2004. CK2 phosphorylation of Bdp1 executes cell cycle-specific RNA polymerase III transcription repression. *Molecular cell* **16**: 81-92.

Huber F, Bunina D, Gupta I, Khmelinskii A, Meurer M, Theer P, Steinmetz LM, Knop M. 2016. Protein Abundance Control by Non-coding Antisense Transcription. *Cell Rep* **15**: 2625-2636.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* **12**: 99-110.

Jackson RJ, Hellen CU, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**: 113-127.

Jadaliha M, Gholamalamdari O, Tang W, Zhang Y, Petracovici A, Hao Q, Tariq A, Kim TG, Holton SE, Singh DK et al. 2018. A natural antisense lncRNA controls breast cancer progression by promoting tumor suppressor gene mRNA stability. *PLoS Genet* **14**: e1007802.

Jarrous N. 2017. Roles of RNase P and Its Subunits. *Trends Genet* **33**: 594-603.

Johnson SA, Dubeau L, Johnson DL. 2008. Enhanced RNA polymerase III-dependent transcription is required for oncogenic transformation. *J Biol Chem* **283**: 19184-19191.

Johnson SS, Zhang C, Fromm J, Willis IM, Johnson DL. 2007. Mammalian Maf1 is a negative regulator of transcription by all three nuclear RNA polymerases. *Molecular cell* **26**: 367-379.

Jonkers I, Kwak H, Lis JT. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3**: e02407.

Jonkers I, Lis JT. 2015. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**: 167-177.

Kapp LD, Lorsch JR. 2004. The molecular mechanics of eukaryotic translation. *Annu Rev Biochem* **73**: 657-704.

Kassavetis GA, Prakash P, Shim E. 2010. The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. *J Biol Chem* **285**: 2695-2706.

Kassiotis G. 2014. Endogenous retroviruses and the development of cancer. *J Immunol* **192**: 1343-1349.

Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J et al. 2005. Antisense transcription in the mammalian transcriptome. *Science* **309**: 1564-1566.

Keam SP, Hutvagner G. 2015. tRNA-Derived Fragments (tRFs): Emerging New Roles for an Ancient RNA in the Regulation of Gene Expression. *Life (Basel)* **5**: 1638-1651.

Khoo SK, Wu CC, Lin YC, Chen HT. 2018. The TFIIE-related Rpc82 subunit of RNA polymerase III interacts with the TFIIB-related transcription factor Brf1 and the polymerase cleft for transcription initiation. *Nucleic Acids Res* **46**: 1157-1166.

Kieckhaefer J, Lukovac S, Ye DZ, Lee D, Beetler DJ, Pack M, Kaestner KH. 2016. The RNA polymerase III subunit Polr3b is required for the maintenance of small intestinal crypts in mice. *Cell Mol Gastroenterol Hepatol* **2**: 783-795.

Kirchner S, Ignatova Z. 2015. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat Rev Genet* **16**: 98-112.

Korde A, Rosselot JM, Donze D. 2014. Intergenic transcriptional interference is blocked by RNA polymerase III transcription factor TFIIIB in Saccharomyces cerevisiae. *Genetics* **196**: 427-438.

Korneev SA, Maconochie M, Naskar S, Korneeva EI, Richardson GP, O'Shea M. 2015. A novel long non-coding natural antisense RNA is a negative regulator of Nos1 gene expression. *Sci Rep* **5**: 11815.

Kramerov DA, Vassetzky NS. 2011. SINEs. *Wiley Interdiscip Rev RNA* **2**: 772-786.

Krude T, Christov CP, Hyrien O, Marheineke K. 2009. Y RNA functions at the initiation step of mammalian chromosomal DNA replication. *J Cell Sci* **122**: 2836-2845.

Kumar A. 2009. An overview of nested genes in eukaryotic genomes. *Eukaryot Cell* **8**: 1321-1329.

Kumar P, Kuscu C, Dutta A. 2016. Biogenesis and Function of Transfer RNA-Related Fragments (tRFs). *Trends Biochem Sci* **41**: 679-689.

Kwak H, Lis JT. 2013. Control of transcriptional elongation. *Annu Rev Genet* **47**: 483-508.

Lai WKM, Pugh BF. 2017. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol* **18**: 548-562.

Lee J, Moir RD, Willis IM. 2009. Regulation of RNA polymerase III transcription involves SCH9-dependent and SCH9-independent branches of the target of rapamycin (TOR) pathway. *The Journal of biological chemistry* **284**: 12604-12608.

