Aus dem Lehrstuhl für Molekularbiologie im Biomedizinischen Centrum (BMC),
Institut der Ludwig-Maximilians-Universität München
Direktor: Prof. Dr. Peter B. Becker

# Transcriptional regulation of endogenous retroviruses by chromatin and RNA processing


Dissertation zum Erwerb des Doktorgrades der
Naturwissenschaften (Dr. rer. nat.) an der
Medizinischen Fakultät der
Ludwig-Maximilians-Universität München
vorgelegt von

**Irina Shcherbakova**

aus Arkhangelsk, Russland

2022

Mit Genehmigung der Medizinischen Fakultät

der Universität München

| | |
|---|---|
| Betreuer: | Prof. Dr. rer. nat. Gunnar Schotta |
| Zweitgutachter: | Prof. Dr. rer. nat. Stefan Stricker |
| Dekan: | Prof. Dr. med. Thomas Gudermann |
| Tag der mündlichen Prüfung: | 30.03.2023 |

**Eidesstattliche Versicherung**

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Thema

**“Transcriptional regulation of endogenous retroviruses by chromatin and RNA processing”**

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Munich, 28.06.2023 Irina Shcherbakova

_____ _____

Ort, Datum Unterschrift, Irina Shcherbakova

While carrying out my PhD thesis I collaborated with colleagues to support scientific projects, which led to the following publications:

Srinivasachar Badarinarayan, S., **Shcherbakova, I**., Langer, S., Koepke, L., Preising, A., Hotter, D., Kirchhoff, F., Sparrer, K.M.J., Schotta, G., Sauter, D. HIV-1 infection activates endogenous retroviral promoters regulating antiviral gene expression. Nucleic Acids Res. 2020 Nov 4; 48(19):10890-10908.

Wang, Z., Fan, R., Russo, A., Cernilogar, F.M., Nuber, A., Schirge, S., **Shcherbakova, I.,** Dzhilyanova, I., Ugur, E., Anton, T., Richter, L., Leonhardt, H., Lickert, H., Schotta, G. Dominant role of DNA methylation over H3K9me3 for IAP silencing in endoderm. Nature Communications. 2022 Sep 19; 13(1):5447.

Angerilli, A., Tait, J., Berges J., **Shcherbakova, I**., Pokrovsky, D., Schauer, T., Smialowski, P., Hsam, O., Mentele, E., Nicetto, D., Rupp, R.A.W. The histone H4K20 methyltransferase SUV4-20H1/KMT5B is required for multiciliated cell differentiation in Xenopus. Life Sci Alliance. 2023 Apr 28; 6(7):e202302023.

# I. Table of contents

# II. Abstract

Several reports have shown that post-transcriptional silencing machinery in *D. melanogaster, C. elegance, C. cerevisiae* targets repetitive elements' transcripts and actively degrades them. Such pathways include spliceosome recognition of aberrant transcripts and their subsequent degradation by the exosome complex coupled with heterochromatin formation. To date, the importance of post-transcriptional pathways in mammalian ERV regulation is poorly understood.

We are the first who documented that ERV transcripts have fast turnover in humans, suggesting that these loci are actively degraded by post-transcriptional silencing pathways.

We found that treatment with DNA methylation inhibitor 5-Aza results in prominent H3K9me3 reduction on ERV loci, which could be explained by the interconnection of DNA methylation and H3K9me3 repressing pathways. Therefore, H3K9me3 reduction could additionally contribute to ERV derepression upon 5-Aza treatment. Moreover, we found that 5-Aza and Panobinostat treatments led to transcriptional ERV derepression, rather than increased ERV transcripts stability. In addition, 5-Aza treatment of leukemic cells results in striking upregulation of a few specific ERV loci with coding potential, which could potentially induce viral mimicry and contribute to the immunomodulatory activity of the 5-Aza epigenetic drug.

Interestingly, knockdowns of both splicing factors SF3A3 and SRSF3 and exosome exonuclease DIS3 result in transcriptional derepression of ERV loci, suggesting that their function might be coupled with transcriptional repressing machinery. These findings suggest that other transcription repressive histone modifications might be linked to studied splicing factors and exonuclease DIS3, such as, for example, H4K20me3 and H3K27me3. Notably, the vast majority of upregulated ERVs already have basal transcription under normal condition. This suggests that basal transcription might be beneficial for the cell since it helps to recruit repressive complexes to ERVs, preventing their high expression. Moreover, SF3A3, SRSF3, and DIS3 knockdowns lead to an increased proportion of bidirectionally transcribed ERVs and high expression of full-length ERV loci, which could potentially contribute to dsRNA production and viral mimicry inducing an immune response in cancer cells. Therefore, possible production of ERV-derived dsRNA upon reduction of splicing factors SRSF3 might contribute to previously observed immunomodulatory effect and associated cancer cells death.

Finally, we found that the knockdown of splicing factor SF3A3 leads to decreased stability of ERVs. Moreover, it results in particularly strong transcript stabilization of H3K9me3-covered loci, such as ERVs and KRAB-ZNF genes. However, DIS3 knockdown does not affect the stability of these loci. This finding implies that splicing factor SF3A3 might specifically target

transcripts originating from H3K9me3 covered loci for subsequent degradation independently from the exosome exonuclease DIS3 degradation pathway.

# III. Zusammenfassung

Mehrere Berichte haben gezeigt, dass die posttranskriptionelle Silencing-Maschinerie bei *D. melanogaster, C. elegance, C. cerevisiae* auf die Transkripte sich wiederholender Elemente abzielt und diese aktiv abbaut. Solche Wege umfassen die Spleißosom-Erkennung von aberranten Transkripten und ihren anschließenden Abbau durch den Exosom-Komplex gekoppelt mit der Heterochromatin-Bildung. Bis heute ist die Bedeutung von posttranskriptionellen Signalwegen bei der ERV-Regulation von Säugetieren nur unzureichend verstanden.

Wir sind die Ersten, die dokumentiert haben, dass ERV-Transkripte beim Menschen einen schnellen Umsatz aufweisen, was darauf hindeutet, dass diese Loci aktiv durch posttranskriptionelle Silencing-Pfade abgebaut werden.

Wir fanden heraus, dass die Behandlung mit dem DNA-Methylierungsinhibitor 5-Aza zu einer deutlichen Reduktion von H3K9me3 an den ERV-Loci führt, was durch die Verbindung von DNA-Methylierung und H3K9me3-Repressionswegen erklärt werden könnte. Daher könnte eine H3K9me3-Reduktion zusätzlich zur ERV-Derepression nach 5-Aza-Behandlung beitragen. Darüber hinaus fanden wir, dass 5-Aza- und Panobinostat-Behandlungen eher zu einer transkriptionellen ERV-Derepression als zu einer erhöhten Stabilität der ERV-Transkripte führten. Darüber hinaus führt die 5-Aza-Behandlung von Leukämiezellen zu einer bemerkenswerten Hochregulation einiger weniger spezifischer ERV-Loci mit kodierendem Potenzial, die möglicherweise virale Mimikry induzieren und zur immunmodulatorischen Aktivität des epigenetischen 5-Aza-Medikaments beitragen könnte.

Interessanterweise führen Knockdowns der beiden Spleißfaktoren SF3A3 und SRSF3 sowie der Exosomen-Exonuklease DIS3 zu einer transkriptionellen Derepression der ERV-Loci, was darauf hindeutet, dass ihre Funktion mit der transkriptionellen Repressionsmaschinerie gekoppelt sein könnte. Diese Ergebnisse legen nahe, dass andere transkriptionsrepressive Histonmodifikationen mit untersuchten Spleißfaktoren und der Exonuklease DIS3 in Verbindung gebracht werden könnten, wie beispielsweise H4K20me3 und H3K27me3. Bemerkenswert ist, dass die überwiegende Mehrheit der hochregulierten ERVs unter normalen Bedingungen bereits eine basale Transkription aufweist. Dies deutet darauf hin, dass die basale Transkription für die Zelle von Vorteil sein könnte, da sie dabei hilft, repressive Komplexe für ERVs zu rekrutieren und ihre hohe Expression zu verhindern. Darüber hinaus führen SF3A3-, SRSF3- und DIS3-Knockdowns zu einem erhöhten Anteil an bidirektional transkribierten ERVs und einer hohen Expression von ERV-Loci in voller Länge, die möglicherweise zur dsRNA-Produktion und viralen Mimikry beitragen könnten, was eine

Immunantwort in Krebszellen auslöst. Daher könnte eine mögliche Produktion von ERV-abgeleiteter dsRNA nach Reduktion des Spleißfaktors SRSF3 zu einer zuvor beobachteten immunmodulatorischen Wirkung und dem damit verbundenen Absterben von Krebszellen beitragen.

Schließlich fanden wir heraus, dass der Knockdown des Spleißfaktors SF3A3 zu einer verringerten Stabilität von ERVs führt. Darüber hinaus führt es zu einer besonders starken Transkriptstabilisierung von H3K9me3-bedeckten Loci, wie ERVs und KRAB-ZNF-Genen. Der DIS3-Knockdown hat jedoch keinen Einfluss auf die Stabilität dieser Loci. Dieser Befund deutet darauf hin, dass der Spleissfaktor SF3A3 spezifisch auf Transkripte abzielen könnte, die von H3K9me3-bedeckten Loci stammen, sowie für deren nachfolgenden Abbau unabhängig vom Abbauweg der Exosomen-Exonuklease DIS3.

# 1. Introduction

## 1.1 Epigenetic mechanisms in transcriptional regulation

Epigenetics is the study of heritable modifications that alter gene activity without changing the DNA sequence. DNA methylation and chromatin modification are the main contributors to epigenetics. Chromatin is the complex of proteins, called histones, which package DNA in a compact form inside the cell nucleus. Four core histones (H3, H4, H2A, H2B) form the nucleosome, around which 147 base pairs of DNA are wrapped. The H3 and H4 histones have flexible unstructured “tails”, which can be modified at several places with modifications, such as methylation, acetylation, phosphorylation, ubiquitylation, and sumoylation. These histone modifications affect gene expression by altering chromatin structure or recruiting various histone modifiers. Specific histone modifications can cause either chromatin compaction leading to gene silencing or chromatin loosening, associated with active transcription. Compacted chromatin is called heterochromatin, while open chromatin is called euchromatin. Generally, euchromatin is associated with high levels of acetylation, for example, acetylation of the histone 3 tail at lysine 9 (H3K9). H3 and H4 histone acetylation is performed by histone acetyltransferases. In addition, histone acetylation can be actively removed by enzymes called histone deacetylases (HDACs). More than 20 histone acetyltransferases and 10 deacetylases have been identified in humans. Heterochromatic state, in contrast, is associated with low levels of acetylation and high levels of methylation modifications, such as methylation of the histone 3 tail at lysine 9 (H3K9), histone 3 tail at lysine 27 (H3K27), and histone 4 tail at lysine 20 H4K20. Although, some methylation modifications of histones are associated with active transcription (Kouzarides, 2007).

Heterochromatin can be subdivided into two types: facultative and constitutive. Facultative heterochromatin dynamically silences cell-type-specific genes. In contrast, constitutive heterochromatin is believed to occur across the same genomic regions in different cell types. Specifically, constitutive heterochromatin silences highly repeated DNA sequences such as tandem-repeat satellites at pericentromeric regions and telomeres, transposable elements, and endogenous retroviruses. Facultative heterochromatin is associated with DNA methylation and H3K27me3, placed by DNA methyltransferases (DNMTs) and polycomb repressive complex (PRC2), respectively. Constitutive chromatin is associated with H3K9me3 and H4K20me3 and placed by various methyltransferases (Saksouk et al., 2015). Although H3K9me3 is believed to be associated with constitutive heterochromatin, this mark was demonstrated to be dynamic and allow basal transcription under certain conditions (Bühler and Moazed, 2007; Nicetto and Zaret, 2019; Vakoc et al., 2005). Taken together, chromatin

modifications are dynamic, allowing modulation of accessible and closed chromatin states, which regulates the transcription activity of affected genomic loci (Yang and Seto, 2007).

## 1.2 Transposable elements

Transposable elements (TE) are low-complexity DNA sequences that can change their position within a genome. TEs are present in most known organisms. In eukaryotes, TEs usually represent a substantial fraction of the genome. Specifically, in humans, TEs make up 41% of the genome (Guffanti et al., 2019). TEs are divided into two classes based on their ability to move in the genome: (1) retrotransposons moving through a copy and paste mechanism and (2) DNA transposons moving through the cut and paste mechanism. In turn, retrotransposons are subdivided based on the presence of long terminal repeat (LTR) – into LTR and non-LTR retrotransposons. The integration of LTR retrotransposons is catalyzed by a retroviral-like integrase, whereas non-LTR retrotransposons, which consist of long and short interspersed nuclear elements (LINEs and SINEs), are catalyzed by reverse transcriptase (Bourque et al., 2018).

LTR retrotransposons are remnants of retroviral infections, which have been acquired by the genome through various exogenous retroviruses integrations to a primate germline over the last millions of years (Bannert and Kurth, 2006). Hence, LTR retrotransposons are also called endogenous retroviruses (ERVs). About 10 % of mammalian genomes consist of ERVs. Specifically, 8% of the human genome is composed of endogenous retroviruses (Lander et al., 2001). ERVs, like exogenous retroviruses, have a typical proviral structure compromising of gag, pro, pol, and env genes flanked by LTRs. The pol, pro, and env genes encode for reverse transcriptase, protease, and envelope surface proteins, respectively. The gag gene encodes viral core structural proteins, such as matrix, capsid, and nucleocapsid. LTR regions have a regulatory significance for viral genes' expression and can act as promoters, enhancers, and polyadenylation signals (Grandi and Tramontano, 2018). Over time, ERVs acquired mutations and truncations, preserving original provirus features to a highly variable extent. In the end, human ERVs lost their ability to transpose, but some ERV loci are still transcriptionally active and able to produce viral proteins and virus-like particles (Johnson, 2019).

Based on the phylogenetic relationship and sequence similarity, human ERVs can be classified into several groups or families, such as HERV-H, HERV-K, HERV-Fc, HERV-W, HERV-L. HERV-K is one of the best-preserved families of human endogenous retroviruses. Many HERV-K loci preserve relatively intact open reading frames and are known to produce viral proteins (Hohn et al., 2013). HERV-K, together with HERV-Fc, is integrated into the genome most recently. HERV-Fc is a very small family of endogenous retroviruses, which is

able to produce gag and env proteins (Bénit et al., 2003; Laska et al., 2013). HERV-H is the most abundant family, with over 1000 copies. Interestingly, the gag gene is relatively complete in the majority of full-length HERV-H loci, while pol and env proteins are usually fragmented (Gemmell et al., 2019). HERV-W is another family of human ERVs which is able to produce viral proteins, namely envelope proteins (Ruprecht and Mayer, 2019). HERV-L is one of the most ancient groups, which does not show remarkable activity (Vargiu et al., 2016). Besides sequence variability between families, many ERV loci exhibit high variability inside families, making them a rich source of host genetic variability.

### 1.3 Physiological role of endogenous retroviruses

Before, ERVs were considered by many as selfish elements not useful for the cell. Often, they were described as “junk DNA” (Nelson et al., 2003). Later, it became clear that ERVs had a long co-evolution with the host genome, which led to ERVs co-option by the host to fulfil new cellular functions, both coding and regulatory. Nowadays, ERVs are considered an important source of genomic variation for the evolution of many species, including humans (Jern and Coffin, 2008).

Endogenous retroviruses can serve as alternative promoters, enhancer elements, or alternative splice sites. ERVs often serve as enhancers and promoters for the nearby genes because of their LTR region. LTR regions of endogenous retroviruses have a high amount of transcription factor binding sites (TFBS) and bidirectional transcription start sites (TSS), which are not only able to regulate ERV expression but also an expression of nearby host genes (Payer and Burns, 2019). Some ERVs still contain splice donor or acceptor sites, which can cause alternative splicing of a coding gene if such ERV is inside the gene or in close vicinity. In addition, transcriptionally active ERVs with splice sites are able to form fusion and chimeric transcripts (Mitsuhashi et al., 2021; Peaston et al., 2004).

Advances in next-generation sequencing (NGS) techniques allowed the investigation of repetitive elements, including ERVs, on a genome-scale level. Up to 30% of all active promoters in human cell lines overlap with repetitive elements (Lynch et al., 2015). In addition, up to 20 % of binding sites of many transcriptional factors are derived from LTR regions (Sundaram et al., 2014). For example, around 30% of the tumor suppressor protein (TP53) binding sites are derived from LTR elements of ERVs (Wang et al., 2007).

ERVs were shown to regulate the immune response. Similar to TP53, binding sites of interferon-gamma (IFNγ) - inducible transcription factors are enriched in several human ERV families (Chuong et al., 2016). Interestingly, some current exogenous retroviruses, including HIV, possess IFN-inducible cis-regulatory elements (Sgarbanti et al., 2008). Therefore, viral

molecular adaptations for transcription and replication over time could have been recycled by the host to shape immune response (Randall and Goodbourn, 2008). In collaboration with Sauter's lab, we illustrated another example of how human cells make use of endogenous retroviruses to improve immune responses against HIV1. We showed that HIV1 infection induces activation of multiple ERV copies in primary CD4+ T cells, including upregulation of LTR elements upstream of antiviral genes GBP2 and GBP5, which regulate inhibition of viral replication. Upregulation of these LTR elements leads to increased cytokine responsiveness of GBP2 and GBP5 genes (Srinivasachar Badarinarayan et al., 2020).

### 1.4 Endogenous retroviruses as therapeutic targets in cancer

Recent developments in the field of cancer immune-therapies revealed ERVs as promising targets. Two main areas of ERV – based cancer treatments currently are under investigation: (1) ERV induces a state of viral mimicry, and (2) ERV proteins act as tumor-associated antigens and are targets for cancer cells elimination.

The viral mimicry approach is based on increased production of ERV double-stranded RNA (dsRNA) or ERV proviral proteins upon drug-induced ERV upregulation. dsRNA activates interferon responses through the RNA-sensing innate immune network, and proviral proteins serve as antigens, which mimic viral infection and lead to immunogenic cancer cells death (Attermann et al., 2018; Cuellar et al., 2017; Lee et al., 2020a; Vergara Bermejo et al., 2020). ERV upregulation is associated with epigenetic-based drugs - DNA demethylating and histone deacetylase inhibiting (HDACi), such as 5-Aza and Panobinostat (Attermann et al., 2018). In addition, increased production of ERV-derived dsRNA and proteins was observed upon cancer cells treatments with 5-Aza (Chiappinelli et al., 2015; Krug et al., 2019; Roulois et al., 2015). These drugs are utilized to treat various cancers, such as colorectal, ovarian, leukemia, and hepatocellular carcinoma (Ishak et al., 2018). Drug-induced viral mimicry could contribute to potent immunomodulatory activity caused by these epigenetic drugs. However, chromatin alterations on ERVs and ERV transcript stability were not addressed upon treatments with epigenetic anticancer drugs

Some cancers have dysregulated epigenetic profiles, which are associated with the already elevated production of specific ERV proteins (Grabski et al., 2019). The utilization of such endogenous retroviral proteins was successfully tested, for example, in chimeric antigen receptor T-cell therapy. This strategy was based on HERV-K-specific sequence introduction into T-cell receptor and recognition of HERV-K envelope protein inside cancer cells. T- cells with HERV-K chimeric antigen receptor were able not only to inhibit breast and melanoma cancers growth but also prevent metastasis to other organs (Krishnamurthy et al., 2015; Zhou et al., 2015).

ERVs activation on transcriptional and protein levels often correlates with autoimmune diseases, neurological disorders, and some cancers (Kassiotis, 2014; Kury et al., 2018; Volkman and Stetson, 2014). Albeit it is unclear if a causative link between ERV upregulation and such diseases exist, undoubtedly, proper regulation of ERV activity is crucial for cell function. Thus, studying the regulatory mechanisms of endogenous retroviruses will not only help better understand their control in normal development but also shed light on the mechanisms of ERV dysregulation in disease states.

## 1.5 Silencing mechanisms of endogenous retroviruses

Cells have developed various strategies to sense ERVs and restrict their excessive activity by (1) establishing repressive chromatin across retroviral sequences and (2) targeting and eliminating viral RNA post-transcriptionally. In mammals, transcriptional ERV silencing through heterochromatin formation was investigated to a greater extent compared to post-transcriptional ERV silencing. Although, studies focusing on other model organisms, such as *D. melanogaster, C. elegance, C. cerevisiae*, have shown that post-transcriptional silencing machinery actively targets and degrades TE transcripts (Eberle et al., 2015; Fischer et al., 2013; Reyes-Turcu et al., 2011; Sugiyama et al., 2016). An emerging number of studies revealed that in mammals similar pathways act on retrotransposons as well (Berrens et al., 2017; Carter et al., 2020b; Li et al., 2014b).

Presumably, transcriptional and post-transcriptional silencing mechanisms were developed by host cells to restrict exogenous retroviruses, but later these mechanisms were utilized for silencing of newly integrated endogenous retroviruses. In the past, ERVs were considered to be largely inactive. With the advance in NGS techniques, regulation of ERV expression was proven to be more dynamic than initially assumed. As was discussed above, ERV copies show cell type-specific activity similar to genes. ERVs are co-opted by host cells to fulfill various physiological functions. Moreover, their activity correlates with various disease states. Therefore, the silencing of endogenous retroviruses can be specifically relieved in certain cell types or upon exogenous stimuli.

### 1.5.1 Transcriptional silencing mechanisms

ERV silencing mechanisms in mammalian cells have been mainly studied in the mouse. Human ERVs are truncated and mutated to a greater extent as compared to a mouse. However, silencing mechanisms appear to be similar. A large proportion of ERVs is silenced by the establishment of repressive heterochromatin. Heterochromatin formation on ERVs is associated with DNA methylation and repressive posttranslational histone modifications, such as H3K9 or H4K20 trimethylations (Day et al., 2010; Machida et al., 2018). In addition,

H3K9me3 has been found to be associated with both exogenous and endogenous retroviruses (Wang et al., 2016; Zhu et al., 2018). Five H3K9-specific methyltransferases were described in mammals - SETDB1, SUV39H1, SUV39H2, GLP, and G9a. SETDB1 and SUV39H1/2 enzymes contribute to the formation of both H3K9me2 and H3K9me3, while GLP and G9a regulate H3K9me1 and H3K9me2 (Rea et al., 2000; Schultz et al., 2002). According to studies performed in mice, SETDB1 is highly enriched on ERVs and plays a central part in H3K9me3-dependent ERV repression, compared to other H3K9 methyltransferases (Karimi et al., 2011; Matsui et al., 2010). SETDB1, together with corepressor TRIM28, is recruited to ERVs by Krüppel-associated box zinc fingers (KRAB-ZNF) proteins. These KRAB-ZNF proteins recognize ERV sequences by their DNA-binding tandem arrays of zinc fingers and recruit repressor proteins to retroviral loci through the KRAB repression domain (Gautam et al., 2017).

KRAB-ZNFs is one of the largest transcriptional factor families in mice and humans, consisting of approximately 200 and 300 genes, respectively. Most KRAB-ZNFs binds preferentially to endogenous retroviruses (Helleboid et al., 2019). KRAB-ZNFs showed strikingly rapid evolution in mammals. Interestingly, a strong correlation was observed between the number of ERVs and the number of KRAB-ZNFs genes in 16 mammalian genomes (Thomas and Schneider, 2011), which points out to a high possibility that they co-evolve with each other and ERVs were the main drivers of KRAB-ZNF diversification. The subset of KRAB-ZNFs has an additional SCAN domain. Interestingly, it has striking homology to a part of the gag capsid from endogenous retroviruses. In addition, it is structurally similar to the HIV capsid. SCAN domain could be derived from retrotransposon and used in the past to target KRAB-ZFPs to retroviral capsids. However, the human SCAN domain from KRAB-ZNFs is not predicted to bind ERVs (Emerson and Thomas, 2011; Ivanov et al., 2005). Therefore, the SCAN domain function in humans is unclear. Interestingly, the KRAB-ZNFs chromatin landscape is distinct from the rest of the genome. Specifically, many KRAB- ZNFs have active histone modifications in promoter regions and, at the same time, repressive - H3K9me3 and H4K20me3 - histone modifications in 3’ prime regions (Rao et al., 2014; Vogel et al., 2006). It is an interesting feature, taking into account that KRAB-ZNFs are themselves responsible for H3K9me3 formation at retroviral loci.

Another important player in transcriptional ERV silencing and heterochromatin formation is DNA methylation. DNA methylation often does not function in isolation to silence transposable elements. The methyl-CpG-binding domain (MBD) protein family binds to methylated DNA and recruits various chromatin remodelers associated with silencing, including H3K9me3-related proteins (Espada et al., 2004b; Fujita et al., 2003; Rose and Klose, 2014) Therefore, DNA methylation could direct H3K9 methylation position through MBD proteins. Interestingly, H3K9

methylation also can affect DNA methylation maintenance, for example, through UHRF1-mediated DNA methyltransferases (DNMT1) recruitment to H3K9me3 regions (Liu et al., 2013; Zhao et al., 2016). The predominant importance of either H3K9me3 or DNA methylation in ERV silencing is context-dependent and may differ in different cell types (Geis and Goff, 2020; Karimi et al., 2011; Kato et al., 2018; Sharif et al., 2016). In addition, the evolutionary age of ERVs correlates with a mechanism of their silencing. Specifically, young ERVs are CpG-rich, while evolutionary older ERVs have lower CpG content. In this regard, young ERVs are predominantly suppressed by DNA methylation, while older ERVs - by H3K9me3 (Ohtani et al., 2018).

### 1.5.2 Post-transcriptional silencing mechanisms

Methylation of H3K9 is a highly conserved mechanism of heterochromatin-based transposable elements silencing from unicellular to multicellular organisms, including not only mammals but also *S. pombe*, *D. melanogaster,* and *C. elegans* (Ahringer and Gasser, 2018; Lee et al., 2020c; Penke et al., 2016). Interestingly, yeast, drosophila, and worm species actively utilize post-transcriptional machinery to silence retrotransposons. Emerging studies report the presence of similar pathways in mammals. Although still, post-transcriptional silencing mechanisms of endogenous retroviruses in humans are poorly understood. In *S. pombe*, *D. melanogaster,* and *C. elegans,* transposon transcripts are sensed through various pathways, such as small spliceosome recognition, nonsense-mediated decay, RNA interference, and chromatin landscape of transposon loci. Then, after recognition by surveillance machinery, such transcripts are usually degraded by exosome - 3′–5′ exonuclease complex (Berrens et al., 2017; Dumesic et al., 2013; Mendell et al., 2004; Reyes-Turcu et al., 2011; Robert et al., 2005; Warf et al., 2012; Zhang et al., 2012).

To identify RNA processing genes involved in ERV silencing in mammals, we re-examined ERV silencing screens in mice from externally published data sets (Chelmicki et al., 2021; Fukuda et al., 2018; Yang et al., 2015) and data from our lab (Sadic et al., 2015). These genome-wide screens are based on CRISPR-mediated knockdown or siRNA-mediated knockout of large gene subsets in the cell line of interest. Then genes, which reduction is associated with the desired phenotype are selected. In these ERV silencing screens, the derepression of ERV-based reporters served as a read-out. Interestingly, after re-examining ERV silencing screens, we detected many RNA processing factors. Top hits appeared to be splicing regulators, such as SF3A3, SRSF1, SRSF9, and SRSF11, together with nuclear exosome exonuclease RRP6, suggesting that splicing factors and nuclear exosome could affect ERV transcript levels in mammals (Figure 1.A).

To investigate if RNA processing factors would associate with known chromatin-associated ERV silencing factors, we examined MORC3 and ATF7IP immunoprecipitation mass spectrometry (IP-MS) experiments in mouse embryonic stem cells (Groh et al., 2021). IP-MS method detects the protein of interest with all other proteins associated with it. The first protein - ATF7IP, is an interaction partner of SETDB1 and is crucial for ERV silencing in mice (Tsusaka et al., 2019). The second protein is MORC3, a critical regulator of DAXX-mediated H3.3 incorporation into ERV regions (Groh et al., 2021). Surprisingly, many splicing factors and the nuclear exosome exonuclease DIS3 were highly enriched in both experiments (Figure 1.B). Suggesting that chromatin-associated ERV repressor factors interact with both splicing and nuclear exosome complex in mice.

To investigate the role of human splicing factors in the silencing of heterochromatin-associated retrotransposons, namely endogenous retroviruses, we focused on SF3A3 and SRSF3 genes. Both splicing factors interact with chromatin-associated ERV silencers ATF7IP and MORC3. In addition, SF3A3 deletion and deletion of several SRSF genes leads to ERV derepression, according to ERV silencing screens in mice (Figure 1).

Therefore, splicing factors and nuclear exosome might be involved in post-transcriptional ERV silencing in mammals. Indeed, spliceosome machinery and exosome complex were found to be part of transposon repressing pathways in other model organisms, which will be reviewed in the following chapters.

A RNA processing factors detected in ERV silencing screens

| Gene | ERV silencing screen |
|---|---|
| ***Spliceosome components*** | |
| Bcas2 | Yang 2015 |
| Khsrp | Groh |
| Prpf3 | Groh |
| Sf3a3 | Groh; Chelmicki 2021 |
| Srsf1 | Yang 2015; Chelmicki 2021 |
| Srsf3 | Chelmicki 2021 |
| Srsf7 | Chelmicki 2021 |
| Srsf9 | Groh |
| Srsf11 | Yang 2015 |
| Xab2 | Fukuda 2018 |
| ***Nuclear exosome components*** | |
| Rrp6 | Groh; Yang 2015 |
| Rrp42 | Chelmicki 2021 |

B RNA processing factors detected in IP-MS of chromatin regulators

| Gene | Morc3 IP (fold enrichment) | Atf7ip IP (fold enrichment) |
|---|---|---|
| ***Spliceosome components*** | | |
| Srsf3 | 328 | 284 |
| Srsf7 | 51 | 33 |
| Srsf2 | 21 | 13 |
| Sf3a3 | 36 | 23 |
| Sf3b5 | 29 | 21 |
| Sf3a2 | 12 | 19 |
| Sf3b4 | 28 | 12 |
| ***Nuclear exosome components*** | | |
| Dis3 | 92 | 98 |

**Figure 1. RNA processing components interact with heterochromatin proteins and their deletion increase transcript abundances of mammalian ERVs.**

A. RNA processing factors detected in mammalian ERV silencing factors screens.

B. IP-MS experiments for Morc3 and Atf7ip - two ERV repressors associated with heterochromatin - revealed strong enrichment for splicing factors and the nuclear exosome.

### 1.6 Transcripts processing

mRNA, produced from DNA template as nascent transcripts, undergo complex processing before they become mature mRNA. Such processing, performed by many protein complexes, typically consists of splicing, 5' capping, and polyadenylation of RNA tail (poly (A) tailing), which results in stable mature and ready for translation mRNA. Then, mRNA is imported into the cytoplasm to be translated into protein by a ribosome. At all steps, transcripts quality is checked by RNA surveillance machinery, which targets aberrant transcripts for degradation (Coppin et al., 2018).

Splicing is an important step of mature mRNA production. Initial RNA, transcribed from DNA template, needs to undergo removal of intron sequences and exons joining through splicing. Alternative splicing is a common phenomenon. Precursor mRNA (pre-mRNA) molecule splices at different splice sites called junctions. As a result, a variety of mature mRNA molecules or transcript isoforms with different combinations of exons are produced (Gallego-Paez et al., 2017). Splicing of pre-mRNA occurs in the nucleus co-transcriptionally and is catalyzed by a highly dynamic, in terms of both conformation and composition, spliceosome complex (Will and Lührmann, 2011).

The spliceosome is highly conserved complex across diverse eukaryotes (Jurica and Moore, 2003). The spliceosome is a huge complex composed of small nuclear RNAs and dozens of proteins, such as small nuclear ribonucleoproteins (snRNPs), RNA helicases, and various splicing factors, such as serine-/arginine-rich splicing factors (SRSFs) and splicing factors 3 (SF3s) (Zhang et al., 2018). First, snRNPs spliceosome subunits assemble on pre-mRNA. Then, the spliceosome is activated by conformational rearrangements to perform an RNA splicing reaction. This reaction is regulated by spliceosome-associated proteins, such as SRSF and SF3 splicing factors. Through the process, different proteins are recruited ensuring splicing accuracy and efficiency (Haselbach et al., 2018).

Efficiently spliced coding gene transcripts without defects are degraded at the end of their useful life. Some of them, such as ribosomal RNA and house-keeping genes, are long-living transcripts (Lekas et al., 2000; Piir et al., 2011), and others, such as some stress-response gene RNA, are short-living (Uhlitz et al., 2017). The quality and accuracy of produced RNAs are controlled by surveillance machinery. Deleterious transcripts are detected by the RNA surveillance machinery and degraded faster through degradation pathways to avoid the production of useless or toxic proteins. Despite the complexity of RNA degradation pathways, the basics of these mechanisms are shared among the vast majority of eukaryotes.

RNA is degraded by enzymes called ribonucleases. Ribonucleases can be categorized into endo- and exoribonucleases called also endonucleases and exonucleases, respectively. Endonucleases have specificity to various RNA conformations, such as double-stranded RNA (dsRNA) or RNA-DNA hybrids. Exonucleases degrade RNA from the transcript beginning (5′) or from the transcript end (3′) utilizing various guiding complexes (Murchison, 2013).

Ribonucleases are associated with many adaptor complexes, which facilitate the recognition of specific RNA substrates, such as RNA helicases and RNA-binding proteins. In humans, RNA helicases, such as MTR4 and UPF1, translocate alongside RNA and unwind secondary structures to help exosome complex degrade RNAs. Alternatively, helicases act as “place markers” and recruit the degradation machinery (Schuch et al., 2014; Singh et al., 2019).

Recently, nuclear exosome targeting (NEXT) and poly(A) tail exosome targeting (PAXT) complexes were identified as important nuclear exosome cofactors in humans. Both complexes contain RNA helicase MTR4, which directly interacts with the exosome and helps to unwind targeted transcripts (Schuch et al., 2014). In addition to MTR4, PAXT contains two zinc-finger proteins, ZFC3H1 and ZC3H3, as well as poly(A) RNA-binding protein PABPN1 and possibly another two RNA binding proteins - RBM26 and RBM27. RBM26 and RBM27guide the exosome towards adenylated, longer transcripts (Silla et al., 2020). NEXT complex consists of MTR4, Zink-finger ZCCHC8, and RNA binding protein RBM7. This complex directs exosome toward unprocessed transcripts, such as promoter upstream transcripts, enhancer RNA, prematurely terminated RNAs, etc. (Garland et al., 2019; Meola et al., 2016). Interestingly, the NEXT complex is stabilized by and interacts with splicing factor SRSF3. Specifically, SRSF3 directly binds to mRNA targets, such as viral intronless mRNA, and interacts with nuclear exosome through the NEXT complex, which leads to RNA degradation (Mure et al., 2018). Taken together, in humans, RNA helicase MTR4 plays a central role in targeting nuclear exosome to distinct RNA species by forming NEXT and PAXT complexes. In addition, the NEXT function is coupled with the splicing factor SRSF3.

A large number of small non-coding RNAs (microRNAs or miRNA) regulates transcript processing as well. Briefly, miRNAs are transcribed from DNA sequences as miRNA precursors and processed into mature miRNAs. miRNAs also arise from spliced-out mRNA introns (Westholm and Lai, 2011). As a rule, miRNA binds to the 3′ untranslated region (3′ UTR) of target mRNAs to induce mRNA degradation. The nuclear phase of miRNA processing is a co-transcriptional event. It is mediated by the Microprocessor complex, which comprises endonuclease DROSHA and additional RNA binding proteins (O'Brien et al., 2018). Interestingly, the SRSF3 splicing factor recruits DROSHA in a sequence-specific manner and promotes miRNA processing by the Microprocessor complex (Kim et al., 2018). Then, pre-

processed by DROSHA, miRNAs are exported into the cytoplasm and transformed into mature miRNAs by endonuclease DICER. Produced small RNAs bind Argonaute proteins, central proteins of the RNA-induced silencing complex, and guide this complex to RNA transcripts for degradation (Michlewski and Cáceres, 2019).

### 1.6.1 Transcript stability regulation

Not only aberrant and cryptic transcripts are short living. Efficiently formed mRNAs exhibit different half-lives as well. The stability of such mRNA may depend on various RNA modifications, length of 3' poly(A) tail, alternative polyadenylation, and elongation dynamic of RNA polymerase II (RNAPII). RNAPII elongation slowdown results in increased deposition of N6-methyladenine (m6A), the most abundant post-transcriptional RNA modification associated with reduced RNA stability. CCR4–NOT deadenylase complex was shown to occupy m6A marked RNA and shorten their poly(A) tails (Du et al., 2016; Slobodin et al., 2020). Transcripts with shortened poly(A) tails, in turn, provide fewer binding sites for the poly(A)-binding proteins (PABPs), stabilizing bound by them RNA (Mangus et al., 2003). Taken together, the listed factors provide flexible tools for transcript stability regulation.

Alternative poly(A) sites are found in 70 % of human protein-coding genes (Derti et al., 2012). Depending on the usage of specific poly(A) sites, distinct 3′ UTRs can be formed. This process is called alternative polyadenylation. Alternative polyadenylation indirectly controls mRNA stability through the length of 3′ UTR cis-regulatory regions. miRNAs and RNA-binding proteins target these regions, which leads to either RNA degradation or stabilization (Barreau et al., 2005; Fabian et al., 2010). In addition, alternative polyadenylation and alternative splicing are likely to compete. Specifically, splicing factors were found to bind alternative polyadenylation sites blocking their usage (Shi and Manley, 2015; Tian et al., 2007).

### 1.6.2 Recognition and elimination of aberrant transcripts

Defective or aberrant transcripts, such as transcripts with premature termination codon, aberrant splicing, and inefficient 3′ end formation and polyadenylation are subjects for the degradation by RNA surveillance machinery (Glaunsinger and Lee, 2010; Hilleren et al., 2001; Jiao et al., 2010; Niemelä et al., 2014; Tudek et al., 2018). As a common principle, RNA processing and RNA decay pathways compete. Various RNA binding proteins occupy properly processed RNA transcripts. The absence of such proteins on aberrantly formed or processed transcripts allows RNA decay pathways to target these transcripts (Schmid and Jensen, 2018).

For example, premature termination codon inside transcript sequence is an aberrant feature, which could originate from transcriptional or splicing errors and mutations (Peltz et al., 1993).

The UPF1, which binds to transcripts promiscuously, accumulates on premature termination codons and gets phosphorylated. Phosphorylated UPF1 inhibits mRNA translation and recruits deadenylation and decapping complexes to such transcripts. Finally, decapped and deadenylated RNA is degraded by either exonuclease XRN1 or cytoplasmic exosome exonuclease DIS3L (Kurosaki et al., 2019).

Interestingly, alternative splicing is connected with nonsense-mediated transcript decay. Specifically, one-third of all the alternative splicing events lead to the inclusion of a premature termination codon, which results in transcripts degradation by nonsense-mediated decay (Lewis et al., 2003). In mammals, exonuclease XRN1 is the central component of the nonsense-mediated decay pathway, which degrades transcripts in the cytoplasm. The exonuclease XRN2, contrarily, co-transcriptionally degrades pre-mRNA upon inhibited splicing and 3′ end processing in the nucleus (Davidson et al., 2012).

RNA aroused from pervasive transcription are also targets of surveillance pathways (Porrua and Libri, 2015). In mammals, pervasive transcripts are antisense short-living transcripts originating from active promoters and gene enhancers. Active promoters often drive transcription initiation in both directions, but as a rule, only one direction results in gene expression. Such opposite-direction RNAs are called promoter upstream transcripts. In addition, enhancers generate similar bidirectional unstable transcripts (Kim et al., 2015). Such short pervasive transcripts have polyadenylation signals on the 3‘ prime region close to the 5’ prime cap. This feature is not typical for the coding genes. Therefore, it is recognized by a cap-binding complex and associated proteins, which leads to recruitment of the NEXT complex and degradation by nuclear exosome (Andersen et al., 2013). If bidirectional transcription happens convergently, it may lead to the production of dsRNA. Such dsRNA can be sensed and processed by endonuclease Dicer into small interfering RNAs. Then small interfering RNAs are loaded into Argonaute protein complex (AGO2 in mammals) and target complementary RNA, followed by cleavage (Svoboda, 2014).

Transposable elements’ transcripts are often transcribed bidirectionally and form dsRNA or are poorly spliced, which triggers their elimination by the surveillance machinery (Dumesic and Madhani, 2013; Ugolini and Halic, 2018). Turnover of transposable elements, including endogenous retroviruses, is poorly characterized. To date, genome-scale transcript turnover of endogenous retroviruses in mammals was not assessed. Although, few studies focusing on long non-coding RNA stability in mice and humans concluded that, in general, non-coding RNAs are less stable than gene transcripts with some locus-specific variations (Clark et al., 2012; Maekawa et al., 2015). Most human ERVs lost their abilities to produce proteins and therefore are non-coding. In addition, some protein-coding ERV loci may still have a harmful

effect on the host cell functions. Taken together, we could expect that ERV transcripts would show different stability compared to coding genes. With current advances in transcript turnover detection protocols, sequencing methods, and data processing tools, we are able to perform an accurate and efficient calculation of ERV transcription and ERV transcripts turnover in humans.

### 1.6.3 Role of splicing factors in transposon silencing

Transposon transcripts often have suboptimal splicing properties in contrast to host genes. It happens because transposons had limited time to adapt to new splicing preferences of the host's genes, namely to particular intron size, splice site, and splicing enhancer sequences (Ast, 2004; Kupfer et al., 2004; Warnecke et al., 2008). One of the possible mechanisms of transposon repression by splicing factors is that transposon-derived poorly spliced transcripts stall in spliceosomes, promoting the targeting of transposon transcripts by other degradation complexes (Dumesic and Madhani, 2013). For example, in *C. elegans,* a striking amount of RNA splicing factors were identified by genetic screens as transposon silencers (Robert et al., 2005). The proposed mechanism is that incompletely spliced messenger RNA (mRNA) precursors trigger siRNA production (Warf et al., 2012). In yeast *C. neoformans,* transposon-derived poorly spliced transcripts stall in spliceosomes, promoting siRNA production and transposon transcripts targeting (Dumesic et al., 2013). Dumesic et al. identified a Spliceosome-coupled and nuclear RNAi (SCANR) complex required for siRNA synthesis from intron-containing mRNA precursors. Authors proposed that the SCANR complex is in kinetic competition with splicing, and stalled in the spliceosome poorly spliced transcripts trigger siRNA production (Dumesic et al., 2013). In *Drosophila,* studies on piRNA silencing suggest that splicing and nuclear export factor (UAP56) recognizes incompletely spliced piRNA precursors from heterochromatic loci, which leads to mature piRNA production and transposon silencing during germline development (Zhang et al., 2012). Therefore, different splicing subunits not only participate in splicing themselves but also may act as surveillance machinery, protecting the host genome from the excessive activity of retrotransposons.

Interestingly, splicing factors participate in the heterochromatic transcript silencing in fission yeast. Specifically, splicing factors deletion results in centromeric repeats derepression and disruption of heterochromatin. Splicing factors directly interact with RNAi machinery and centromeric chromatin. Authors proposed that spliceosome complexes provide a platform for centromere repeat's siRNA generation and mediate the recruitments of the H3K9 methyltransferase complex (CLRC) to such heterochromatic regions (Bayne et al., 2008; Chinen et al., 2010). Another study suggested that some splicing factors might contribute to heterochromatin assembly indirectly by RNAi and alternative splicing of heterochromatin

factors (Kallgren et al., 2014). However, the most recent study illustrated that one of the splicing factors in *S. pombe* binds nascent heterochromatic transcripts and serves as a platform to recruit chromatin remodelers to promote heterochromatin assembly. Notably, the splicing of heterochromatin factors was not affected. (Weigt et al., 2021).

#### *1.6.3.1 SF3A3 and SRSF3 splicing factors characterization*

SRSF (serine / arginine-rich splicing factors, also known as SR proteins) are mRNA-binding proteins and splicing factors with diverse cellular functions in various mRNA regulation pathways (Änkö, 2014). In addition to splicing, SR proteins regulate transport, stability, RNA polyadenylation of mRNA, and miRNA processing (Lemaire et al., 2002; Muller-McNicoll et al., 2016). Interestingly, a recent study showed that the SRSF3 splicing factor depletion leads to striking derepression of SINE retrotransposons in mouse oocytes (Do et al., 2018). Although, the mechanism behind SINE derepression upon SRSF3 depletion is unclear.

SRSF3 plays a role in alternative polyadenylation of transcripts. Shen et al. observed increased usage of proximal polyadenylation sites and, consequently, the shortening of transcripts 3′ UTRs in both humans and mice upon SRSF3 knockdown. 3′ UTRs shortening was associated with increased amounts of corresponding transcripts and proteins. Authors hypothesized that 3′ UTR shortening resulted in the exclusion of miRNA-target sites. Therefore, shortened transcripts escaped 3′ UTR-specific miRNA – mediated degradation (Shen et al., 2019a). In addition, SRSF3 was demonstrated to regulate miRNA processing by facilitating primary microRNA transcripts recognition and processing into mature microRNA. Specifically, SRSF3 co-binds primary miRNA transcripts and induces miRNA processing by the microprocessor-protein complex (Kim et al., 2018). However, the underlying SRSF3-mediated stimulation mechanism is unknown. Taken together, SRSF3 was shown to facilitate transcript silencing through either miRNA-target sites inclusion by alternative polyadenylation or promotion of the microprocessor-mediated miRNA production.

In addition, SRSF3 is involved in transcript transport. mRNA transport from the nucleus to the cytoplasm is mediated by nuclear export factor 1 (NXF1), which requires several adaptor proteins to recognize and bind mRNA. SR protein family was shown to serve as adaptors for NXF1. These proteins bind the last exons of mRNA and shuttle between the nucleus and the cytoplasm (Hargous et al., 2006). SRSF3 is the most potent NXF1 adaptor among other SR proteins. Notably, many SRSF3 - NXF1 target transcripts were alternatively spliced or polyadenylated (Muller-McNicoll et al., 2016). Therefore, the SRSF3 splicing factor may play a role in coupling pre-mRNA processing to mRNA export.

SRSF3 is an oncogene. It is frequently overexpressed in several cancers and associated with poor cancer prognosis (Corbo et al., 2013; Jia et al., 2010). SRSF3 knockdown in cancer cells, on the one hand, induces autophagy, apoptosis, and cellular senescence - an irreversible proliferation arrest, which prevents cancer (Kim et al., 2014; Tang et al., 2013; Zhou et al., 2019). On the other hand, SRSF3 silencing disrupts drug resistance and inhibits the proliferation, migration, and metastasis of tumor cells (He et al., 2004; Kuranaga et al., 2018; Song et al., 2019). Interestingly, the number of potential anticancer drugs inducing cytotoxicity in tumor cells, such as Amiodarone, Amiloride, Caffeine, and Theobromine, was shown to decrease SRSF3 expression (Chang et al., 2017; Chang et al., 2018; Lu et al., 2014; Sen et al., 2013).

It is worth SF3A3 is a potential target for anticancer treatment as well. Specifically, silencing of SF3A3 effectively kills non-small cell lung cancer (NSCLC). At the same time, cytotoxicity for normal lung fibroblasts was much weaker. The proposed mechanism is the activation of tumor suppressor p53 upon SF3A3 protein level reduction. However, the authors noted that induction of cell death and activation of the p53 pathway were rather unrelated phenomena (Siebring-van Olst et al., 2017). Taken together, splicing factors SF3A3 and SRSF3 are not only involved in various mRNA processing pathways but also are promising therapeutic targets in anticancer treatment. SRSF3 silencing was found to be a particularly potent inhibitor of many tumors.

### 1.6.4 Role of nuclear exosome complex in transposon silencing

Nuclear exosome components are other important post-transcriptional regulators, which were identified as ERV silencers in mouse screens and interactors with chromatin-associated ERV repressors (Figure 1).

The exosome is a major 3’ degradation multiprotein complex, which is involved in the degradation of stable transcripts, as well as RNA quality control through degradation of defective transcripts. The nuclear exosome comprises nine catalytically inactive subunits forming a barrel-like structure. Two catalytically active exonucleases, EXOSC10 (RRP6) and DIS3 are equipped at opposite ends of the barrel-like protein structure (Zinder and Lima, 2017). The nuclear RNA exosome components are highly conserved from lower to higher eukaryotes. The nuclear exosome regulates the maturation and degradation of a wide range of RNA species. It has broad functions in nuclear RNA surveillance, degrading a variety of aberrant transcripts, for example, intron-retained and polyadenylation-defective pre-mRNAs (Lemieux et al., 2011; Milligan et al., 2005; Schneider et al., 2012). The nuclear exosome controls RNA amounts generated by cryptic and pervasive transcription. Specifically, it

degrades bidirectional transcripts arising due to antisense transcription from gene promoters, which in humans are called promoter upstream transcripts, as well as enhancer RNA (Pefanis et al., 2015; Richard and Manley, 2013). Interestingly, inefficiently spliced and polyadenylated lincRNAs are transcriptionally stabilized upon nuclear exosome component knockdown in humans. Authors suggested that, independently of miRNA, the Microprocessor complex subunit (DGCR8) helps recruit the exosome to such aberrantly spliced and polyadenylated lincRNAs for their subsequent degradation (Schlackow et al., 2017).

RNA exosome silences repeats and helps to preserve heterochromatin in *S. cerevisiae, S. pombe,* and *Drosophila*. In *S. cerevisiae,* repeat transcripts arising from heterochromatic loci are targeted and degraded by RNA exosome (Vasiljeva et al., 2008). In *S. pombe*, deletion of nuclear exosome component leads to a substantial increase in centromere repeat transcripts and heterochromatin level reduction across pericentromeric domains (Reyes-Turcu et al., 2011). Recent studies demonstrated that cleavage and polyadenylation RNA-binding proteins facilitate RNA decay by recruiting exosome to centromere repeats and LTR retrotransposons (Lee et al., 2020c; Mallet et al., 2017). In *Drosophila,* the homolog of histone methyltransferase SUV39h1 targets the nuclear exosome to a subset of heterochromatic loci to degrade non-coding RNAs, including LTR transposons. In addition, exosome component deletion leads to less compacted heterochromatin (Eberle et al., 2015).

As a rule, the nuclear exosome complex needs mediator proteins to target specific transcripts. In mammals, NEXT and PAXT complexes direct a subset of RNAs for exosomal degradation. Interestingly, NEXT, which directs the exosome toward shorter and unprocessed transcripts, directly interacts with the SRSF3 and SF3B family of splicing factors (Falk et al., 2016; Mure et al., 2018). In addition, NEXT proteins interact not only with splicing factors but also with other transcript-processing complexes, such as nuclear cap-binding and pre-mRNA 3′ processing complexes (Andersen et al., 2013; Shi et al., 2009). Notably, PAXT was shown to influence heterochromatin occupancy through the PRC2 complex. Namely, PAXT component deletion leads to decreased PRC2 stability and binding to chromatin and low PRC2-directed H3K27me3 (Garland et al., 2019).

Several studies suggested exosome-mediated transposon and viral transcripts degradation in mammals. NEXT complex was shown to target LINEs and LTR-retrotransposon transcripts, mediating their degradation by exosome in mouse embryonic stem cells (Garland et al., 2019). Moreover, another study showed that SRSF3 directly interacts with both NEXT and RNA exosome complexes and, in addition, stabilizes the NEXT complex. Specifically, SRSF3 targets viral intronless mRNA and recruits NEXT complex with the nuclear exosome, which leads to viral mRNA decay (Mure et al., 2018).

Taken together, nuclear exosome complex in both lower and higher eukaryotes, including mammals, is involved in the degradation of various unwanted transcripts, such as intron-retained, polyadenylation-defective pre-mRNAs, viral mRNA, heterochromatin-associated, bidirectional, and transposon transcripts. In mice, nuclear exosome degrades transposon transcripts through exosome targeting NEXT complex. Moreover, in humans, the splicing factor SRSF3 interacts with the NEXT complex to target viral RNA for exosomal degradation. Therefore, in humans, post-transcriptional ERV silencing might depend on nuclear exosome and splicing factors more than was anticipated.

### 1.7 Aim of the thesis

To date the importance of post-transcriptional pathways in mammalian ERV transcripts regulation is poorly understood. Given the ERVs importance not only in the physiological process but also in various diseases, such as cancer, proper characterization of ERV transcriptional activity and transcripts turnover is an essential task that would significantly benefit advances in the field of cancer treatment. In this work, we focused on three potential mammalian post-transcriptional ERV silencers – splicing factors SF3A3, SRSF3, and nuclear exosome exonuclease DIS3. These proteins were identified in mammalian ERV silencing screens as potent ERV repressors and were shown to interact with known chromatin-associated ERV silencers.