Lee JY, Ji Z, Tian B. 2008. Phylogenetic analysis of mRNA polyadenylation sites reveals a role of transposable elements in evolution of the 3'-end of genes. *Nucleic Acids Res* **36**: 5581-5590.

Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L, Washburn MP, Conaway JW, Conaway RC, Shilatifard A. 2010. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol Cell* **37**: 429-437.

Liu F, Clark W, Luo G, Wang X, Fu Y, Wei J, Wang X, Hao Z, Dai Q, Zheng G et al. 2016. ALKBH1-Mediated tRNA Demethylation Regulates Translation. *Cell* **167**: 816-828.e816.

Lockwood WW, Chari R, Coe BP, Thu KL, Garnis C, Malloff CA, Campbell J, Williams AC, Hwang D, Zhu CQ et al. 2010. Integrative genomic analyses identify BRF2 as a novel lineage-specific oncogene in lung squamous cell carcinoma. *PLoS Med* **7**: e1000315.

Lukoszek R, Mueller-Roeber B, Ignatova Z. 2013. Interplay between polymerase II- and polymerase III-assisted expression of overlapping genes. *FEBS Lett* **587**: 3692-3695.

Lund RJ, Rahkonen N, Malonzo M, Kauko L, Emani MR, Kivinen V, Narva E, Kemppainen E, Laiho A, Skottman H et al. 2017. RNA Polymerase III Subunit POLR3G Regulates Specific Subsets of PolyA(+) and SmallRNA Transcriptomes and Splicing in Human Pluripotent Stem Cells. *Stem Cell Reports* **8**: 1442-1454.

Lunyak VV, Prefontaine GG, Nunez E, Cramer T, Ju BG, Ohgi KA, Hutt K, Roy R, Garcia-Diaz A, Zhu X et al. 2007. Developmentally regulated activation of a

SINE B2 repeat as a domain boundary in organogenesis. *Science* **317**: 248-251.

Makalowska I, Lin CF, Makalowski W. 2005. Overlapping genes in vertebrate genomes. *Comput Biol Chem* **29**: 1-12.

Mange F, Praz V, Migliavacca E, Willis IM, Schutz F, Hernandez N, Cycli XC. 2017. Diurnal regulation of RNA polymerase III transcription is under the control of both the feeding-fasting response and the circadian clock. *Genome Res* **27**: 973-984.

Mariner PD, Walters RD, Espinoza CA, Drullinger LF, Wagner SD, Kugel JF, Goodrich JA. 2008. Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol Cell* **29**: 499-509.

Marshall L, White RJ. 2008. Non-coding RNA production by RNA polymerase III is implicated in cancer. *Nat Rev Cancer* **8**: 911-914.

Martens JA, Laprade L, Winston F. 2004. Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature* **429**: 571-574.

Mauck JC, Green H. 1974. Regulation of pre-transfer RNA synthesis during transition from resting to growing state. *Cell* **3**: 171-177.

Melnikov S, Ben-Shem A, Garreau de Loubresse N, Jenner L, Yusupova G, Yusupov M. 2012. One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol* **19**: 560-567.

Michels AA, Robitaille AM, Buczynski-Ruchonnet D, Hodroj W, Reina JH, Hall MN, Hernandez N. 2010. mTORC1 directly phosphorylates and regulates human MAF1. *Molecular and cellular biology* **30**: 3749-3757.

Mishra S, Maraia RJ. 2018. RNA polymerase III subunits C37/53 modulate rU:dA hybrid 3' end dynamics during transcription termination. *Nucleic Acids Res*.

Modarresi F, Faghihi MA, Lopez-Toledano MA, Fatemi RP, Magistri M, Brothers SP, van der Brug MP, Wahlestedt C. 2012. Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. *Nat Biotechnol* **30**: 453-459.

Moir RD, Lee J, Haeusler RA, Desai N, Engelke DR, Willis IM. 2006. Protein kinase A regulates RNA polymerase III transcription through the nuclear localization of Maf1. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 15044-15049.

Moqtaderi Z, Wang J, Raha D, White RJ, Snyder M, Weng Z, Struhl K. 2010. Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nat Struct Mol Biol* **17**: 635-640.

Mroczek S, Dziembowski A. 2013. U6 RNA biogenesis and disease association. *Wiley Interdiscip Rev RNA* **4**: 581-592.

Nadal-Ribelles M, Sole C, Xu Z, Steinmetz LM, de Nadal E, Posas F. 2014. Control of Cdc28 CDK1 by a stress-induced lncRNA. *Mol Cell* **53**: 549-561.