Therefore, the aims of the thesis are following:

- characterize ERV transcript stability in human cell lines with transient transcriptome sequencing (TT-seq) protocol, utilizing 4sU metabolic labeling
- investigate genome-wide ERV loci transcription coupled with H3K9me3 alterations and transcript stability changes upon anticancer epigenetic drug treatments with 5-Aza and Panobinostat
- study the influence of splicing factors SF3A3, SRSF3, and nuclear exosome exonuclease DIS3 knockdowns on ERV transcriptional activity and turnover.
- finally, check possible crosstalk between splicing factors and exosome-induced changes in ERV transcription or turnover and H3K9me3.

# 2. Results

## 2.1 TT-seq protocol allows ERV transcripts turnover calculation and efficiently detects short-living transcripts

### 2.1.1 Detection of endogenous retroviral transcription and turnover

One of the main aims of the thesis is to investigate ERV transcript stability and transcripts stability changes in human cell lines upon anticancer epigenetic drug treatments, as well as upon knockdowns of splicing factors SF3A3, SRSF3, and nuclear exosome exonuclease DIS3.

Genome-wide transcript turnover can be investigated by such methods, as transcription inhibition and metabolic RNA labeling. Transcription inhibitors, such as actinomycin D, are used to block transcription via polymerase II inhibition, which allows RNA decay measurement over time. The disadvantage of this method is that global transcription rate reduction may lead to transcript stabilization (Pelechano and Pérez-Ortín, 2008) which would cause transcript stability overestimation. Metabolic labeling does not inhibit transcription. This method is based on the introduction of labeled nucleotides into newly synthesized RNA within a short time frame. Then, transcript turnover is quantified by comparing labeled and unlabeled RNA fractions. Initially, radioactively labeled nucleotides were utilized. Later, non-radioactively modified nucleotides were used, such as 5-bromouracil (BrU) and 4-thiouracil (4sU). BrU method is based on immunoprecipitation of labeled RNA, while 4sU – on labeled RNA biotinylation and isolation by streptavidin beads. Nowadays, the 4sU method is predominantly used because of relatively easy implementation and low cost (Russo et al., 2017). In addition, methods based on nascent transcripts separation and enrichment, such as the 4sU protocol, allow more sensitive detection of unstable non-coding transcripts compared to traditional RNA-seq methods. However, traditional RNA-seq methods detect only steady-state RNA. Therefore, we chose TT-seq - 4sU-based method to detect nascent RNA fraction and calculate transcripts stability.

Despite ERVs being repetitive elements, in humans, they diverge from each other due to mutations and truncations. It permits studying ERVs on a single loci level. Standard 50 base pairs single end sequencing mode allows identification of over 50 % transposable loci with uniquely mapped reads. The paired-end mode further improves these results with over 70 % unique transposable loci identification (Sexton, 2019). In addition, proper assignment of transcript reads to ERV reference is another important step, which can significantly affect transcription estimations. According to the literature, the STAR aligner is one of the most

efficient software for the mapping of transposable elements. Specifically, the usage of STAR is advantageous over other aligners for mapping paired-end reads (Teissandier, 2019).

Taken together, current advances in transcript turnover detection protocols, sequencing methods, and data processing tools allow accurate and efficient calculation of ERV transcription and turnover in humans.

### 2.1.2 TT-seq protocol design and quality control

TT-seq measures nascent and steady-state transcript levels at the same time, which allows not only to detect unstable transcripts but also calculate transcript stability (turnover) by comparing the amount of nascent and steady transcripts for each locus of interest (Rädle et al., 2013; Schwalb et al., 2016). Briefly, the TT-seq method is based on metabolic labeling of newly synthesized transcripts with a uridine analog, 4-thiouridine (4sU), which allows thiol-specific biotinylation, followed by separation of labeled nascent and unlabeled stable transcripts by streptavidin magnetic beads (Figure 2.1A).

To assess nascent RNA separation, we compared ratios of nascent to steady transcripts in well-known genes with fast and slow turnover – in MYC and GAPDH, respectively (Lekas et al., 2000; Snyder and Miller, 1992). If the amount of nascent RNA from a gene is much higher than steady, we regard this transcript as unstable. In contrast, stable transcripts are characterized by a higher amount of steady RNA. In line with previous findings, MYC showed a higher amount of nascent transcripts, while GAPDH had a higher amount of steady transcripts (Figure 2.1B). Thus, by investigating genes with known transcript turnover in our dataset, we confirmed successful nascent RNA labeling and separation of labeled and unlabeled RNA.

In the TT-seq protocol, steady-state transcripts can be measured based on either total RNA aliquot taken before separation, or unlabeled fraction, taken after separation (Figure 2.1A). In the THP-1 cells, an unlabeled (pre-existing) RNA fraction was collected to measure steady-state transcripts. In HeLa cells, total RNA fraction was collected to measure steady-state transcripts. Since only a small fraction of the total RNA fraction comprises nascent RNA, both ways of measuring steady-state transcripts amount are comparable, as shown by nascent and steady transcripts amount from MYC and GAPDH genes in both cell lines (Figure 2.1B).

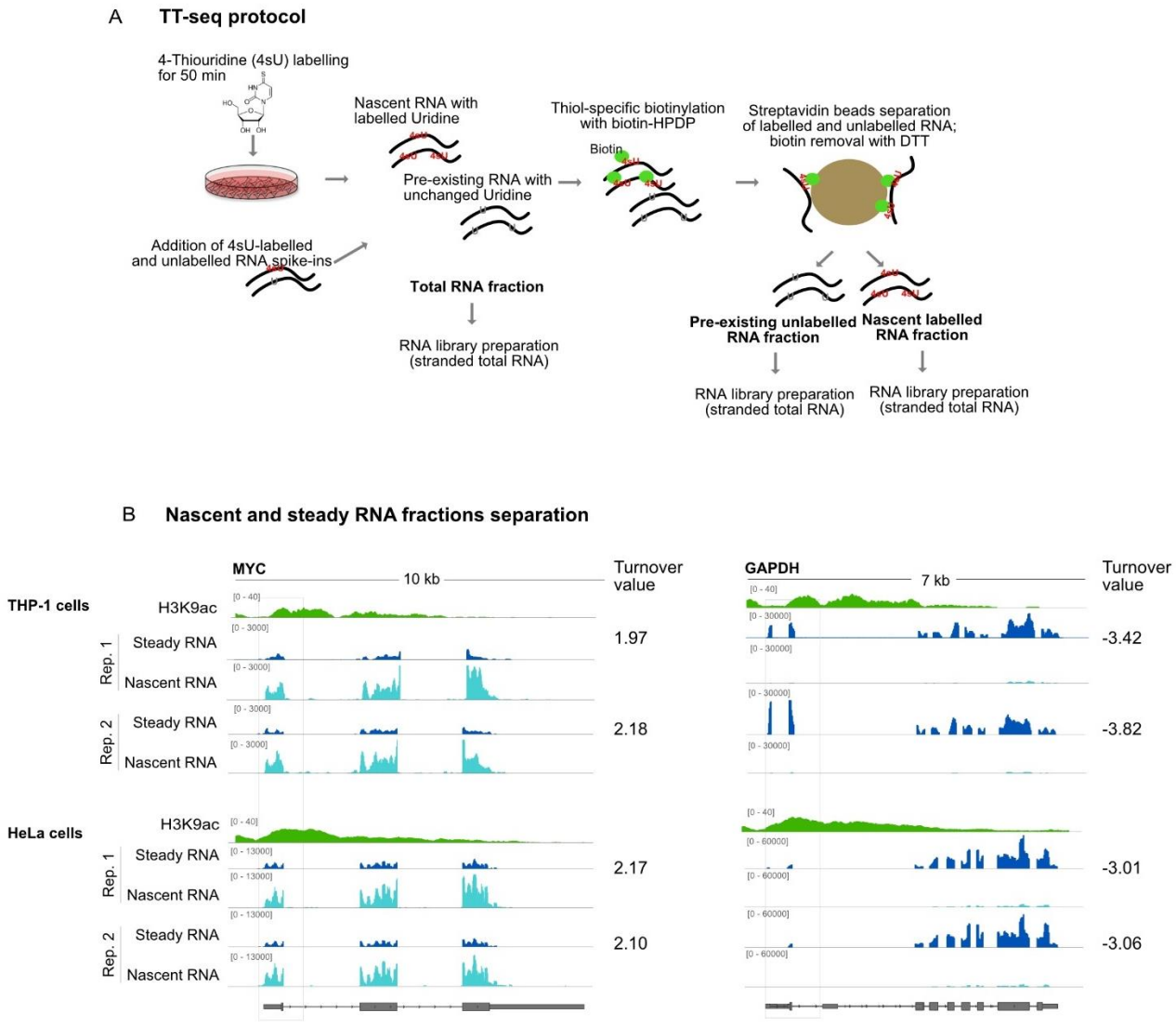


**Figure 2.1. TT-seq protocol overview and quality control of efficient fractions separation**

A. TT-seq protocol overview.

B. Quality control of RNA fractions separation by visualization of sequenced fractions for known fast- and slow-degrading genes. Transcript turnover values are noted. Genome browser view of nascent and steady RNA-seq for MYC gene with fast transcript turnover and GAPDH gene with slow transcript turnover in untreated THP-1 and HeLa cells.

### 2.1.3 Transcript abundances of nascent versus steady RNA and transcript turnovers

To test the hypothesis of faster ERV turnover, compared to coding genes, we measured nascent and steady transcripts for all coding genes, non-coding genes (lincRNA), and ERV loci in THP-1 and HeLa cells. Only ERVs containing H3K9ac peaks in any condition for each cell line were selected for the downstream analysis to filter out ERVs with read-through transcription. We found that, on average, genes had similar amounts of nascent and steady RNA (Figure 2.2.A). In contrast, ERV transcripts clearly had more nascent than steady transcripts in both cell lines, suggesting a high degradation rate (Figure 2.2A).

Then, transcripts turnovers were plotted for coding genes, lincRNA, and ERV loci (Figure 2.2B). Transcript turnover for each locus was calculated by subtracting steady RNA normalized log2 transformed read counts from nascent RNA normalized log2 transformed read counts. lincRNA transcripts were rather unstable in HeLa cells with an average turnover level close to ERVs turnover. While in THP-1 cells, lincRNAs, on average, had stability close to coding genes. Different overall lincRNA turnover levels between cell lines could be due to lincRNA transcriptional heterogeneity. Depending on the cell line, different lincRNAs are transcribed (Djebali et al., 2012) with higher or lower transcript stability. ERVs, in contrast, in both cell lines had approximately twice-faster transcripts turnover compared to transcripts turnover of coding genes.

Taking together, our hypothesis about fast ERV transcripts turnover was confirmed, suggesting that post-transcriptional silencing machinery targets repetitive elements' transcripts and actively degrade them in humans.

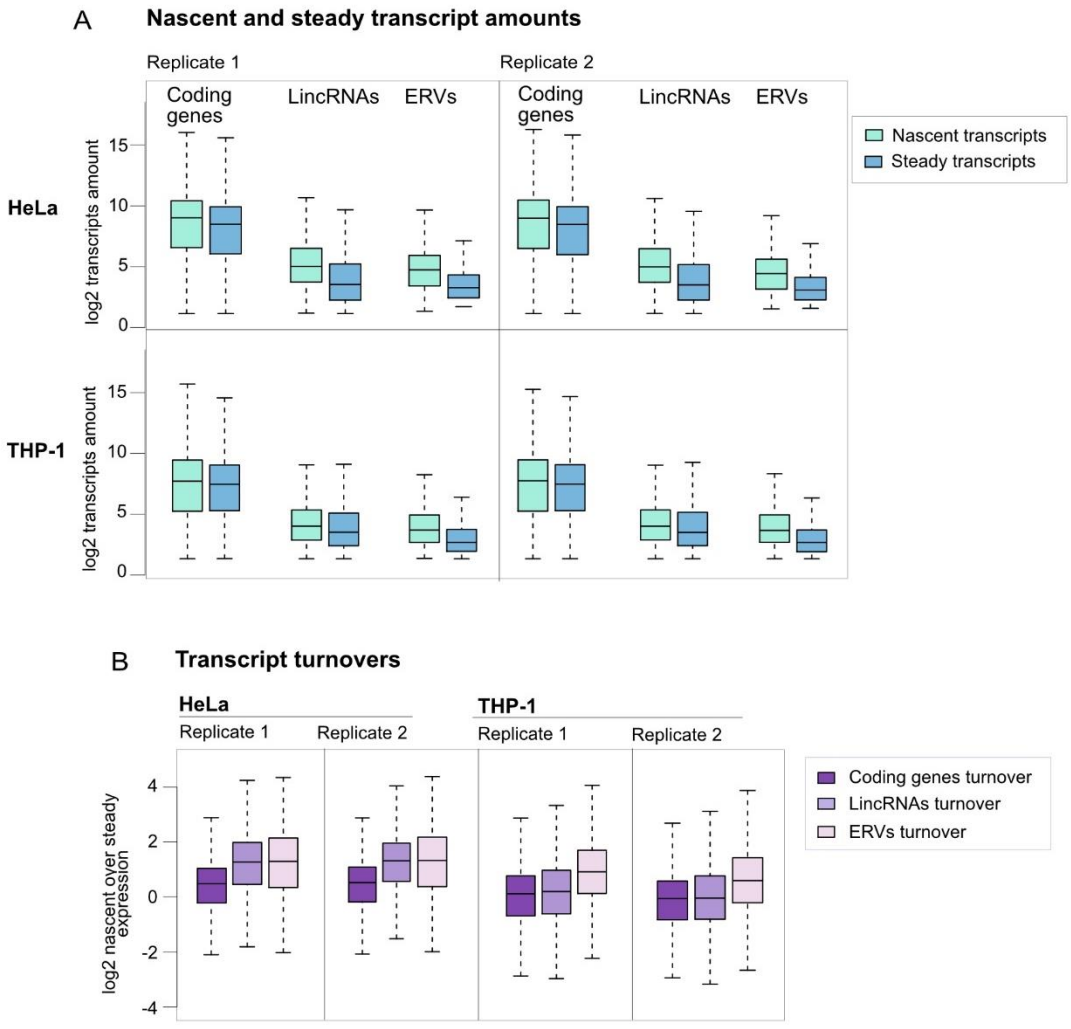


**Figure 2.2. ERV transcripts have a faster turnover compared to coding genes.**

A. Boxplots of log2 nascent and steady transcript amounts assigned to coding genes, lincRNAs, and ERV loci in untreated cells. ERV loci and genes with normalized log2 expression values more than 3 on average for both fractions are selected to calculate turnover accurately.

B. Boxplots of transcripts turnovers assigned to coding genes, lincRNAs, and ERV loci in untreated cells. ERV loci and genes with normalized log2 expression values more than 3 on average for both fractions are selected to calculate turnover accurately.

### 2.1.4 TT-seq method allows identifying at least twice more active ERVs compared to the traditional RNA-seq method

As we showed before, ERVs have less amount of steady transcripts compared to nascent transcripts, which could lead to underestimation of their activity by traditional RNA-seq. Due to the ability to enrich nascent transcripts, the TT-seq method allows the detection of rapidly degraded ERV transcripts. The following plot illustrates that more than half of transcriptionally active ERVs are not detected if we rely only on steady-state RNA sequencing (Figure 2.3). Thus, the TT-seq method is a powerful tool to investigate ERV activity since it detects not only steady transcripts but also nascent RNA. Therefore, in ERVs-focused studies, TT-seq clearly outperforms the traditional RNA-seq method, which measures only steady RNA

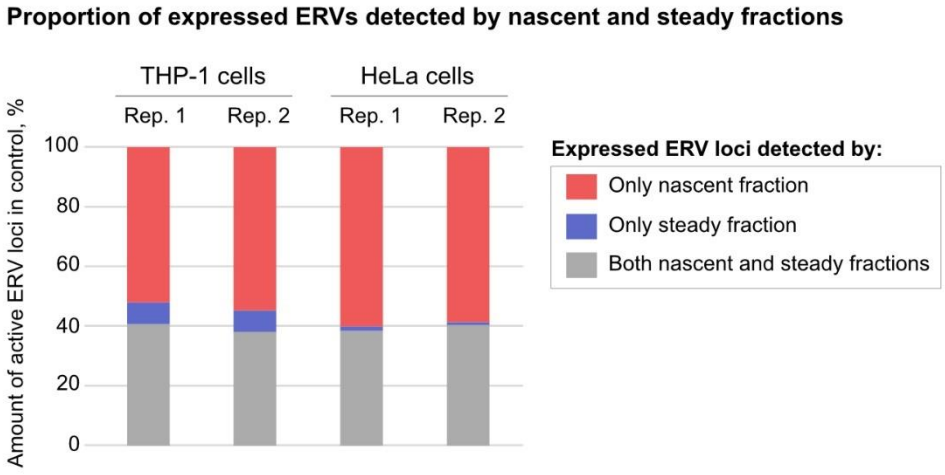


**Figure 2.3. TT-seq method allows identifying much more transcriptionally active ERV copies compared to the traditional RNA-seq method**

Proportions of transcriptionally active uniquely mapped ERV copies with H3K9ac peaks, detected by both fractions (gray), by only a steady fraction (blue), by only a nascent fraction (red) in untreated THP-1 and HeLa cells. ERV loci with log2 normalized expression value > 4 counted as expressed.

## 2.2 Influence of anti-cancer epigenetic drugs 5-Aza and Panobinostat on transcriptome and epigenome

### 2.2.1 DNA methylation and histone modification changes upon treatment

As a next step, we wanted firstly (1) to measure ERV transcriptional activity by TT-seq through perturbing HERV silencing, and secondly (2) to determine the stability of HERV transcripts upon increasing their transcription rate. For this approach, we treated cells with 5-Aza and Panobinostat reagents, also used in cancer treatments, which are known to induce ERV activity in human leukemia (Conti et al., 2016; Daskalakis et al., 2018). 5-Aza is a DNA methylation inhibitor, which reduces DNA methylation, while Panobinostat is an HDAC inhibitor, which increases histone acetylation.

We selected drug concentrations and treatment durations in a way that is enough to change epigenome and activate ERVs, but, at the same time, cell viability is not notably affected yet by the end of the treatments to avoid excessive cellular stress. 500 nM final concentration of 72 hours 5-Aza treatment was already used in our lab to successfully induce ERV activity in the mammalian cell culture. Panobinostat is associated with induction of H3 and H4 acetylation, such as H3K9ac and H4K8ac (Saitoh et al., 2019; Scuto et al., 2008). Exact Panobinostat concentration, which is enough to increase histone acetylation in our cell lines, was selected using titration experiment with antibody against H3K9ac. Western blot showed that 10 nM of Panobinostat was already enough to increase the H3K9ac signal in both cell lines (Figure 2.4.A). Therefore, 5-Aza and Panobinostat were added to cell cultures in total concentrations of 500 nM and 10 nM, respectively. The experiment was as follows: leukemic cell lines THP-1 and KG1a were treated with 5-Aza or Panobinostat, then right away after treatments, samples were collected for TT-seq, chromatin ChIP, RRBS, and histone mass spectrometry (Figure 2.4.B).

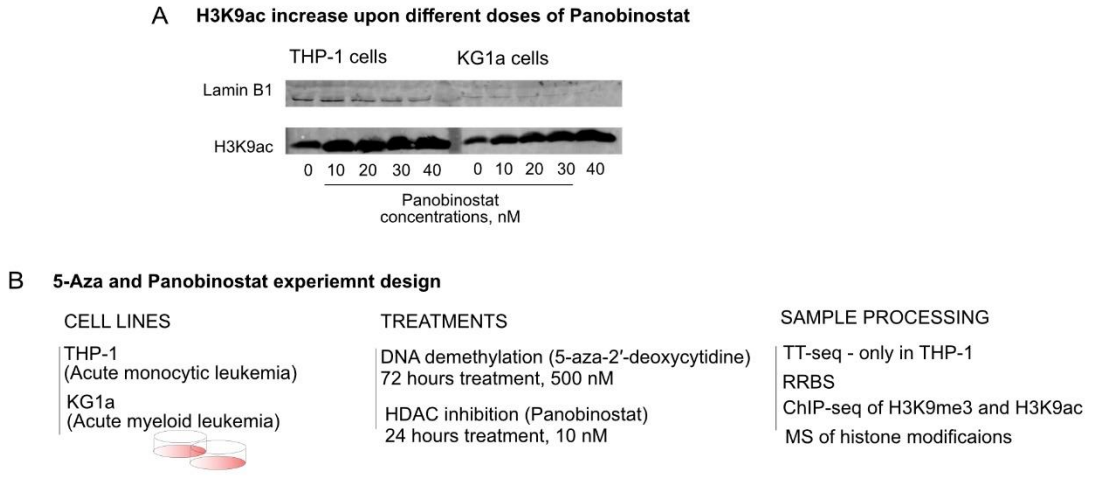


**Figure 2.4. Experiment design for 5-Aza and Panobinostat treatments**

A. Protein expression analysis by western showed an increase of H3K9ac in THP-1 and KG1a cells upon 24 hours Panobinostat treatment. 10 nM concentration of Panobinostat was chosen for the following experiments. Lamin B1 was used as a loading control.

B. Experiment design of 5-Aza and Panobinostat treatments.

To validate 5-Aza and Panobinostat treatments, we measured DNA methylation using Reduced Representation Bisulfite Sequencing (RRBS) and histone acetylation by mass-spec analysis of histone modifications. We found that upon 5-Aza treatment, DNA methylation was globally reduced from 80% to around 40% in the THP-1 cell line and from 90 % to 75 % in the KG1a cell line, while Panobinostat treatment did not affect DNA methylation in any of the cell lines (Figure 2.5A). Panobinostat, in contrast, greatly increased the acetylation of H3 and H4 histones. Specifically, H3K9ac overall increased 6 times compared to control in the THP-1 cell line and 3.2 times in the KG1a cell line; while 5-Aza treatment did not lead to an overall increase in measured acetylation signals (Figure 2.5B).

To assess loci-specific changes of active histone mark H3K9ac and repressive histone mark H3K9me3 on ERVs upon treatments, we compared the amount of ERV loci with ChIP H3K9ac and H3K9me3 peaks in treatments to controls. We observed that inhibition of histone deacetylases resulted in an increased amount of ERV loci with H3K9ac peak from 10k to 15k in the THP-1 cell line and from 6k to 14k in the KG1a cell line (Figure 2.5C). Venn diagram of H3K9ac-ERV loci overlap in treatments and control illustrates gain of new peaks in both cell lines upon Panobinostat treatment. Interestingly, DNA methylation removal by 5-Aza treatment led to a striking reduction of H3K9me3 peaks on ERV loci, while overall H3K9me3 signal measured by mass-spec was somewhat increased by 1.5 – 2-fold changes compared to control (Figure 2.5B, C).

Taking together, we confirmed the strong effect of 5-Aza and Panobinostat treatments on respective epigenetics marks in both cell lines. However, the THP-1 cell line showed stronger DNA methylation reduction and an increase in the H3K9ac signal. Therefore, we performed a TT-seq experiment only in the THP-1 cell line.

Experiments performed in 5-Aza and Panobinostat treated cells are not replicated, except mass-spectrometry data. Therefore, conclusions, which are made in this chapter, should be confirmed with additional replicates.

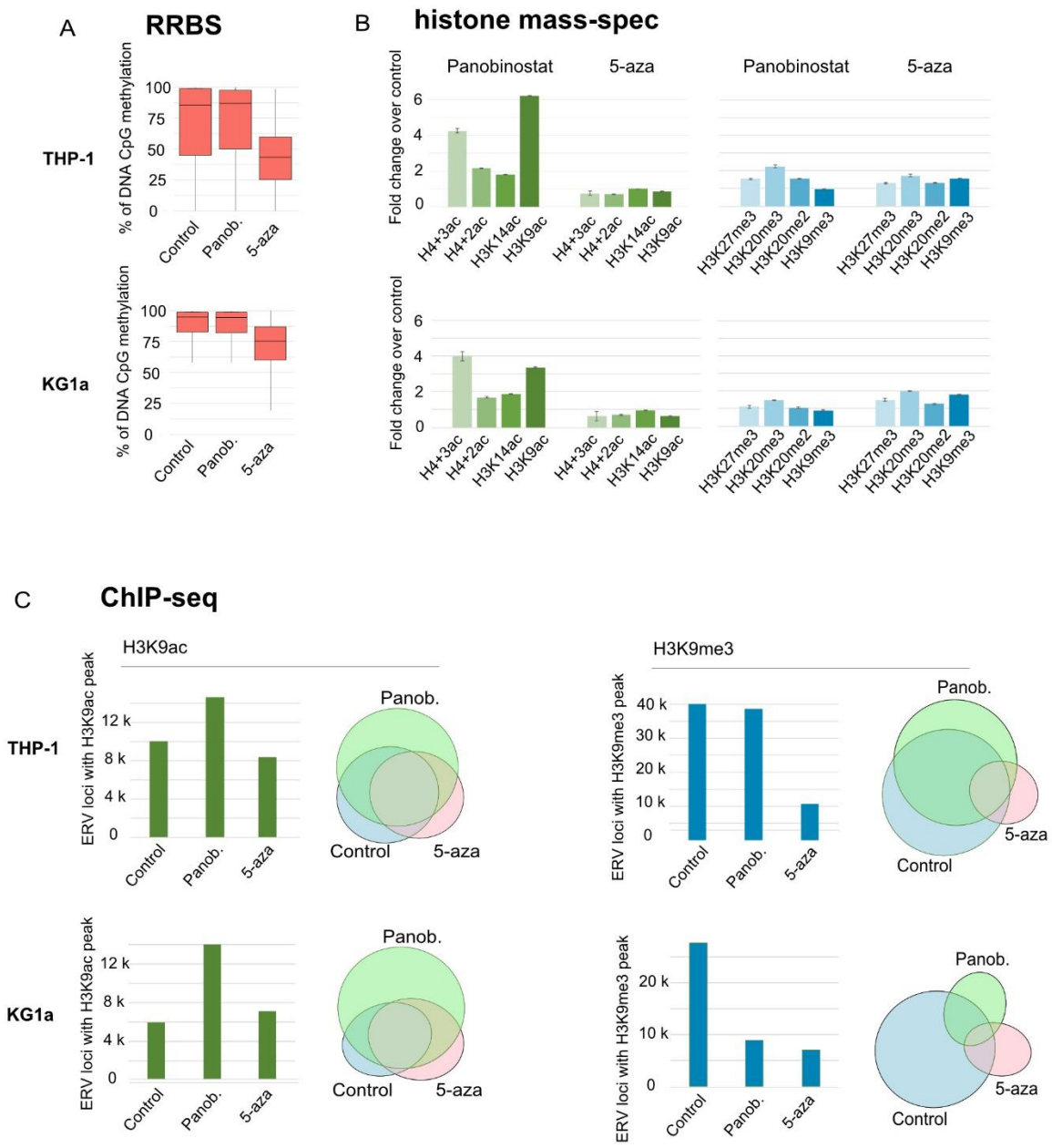


**Figure 2.5. DNA methylation and histone modification changes upon treatments**

A-B. Treatment of THP-1 and KG1a cells with 5-Aza and Panobinostat results in changes in DNA methylation and histone acetylation, respectively. (A) RRBS analysis to document DNA methylation changes (n=1); (B) quantitative mass-spec analysis for histone modifications (n=3, mean values ± SD).

C. Treatment of THP-1 and KG1a cells with 5-Aza results in a striking reduction of H3K9me3 peaks on ERV loci, while inhibition of histone deacetylases results in an increase of H3K9ac peaks on ERV loci compared to control, as documented by ChIP-seq data (n=1).

### 2.2.2 Gene expression changes upon treatments

To confirm previous results that 5-Aza and Panobinostat treatments lead to immune pathways activation (Ebelt et al., 2020; Mazzone et al., 2017a; Medon et al., 2017), we characterized expression patterns of coding genes in control and treated cells, Both 5-Aza and Panobinostat

treatments resulted in a similar level of gene deregulation (Figure 2.6A). According to over-representation analysis (ORA) against Reactome pathways database, both treatments led to upregulation of genes enriched in immune system-related pathways, such as Interleukin-4 and Interleukin-13 signaling and innate immune system pathways. In addition, a pathway of immunoregulatory interactions between Lymphoid and non-Lymphoid cells was enriched in upregulated genes upon 5-Aza treatment (Figure 2.6B). For example, upon 5-Aza treatment were upregulated interleukins (Interleukin 1 beta - IL1B, Interleukin 1 receptor type 1 - IL1R1, Interleukin 2 receptor subunit gamma - IL2RG), used by Immune system cytokines (Chemokine ligand 2 - CCL2, Tumor necrosis factor superfamily member 14 - TNFSF14, Tumor necrosis factor - TNF), a surface receptor that amplifies inflammatory processes (Triggering receptor expressed on myeloid cells 1 - TREM1). Upon Panobinostat treatment, were upregulated surface antigens (Fc fragment of IgA receptor - FCAR, Integrin alpha M - ITGAM, lymphocyte function-associated antigen 1 – ITGAL, Integrin subunit alpha X - ITGAX), and TREM1. The abovementioned genes are marked in Figure 2.6A.

These results are in line with previous studies, where 5-Aza and Panobinostat have shown a potent immunomodulatory activity helping cancers treatments (Ebelt et al., 2020; Mazzone et al., 2017a; Medon et al., 2017).

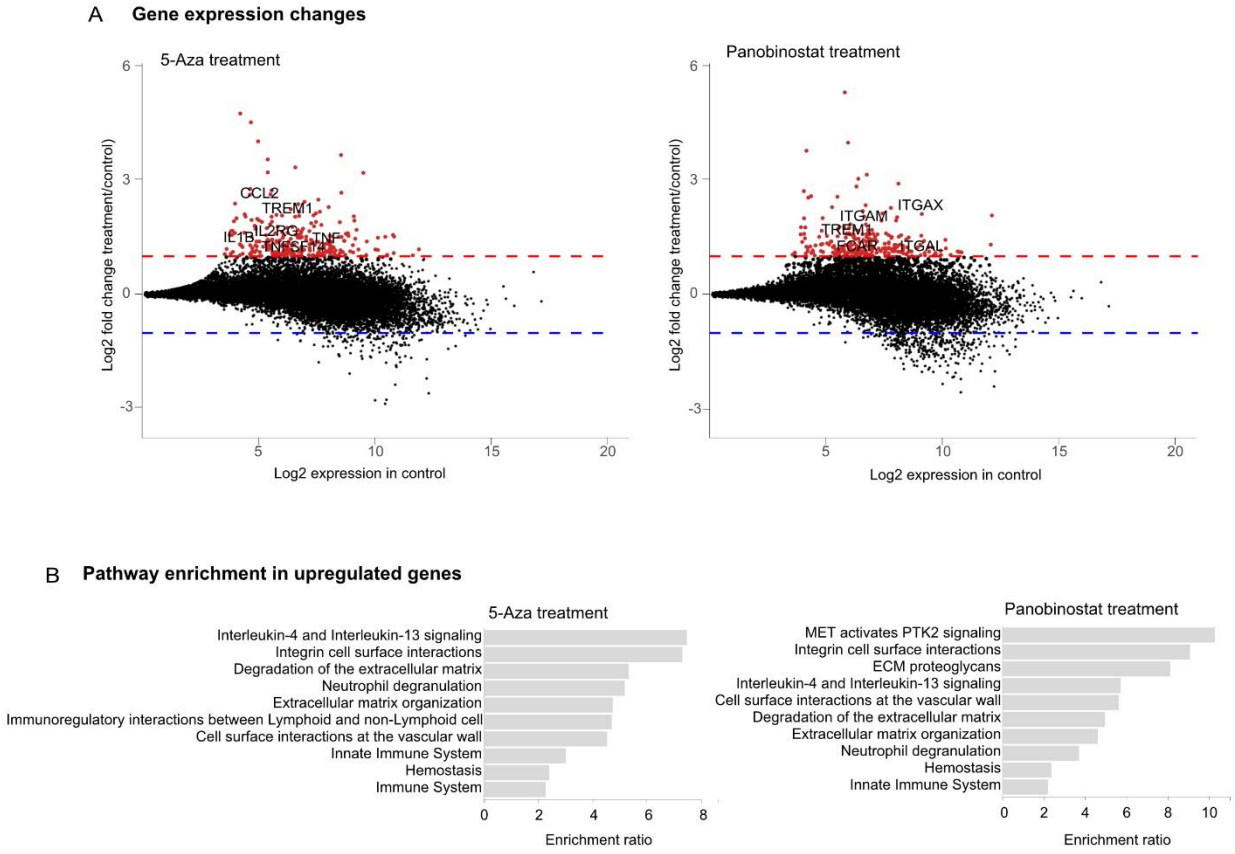


**Figure 2.6. Differential gene expression in cells upon 5-Aza and Panobinostat treatments**

A. Dot plots, showing log2 steady expression in control versus a log2-fold change of protein-coding genes in treatments versus control cells (n=1). Genes with expression changed more than 1 log2 fold change are colored with red. Selected genes involved in the immune response are labeled with corresponding names.

B. Over-representation analysis (ORA) for Reactome pathway enrichment in significantly upregulated genes. Shown is the enrichment ratio of pathways in upregulated genes relative to all coding genes.

### 2.2.3 5-Aza treatment leads to derepression of HERV-Fc1 full ERV loci with extremely high expression level

ERV upregulation upon 5-Aza and HDACi treatments was already detected in various cell lines. Specifically, the HERV-Fc1 family showed a strong increase in transcripts amounts upon both types of anti-cancer treatments with HERV-Fc1 gag protein production upon 5-Aza treatment (Conti et al., 2016; Daskalakis et al., 2018; Laska et al., 2013). Although, these previous reports are based only on a family-level analysis of ERVs. Therefore, ERV activity on a single loci level is unknown. In addition, the association of ERV transcriptional changes with repressive histone marks and ERV transcript stabilities so far were not addressed. First, we addressed ERV families derepression and its connection with H3K9me3 reduction upon 5-Aza and Panobinostat treatments (Figure 2.7). Both treatments resulted in increased expressions of ERV families consistent with reports of ERV upregulation upon 5-Aza and HDACi treatments. Interestingly, we observed a substantial decrease in H3K9me3 signal on ERVs upon 5-Aza treatment, but not upon Panobinostat (Figure 2.7). In addition, HERV-Fc1 and closely related families were strongly upregulated and showed H3K9me3 signal reduction upon 5-Aza. Panobinostat led to similar H3K9me3 reduction on HERV-Fc1 and HERV-Fc2 families but milder transcriptional upregulation (Figure 2.7).

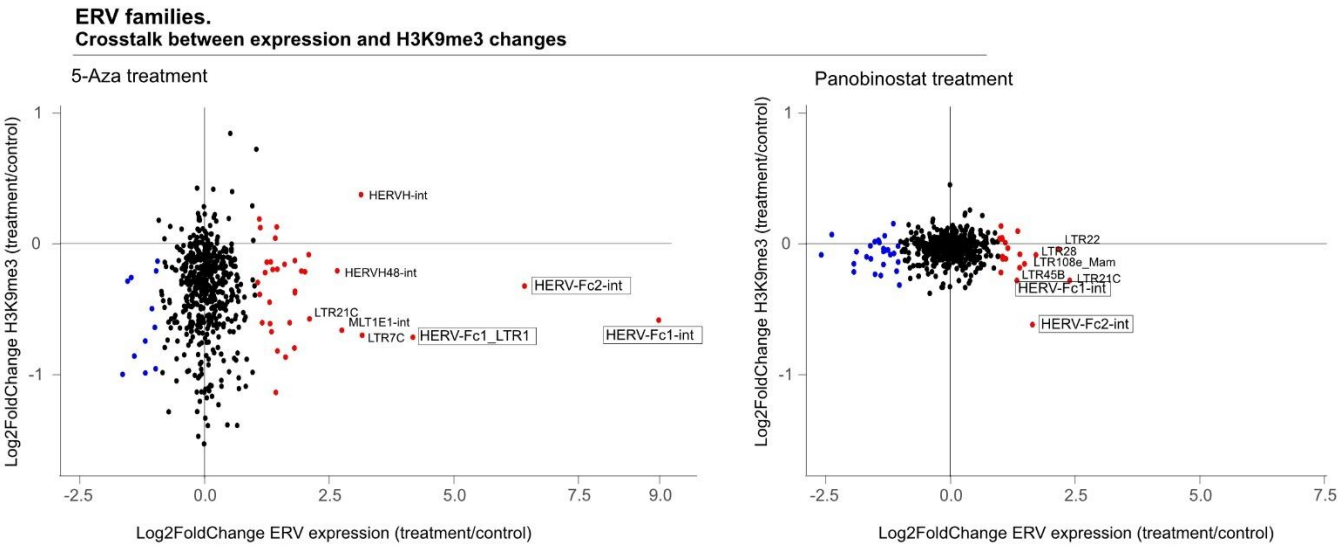


**Figure 2.7. Differential ERV family's expression and chromatin changes upon treatments with 5-Aza and Panobinostat**

ERV family analysis by plotting log2 fold-change of H3K9me3 signal (normalized by input) in the y-axis against log2 fold-change of expression in x-axis for treatment vs. control cells (n=1). ERV families with notably changed expression (log2 fold change > 1) are colored (red - increased expression in treated cells, blue - decreased expression in treated cells). The few most upregulated ERV families are labeled. Upregulated Fc1 and Fc2 ERV families are framed.

Next, we investigated HERV-Fc1 and Fc2 family activities on a single loci level. On a loci level, only one copy of the HERV-Fc1 family was strongly derepressed (7q11.1). This locus has the size and structure of a full-length ERV element (Figures 2.8A and 2.9A). We use only unique mapped reads for the single loci analysis and multi-mapped for the ERV families. Therefore, we checked if derepressed HERV-Fc1 elements with low mappability are present, which could be detected only with multi-mapped reads but not with unique ones. In this case, they would be undetected in a single ERV loci analysis. Therefore, we aligned both uniquely mapped reads and multi-mapped reads to HERV-Fc families to assess mappability levels. Only 6 % of reads annotated to the expressed HERV-Fc1 family in 5-Aza treatments were lost due to unique mapping (Figure 2.8B). Hence, HERV-Fc families have a high mappability rate. Interestingly, a single HERV-Fc1 locus with extremely high overexpression in 5-Aza treated cells is responsible for 91 % of the overall HERV-Fc1 family expression (Figure 2.8B). To assess the expression level of this HERV-Fc1 locus, we plotted expression densities of ERVs and coding genes. As expected, expressed ERV loci (orange distribution) on average had much weaker expressions compared to coding genes (black distribution) (Figure 2.8C). Interestingly, the highly upregulated HERV-Fc1 locus (orange circle) had expression values above average even compared to coding genes in 5-Aza treated cells.

In general, only a few ERV copies are derepressed upon Panobinostat treatment, while more are derepressed upon 5-Aza (Figure 2.8A). Interestingly, several highly upregulated upon 5-Aza copies have the size and structure of full-length HERVH and HERV3 elements with high expression values (Figures 2.8A and 2.9C, D). In Panobinostat treated cells, HERV-Fc1 or Fc2 loci were not among highly derepressed ERVs. While on a family level, they showed upregulation. This could be due to mild upregulation of many HERV-Fc1 and Fc2 loci, which together contributed to increased family expression.

Next, we assessed if upregulation of ERV loci in treated cells is associated with H3K9me3 change. We found that 70 % of upregulated ERV loci exhibited a reduction of H3K9me3 repressive mark in 5-Aza treated cells, while only 30 % - in Panobinostat treated cells (Figure 2.8.D). Pointing out that ERV upregulation upon 5-Aza treatment might be connected with reduction of repressive histone mark H3K9me3. Some specific examples of upregulated ERV loci upon treatments and associated H3K9me3 profiles are shown in Figure 2.9.

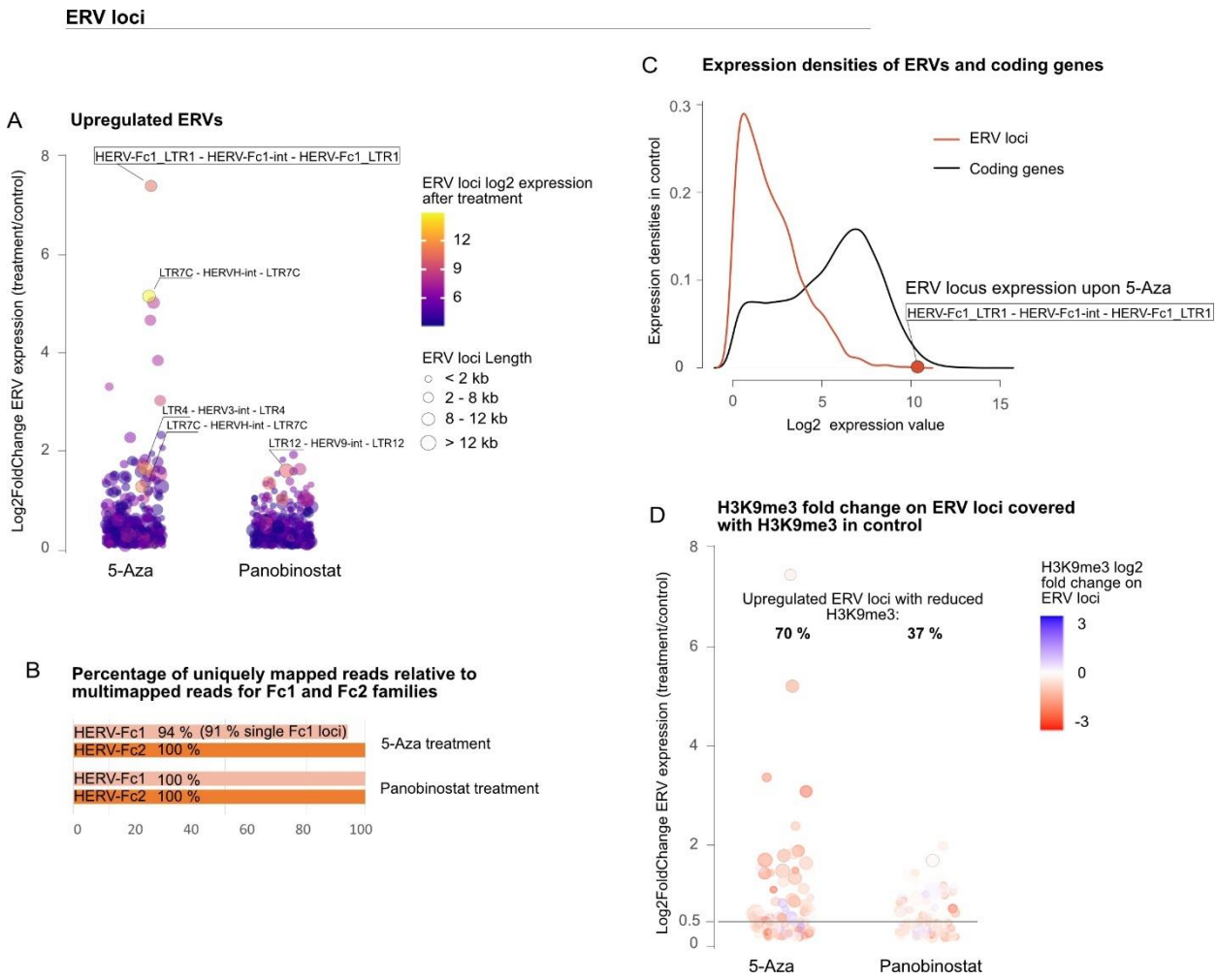


**Figure 2.8. Differential ERV loci expression and chromatin changes upon treatments with 5-Aza and Panobinostat**

A. Dot plot, showing positive, more than 0.1, log2 fold-change of ERV loci expressions for treatment vs. control cells (n=1). The length of corresponding ERV loci is coded by circle sizes. ERV loci log2 expression in treatments is coded by color. ERV loci with log2 expression > 10 are labeled. Full HERV-Fc1 locus is highly upregulated and expressed in 5-Aza treated cells.

B. Bar plot, showing the percentage of uniquely mapped RNA reads relative to multi-mapped RNA reads for HERV-Fc1 and HERV-Fc2 families in treated cells. In addition, the percentage of highly upregulated HERV-Fc1 loci expression relative to entire HERV-Fc1 family expression in 5-Aza treated cells is marked.

C. Density plot, showing log2 expression distributions for expressed ERV loci and coding genes in control. Plotted expression values for ERV loci and genes are more than 0 log2. HERV-Fc1 locus from the previous plot is marked.

D. Dot plot, showing positive, more than 0.1, log2 fold-change of ERV loci expressions for treatment vs. control cells (n=1). Out of those, only ERV loci with H3K9me3 enrichment in control are plotted. ERV loci are counted as loci with reduced H3K9me3 if H3K9me3 enrichment reduced more than 0.5 log2 fold change in treatments compared to control. The length of corresponding ERV loci is coded by circle sizes. Log2 fold change of H3K9me3 enrichment in treatments compared to control is coded by color.

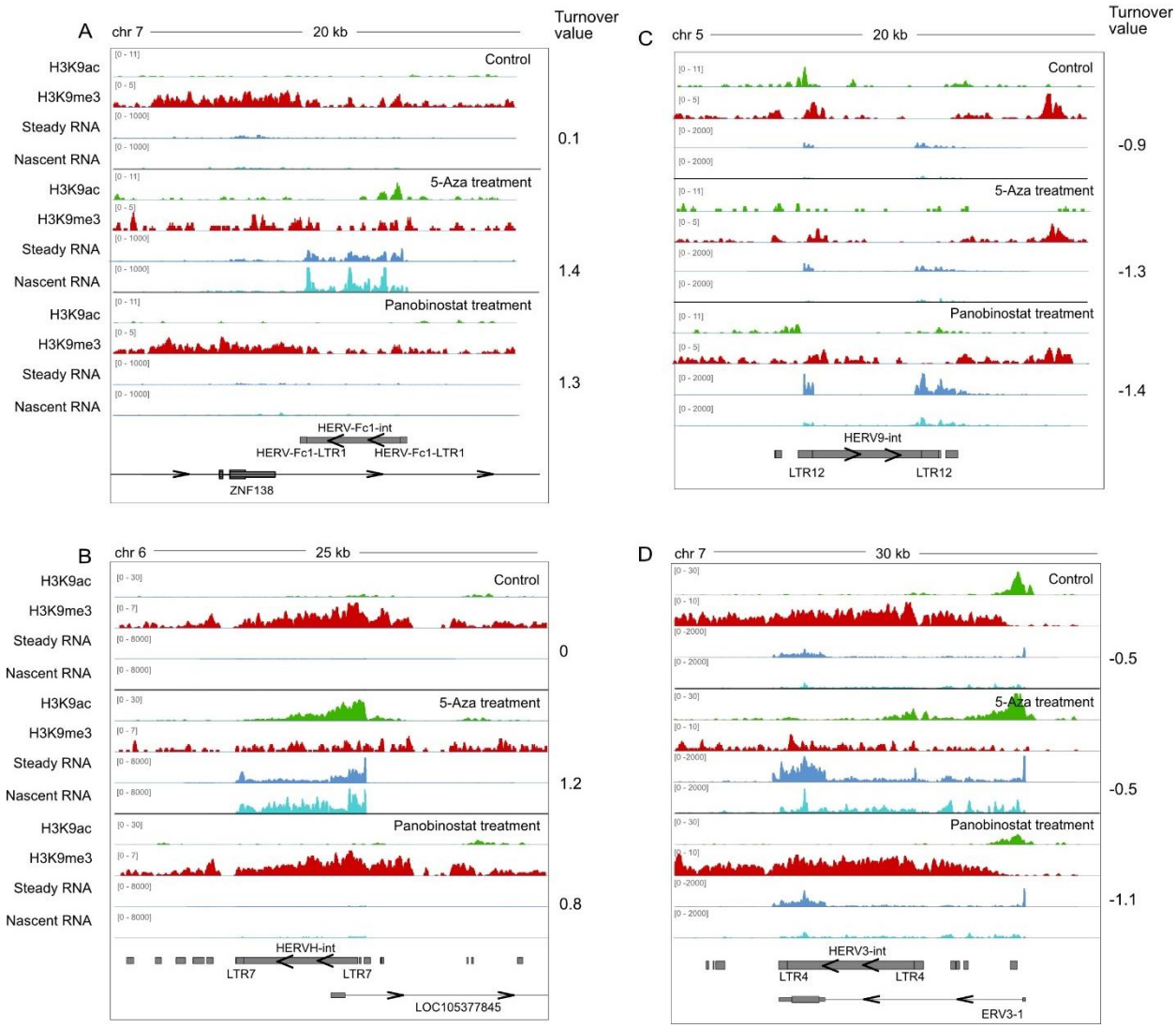


**Figure 2.9. Examples of ERV loci upregulation upon 5-Aza and Panobinostat treatments. Upregulation upon 5-Aza is often associated with a prominent H3K9me3 reduction.**

A. Genome browser view of nascent, steady RNA, active H3K9ac, and repressive H3K9me3 histone marks for upregulated HERV-Fc1 full-length loci.

B. Genome browser view of nascent, steady RNA, active H3K9ac, and repressive H3K9me3 histone marks for upregulated HERVH full-length loci.

C. Genome browser view of nascent, steady RNA, active H3K9ac, and repressive H3K9me3 histone marks for upregulated HERV9 full-length loci.

D. Genome browser view of nascent, steady RNA, active H3K9ac, and repressive H3K9me3 histone marks for upregulated HERV3 full-length loci.

Increased production of ERV-derived dsRNA was already reported upon cancer cells treatments with 5-Aza (Chiappinelli et al., 2015). dsRNA activates interferon responses through the RNA-sensing innate immune network mimicking viral infection and leading to immunogenic cancer cells death (Attermann et al., 2018). ERV-derived dsRNA could be produced from bidirectionally transcribed ERV loci (Sadeq et al., 2021). Therefore, we checked how many ERV loci were transcribed bidirectionally in our data set upon 5-Aza and

Panobinostat anti-cancer treatments. ERVs were assigned as bidirectional if the transcript amount from both strands was more than 2 log2-normalized nascent or steady RNA expression levels. Then, we calculated how many actively transcribed loci were bidirectional and plotted them as a percentage of all actively transcribed loci in each condition. In untreated THP-1 cells, 9 % of all actively transcribed ERV loci were bidirectional. Upon 5-Aza treatment amount of bidirectionally transcribed loci slightly increased to 10 %. Panobinostat treatment did not lead to a notable increase in the amount of bidirectionally transcribed loci (Figure 2.10A). The next plot shows how many out of upregulated ERV loci were bidirectionally transcribed. We calculated how many upregulated ERVs had bidirectional transcription and plotted them as a percentage of upregulated loci. 24 and 8 % of upregulated loci were transcribed bidirectionally upon 5-Aza and Panobinostat treatments, respectively (Figure 2.10B). In summary, 5-Aza treatment led to a slight increase in bidirectionally transcribed ERV transcripts but not Panobinostat treatment. In addition, taking into account a very mild effect and only one replicate, this conclusion should be taken with caution.

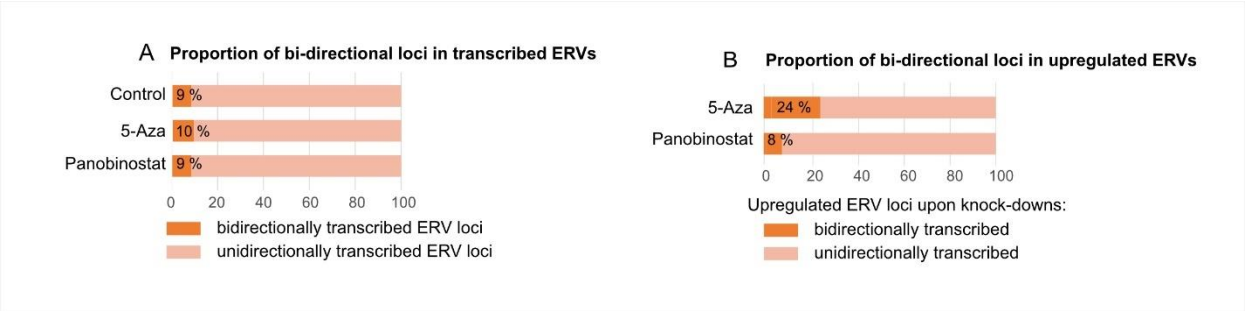


**Figure 2.10. Differential ERV loci expression and chromatin changes upon treatments with 5-Aza and Panobinostat**

A. Barplots, showing the proportion of transcribed ERV loci in each condition with bidirectional transcription versus transcribed ERV loci with unidirectional transcription. Actively transcribed ERVs in each treatment were selected if log2 nascent expression > 4.

B. Barplots, showing the proportion of upregulated ERV loci with bidirectional transcription versus upregulated ERV loci with unidirectional transcription for both treatments. ERV loci were counted as upregulated if nascent expression in treated cells was more than 4 log2-normalized value and log2 fold change between knockdown and control more than 0.5.