Nedialkova DD, Leidel SA. 2015. Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. *Cell* **161**: 1606-1618.

Neugebauer KM. 2002. On the importance of being co-transcriptional. *J Cell Sci* **115**: 3865-3871.

Oler AJ, Alla RK, Roberts DN, Wong A, Hollenhorst PC, Chandler KJ, Cassiday PA, Nelson CA, Hagedorn CH, Graves BJ et al. 2010. Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. *Nat Struct Mol Biol* **17**: 620-628.

Orioli A, Praz V, Lhote P, Hernandez N. 2016. Human MAF1 targets and represses active RNA polymerase III genes by preventing recruitment rather than inducing long-term transcriptional arrest. *Genome Res* **26**: 624-635.

Orphanides G, Reinberg D. 2002. A unified theory of gene expression. *Cell* **108**: 439-451.

Pattanakitsakul S, Zheng JH, Natsuume-Sakai S, Takahashi M, Nonaka M. 1992. Aberrant splicing caused by the insertion of the B2 sequence into an intron of the complement C4 gene is the basis for low C4 production in H-2k mice. *J Biol Chem* **267**: 7814-7820.

Pavon-Eternod M, Gomes S, Geslain R, Dai Q, Rosner MR, Pan T. 2009. tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res* **37**: 7268-7280.

Pelechano V, Steinmetz LM. 2013. Gene regulation by antisense transcription. *Nat Rev Genet* **14**: 880-893.

Persson H, Kvist A, Vallon-Christersson J, Medstrand P, Borg A, Rovira C. 2009. The non-coding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs. *Nat Cell Biol* **11**: 1268-1271.

Peterlin BM, Brogie JE, Price DH. 2012a. 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley interdisciplinary reviews RNA* **3**: 92-103.

Peterlin BM, Brogie JE, Price DH. 2012b. 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdiscip Rev RNA* **3**: 92-103.

Peterlin BM, Price DH. 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* **23**: 297-305.

Plotkin JB, Robins H, Levine AJ. 2004. Tissue-specific codon usage and the expression of human genes. *Proc Natl Acad Sci U S A* **101**: 12588-12591.

Policarpi C, Crepaldi L, Brookes E, Nitarska J, French SM, Coatti A, Riccio A. 2017. Enhancer SINEs Link Pol III to Pol II Transcription in Neurons. *Cell Rep* **21**: 2879-2894.

Ponicsan SL, Kugel JF, Goodrich JA. 2010. Genomic gems: SINE RNAs regulate mRNA production. *Curr Opin Genet Dev* **20**: 149-155.

Raha D, Wang Z, Moqtaderi Z, Wu L, Zhong G, Gerstein M, Struhl K, Snyder M. 2010. Close association of RNA polymerase II and many transcription factors with Pol III genes. *Proc Natl Acad Sci U S A* **107**: 3639-3644.

Rak R, Dahan O, Pilpel Y. 2018. Repertoires of tRNAs: The Couplers of Genomics and Proteomics. *Annu Rev Cell Dev Biol* **34**: 239-264.

Reina JH, Azzouz TN, Hernandez N. 2006. Maf1, a new player in the regulation of human RNA polymerase III transcription. *PLoS One* **1**: e134.

Renaud M, Praz V, Vieu E, Florens L, Washburn MP, l'Hote P, Hernandez N. 2014. Gene duplication and neofunctionalization: POLR3G and POLR3GL. *Genome Res* **24**: 37-51.

Rib L, Villeneuve D, Minocha S, Praz V, Hernandez N, Guex N, Herr W, Cycli XC. 2018. Cycles of gene expression and genome response during mammalian tissue regeneration. *Epigenetics Chromatin* **11**: 52.

Rubin CM, Kimura RH, Schmid CW. 2002. Selective stimulation of translational expression by Alu RNA. *Nucleic Acids Res* **30**: 3253-3261.

Saitsu H, Osaka H, Sasaki M, Takanashi J, Hamada K, Yamashita A, Shibayama H, Shiina M, Kondo Y, Nishiyama K et al. 2011. Mutations in POLR3A and POLR3B encoding RNA Polymerase III subunits cause an autosomal-recessive hypomyelinating leukoencephalopathy. *Am J Hum Genet* **89**: 644-651.