### 2.2.4 ERV transcript turnovers majorly do not change upon 5-Aza and Panobinostat treatments

To investigate transcript stability changes upon 5-Aza and Panobinostat treatments, we plotted ERV and coding genes turnovers, calculated as “log2 nascent expression - log2 steady expression” for each ERV loci and coding genes. According to our data, turnover of ERV and gene transcripts overall did not show notable changes upon treatments (Figure 2.11A). Therefore, an increased amount of ERV transcripts upon 5-Aza and Panobinostat anti-cancer treatments is a rather consequence of an increased ERV transcription but not stabilization of ERV-derived transcripts.

Although overall ERV transcript turnovers did not change upon treatments with epigenetic drugs, some ERV loci showed turnover changes (Figure 2.9). Therefore, we checked if ERV loci respond differentially to the 5-Aza and Panobinostat treatments in terms of turnover changes (Figure 2.11B,C). Based on the plots, we can conclude that in many cases (63 %) ERV loci with decreased turnover upon 5-Aza also show a decrease in turnover upon Panobinostat treatment and vice versa (44 %) (Figure 2.11C). The same we can conclude about ERV loci with increased transcript turnover (Figure 2.11B). Figure 2.9, panels A, B, C illustrate this phenomenon. Therefore, a subset of ERV loci exhibits an increase or decrease of turnover upon both treatments in a similar manner independently of induced chromatin changes. Although, some ERV loci still show transcript turnover increase or decrease upon one treatment and absence of turnover change in another treatment (Figure 2.11B,C).

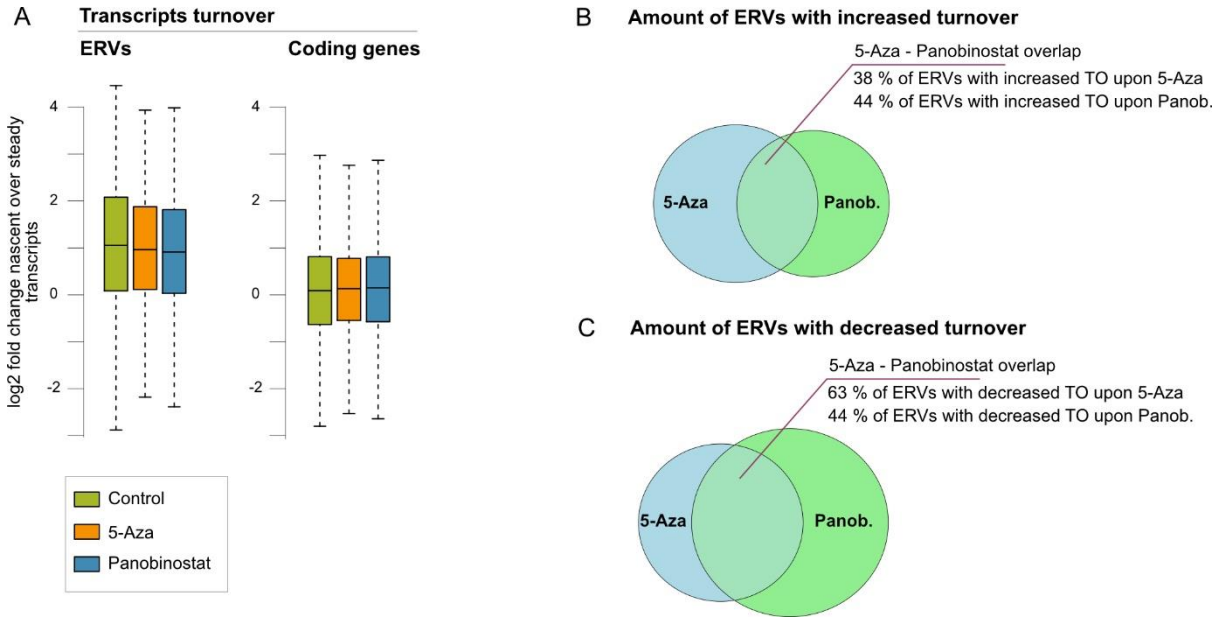


**Figure 2.11. ERV and gene turnover changes upon treatments**

A. Boxplots, showing transcripts turnovers of ERV loci and coding genes in control and upon treatments. Turnover of ERV transcripts slightly decreases upon treatments, while gene turnover on average is unchanged. ERV loci and genes were plotted, if they had log2 expression values more than 3 for both nascent and steady fractions for all conditions on average for robust turnover calculations (n=1).

B. Overlap of ERV loci with increased turnover upon 5-Aza and Panobinostat treatments. ERV loci were taken as an input for the plot if they had log2 expression values more than 3 for both nascent and steady fractions for all conditions on average. ERVs counted as loci increased turnover if turnover fold change between the treatment and control was higher than 0.5.

C. Overlap of ERV loci with decreased turnover upon 5-Aza and Panobinostat treatments. ERV loci were taken as an input for the plot if they had log2 expression values more than 3 for both nascent and steady fractions for all conditions on average. ERVs counted as loci with decreased turnover if turnover fold change between the treatment and control was lower than -0.5.

## 2.3 Knockdowns of splicing factors and exosome exonuclease lead to broad ERV upregulation and transcript turnover changes

### 2.3.1 Confirmation of efficient knockdowns

To investigate the influence of DIS3, SF3A3, and SRSF3 on ERVs transcriptional activity in humans, we utilized a knockdown approach with pre-designed Silencer Select siRNAs. We chose the knockdown approach since complete depletion of selected proteins by knock-out could lead to excessive cellular stress and faster cell death. 20-22 nt long siRNAs were transiently delivered into the cells using transfection reagent. Choice of transfection kit can affect knockdown efficiency. Therefore, we compared the performance of Lipofectamine RNAiMAX and JetPRIME transfection kits using DIS3-targeting siRNA. Manufacture-advised final siRNA concentrations range from 10 nM to 50 nM. Therefore, we have tested two concentrations: 20 nM and 40 nM. 72 hours (standard siRNA treatment time) post-transfection with siDIS3, RT-qPCR was performed on collected cells to assess (1) performance of transfection kits and (2) siRNA concentrations influence on the knockdown efficiency. Test experiment showed that siRNA delivery with Lipofectamine RNAiMAX reagent led to more efficient target transcript reduction in HeLa cells (Figure 2.12A). 20 nM concentration of siRNA was chosen over 40 nM since the difference in transcript reduction using 20 nM was minimal (Figure 2.12A). In addition, a good practice is to use less amount of siRNA to reduce off-target effects.

The experiment was as follows: cervical cancer HeLa cell line was treated with 20 nM negative control siRNA and siRNAs targeting DIS3, SF3A3, SRSF3, SETDB1. SETDB1 is a known ERV silencer in mammals. Therefore, its knockdown was performed to compare the mode of ERV derepression upon knockdown of splicing factors and nuclear exosome exonuclease. siRNAs, except negative control siRNA, were ordered in a set of 2 for each gene. Two replicates of siRNA treatments were collected, where each replicate was treated with a different set of siRNAs. After 72 hours of siRNA transfections, samples for TT-seq and chromatin ChIP were collected (Figure 2.12B) together with quality control samples for RT-qPCR and western blot.

Next, we performed TT-seq on knockdown samples to assess amounts of nascent and steady transcripts and transcripts turnover. Nascent transcripts comprise only a small proportion of total transcripts. Poor streptavidin beads separation of 4sU labeled nascent RNA fraction and unlabeled steady RNA can lead to labeled fraction contamination with unlabeled fraction and create a bias in the readout. Therefore, the separation efficiency of labeled and unlabeled transcripts is a crucial step in the TT-seq protocol, which needs to be controlled. For this, we used 4sU RNA spike-ins and RNA spike-ins with standard uridine, which were synthesized based on an in-vitro transcription of 4 spike-ins with a unique sequence from ERCC spike-in

mix with either standard dNTPs or dNTPs plus 4-Thio-UTP. To quality control efficiency of 4sU incorporation, all spike-ins were biotinylated with thiol-specific biotin and incubated in a dot blot with biotin-specific streptavidin-HRP followed by electrochemiluminescence visualization. As expected, 4sU spike-ins #3 and #5 were successfully visualized in the dot blot, while unmodified spike-ins #4 and #6 were not bound by thiol-specific biotin and did not show any signal (Figure 2.12C). 4sU labeled and unlabeled RNA spike-ins were mixed in equal amounts and added into each sample before RNA extraction. After nascent and steady RNA sequencing, labeled and unlabeled spike-ins read counts were calculated for both RNA fractions in each sample (Figure 2.12D). As expected, in total RNA fractions, all 4 spike-ins were represented equally, while in nascent (labeled) fractions, unlabeled spike-ins were heavily underrepresented. Particularly, the proportion of retained unlabeled RNA spike-ins in nascent fractions varied between 0.13 – 1.8 % (Table 2.1).

**Table 2.1. Percentage of 4sU-labeled RNA fraction contamination by 4sU-unlabeled RNA, measured by spike-ins.**

| Sample | Replicate | Amount of retained 4sU-unlabeled RNA spike-ins in 4sU-labeled fraction after separation, % | Sample | Replicate | Amount of retained 4sU-unlabeled RNA spike-ins in 4sU-labeled fraction after separation, % |
|---|---|---|---|---|---|
| siCtrl | 1 | 1.06 | siCtrl | 2 | 0.65 |
| siDIS3 | 1 | 0.98 | siDIS3 | 2 | 0.13 |
| siSF3A3 | 1 | 1.13 | siSF3A3 | 2 | 1.72 |
| siSRSF3 | 1 | 1.80 | siSRSF3 | 2 | 0.14 |
| siSETDB1 | 1 | 1.14 | siSETDB1 | 2 | 0.42 |

To control targeted transcripts reduction, RT-qPCR was performed for both replicates. For each knockdown, not only targeted gene expressions were assessed but also expression levels of other genes investigated in the experiment. RT-qPCR shows strong targeted transcripts reduction upon siRNA knockdowns. On average upon knockdowns DIS3 expression was reduced on 85 %, SF3A3 – on 89 %, SRSF3 – on 83 %, and SETDB1 – on 64 % compared to control (Figure 2.12E). Figure 2.12F depicts steady transcripts per kilobase million (TPM) expression for the same genes relative to control taken from TT-seq data. As expected, RT-qPCR and TT-seq showed comparable expression reduction of targeted by siRNAs genes. RT-qPCR and sequencing data showed different relative expressions of SETDB1 upon siSF3A3 treatment since the RT-qPCR primer pair targets the 13$^{th}$ exon of SETDB1, which is not included in the short SETDB1 isoform. SETDB1 total transcript amount was not reduced upon SF3A3 knockdown, instead, short SETDB1 isoform was predominantly

expressed upon knockdown of SF3A3 (Figure 2.12G). Western blot illustrated notable protein level reduction of targeted proteins (Figure 2.12H). Similar to the RT-qPCR set-up, not only targeted protein levels were assessed for each knockdown but also protein levels of other genes investigated in the experiment. On a protein level, SETDB1 reduction upon SF3A3 knockdown was not observed (Figure 2.12H). To exclude the influence of knocked down genes on protein levels of another important transcriptional ERV silencer - TRIM28 - it was assessed in western blot in addition to other proteins. Western blot showed an absence of notable TRIM28 protein reduction upon assessed knockdowns.

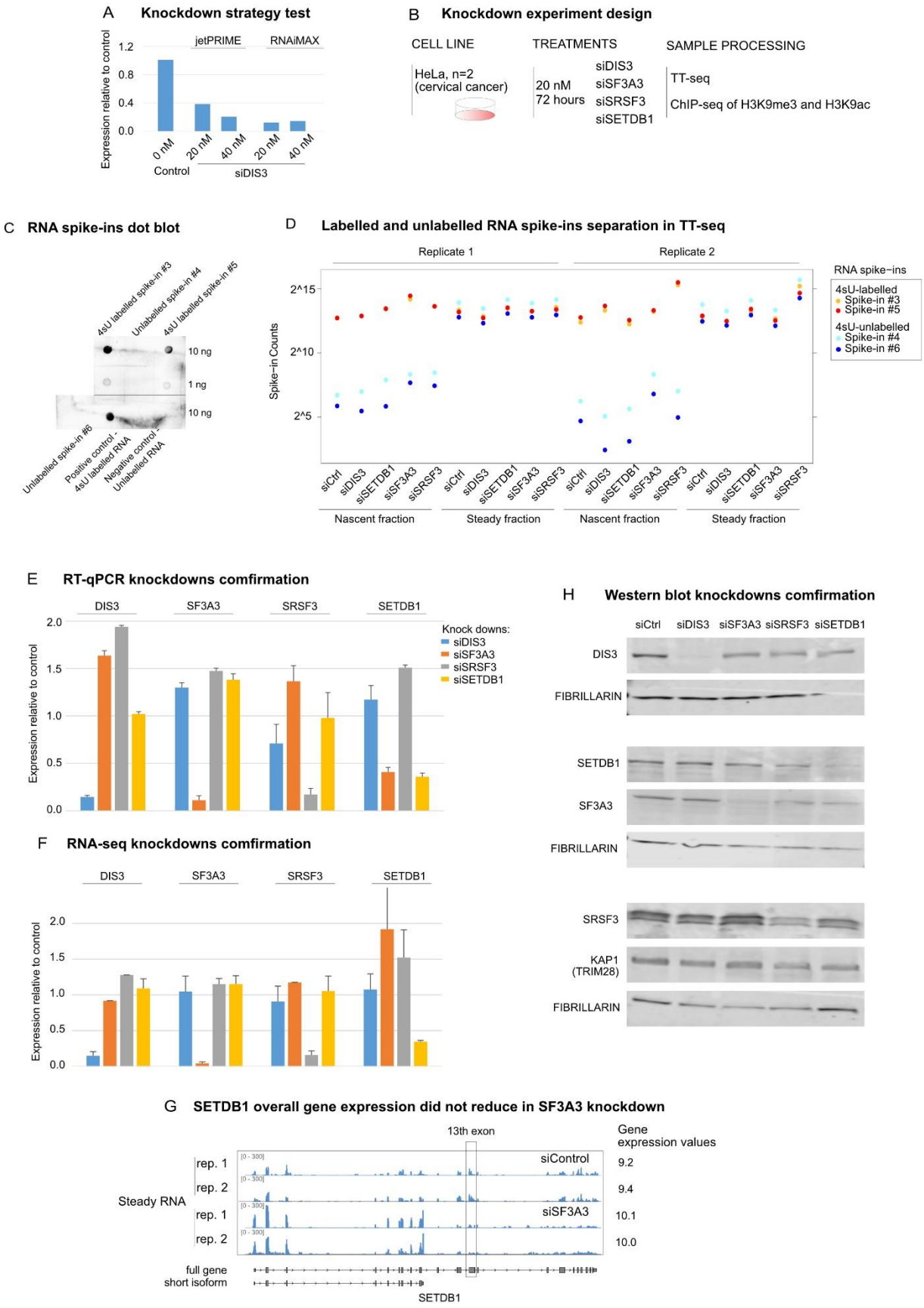


**Figure 2.12. Confirmed gene expression and proteins reduction of corresponding genes upon knockdowns**

A. Knockdown strategy test (n=1).

B. Experiment design of siRNAs knockdowns.

C. Dot blot, confirming successful in-vitro synthesis of 4sU-containing RNA spike-ins (#3 and #5). Labeled and Unlabelled RNA spike-ins were biotinylated with biotin–HPDP and visualized using streptavidin HRP. 4sU-labeled and biotinylated RNA spike-ins show strong signal after thiol-specific biotinylation, while RNA spike-ins with normal uridine show no signal.

D. Successful nascent and steady RNA fractions separation, measured by RNA spike-ins. RNA spike-in counts of four RNA spike-ins (two 4sU-labelled and two 4sU-unlabelled RNA species) were detected in each sequenced sample. RNA spike-ins were mixed in equal amounts and added to each RNA sample before fractions separation.

E. Expression levels of DIS3, SF3A3, SRSF3, and SETDB1 by RT-qPCR in siDIS3, siSF3A3, siSRSF3, and siSETDB1 knockdown HeLa cells compared to control samples. For each replicate, expression was normalized by that of GAPDH. Error bars represent standard deviations of 2 biological replicates.

F. Steady transcripts per kilobase million (TPM) expression levels of DIS3, SF3A3, SRSF3, and SETDB1 by TT-seq in siDIS3, siSF3A3, siSRSF3, and siSETDB1 knockdown HeLa cells compared to control samples. Error bars represent standard deviations of 2 biological replicates.

G. Genome browser view of steady RNA-seq for the SETDB1 gene in control and in SF3A3 knockdown. Full SETDB1 gene, shorter isoform, and 13$^{th}$ exon are noted. Full gene log2 transformed normalized expression values for each replicate are written in the plot.

H. Protein expression analysis by western blots confirms the reduction of DIS3, SF3A3, SRSF3, and SETDB1 proteins in corresponding knockdowns. Fibrillarin was used as a loading control. In addition, TRIM28 protein levels were confirmed to be unchanged upon knockdowns.

### 2.3.2 Gene expression changes upon siRNA knockdowns

Next, we characterized expression patterns of coding genes between control and siRNA knockdowns. All knockdowns showed significant transcripts reduction targeted by siRNAs genes (Figure 2.13A). Specifically, in DIS3 knockdown, DIS3 expression fold change between control and siDIS3 treatment was -5.3; in SF3A3 knockdown, SF3A3 expression fold change was -7.5; in SRSF3 knockdown, SRSF3 expression fold change was -5; in SETDB1 knockdown, SETDB1 expression fold change was -3.3.

SF3A3 knockdown led to the highest gene deregulation compared to the rest of knockdowns. The up-regulated gene sets in SF3A3 knockdown were enriched for various RNA processing pathways, such as mRNA splicing, processing of capped intron-containing mRNA, mRNA metabolism, and rRNA processing (Figure 2.13B). In particular, DEAH-box helicases (DHXs), Sm-like proteins (LSMs), pre-mRNA processing factors (PRPFs) together with SF3B and SRSF gene clusters were widely represented in upregulated mRNA splicing pathway. Small nuclear ribonucleoproteins (SNRPs) and ribosomal proteins L and S (RPLs and RPSs) were enriched in the rRNA processing pathway. Upon SF3A3 knockdown, many genes involved in RNA transport and RNA maturation were upregulated, such as cleavage and polyadenylations factors, tRNA methyltransferases (TRMTs) (Supplementary table). Accumulation of unspliced pre-mRNAs was observed before upon SF3A proteins knockdowns (Tanackovic and Krämer, 2005). Therefore, upregulation of genes involved in mRNA splicing could be a compensatory response mechanism to the deletion of an important component of splicing and RNA processing machinery.

SRSF3 knockdown also resulted in a notable amount of deregulated genes, where upregulated genes were enriched for immune-related pathways (Figure 2.13B). Specifically, interleukin-1 alpha and beta (IL1a and ILb), interleukin-16 (IL16), and interleukin-18 (IL18) are upregulated by knockdown (Supplementary table). In addition, several core apoptotic genes were significantly upregulated, such as caspase-9 (CASP9), cathepsin L (CTSL), tumor necrosis factor receptor-associated factor 1 (TRAF1), and DNA damage-inducible transcript 3 (DDIT3). This observation is in line with previous studies showing that SRSF3 is an oncogene, and its silencing is associated with an increase in the protein synthesis of immune mediators and cancer cell death (Guo et al., 2020; Jimenez-Vacas et al., 2020; Zhou et al., 2020).

DIS3 and SETDB1 knockdowns led to significant deregulation of a small subset of genes. Therefore, pathway enrichment analysis was not performed for these knockdowns. Interestingly, Chemokine ligand 20 (CCL20) is upregulated not only in siSRSF3 cells but in all knockdowns. Previous reports showed that CCL20 expression is induced by dsRNA and may play a potential role in antiviral responses (Ghosh et al., 2009; Liu et al., 2019).

Next, we addressed the influence of splicing factors knockdowns on alternative splicing events and subsequently on differential isoforms representation. For this, we calculated isoforms expressions in all conditions and scaled the cumulative expression of all isoforms for each gene to 1. In this way, we obtained the expression ratio for each isoform. Then, we visualized isoforms expression ratios between knockdowns and control by volcano plot. On the x-axis, we plotted relative enrichment of isoforms ratios in knockdowns over control and on the y-axis - log10 respective adjusted p-values. In addition, isoforms were marked with red if they came from genes detected as differentially spliced. As expected, splicing factors knockdowns resulted in the highest amount of differentially spliced genes. Specifically, both SF3A3 and SRSF3 knockdowns led to 555 and 278 differentially spliced genes, respectively, while DIS3 and SETDB1 knockdowns resulted in 36 and 9 differentially spliced genes, respectively (Figure 2.13C). These data confirmed the main function of SF3A3 and SRSF3 as splicing factors. Differentially spliced and expressed gene sets in SF3A3 knockdown were enriched for highly similar pathways. Namely, differentially spliced gene set in SF3A3 knockdown was enriched for various RNA processing pathways, such as mRNA splicing, processing of capped intron-containing mRNA, mRNA metabolism, and rRNA processing (Figure 2.13D). Differentially spliced gene set in SRSF3 knockdown was enriched for divergent pathways, compared to differentially expressed genes. Namely, differentially spliced gene set in SRSF3 knockdown was enriched for translational initiation, mRNA metabolism, and rRNA processing

pathways (Figure 2.13D).

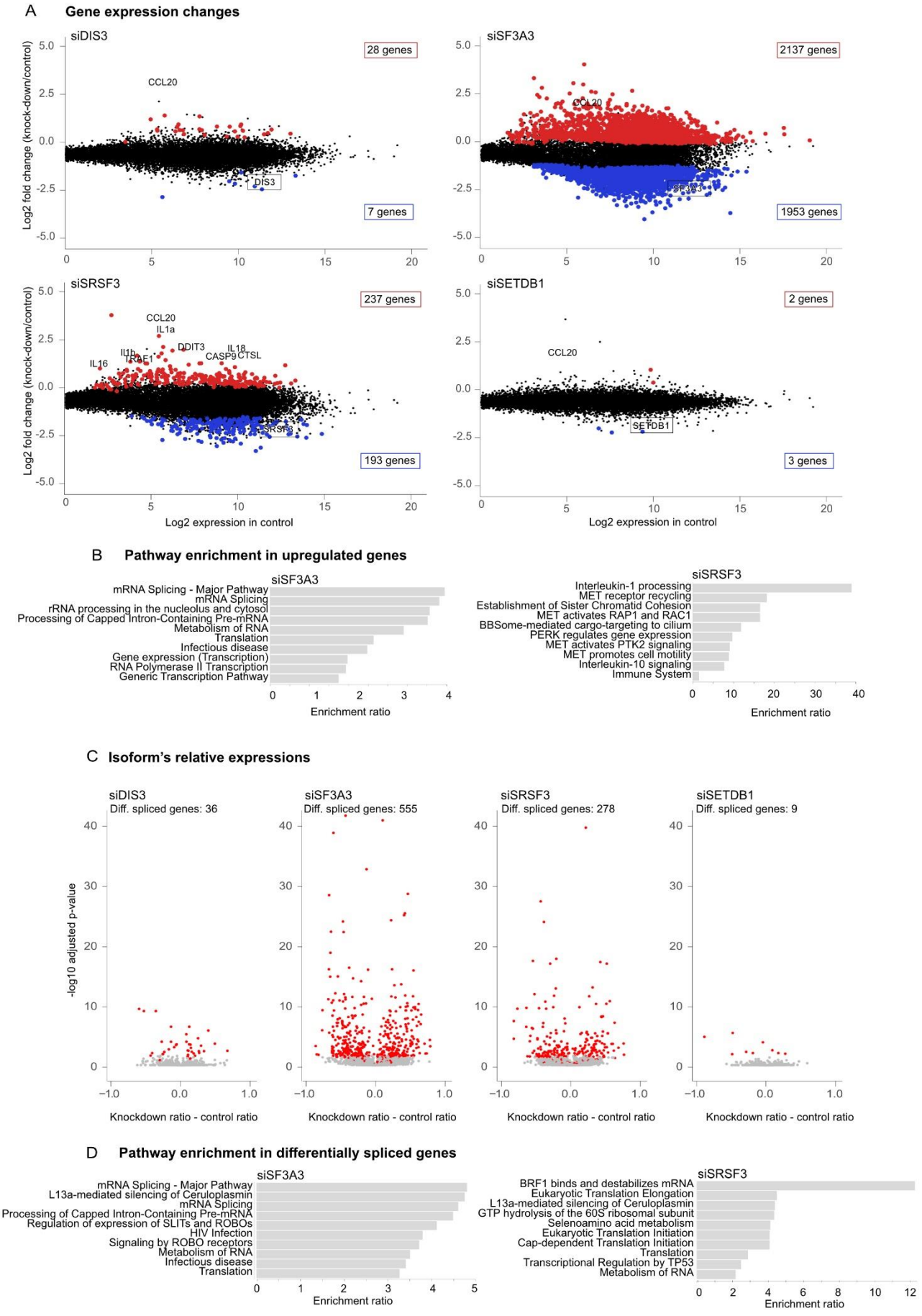

**Figure 2.13. Differential gene and isoform expression in knockdown cells**

A. Dot plots, showing log2 steady expression in control versus a log2-fold change of protein-coding genes in knockdown versus control cells after 3 days of siRNA treatment. Genes with significantly changed expression (adjusted p-value < 0.05; n=2) are coloured with red - increased expression, blue - decreased expression. Knockdown target genes are labeled and marked with a black box. A number of significantly upregulated and downregulated genes are marked red and blue boxes, respectively.

B. Over-representation analysis (ORA) for Reactome pathway enrichment in significantly upregulated genes. Shown is the enrichment ratio of pathways in upregulated genes relative to all coding genes.

C. Volcano plot, showing on x-axis relative enrichment of isoforms ratios in knockdowns over control and on the y-axis -log10 adjusted p-values (n=2). In addition, isoforms were marked with red if they came from genes detected as differentially spliced.

D. Over-representation analysis (ORA) for Reactome pathway enrichment in significantly differentially spliced genes relative to all coding genes.

### 2.3.3 Knockdown of splicing factors and exosome exonuclease leads to broad ERVs upregulation and increased proportion of expressed bidirectionally transcribed ERVs

Next, we evaluated ERV transcriptional deregulation upon knockdowns. Knockdowns of DIS3, SF3A3, and SRSF3 showed striking ERV up-regulation, while the amount of downregulated ERV loci was considerably lower (Figure 2.14A). Moreover, splicing factors knockdown resulted in the strongest ERV up-regulation. Interestingly, SETDB1 knockdown led to much less ERV copy derepression compared to knockdowns of splicing factors and nuclear exosome exonuclease (Figure 2.14A). As we expected that SETDB1 would lead to stronger ERV derepression, we assessed publicly available SETDB1 knockout data in HeLa cells from the NCBI database to rule out the possibility of inefficient SETDB1 knockdown (Timms et al., 2016). We processed data from knockout RNA-seq using the same pipeline we used in our study. Both genes and ERVs showed a similar level of deregulation compared to our data (Figure 2.15), pointing out that SETDB1 reduction leads to mild gene and ERV deregulation in HeLa cells.

To assess overlap of SF3A3, SRSF3, and DIS3 function in ERV derepression, we plotted proportion overlap of upregulated ERVs sheared between knockdowns in a Venn diagram. Interestingly, we found that 70 % of upregulated ERVs upon DIS3 knockdown are also upregulated in SRSF3 knockdown, with 26 % of upregulated ERVs in SRSF3 knockdown also upregulated in DIS3 reduced cells (Figure 2.14B). Overlap between DIS3 and SF3A3 derepressed ERVs is lower than in DIS3 – SRSF3 pair. Specifically, 40 % of upregulated ERVs upon DIS3 knockdown are upregulated in SF3A3 knockdown, with 13 % of upregulated ERVs in SF3A3 knockdown also upregulated in DIS3 reduced cells. Therefore, we detected a major overlap between upregulated ERVs upon DIS3 and SRSF3 knockdown, while the overlap between DIS3 and SF3A3 knockdowns was substantially lower.

In several model organisms, such as *S. pombe, D. melanogaster, C. elegans*, low or basal transcription level from repetitive regions was shown to help to recruit repressive complexes and prevent high expression of such loci. Such transcriptional events are poorly explored in mammals. Therefore, we checked whether upregulated ERV loci in our data already had basal expression in control. To answer it, we calculated the amount of upregulated loci. ERV loci were counted as upregulated if nascent expression in knockdown was more than 5 log2-normalized value and log2 fold change between knockdown and control more than 1. Then, we calculated how many upregulated loci were already expressed in control and plotted them as a percentage of upregulated loci (Figure 2.14C). Interestingly, the vast majority of upregulated ERV loci already had basal transcription in control. This finding suggested that transcriptionally active ERV loci, in most cases, are not tightly repressed but produce transcripts at a certain level. Which could help silencing machinery to sense such loci and prevent their active transcription (Aravin et al., 2008; Carter et al., 2020a; Li et al., 2014a).

Next, we investigated if RNA processing proteins knockdowns resulted in the derepression of ERV loci with coding potential. For that, we assessed expression levels and sizes of upregulated ERV loci. Many full ERVs are upregulated and highly expressed upon SF3A3 knockdown (Figure 2.14D). Similar to volcano plots, this plot showed that splicing factors had the highest amount of upregulated ERV loci. All knockdowns led to upregulation of ERVs subset with expressions on the level of coding genes while having length and structure of full ERV element – internal part with retroviral genes flanked by LTR promoters (Figure 2.14D).

Next, we checked how many ERV loci were transcribed bidirectionally. In particular, we calculated how many actively transcribed loci were bidirectional and plotted them as a percentage of all actively transcribed loci in each condition (Figure 2.14E). In untreated HeLa cells, 27 % of all actively transcribed ERV loci were bidirectional. Interestingly, all knockdowns led to increased production of bidirectional transcripts, especially siSRSF3, upon which 45 % of all transcribed ERV loci were bidirectional (Figure 2.14E). Next, we wanted to distinguish if upregulated ERVs were enriched with bidirectionally transcribed loci or ratios between uni- and bidirectional ERV loci in knockdowns were comparable to the control situation. For that, we assessed how many bidirectionally transcribed ERV loci are among upregulated loci. In particular, we calculated how many upregulated ERV loci had bidirectional transcription in knockdowns and plotted them as a percentage of upregulated loci (Figure 2.14F). Upon knockdowns, 40 to 60 % of upregulated loci were transcribed bidirectionally. Potentially, bidirectional transcription from ERV loci upon splicing factors and nuclear exosome exonuclease knockdowns could lead to increased double-stranded RNA production.

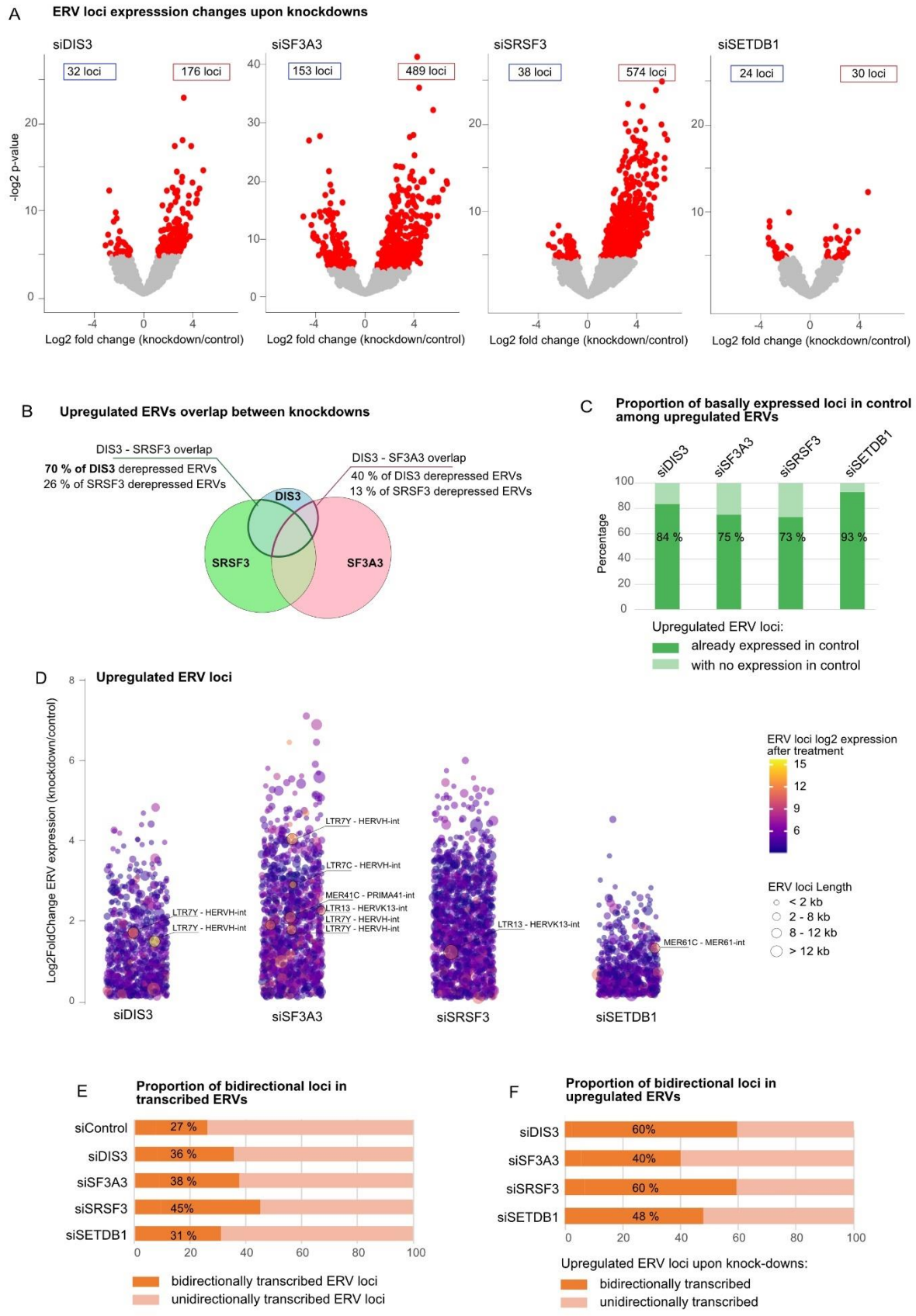


**Figure 2.14. Differential ERV loci expression in knockdown cells**

A. Volcano plots, showing a log2-fold change of ERV loci expression in knockdown versus control cells in the x-axis and –log2 p-values in the y-axis. ERV loci with significantly changed expression (p-value < 0.05;

n=2) are coloured with red. A number of significantly upregulated and downregulated ERV loci are marked red and blue boxes respectively.

B. Overlap of upregulated ERVs, sheared between SF3A3, SRSF3, and DIS3 knockdowns.

C. Barplots, showing the proportion of upregulated ERV loci with expression in control versus upregulated ERV loci silent in control (n=2, mean values are plotted). ERV loci are counted as expressed in control if log2 expression value > 2, and as silent if log2 expression value < 2. ERV loci were counted as upregulated if nascent expression in knockdown was more than 5 log2-normalized value and log2 fold change between knockdown and control more than 1.

D. Dotplot, showing log2 fold-change of ERV loci expressions for knockdowns vs. control cells. Length of corresponding ERV loci is coded by circle sizes (n=2, mean values are plotted). ERV loci log2 expression in treatments is coded by color. Some full ERV loci with log2 expression > 10 are labeled. ERV with log2 nascent expression in knockdowns > 3 and log2 fold change between knockdown and control > 0.1 are plotted.

E. Barplots, showing the proportion of transcribed ERV loci in each condition with bidirectional transcription versus transcribed ERV loci with unidirectional transcription (n=2, mean values are plotted). Actively transcribed ERVs in each condition were selected if log2 nascent expression > 5.

F. Barplots, showing the proportion of upregulated ERV loci with bidirectional transcription versus upregulated ERV loci with unidirectional transcription for each knockdown (n=2, mean values are plotted). ERV loci were counted as upregulated if nascent expression in knockdown was more than 5 log2-normalized value and log2 fold change between knockdown and control more than 1.

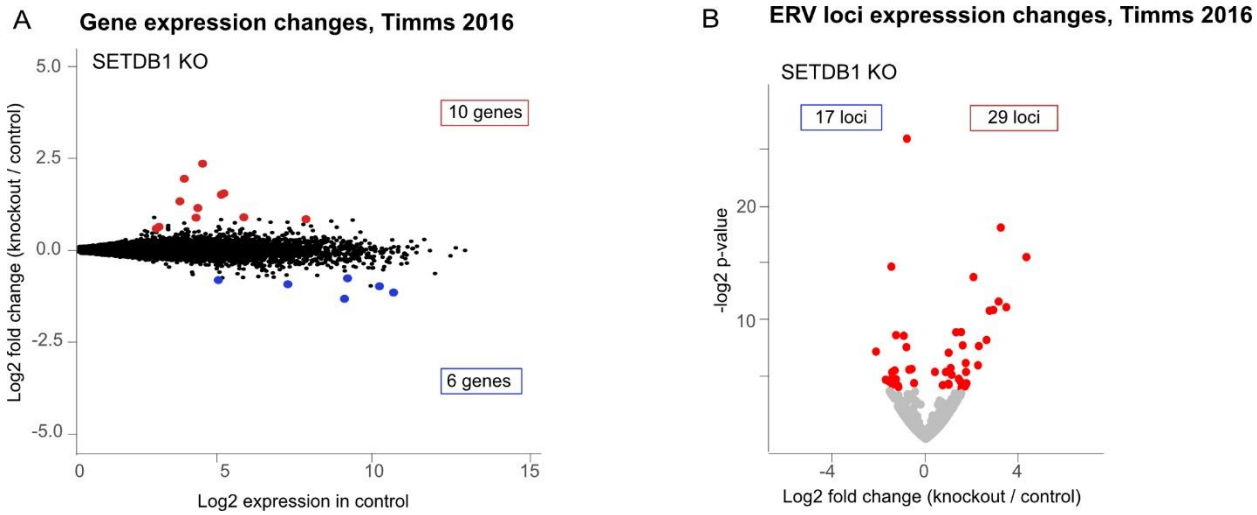


**Figure 2.15. Differential gene and ERV loci expression in SETDB1 knockout HeLa cells from the publicly available data set.**

A. Dot plot, showing log2 steady expression in control versus a log2-fold change of protein-coding genes in knockout versus control cells. Genes with significantly changed expression (adjusted p-value < 0.05; n=2) are coloured with red - increased expression, blue - decreased expression. Some of significantly upregulated and downregulated genes are marked red and blue boxes, respectively.

B. Volcano plot, showing a log2-fold change of ERV loci expression in knockdown vs. control cells in the x-axis and –log2 p-values in the y-axis. ERV loci with significantly changed expression (p-value < 0.05; n=2) are coloured with red. Some of significantly upregulated and downregulated ERV loci are marked red and blue boxes, respectively.

### 2.3.4 H3K9me3 and ERV expression changes upon knockdowns

To address repressive chromatin mark changes on ERV loci upon knockdowns, we first calculated the amount of ERV loci overlapped with ChIP H3K9me3 peaks (Figure 2.16A). Then, we selected only ERV loci overlapped with H3K9me3 peaks and calculated the amount of ERVs with notable H3K9me3 reduction upon knockdowns. Percentage of ERV loci with

reduced H3K9me3 was calculated based on a cutoff of -0.5 log2 fold-change H3K9me3 signal intensity reduction between knockdowns and control (Figure 2.16B). Figure 2.16A depicts that SETDB1 knockdown led to a broad reduction of ERVs amount with H3K9me3 peaks. In addition, SETDB1 knockdown led to H3K9me3 peak intensity reduction on half of ERVs covered with H3K9me3 in control (Figure 2.16B), in accordance with SETDB1 main function as histone methyltransferase. SRSF3 knockdown led to a notable reduction of both H3K9me3 peaks and ChIP signal intensity on ERV loci. SF3A3 and DIS3 knockdowns resulted in mild H3K9me3 peak intensity reduction, while the amount of ERV loci with H3K9me3 peaks did not reduce upon these knockdowns.

To investigate crosstalk between repressive chromatin mark H3K9me3 and ERV upregulation, we selected upregulated ERV loci covered with H3K9me3 in control. Then, we calculated how many of such loci showed H3K9me3 reduction and plotted them as a percentage of upregulated loci (Figure 2.16C). The majority (81 %) of upregulated ERVs upon SETDB1 knockdown had reduced H3K9me3 occupancy. Although, only a minority of upregulated ERVs upon rest of knockdowns showed notable K9me3 reduction – 5, 24, and 26 % for siDIS3, SF3A3, and SRSF3, respectively. An IGV browser example illustrates the upregulation of ERV loci associated with an increase in active histone mark H3K9ac at the LTR promoter (Figure 2.16D). At the same time, repressive histone mark H3K9me3 did not show a notable reduction upon ERVs derepression, suggesting that their derepression might not depend on K9me3 removal in these knockdowns.

In summary, although all knockdowns led to reduced H3K9me3 to a different extent, H3K9me3 depletion seems not a strict prerequisite for the ERV derepression.

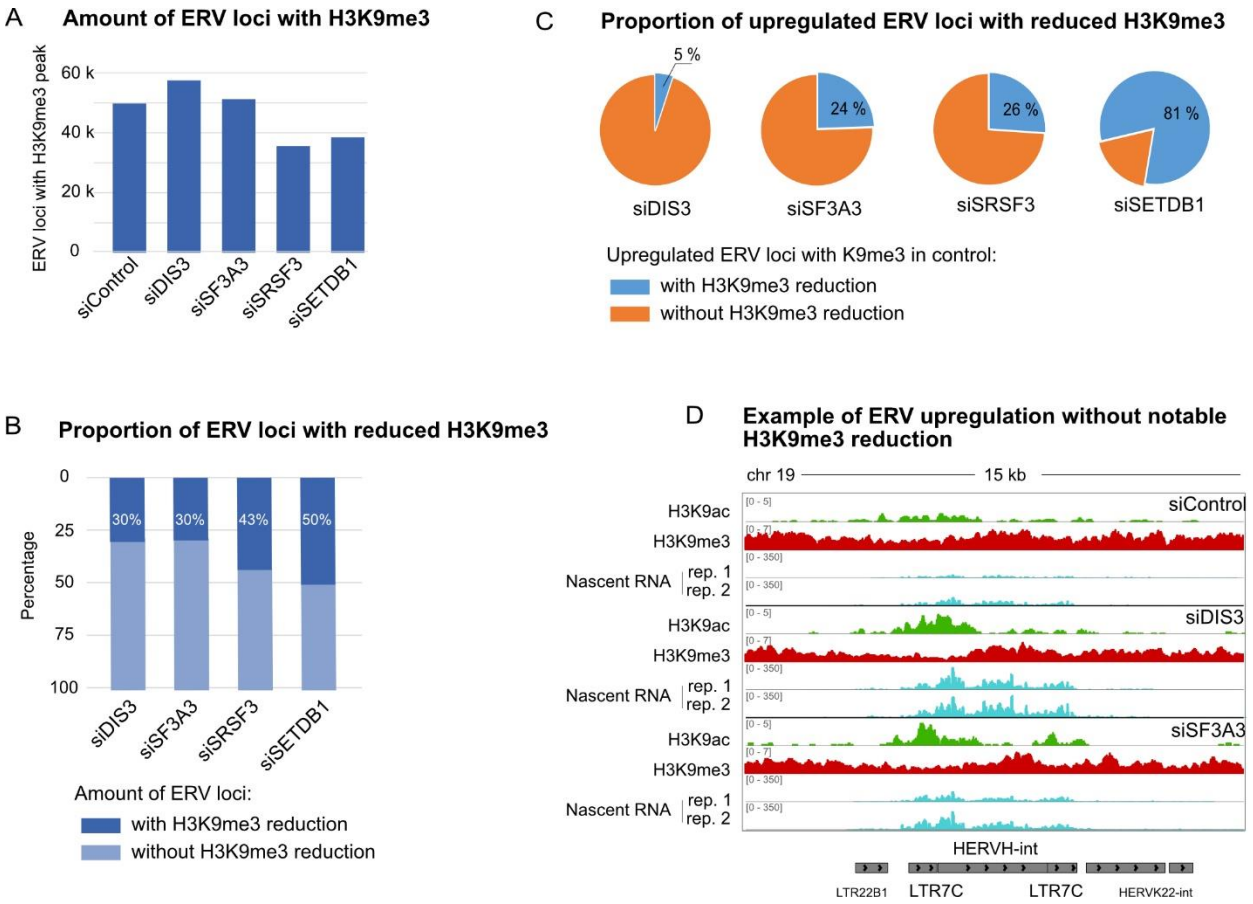


**Figure 2.16. H3K9me3 and expression changes on ERVs upon knockdowns**

A. Barplots of ERV loci amount with H3K9me3 ChIP-seq peaks. SETDB1 and SRSF3 knockdowns resulted in a notable reduction of H3K9me3 peaks amount on ERV loci.

B. Barplots of ERV loci with H3K9me3 peak in control, divided on ERVs with H3K9me3 reduction and ERVs without notable K9me3 reduction. The reduction threshold is -0.5 log2 fold change between knockdown and control H3K9me3 signal intensity. SETDB1 knockdown led to K9me3 reduction on half of ERVs. SRSF3 knockdown led to a notable reduction of ChIP signal intensity on ERVs as well, while SF3A3 and DIS3 knockdowns resulted in only mild H3K9me3 reduction on ERVs.

C. Piecharts of upregulated ERV loci with H3K9me3 ChIP-seq peaks in control. ERV loci were counted as upregulated if nascent or steady expression in knockdown was more than 5 log2-normalized value and log2 fold change between knockdown and control more than 0.5. H3K9me3 reduction threshold is -0.5 log2 fold change between knockdown and control H3K9me3 signal intensity. The majority of upregulated ERVs upon SETDB1 knockdown had reduced H3K9me3 occupancy, while the majority of upregulated ERVs upon the rest of knockdowns did not show notable K9me3 reduction.

D. Genome browser view of nascent RNA, H3K9ac, and H3K9me3 ChIP for an example of upregulated ERV loci upon DIS3 and SF3A3 knockdowns. Transcription and active histone mark H3K9ac increased at shown ERV locus, while repressive mark H3K9me3 did not show a notable reduction.

### 2.3.5 ERV transcripts are stabilized upon DIS3 knockdown

The balance between nascent RNA synthesis and the amount of steady transcripts significantly depends on RNA processing pathways. Therefore, we assessed ERVs' nascent and steady transcript amounts change upon knockdowns of splicing factors and nuclear exosome complex. The same comparison we carried out for coding genes (Figure 2.17A). Overall, SF3A3 knockdown resulted in an increased amount of newly synthesized RNA and a moderate increase of steady transcripts. Better illustrate transcript stability changes, we

plotted ERV and coding genes turnovers, calculated as “log2 nascent expression - log2 steady expression” for each locus (Figure 2.17B). Surprisingly, the SF3A3 splicing factor acts as a strong ERV transcriptional repressor, but its deletion is not associated with the stabilization of ERV-derived transcripts, as we expected (Figure 2.16A, B). SRSF3 knockdown resulted in increased both nascent and steady RNA amounts. Similar to the SF3A3 splicing factor, SRSF3 knockdown did not lead to ERV transcripts stabilization highlighting the fact that both splicing factors investigated in this study do not destabilize ERV-derived transcripts but rather act as potent ERV repressors on a transcriptional level.

In contrast to splicing factors knockdowns, in siDIS3 treated cells, nascent transcripts amount did not markedly increase, while the amount of stable transcripts increased substantially (Figure 2.17A, B). Gene turnovers do not show a notable difference between control and DIS3 knockdown, suggesting that ERV transcripts elimination is one of the central functions of DIS3. Therefore, we could confirm our hypothesis that nuclear exosome exonuclease DIS3 actively eliminates ERV transcripts, in accordance with the main DIS3 function as a central component of the nuclear transcript degradation complex. Nevertheless, in addition to ERV transcripts degradation function, DIS3 is also a transcriptional ERV repressor, although to a lesser extent than SF3A3 and SRSF3 splicing factors (Figure 2.17A).

Figure 2.17C depicts a representative example of ERV loci, illustrating transcript stabilization upon DIS3 knockdown. Despite the increase in nascent transcription is observed for all knockdowns, the proportion of stable compared to nascent transcripts substantially increased in DIS3 knockdown. SRSF3 knockdown did not lead to an increased amount of stable transcripts from the same ERV loci, while SF3A3 knockdown led to an increased transcript instability.

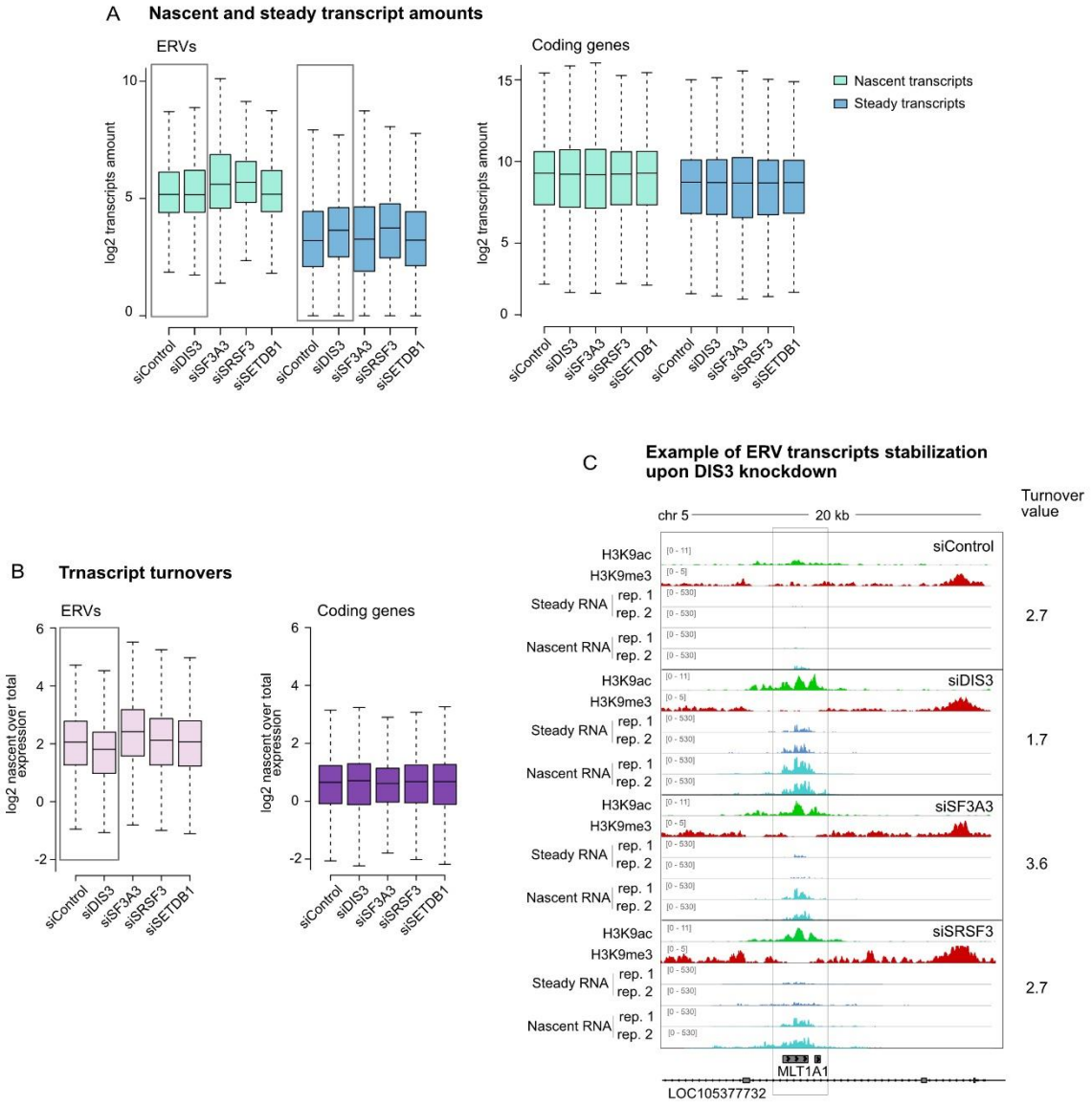


**Figure 2.17. ERV transcripts are stabilized upon DIS3 knockdown**

A. Boxplots, showing log2 transcript amounts from ERV loci and coding genes in control and knockdown cells. ERV loci and genes were plotted, if they had log2 expression values more than 3 for both nascent and steady fractions for all conditions on average for robust turnover calculations (n=2, mean values are plotted). Upon DIS3 knockdown, the total fraction of ERV transcripts is overall increased to a much greater extent than the nascent fraction.