Sasaki T, Nishihara H, Hirakawa M, Fujimura K, Tanaka M, Kokubo N, Kimura-Yoshida C, Matsuo I, Sumiyama K, Saitou N et al. 2008. Possible involvement of SINEs in mammalian-specific brain formation. *Proc Natl Acad Sci U S A* **105**: 4220-4225.

Schmitt BM, Rudolph KL, Karagianni P, Fonseca NA, White RJ, Talianidis I, Odom DT, Marioni JC, Kutter C. 2014. High-resolution mapping of transcriptional dynamics across tissue development reveals a stable mRNA-tRNA interface. *Genome Res* **24**: 1797-1807.

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

Scott PH, Cairns CA, Sutcliffe JE, Alzuherri HM, McLees A, Winter AG, White RJ. 2001. Regulation of RNA polymerase III transcription during cell cycle entry. *J Biol Chem* **276**: 1005-1014.

Shearwin KE, Callen BP, Egan JB. 2005. Transcriptional interference--a crash course. *Trends Genet* **21**: 339-345.

Shilatifard A, Conaway RC, Conaway JW. 2003. The RNA polymerase II elongation complex. *Annual review of biochemistry* **72**: 693-715.

Sims RJ, 3rd, Belotserkovskaya R, Reinberg D. 2004. Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18**: 2437-2468.

Smale ST, Kadonaga JT. 2003. The RNA polymerase II core promoter. *Annual review of biochemistry* **72**: 449-479.

Smit AF, Riggs AD. 1995. MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. *Nucleic Acids Res* **23**: 98-102.

Sobala A, Hutvagner G. 2011. Transfer RNA-derived fragments: origins, processing, and functions. *Wiley Interdiscip Rev RNA* **2**: 853-862.

Sonawane AR, Platig J, Fagny M, Chen CY, Paulson JN, Lopes-Ramos CM, DeMeo DL, Quackenbush J, Glass K, Kuijjer ML. 2017. Understanding Tissue-Specific Gene Regulation. *Cell Rep* **21**: 1077-1088.

Sonenberg N, Hinnebusch AG. 2009. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**: 731-745.

Su WY, Li JT, Cui Y, Hong J, Du W, Wang YC, Lin YW, Xiong H, Wang JL, Kong X et al. 2012. Bidirectional regulation between WDR83 and its natural antisense transcript DHPS in gastric cancer. *Cell Res* **22**: 1374-1389.

Thiffault I, Wolf NI, Forget D, Guerrero K, Tran LT, Choquet K, Lavallee-Adam M, Poitras C, Brais B, Yoon G et al. 2015. Recessive mutations in POLR1C cause a leukodystrophy by impairing biogenesis of RNA polymerase III. *Nat Commun* **6**: 7623.

Treangen TJ, Salzberg SL. 2011. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* **13**: 36-46.

Tufarelli C, Stanley JA, Garrick D, Sharpe JA, Ayyub H, Wood WG, Higgs DR. 2003. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet* **34**: 157-165.

Uhler JP, Hertel C, Svejstrup JQ. 2007. A role for noncoding transcription in activation of the yeast PHO5 gene. *Proc Natl Acad Sci U S A* **104**: 8011-8016.

Upadhya R, Lee J, Willis IM. 2002. Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. *Molecular cell* **10**: 1489-1494.

Van Bortle K, Phanstiel DH, Snyder MP. 2017. Topological organization and dynamic regulation of human tRNA genes during macrophage differentiation. *Genome Biol* **18**: 180.

van Werven FJ, Neuert G, Hendrick N, Lardenois A, Buratowski S, van Oudenaarden A, Primig M, Amon A. 2012. Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. *Cell* **150**: 1170-1181.

van Zon A, Mossink MH, Houtsmuller AB, Schoester M, Scheffer GL, Scheper RJ, Sonneveld P, Wiemer EA. 2006. Vault mobility depends in part on microtubules and vaults can be recruited to the nuclear envelope. *Exp Cell Res* **312**: 245-255.

Vannini A, Cramer P. 2012. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol Cell* **45**: 439-446.

Vassetzky NS, Kramerov DA. 2013. SINEBase: a database and tool for SINE analysis. *Nucleic Acids Res* **41**: D83-89.

Vigetti D, Deleonibus S, Moretto P, Bowen T, Fischer JW, Grandoch M, Oberhuber A, Love DC, Hanover JA, Cinquetti R et al. 2014. Natural antisense transcript for hyaluronan synthase 2 (HAS2-AS1) induces transcription of HAS2 via protein O-GlcNAcylation. *J Biol Chem* **289**: 28816-28826.