B. Boxplots, showing transcripts turnovers of ERV loci and coding genes in control and knockdown cells. ERV loci and genes were plotted, if they had log2 expression values more than 3 for both nascent and steady fractions for all conditions on average for robust turnover calculations (n=2, mean values are plotted). Only upon DIS3 knockdown turnover of ERV transcripts was decreased.

C. Genome browser view of nascent, total RNA-seq, H3K9ac and H3K9me3 ChIP for 3 examples of upregulated ERV loci upon DIS3, and SF3A3 knockdowns. Transcript turnover values (average of 2 replicates) are noted. Upregulated upon siDIS3 ERV loci have stable transcripts, while the same loci upregulated upon splicing factor knockdown are degrading fast.

### 2.3.6 SF3A3 knockdown leads to ERV and KRAB zinc-finger genes transcript stabilization in the presence of H3K9me3

To investigate if ERV transcript turnover changes upon knockdowns correlate with H3K9me3 presence on ERVs, we sub-selected ERV loci with H3K9me3, based on the presence of K9me3 peak or enrichment of K9me3 signal over input (more than 1.5 for K9me3 covered). We plotted turnover changes separately for H3K9me3-covered ERV loci and overall ERV loci (Figure 2.18A). Surprisingly, upon SF3A3 knockdown ERVs with H3K9me3 showed opposite response to ERVs - ERV turnovers were stabilized in the presence of H3K9me3, while overall ERVs showed an increase in turnover. To understand ERV nascent and steady transcripts changes upon SF3A3 knockdown, we compared their amounts depending on H3K9me3 presence (Figure 2.18C). Upon SF3A3 knockdown, overall ERVs showed a higher increase of nascent RNA compared to steady, while ERVs with H3K9me3 had a higher increase of stable transcripts compared to nascent transcripts. Taking together, in the presence of H3K9me3, SF3A3 knockdown led to ERV transcripts stabilization, while transcript stabilization was not observed in overall ERVs.

Interestingly, the majority of active ERVs with repressive mark showed H3K9me3 coverage at 3’prime region. This feature is typical for many zinc-finger genes (Blahnik et al., 2011; Valle-García et al., 2016). Therefore, we prepared the same plot illustrating turnover changes for overall genes and KRAB-ZNFs with H3K9me3. We detected strong stabilization of KRAB-ZNFs with H3K9me3 only upon SF3A3 knockdown, while overall genes turnover values did not change (Figure 2.18B). Then we plotted nascent and steady transcripts changes upon SF3A3 knockdown for genes KRAB-ZNFs with H3K9me3. We found that H3K9me3-covered ZNFs had an increase in steady RNA amount compared to overall genes (Figure 2.18D).

Figure 2.18E-F depicts two examples of ERV loci and KRAB-ZNF genes covered with H3K9me3, illustrating transcripts stabilization of these loci upon SF3A3 knockdown, where the amount of stable transcripts substantially increases compared to control.

Taking together, we could hypothesize that the H3K9me3 peak at the 3’prime region promotes increased turnover (transcript degradation) of transcribed entity, mediated by splicing factor SF3A3.

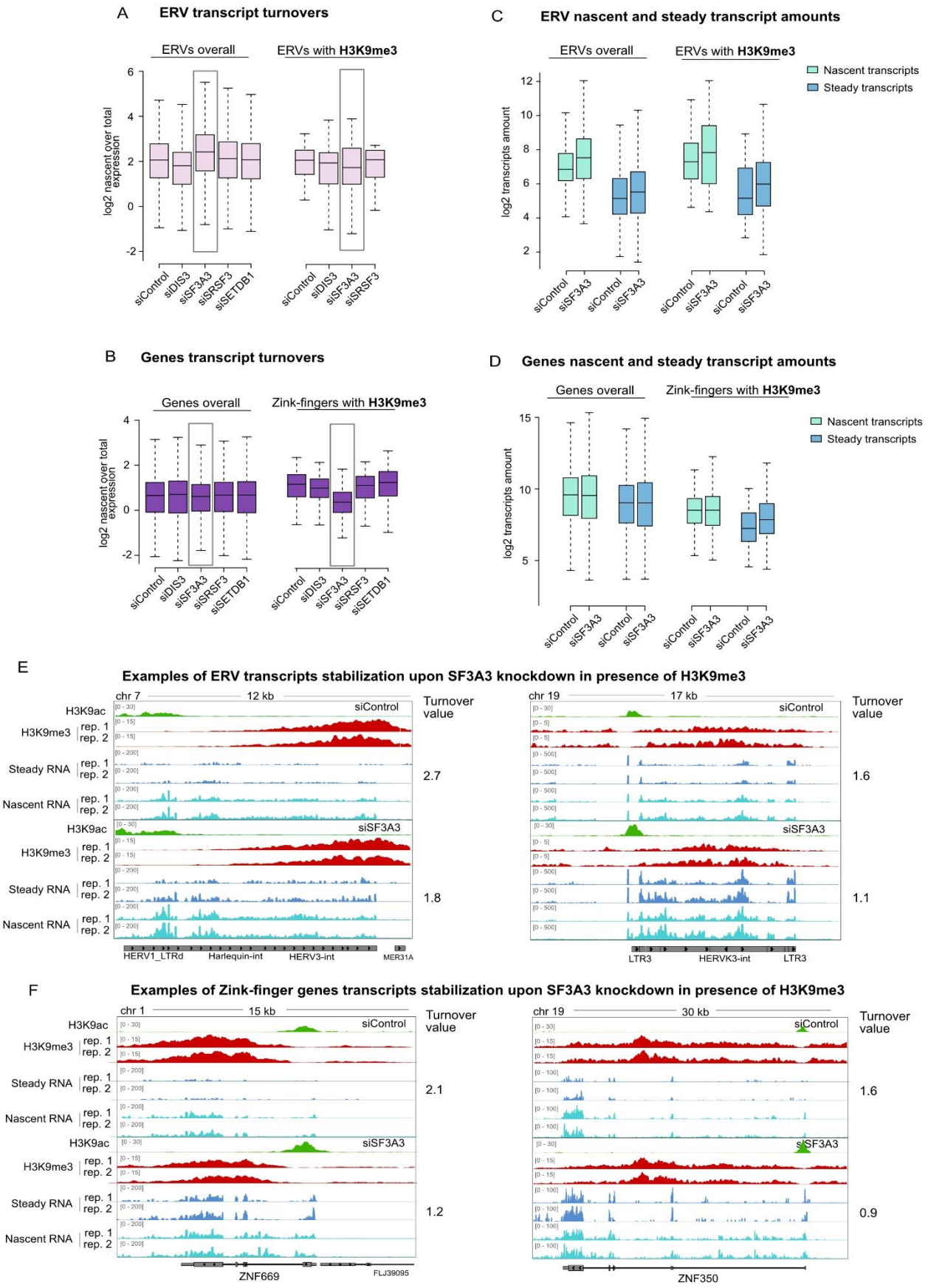


**Figure 2.18. SF3A3 knockdown leads to ERVs and KRAB zinc-finger genes transcript stabilization in the presence of H3K9me3**

A-B. Boxplots showing transcripts turnovers of ERVs and genes in control and knockdowns divided into 2 groups - with H3K9me3 coverage and overall, based on presence of K9me3 peak or enrichment of K9me3 signal over input (more than 1.5 for K9me3 covered) for all conditions plotted. ERV loci and genes were plotted, if they had log2 expression values more than 5 for both nascent and steady fractions for all conditions on average for robust turnover calculations (n=2, mean values are plotted).

C-D. Boxplots showing log2 transcript amounts from ERVs and genes in control and SF3A3 knockdown divided into 2 groups - with H3K9me3 coverage and overall, based on the presence of K9me3 peak or enrichment of K9me3 signal over input (more than 1.5 for K9me3 covered) for all conditions plotted. ERV loci and genes were plotted, if they had log2 expression values more than 5 for both nascent and steady fractions for all conditions on average for robust turnover calculations (n=2, mean values are plotted).

E-F. Genome browser view for 2 examples of stabilizing ERV and KRAB zinc-finger genes transcripts upon SF3A3 knockdown in the presence of H3K9me3. Transcript turnover values (average of 2 replicates) are noted.

# 3. Discussion

## 3.1 TT-seq method identifies fast transcript turnover of human endogenous retroviruses

The transient-transcriptome sequencing method or TT-seq is based on metabolic labeling of newly synthesized nascent RNA, which allows transcript stability calculation by comparing abundances of nascent and stable RNA fractions for each locus of interest. The advantage of this method is that it does not inhibit transcription during labeling time, which leads to minimal effect on cellular processes, compared to other methods based on transcription inhibition or transcription synchronization drugs (Alpert et al., 2017; Wada and Becskei, 2017). In addition, RNA transcription and degradation are linked (Haimovich et al., 2013). Therefore, transcription inhibition could create bias in transcript stability measurements, which is not the case for methods based on metabolic labeling. Therefore, we chose the TT-seq method to investigate the transcriptional stability of ERVs. Nascent and steady RNA fraction separation is a crucial step, which we controlled by synthesizing 4sU-labelled and unlabelled RNA spike-ins and adding them into the samples before separation. Successful fraction separation was confirmed by the negligible unlabelled spike-ins presence in labeled nascent RNA fractions and by steady and nascent RNA amounts for known transcriptionally stable and unstable genes – GAPDH and MYC, respectively.

Transposon transcripts are subject to fast degradation in many eukaryotes, such as *S. pombe, D. melanogaster*, and *C. elegans* (Berrens et al., 2017; Dumesic et al., 2013; Mendell et al., 2004; Reyes-Turcu et al., 2011; Robert et al., 2005; Warf et al., 2012; Zhang et al., 2012). Although in mammals, ERV transcripts turnover was not investigated to date. However, previous reports demonstrated increased transcript turnover of non-coding transcripts in mouse and human cells. Namely, long non-coding RNAs in mouse Neuro-2a and human HeLa cell lines were shown to have unstable transcripts compared to coding genes (Clark et al., 2012; Maekawa et al., 2015). Using our data set, we confirmed the previous finding that lincRNA transcripts were unstable compared to coding genes in HeLa cells. Although in THP-1 cells, lincRNAs, on average, had stability close to coding genes. lincRNAs are known to have cell-type dependent transcriptional activity (Djebali et al., 2012). Therefore, THP-1 and HeLa cell lines can have various sets of active lincRNAs with different transcript stability, which would influence overall stability values for this gene cluster.

Interestingly, ERVs have two times faster transcripts turnover compared to coding genes in both human cell lines – THP-1 and HeLa. On average, genes have similar amounts of nascent and steady RNA, while ERV loci have clearly more nascent transcripts. Therefore, we

confirmed our hypothesis that, in humans, post-transcriptional silencing pathways target and actively degrade ERVs transcripts.

Due to increased ERV turnover, the majority of transcriptionally active ERV loci were detected only by a nascent fraction and not by steady-state RNA. TT-seq method by enrichment and separate processing of nascent fraction allows sensitive detection of rapidly degraded unstable transcripts compared to traditional RNA-seq methods, which measure only steady-state RNA levels. This makes the TT-seq method favorable for studying the transcriptional profile of unstable RNA species such as ERVs.

Taken together, the transient transcriptome sequencing method based on metabolic labeling of nascent RNA allows detecting more unstable transcripts compared to the traditional RNA sequencing technique. Using TT-seq, we detected fast ERV transcript turnover compared to coding genes in both analyzed human cell lines THP-1 and HeLa. In addition, we confirmed fast lincRNA turnover in HeLa cells, although lincRNA transcripts stability seems to be cell-type specific, depending on the subset of active loci. Therefore, we cannot exclude that ERV transcript stability may differ in other cell lines due to cell-type-specific activation of the ERV subset.

### 3.2 Genome-wide ERV loci activity and H3K9me3 repressive mark changes upon 5-Aza and Panobinostat anticancer drug treatments

Anticancer drugs DNA demethylation 5-Aza and inhibitor of histone deacetylases Panobinostat modulate the expression of immune and apoptotic genes, which lead to an increased response to immunotherapy in different cancers (Mazzone et al., 2017b). In line with previous findings, both treatments led to the upregulation of genes enriched in a number of immune system-related pathways. In addition, 5-Aza and HDAC inhibitors are known to activate ERVs and, in the case of 5-Aza, lead to increased production of ERV-derived dsRNA and proteins (Brocks et al., 2017; Chiappinelli et al., 2015; Laska et al., 2013; Roulois et al., 2015). ERV upregulation, in turn, induces a state of viral mimicry in cancer cells and triggers immune activation through sensing of ERV dsRNA or proteins mimicking infection by exogenous viruses (Attermann et al., 2018; Cuellar et al., 2017; Lee et al., 2020b; Vergara Bermejo et al., 2020). However, genome-wide ERV loci activity coupled with chromatin state and ERV transcript stability changes upon treatments with epigenetic anticancer drugs were not assessed.

We confirmed a reduction of respective epigenetic marks upon epigenetic drug treatments with 5-Aza and Panobinostat in two leukemic cell lines - THP-1 and KG1a. Namely, 5-Aza led to notable DNA methylation reduction and Panobinostat – increase in histone acetylation.

Interestingly, DNA demethylation led to a striking decrease of H3K9me3 peaks on ERV loci. At the same time, the overall H3K9me3 signal mildly increased, according to mass spectrometry data. Several studies reported global reorganization of H3K9me3 upon 5-Aza treatment in various cancer cell lines (Jin et al., 2011; Komashko and Farnham, 2010; Lee et al., 2018; Lee et al., 2016). Komashko and Farnham revealed the reduction of H3K9me3 signal in H3K9me3-associated genes upon 5-Aza treatment, while overall H3K9me3 signal increased due to H3K9me3 rearrangements in the genome. Interestingly, some H3K9me3-associated genes, such as KRAB-ZNFs, after 5-Aza treatment switched repressive marks from H3K9me3 to H3K27me3 (Komashko and Farnham, 2010). In addition, global disorganization of nuclear architecture and histone H3 modifications were observed upon knockdown of DNA methyltransferase DNMT1 (Espada et al., 2004a), which also has reduced activity upon 5-Aza treatment (Laska et al., 2013). Therefore, interconnection of DNA and histone methylation repressing pathways, compromised upon DNA methylation removal by 5-Aza, could result in observed H3K9me3 reorganization and H3K9me3 reduction on ERVs in our data set.

According to our data, 5-Aza and Panobinostat treatments resulted in increased expression in a number of ERV families coupled with decreased H3K9me3 occupancy with an especially strong effect in 5-Aza treated cells. HERV-Fc families were upregulated in both treatments. A previous study reported strong upregulation of the HERV-Fc1 family on a transcriptomic level upon 5-Aza and HDAC inhibitors treatments (Laska et al., 2013). Moreover, our data showed that on a loci level, only one copy of the HERV-Fc1 family was strongly derepressed upon 5-Aza with expression values comparable to highly expressed coding genes. This HERV-Fc1 locus has the size and structure of a full-length ERV element with viral genes flanked by LTR promoter sequences. Therefore, this highly expressed coding potent locus 5-Aza treatment could potentially produce proviral proteins. At the same time, none of the ERV loci belonging to the HERV-Fc family had notable upregulation in Panobinostat treated cells. Such discrepancy could be explained by mild upregulation of many HERV-Fc loci contributing to a notable increase in family expression. Indeed, Laska and colleagues found HERV-Fc1 gag protein production only in 5-Aza treated cells but not in cells treated with HDAC inhibitors. Interestingly, a few additional highly 5-Aza upregulated copies have the structure of full-length HERVH and HERV3 elements. Thus, viral proteins potentially can be produced from these copies as well. Therefore, few single ERV loci could contribute to the increased production of proviral proteins in 5-Aza treated cells. As a future perspective, it would be interesting to investigate if identified highly upregulated upon 5-Aza treatment HERV-Fc1 locus is alone responsible for increased proviral protein production and is able to cause immunomodulatory effect in leukemic cells.

ERV-derived dsRNA is another contributor to the viral mimicry state in addition to ERV proteins. LTR promoters are bidirectional. Therefore, ERVs often show bidirectional transcription, which may result in dsRNA production (Sadeq et al., 2021). According to our data set, approximately a quarter of upregulated ERVs upon 5-Aza transcribed bi-directionally, which could potentially give rise to dsRNA. While in Panobinostat treated cells, bidirectional loci represented less than one-tenth of upregulated ERVs. These data are consistent with previous reports showing increased dsRNA production derived from endogenous retroviruses upon 5-Aza treatments.

In addition, we did not detect overall ERV transcripts stabilization upon treatments. Therefore, we can conclude that an increased amount of ERV transcripts upon 5-Aza and Panobinostat anti-cancer treatments is rather a consequence of increased ERV transcription but not stabilization of ERV-derived transcripts. Although considering a small effect and only one analyzed TT-seq replicate of drug-treated cells, this finding needs to be repeated and confirmed in future experiments.

Taken together, we found that 5-Aza treatment resulted in prominent H3K9me3 reduction on ERV loci, which could be explained by the interconnection of DNA methylation and H3K9me3 repressing pathways. In addition, we detected a few highly upregulated ERV loci with coding potential, which could potentially induce viral mimicry and contribute to the immunomodulatory activity of the 5-Aza epigenetic drug. However, ERV derepression, observed in our 5-Aza and Panobinostat treatments, did not result in a strong interferon response in the THP-1 cell line contrary to publications, where these epigenetic drug treatments in other human cell lines led to an immune response activation associated with ERV derepression. Therefore, strong activation of an immune response upon these epigenetic drugs might be cell type-specific, and a challenge will be to predict if cells are responsive to 5-Aza or Panobinostat to elicit IFN signaling.

### 3.3 Exosome exonuclease DIS3 and splicing factors SF3A3 and SRSF3 are novel transcriptional ERV repressors in humans

SF3A3 and SRSF3 are involved not only in splicing regulation but also in various mRNA processing pathways. Moreover, they are promising therapeutic targets in anticancer treatment, increasing the synthesis of immune mediators and leading to cancer cell death (Mure et al., 2018; Shen et al., 2019b; Siebring-van Olst et al., 2017). Particularly SRSF3 silencing was found to be a potent inhibitor of many tumors (Guo et al., 2020; Jimenez-Vacas et al., 2020; Zhou et al., 2020). In line with that, we found that SRSF3 knockdown led to the upregulation of genes enriched in immune-related pathways, namely interleukins and apoptotic genes. SF3A3 knockdown, contrarily, was associated with the upregulation of genes

involved in various RNA processing pathways, such as mRNA splicing, processing of capped intron-containing mRNA, mRNA metabolism, and rRNA processing. Upregulated splicing-related genes were DEAH-box helicases, Sm-like proteins, small nuclear ribonucleoproteins, pre-mRNA processing factors, SF3B, and SRSF splicing factors. Among other upregulated mRNA processing genes were cleavage and polyadenylation factors and tRNA methyltransferases. Previous studies investigating SF3A genes knockdown detected notable accumulation of unspliced pre-mRNAs (Tanackovic and Krämer, 2005). Therefore, upregulation of genes involved in mRNA splicing and processing could be a compensatory mechanism in response to SF3A3 deletion – an important component of splicing and RNA processing machinery. In addition, we detected a notable amount of differentially spliced genes upon both SF3A3 and SRSF3 knockdowns. Taken together, both splicing factors SF3A3 and SRSF3 are involved in alternative gene splicing. Although, their knockdown leads to transcriptional deregulation of different gene subsets, confirming additional functions of SF3A3 and SRSF3, apart from canonical splicing, and their involvement in diverse biological pathways.

Our data showed a striking increase in ERV transcription upon deletion of both splicing factors SF3A3 and SRSF3. An increase in transcription was partially coupled with a decrease in H3K9m3 occupancy on upregulated ERVs with a more prominent effect of H3K9me3 reduction in SRSF3 knockdown cells. Knockdown of nuclear exosome exonuclease DIS3 also led to pronounced ERV upregulation. Although in contrast to splicing factors, DIS3 knockdown did not lead to H3K9me3 reduction. Moreover, an increase in ERV transcription in response to DIS3 knockdown happened undependable of H3K9me3 presence. Namely, ERVs, covered with H3K9me3 in control and upregulated upon DIS3 knockdown, did not show any notable H3K9me3 reduction. Contrary to DIS3 deficient cells, most upregulated ERVs upon SETDB1 knockdown had reduced H3K9me3, pointing out distinct ERV repressive pathways in which SETDB1 and DIS3 are involved.

Notably, despite broad SETDB1-induced H3K9me3 reduction on ERVs, it did not lead to prominent ERV upregulation. The relatively small REV derepression effect of SETDB1 in HeLa cells, which we observed, does not seem like an effect of inefficient knockdown. Whereas publicly available knockout in HeLa cell from Timms study (Timms et al., 2016), where SETDB1 protein was completely gone, showed a similar level of deregulation, compared to our data (Figure 2.15), pointing out that SETDB1 reduction leads to mild gene and ERV deregulation in HeLa cells. Indeed, prevalent ERV repressive pathways were shown to have cell-dependent variability (Geis and Goff, 2020; Karimi et al., 2011; Kato et al., 2018; Sharif et al., 2016). Therefore, in HeLa cells, H3K9me3 might not be the main ERV repressing pathway.

In summary, although all knockdowns overall led to H3K9me3 reduction, reduced H3K9me3 did not necessarily lead to increased ERV expression. We showed that SF3A3 and SRSF3 splicing factors knockdowns led to striking ERV derepression in the human cell line, which was coupled only partially with H3K9me3 histone repressive mark reduction. However, prominent ERVs derepression, undependable of H3K9me3, was observed upon nuclear exosome exonuclease DIS3 knockdown. This finding suggests that transcriptional repression of ERVs by splicing factors SF3A3 and SRSF3 and nuclear exosome DIS3 depends on another repressing pathway, different from H3K9me3, such as H4K20me3 or H4K27me3. In fission yeast, splicing factors participate in heterochromatin formation on repetitive loci by binding to repetitive elements RNA and mediating recruitment of the H3K9 methyltransferase complex and chromatin remodelers (Bayne et al., 2008; Chinen et al., 2010; Weigt et al., 2021). Although, in mammals, the role of splicing factors in heterochromatin formation so far was not assessed. Future studies are needed to investigate the mechanism of transcriptional ERV repression by splicing factors SF3A3, SRSF3, and exonuclease DIS3 in human cells.

The nuclear exosome, in mammals, targets various RNA species through mediator complexes NEXT and PAXT. NEXT complex targets nuclear exosome to short and unprocessed transcripts, as well as to retrotransposon RNAs, such as LINEs and ERVs (Wu et al., 2019). At the same time, NEXT directly interacts with the SRSF3 and SF3B family of splicing factors. (Falk et al., 2016; Mure et al., 2018). The functional role of SF3B splicing factors in transcripts targeting and exosome degradation was not addressed. However, the SRSF3 splicing factor was shown not only to stabilize the NEXT complex but also target viral mRNAs, expressed from the intronless genes, for subsequent degradation by nuclear exosome through NEXT complex mediation (Mure et al., 2018). In addition, SRSF3 depletion increases the amount of SINE transcripts in mouse oocytes (Do et al., 2018). In our data, we found that the majority, approximately two-thirds, of upregulated ERVs from DIS3 depleted cells were also derepressed upon SRSF3 knockdown, while only one-third of upregulated ERVs upon DIS3 knockdown was derepressed in SF3A3 depleted cells. Therefore, the major overlap between upregulated ERVs upon DIS3 and SRSF3 knockdowns suggests that splicing factor SRSF3 might target nuclear exosome to ERV transcripts probably through interaction with NEXT complex. Although, it is surprising that both DIS3 and SRSF3 are strong ERV repressors, which silence ERVs on a transcriptional level and act not only as post-transcriptional regulators. Moreover, the ERV transcript stabilization effect was observed only upon DIS3 but not SRSF3 knockdown. This could be explained by other splicing factors substituting the reduced SRSF3 activity.

Interestingly, the vast majority of upregulated ERVs had already basal transcription in control in all knockdowns. Basal transcription from repetitive and heterochromatic regions was

reported in several model organisms, such as *S. pombe, D. melanogaster, C. elegans,* Although, such transcriptional events are much less explored in mammals. In general, low transcription level from repetitive and heterochromatic regions helps to recruit repressive complexes to such loci preventing their high expression (Bühler and Moazed, 2007; Saksouk et al., 2015). Therefore, in mammals, as in other model organisms, allowing basal ERV transcription might be an adaptation for the cells, which helps to keep transposable elements under control.

Increased dsRNA production from retrotransposon loci, including ERVs, have been already documented in human cancer cell line upon SETDB1 deletion. Increased dsRNA production coincided with type I interferon response and cancer cell apoptosis (Cuellar et al., 2017). In our data set we found an increased proportion of bidirectionally transcribed ERVs upon SETDB1 knockdown, which could give rise to dsRNA. Interestingly, we detected a higher proportion of bidirectionally transcribed ERVs upon the rest of knockdowns with a particularly high elevation in SRSF3 knockdown, where almost half of the expressed ERVs had bidirectional transcription. Moreover, Chemokine ligand 20 (CCL20), known to be induced by dsRNA (Bellak et al., 2005; Ghosh et al., 2009), is upregulated not only in SETDB1 but also in DIS3, SF3A3, and SRSF3 knockdowns. Remarkably, a high amount of bidirectionally transcribed ERVs coincides with immune pathway genes activation in SRSF3 knockdown. However, the classical dsRNA sensors, such as RIG-I, MDA5, or TLR3, and associated proinflammatory cytokines, like interferons, were not upregulated in the knockdowns. This might be due to the necessity in prolonged ERV upregulation and increased dsRNAs to activate dsRNA sensors and subsequent proinflammatory cytokines. Therefore, the contribution of ERV-derived dsRNA to immunomodulatory effect and subsequent cancer cells death upon SRSF3 downregulation needs further investigation. In addition to an increased proportion of bidirectionally expressed ERVs, we identified in DIS3, SF3A3, and SRSF3 knockdowns several highly expressed full-length ERV loci from HERV-H and HERV-K families, which could produce proviral proteins. Such proteins could contribute to the viral mimicry effect and, additionally to dsRNA, induce an immune response in cancer cells. In future experiments, increased production of ERV dsRNA and proteins upon knockdowns of splicing factors SF3A3 and SRSF3 and exonuclease DIS3 could be checked to assess the viral mimicry effect caused by ERV upregulation in cancer cells.