Walter P, Blobel G. 1982. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* **299**: 691-698.

Wang H, Iacoangeli A, Lin D, Williams K, Denman RB, Hellen CU, Tiedge H. 2005. Dendritic BC1 RNA in translational control mechanisms. *J Cell Biol* **171**: 811-821.

Wei Y, Tsang CK, Zheng XFS. 2009. Mechanisms of regulation of RNA polymerase III-dependent transcription by TORC1. *The EMBO journal* **28**: 2220-2230.

Wei YY, Chen HT. 2018. Functions of the TFIIE-Related Tandem Winged-Helix Domain of Rpc34 in RNA Polymerase III Initiation and Elongation. *Mol Cell Biol* **38**.

White RJ. 2004. RNA polymerase III transcription and cancer. *Oncogene* **23**: 3208-3216.

-. 2005. RNA polymerases I and III, growth control and cancer. *Nat Rev Mol Cell Biol* **6**: 69-78.

-. 2008. RNA polymerases I and III, non-coding RNAs and cancer. *Trends Genet* **24**: 622-629.

White RJ, Gottlieb TM, Downes CS, Jackson SP. 1995. Cell cycle regulation of RNA polymerase III transcription. *Mol Cell Biol* **15**: 6653-6662.

Willis IM, Moir RD. 2018. Signaling to and from the RNA Polymerase III Transcription and Processing Machinery. *Annu Rev Biochem*.

Winter AG, Sourvinos G, Allison SJ, Tosh K, Scott PH, Spandidos DA, White RJ. 2000. RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. *Proc Natl Acad Sci U S A* **97**: 12619-12624.

Woiwode A, Johnson SA, Zhong S, Zhang C, Roeder RG, Teichmann M, Johnson DL. 2008. PTEN represses RNA polymerase III-dependent transcription by targeting the TFIIIB complex. *Mol Cell Biol* **28**: 4204-4214.

Wong RC, Pollan S, Fong H, Ibrahim A, Smith EL, Ho M, Laslett AL, Donovan PJ. 2011. A novel role for an RNA polymerase III subunit POLR3G in regulating pluripotency in human embryonic stem cells. *Stem Cells* **29**: 1517-1527.

Wu CC, Lin YC, Chen HT. 2011. The TFIIF-like Rpc37/53 dimer lies at the center of a protein network to connect TFIIIC, Bdp1, and the RNA polymerase III active center. *Mol Cell Biol* **31**: 2715-2728.

Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. 1997. Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature* **389**: 745-749.

Xi Q, Gao N, Zhang X, Zhang B, Ye W, Wu J, Zhang X. 2014. A natural antisense transcript regulates acetylcholinesterase gene expression via epigenetic modification in Hepatocellular Carcinoma. *Int J Biochem Cell Biol* **55**: 242-251.

Yang Z, Zhu Q, Luo K, Zhou Q. 2001. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**: 317-322.

Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, Gil J, Walsh MJ, Zhou MM. 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell* **38**: 662-674.

Yee NS, Gong W, Huang Y, Lorent K, Dolan AC, Maraia RJ, Pack M. 2007. Mutation of RNA Pol III subunit rpc2/polr3b Leads to Deficiency of Subunit Rpc11 and disrupts zebrafish digestive development. *PLoS Biol* **5**: e312.

Yeganeh M, Praz V, Carmeli C, Villeneuve D, Rib L, Guex N, Herr W, Delorenzi M, Hernandez N, Cycli Xc. 2018. Differential regulation of RNA polymerase III genes during liver regeneration. *Nucleic Acids Res*.

Yeganeh M, Praz V, Cousin P, Hernandez N. 2017. Transcriptional interference by RNA polymerase III affects expression of the Polr3e gene. *Genes Dev* **31**: 413-421.

Yu P, Ma D, Xu M. 2005. Nested genes in the human genome. *Genomics* **86**: 414-422.

Zhai W, Comai L. 2000. Repression of RNA polymerase I transcription by the tumor suppressor p53. *Molecular and cellular biology* **20**: 5930-5938.

Zhou Q, Li T, Price DH. 2012. RNA polymerase II elongation control. *Annu Rev Biochem* **81**: 119-143.

Zhu XT, Yuan JH, Zhu TT, Li YY, Cheng XY. 2016. Long noncoding RNA glypican 3 (GPC3) antisense transcript 1 promotes hepatocellular carcinoma progression via epigenetically activating GPC3. *FEBS J* **283**: 3739-3754.