Taken together, we showed that SF3A3 and SRSF3 splicing factors and nuclear exonuclease DIS3 are potent ERV repressors. Their knockdown leads to striking ERV derepression, including increased proportions of bidirectionally transcribed ERVs and high expression of full-length ERV loci, which potentially could contribute to dsRNA production and viral mimicry inducing an immune response in cancer cells. Interestingly, we detected basal transcription in

control cells for the vast majority of upregulated ERVs, suggesting that low transcription level from ERVs helps to recruit repressive complexes to such loci preventing their high expression. In addition, DIS3 and SRSF3 knockdowns show major overlap in derepressed ERVs, implying ERV transcripts recognition by splicing factor SRSF3 and targeting to nuclear exosome for their subsequent degradation, possibly through the NEXT complex.

### 3.4 DIS3 and SF3A3 knockdowns lead to ERV transcripts stabilization but not SRSF3

As we hypothesized, we detected prominent ERV transcripts stabilization upon nuclear exonuclease DIS3 depletion, while general gene transcripts stability was not affected, suggesting preferential DIS3 targeting and degradation of ERV transcripts over coding genes. Previous studies showed that nuclear exosome regulates a variety of aberrant transcripts, such as inefficiently spliced and polyadenylated RNA, cryptic and pervasive bidirectional transcripts (Pefanis et al., 2015; Schlackow et al., 2017; Schneider et al., 2012). Therefore, in our study we illustrated that ERV transcripts are subjects to DIS3-mediated degradation in the human cell line.

In contrast, depletion of splicing factor SRSF3 did not lead to an overall change in ERV transcripts turnover, despite major overlap with the DIS3-mediated pathway. This finding indicates that splicing factor SRSF3 may target ERV transcripts for nuclear exosome degradation. Although other splicing factors, as SF3B splicing family members, could have a similar function and recruit nuclear exosome for ERV transcripts degradation when SRSF3 is depleted, preventing ERV transcripts stabilization upon SRSF3 knockdown. To confirm SRSF3-mediated ERV transcripts decay by the nuclear exosome, occupancy of SRSF3 splicing factor on ERV transcripts, stabilized by DIS3, could be assessed.

SF3A3 depletion leads to an increase in transcription and a decrease in transcripts stability overall for H3K9me3-free ERVs. Pointing out that SF3A3 serves as ERV transcriptional repressor through direct or indirect mechanisms on loci without H3K9me3. In contrast, H3K9me3-covered ERV loci show not only an increase in transcription but also transcripts stabilization upon SF3A3 depletion. The H3K9me3 dependent ERV transcript stabilization was not detected in other knockdowns, suggesting that splicing factor SF3A3 mediates H3K9me3 covered ERV transcripts decay in addition to transcriptional repressing function. This could occur due to either (1) direct recognition of H3K9me3 or H3K9me3-associated proteins by SF3A3 or (2) SF3A3 recognition of specific RNA features and modifications associated with heterochromatin-originated transcripts and subsequent recruitment of RNA degradation machinery. This novel mechanism of SF3A3 recognition and H3K9me3-associated transcripts targeting for subsequent degradation needs to be investigated in future experiments.

Interestingly, stabilized by SF3A3 ERVs have H3K9me3 coverage at the 3'prime region and active acetylation mark in the 5' region. This rare chromatin marks combination is also present in active KRAB zinc-finger genes (Blahnik et al., 2011; Valle-García et al., 2016). Surprisingly, the effect of SF3A3 depletion on transcripts stabilization of H3K9me3 covered loci was even stronger in KRAB-ZNFs than in ERVs. Similar to H3K9me3 covered ERVs, only SF3A3 knockdown, but not DIS3 or SRSF3, showed striking transcripts stabilization of KRAB-ZNFs covered with H3K9me3. Therefore, another degradation machinery, in addition to nuclear exosome exonuclease DIS3, might be responsible for SF3A3-mediated transcripts decay, such as, for example, another nuclear exosome exonuclease RRP6 or ribonucleases coupled with nonsense-mediated RNA decay pathway. Therefore, splicing factor SF3A3 might target transcripts originated from H3K9me3 covered loci, such as ERVs and KRAB-ZNFs, and lead them to decay.

### 3.5 Conclusions and prospects

We are the first who documented that ERV transcripts have fast turnover in humans, suggesting that these loci are actively degraded by post-transcriptional silencing pathways. To detect it, we used the transient transcriptome sequencing method based on metabolic labeling of nascent RNA, which allows not only a calculation of transcripts stability but also more sensitive detection of unstable transcripts compared to the traditional RNA sequencing technique.

Secondly, we found that 5-Aza treatment of leukemic cells results in striking upregulation of a few specific ERV loci with coding potential, which potentially could induce viral mimicry and contribute to the immunomodulatory activity of the 5-Aza epigenetic drug. In addition, treatment with DNA methylation inhibitor 5-Aza results in prominent H3K9me3 reduction on ERV loci, which could be explained by the interconnection of DNA methylation and H3K9me3 repressing pathways. Therefore, H3K9me3 reduction could additionally contribute to ERV derepression upon 5-Aza treatment. Conversely, a treatment with histone deacetylase inhibitor Panobinostat results in milder ERV derepression. Moreover, we found that 5-Aza and Panobinostat treatments led to transcriptional ERV derepression rather than increased ERV transcripts stability.

Thirdly, we showed that SF3A3 and SRSF3 splicing factors and nuclear exonuclease DIS3 are potent transcriptional ERV repressors. Their knockdown leads to striking ERV upregulation, including increased proportions of bidirectionally transcribed ERVs and high expression of full-length ERV loci, which could contribute to dsRNA production and viral mimicry inducing an immune response in cancer cells. Interestingly, it was already shown that induced dsRNA production from ERVs upon SETDB1 deletion increased dsRNA production

coincided with type I interferon response and leukemic cell death (Cuellar et al., 2017). In addition, SRSF3 was found to be an oncogene and its silencing is associated with the successful elimination of various cancers (Guo et al., 2020; Jimenez-Vacas et al., 2020; Zhou et al., 2020). Therefore, possible production of ERV-derived dsRNA upon reduction of splicing factors SRSF3 might contribute to previously observed immunomodulatory effect and associated cancer cells death.

Fourthly, we showed that ERV transcripts overall are subject to active degradation by nuclear exosome exonuclease DIS3 in human cell line. Splicing factor SRSF3, an interactor of exosome targeting complex NEXT, deregulates highly overlapping with DIS3 subset of ERVs. Nevertheless, ERV transcript stability is not affected by SRSF3 knockdown, possibly due to the substitution of lost SRSF3 with other splicing factors.

Interestingly, knockdowns of both splicing factors SF3A3 and SRSF3 and exosome exonuclease DIS3 result in transcriptional derepression of ERV loci, suggesting that their function might be coupled with transcriptional repressing machinery. We detected H3K9me3 reduction on ERV loci upon SRSF3 knockdown, although only a quarter of upregulated ERVs with H3K9me3 in control showed H3K9me3 reduction upon SRSF3 and SF3A3 knockdowns. In addition, ERVs upregulation upon DIS3 knockdown did not show any connection with H3K9me3 reduction. These findings suggest that other transcriptional repressing pathways might be linked to studied splicing factors and exonuclease DIS3, such as, for example, H4K20me3 and H3K27me3. Notably, the vast majority of upregulated ERVs already have basal transcription in control. This suggests that basal transcription might be beneficial for the cells since it helps recruit repressive complexes to ERVs, preventing their high expression.

Finally, we found that knockdown of splicing factor SF3A3 leads to decreased stability of ERVs. Moreover, it results in particularly strong transcript stabilization of H3K9me3 covered loci, such as ERVs and KRAB-ZNF genes. However, DIS3 knockdown does not affect the stability of these loci. This finding implies that splicing factor SF3A3 might specifically target transcripts originated from H3K9me3 covered loci for subsequent degradation independently from exosome exonuclease DIS3 degradation pathway. The schematic model of possible ERV transcriptional and post-transcriptional silencing by the splicing factor SF3A3 is shown in Figure 3.

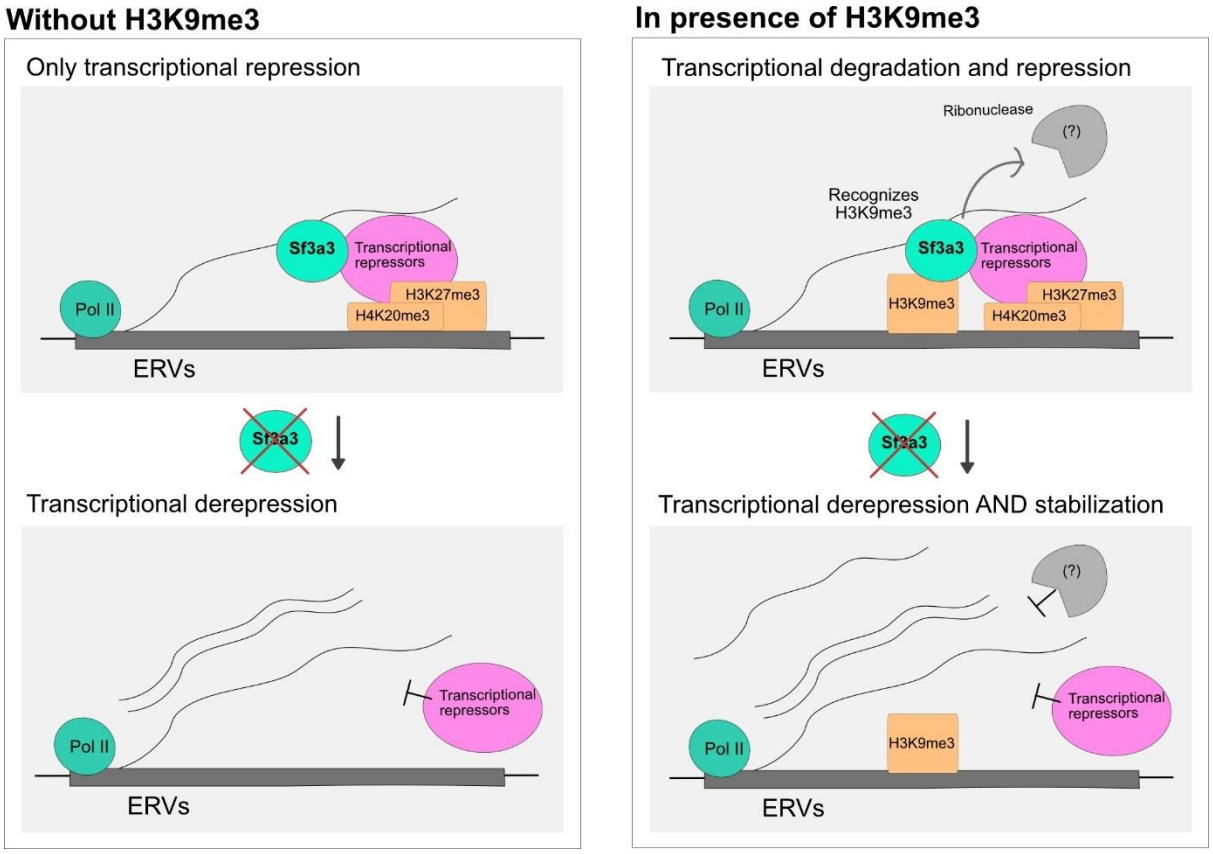


**Figure 3. Model for the transcriptional ERV repression and post-transcriptional ERV degradation mediated by the splicing factor SF3A3**

The schema depicts a possible model for the Sf3a3-associated ERV silencing and ERV transcripts degradation. SF3A3 splicing factor recognizes basally transcribed ERV-derived RNA and recruits transcriptional repressing machinery associated with such repressive marks as H4K20me3 or H3K27me3. Upon Sf3a3 deletion, ERVs are transcriptionally derepressed. In the presence of H3K9me3, SF3A3 not only represses ERV-derived RNA but also targets these transcripts for degradation independently from the DIS3 nuclear exonuclease pathway. Upon Sf3a3 deletion, ERVs transcripts derived from H3k9me3-covered loci are not only transcriptionally derepressed but also stabilized.

In this work, we shed light on anti-cancer drug-associated ERV derepression in terms of ERV loci-specific transcript stability, expression, and epigenetic mark changes. Therefore, our work will be beneficial for the scientific community, investigating the immunomodulatory role of ERV derepression in cancer therapies using 5-Aza and Panobinostat epigenetic drugs.

The main finding of our work is the identification of novel transcriptional ERV repressors – splicing factors SF3A3, SRSF3, and nuclear exosome DIS3. This finding highlights a novel mechanism in mammalian ERV repression, namely transcriptional repression by RNA processing factors, which were thought to act only post-transcriptionally. In addition, our findings suggest that classical ERV repressing epigenetic mark H3K9me3 might not be a necessary prerequisite for transcriptional ERV repression. Thus, future studies are needed to investigate mechanisms of transcriptional ERV repression by splicing factors and nuclear exosome and associated repressive epigenetic marks. Furthermore, identified by us a remarkable association between SF3A3 splicing factor and H3K9me3-associated transcript

degradation suggests the existence of crosstalk between transcriptional repressing machinery and post-transcriptional RNA processing in a mammalian model.

# 4. Materials & Methods

## 4.1 Materials

### 4.1.1 Cell lines

| Line | Description | Derived from | Organism |
|---|---|---|---|
| THP-1 | Monocytic cell line | Acute monocytic leukemia | Homo Sapiens |
| KG1a | Macrophage cell line | Acute myelogenous leukemia | Homo Sapiens |
| HeLa | Epithelial cell line | Cervix epithelioid carcinoma | Homo Sapiens |

### 4.1.2 Tissue culture reagents

| Name | Company |
|---|---|
| RPMI-1640 Medium | Sigma-Aldrich |
| Iscove's Modified Dulbecco's Medium | Sigma-Aldrich |
| Dulbecco's Modified Eagle Medium – high glucose | Sigma-Aldrich |
| Fetal Bovine Serum | Sigma-Aldrich |
| MEM Non-essential Amino Acid solution (100×) | Sigma-Aldrich |
| Penicillin-Streptomycin | Sigma-Aldrich |
| Trypsin (EDTA), 0.05% | Thermo Fisher Scientific |
| Decitabine (5-aza-2′-deoxycytidine) | Sigma-Aldrich |
| Panobinostat (LBH589) | Selleck Chemicals |
| Lipofectamine RNAiMAX transfection reagent | Sigma-Aldrich |
| Opti-MEM Reduced-Serum Medium | Sigma-Aldrich |
| DMSO | Sigma-Aldrich |
| 4-thiourudine | Sigma-Aldrich |

### 4.1.3 Pre-designed Silencer Select siRNAs (Sigma)

| siRNA ID | Target | Sense sequence | Antisense sequence |
|---|---|---|---|
| S229658 | DIS3 | GGAGCAUUACUGAAAAGGAtt | UCCUUUUCAGUAAUGCUCCag |
| S229660 | DIS3 | GAAGAGACUGAAACAGAAtt | UUCUGUUUCAGUCUCUUUCtc |
| S21534 | SF3A3 | CAACUACAACUGUGAGAUUtt | AAUCUCACAGUUGUAGUUGat |
| S21535 | SF3A3 | CGACAUCUCACUCAUGAAAtt | UUUCAUGAGUGAGAUGACGct |
| S12731 | SRSF3 | CGCUGUCUCGGGAGAGAAAtt | UUUCUCUCCCGAGACAGCGat |
| S12733 | SRSF3 | GCAACAAGACGGAAUUGGAtt | UCCAAUUCCGUCUUGUUGCaa |
| S19112 | SETDB1 | GCCGUGACUUCAUAGAGGAtt | UCCUCUAUGAAGUCACGGCag |
| S19110 | SETDB1 | GGACAAUGCAGGAGAUAGAtt | UCUAUCUCCUGCAUUGUCCga |
| 4390843 | Negative Control No. 1 | | |

### 4.1.4 Kits

| Name | Company |
|---|---|
| RNA Clean & Concentrator-25 | Zymo research |
| SuperScript III | Sigma |
| µMACS streptavidin kit | Miltenyi |
| Ovation Universal RNA-Seq System 1-16 | Nugen |
| Ultra II DNA Library prep kit for Illumina | NEB |
| ThruPLEX DNA-seq kit | Takara |
| DNEasy Blood and Tissue Kit | Qiagen |
| EpiTect Bisulfite Kit | Qiagen |
| Methylated NEBNext Multiplex Oligos for Illumina | NEB |
| QIAquick PCR purification kit | Qiagen |

### 4.1.5 Antibodies

#### 4.1.5.1 Primary antibodies

| Antigen (species) | ID | Company | Purpose | Dilution |
|---|---|---|---|---|
| Srsf3 (rabbit) | 306 | Abcam | Western blot | 1:1000 |
| Dis3 (rabbit) | 307 | Abcam | Western blot | 1:1000 |
| Sf3a3 (rabbit) | 308 | Abcam | Western blot | 1:1000 |
| Setdb1 (rabbit) | 300 | ThermoFischer Scientific | Western blot | 1:250 |
| KAP1 (rabbit) | 296 | Bethyl (Biomol) | Western blot | 1:1000 |
| Fibrillarin (mouse) | 304 | Santa Cruz | Western blot | 1:500 |
| Lamin B1 (mouse) | 311 | Santa Cruz | Western blot | 1:1000 |
| H3K9ac (rabbit) | 275 | Active Motif | Western blot | 1:1000 |
| H3K9ac (rabbit) | 275 | Active Motif | ChIP | 3 ug |
| H3K9me3 (rabbit) | 249 | Active Motif | ChIP | 3 ug |

#### 4.1.5.2 Secondary antibodies

| Antigen (species) | ID | Company | Purpose | Dilution |
|---|---|---|---|---|
| IRDye 680RD Goat anti-Mouse IgG | - | Li-COR | Western blot | 1:3000 |
| IRDye® 800CW Goat anti-Rabbit IgG | - | Li-COR | Western blot | 1:3000 |

### 4.1.6 Amplified ERCC spike-ins

| ERCC ID | GeneBank ID | Name | Length (nt) | %GC | 4-Thio-UTPs % | UTPs % |
|---|---|---|---|---|---|---|
| 136 | KC702222.1 | Spike-in 3 | 1,014 | 43 | 10 | 90 |
| 145 | KC702227.1 | Spike-in 4 | 1,015 | 46 | 0 | 100 |
| 92 | KC702203.1 | Spike-in 5 | 1,079 | 52 | 10 | 90 |
| 2 | KC702164.1 | Spike-in 6 | 1,037 | 53 | 0 | 100 |

### 4.1.7 Primers

#### 4.1.7.1 Spike-ins amplification primers

| Name | Sequence 5´-3´ |
|---|---|
| spike-in_#3_amplification_from_cDNA_f | GGTTTCGACGTTTTGAAGGAGGG |
| spike-in_#3_amplification_from_cDNA_r | GTACCCGGGAAAATCCTAGTTC |
| spike-in_#4_amplification_from_cDNA_f | GGACTGTCCTTTCATCCATAAGCGG |
| spike-in_#4_amplification_from_cDNA_r | CGCACGCCGAATGATGAAACG |
| spike-in_#5_amplification_from_cDNA_f | GGGATGTCCTTGGACGGGGT |
| spike-in_#5_amplification_from_cDNA_r | GCTTTCGGAGCAAATCGCG |
| spike-in_#6_amplification_from_cDNA_f | GGCCAGATTACTTCCATTTCCGCC |
| spike-in_#6_amplification_from_cDNA_r | GGGTAAAACGCAAGCACCG |

#### 4.1.7.2 RT-qPCR Primers

| Name | Sequence 5´-3´ |
|---|---|
| GAPDH-F | CGCTCTCTGCTCCTCCTGTT |
| GAPDH-R | CCATGGTGTCTGAGCGATGT |
| MYC-F | CAGCTGCTTAGACGCTGGATT |
| MYC-R | GTAGAAATACGGCTGCACCGA |
| DIS3-F | ACCCTCACTTAAAATAGAAGATACAGT |
| DIS3-R | CCATTAAGGTCCATGTTTGAAGT |
| SF3A3-F | GGCACAAAACTTGGTGGAGT |
| SF3A3-R | CCAGCTTCTCAGATGCCTTC |
| SRSF3-F | ACTTCATAAGCTTGGTGCAT |
| SRSF3-R | ACACCTTTGTGTCACTGTTT |
| SETDB1-F | AAGAGTGTCTGCCCACAGGG |
| SETDB1-R | CATGTTTGGGTCACAATTGCA |

#### 4.1.7.3 qPCR Primers for ChIP

| Name | Sequence 5´-3´ |
|---|---|
| GAPDH-F | GCACGTAGCTCAGGCCTCAAGAC |

| | |
|---|---|
| GAPDH-R | GACTGTCGAACAGGAGGAGCAGAG |
| ZNF180-F | TGATGCACAATAAGTCGAGCA |
| ZNF180-R | TGCAGTCAATGTGGGAAGTC |

### 4.1.8 High-throughput sequencing libraries

| Sample name | ID | Experiment | Sequencing | Index |
|---|---|---|---|---|
| THP-1_control_LF | GS843 | RNAseq | 50 bp SE | AAGCCT |
| THP-1_5aza_LF | GS844 | RNAseq | 50 bp SE | GTCGTA |
| THP-1_panob_LF | GS845 | RNAseq | 50 bp SE | GGAGAA |
| THP-1_control_UF | GS846 | RNAseq | 50 bp SE | CGTAGA |
| THP-1_5aza_UF | GS847 | RNAseq | 50 bp SE | TCAGAG |
| THP-1_panob_UF | GS877 | RNAseq | 50 bp SE | CACAGT |
| THP-1_5aza_H3K9me3 | GS768 | ChIP-seq | 50 bp SE | AGTCAA |
| THP-1_panob_H3K9me3 | GS769 | ChIP-seq | 50 bp SE | AGTTCC |
| THP-1_ctrl_H3K9me3 | GS770 | ChIP-seq | 50 bp SE | ATGTCA |
| THP-1_5aza_H3K9ac | GS771 | ChIP-seq | 50 bp SE | CCGTCC |
| THP-1_panob_H3K9ac | GS772 | ChIP-seq | 50 bp SE | GTCCGC |
| THP-1_ctrl_H3K9ac | GS752 | ChIP-seq | 50 bp SE | GTGAAA |
| KG1A_5aza_H3K9me3 | GS774 | ChIP-seq | 50 bp SE | GTGGCC |
| KG1A_panob_H3K9me3 | GS775 | ChIP-seq | 50 bp SE | GTTTCG |
| KG1A_ctrl_H3K9me3 | GS753 | ChIP-seq | 50 bp SE | CGTACG |
| KG1A_5aza_H3K9ac | GS777 | ChIP-seq | 50 bp SE | GAGTGG |
| KG1A_panob_H3K9ac | GS778 | ChIP-seq | 50 bp SE | ACTGAT |
| KG1A_ctrl_H3K9ac | GS779 | ChIP-seq | 50 bp SE | ATTCCT |
| THP-1_5aza | GS762 | RRBS | 50 bp SE | GGTAGC |
| THP-1_panob | GS763 | RRBS | 50 bp SE | ATGAGC |
| THP-1_ctrl | GS764 | RRBS | 50 bp SE | CAAAAG |
| KG1A _5aza | GS765 | RRBS | 50 bp SE | CAACTA |
| KG1A _panob | GS766 | RRBS | 50 bp SE | CACCGG |
| KG1A _ctrl | GS767 | RRBS | 50 bp SE | CACGAT |
| HeLa_LF_siCtrl.r1 | GS1121 | RNAseq | 50 bp PE | AACCAG |
| HeLa_LF_siDIS3.r1 | GS1122 | RNAseq | 50 bp PE | TGGTGA |
| HeLa_LF_siSF3A3.r1 | GS1123 | RNAseq | 50 bp PE | AGTGAG |
| HeLa_LF_siSRSF3.r1 | GS1124 | RNAseq | 50 bp PE | GCACTA |
| HeLa_LF_siSETDB1.r1 | GS1125 | RNAseq | 50 bp PE | ACCTCA |
| HeLa_TF_siCtrl.r1 | GS1126 | RNAseq | 50 bp PE | GTGCTT |
| HeLa_TF_siDIS3.r1 | GS1127 | RNAseq | 50 bp PE | AAGCCT |
| HeLa_TF_siSF3A3.r1 | GS1128 | RNAseq | 50 bp PE | GTCGTA |
| HeLa_TF_siSRSF3r1 | GS1129 | RNAseq | 50 bp PE | AAGAGG |
| HeLa_TF_siSETDB1.r1 | GS1130 | RNAseq | 50 bp PE | GGAGAA |

| | | | | |
|---|---|---|---|---|
| HeLa_siCtrl H3K9me3. r.1 | GS1146 | ChIP-seq | 50 bp PE | AGTCAA |
| HeLa_siDIS3 H3K9me3. r.1 | GS1147 | ChIP-seq | 50 bp PE | AGTTCC |
| HeLa_siSF3A3 H3K9me3. r.1 | GS1148 | ChIP-seq | 50 bp PE | ATGTCA |
| HeLa_siSRSF3 H3K9me3. r.1 | GS1149 | ChIP-seq | 50 bp PE | CCGTCC |
| HeLa_siSETDB1 H3K9me3. r.1 | GS1150 | ChIP-seq | 50 bp PE | GTCCGC |
| HeLa_siCtrl H3K9ac. r.1 | GS1151 | ChIP-seq | 50 bp PE | GTGAAA |
| HeLa_siDIS3 H3K9ac r.1 | GS1152 | ChIP-seq | 50 bp PE | GTGGCC |
| HeLa_siSF3A3 H3K9ac. r.1 | GS1153 | ChIP-seq | 50 bp PE | GTTTCG |
| HeLa_siSRSF3 H3K9ac. r.1 | GS1154 | ChIP-seq | 50 bp PE | CGTACG |
| HeLa_siSETDB1 H3K9ac. r.1 | GS1155 | ChIP-seq | 50 bp PE | GAGTGG |
| HeLa_LF_siCtrl.r2 | GS1156 | RNAseq | 50 bp PE | AACCAG |
| HeLa_LF_siDIS3.r2 | GS1157 | RNAseq | 50 bp PE | TGGTGA |
| HeLa_LF_siSF3A3.r2 | GS1158 | RNAseq | 50 bp PE | AGTGAG |
| HeLa_LF_siSRSF3.r2 | GS1159 | RNAseq | 50 bp PE | GCACTA |
| HeLa_LF_siSETDB1.r2 | GS1160 | RNAseq | 50 bp PE | ACCTCA |
| HeLa_TF_siCtrl.r2 | GS1161 | RNAseq | 50 bp PE | AGCATG |
| HeLa_TF_siDIS3.r2 | GS1162 | RNAseq | 50 bp PE | GAGTCA |
| HeLa_TF_siSF3A3.r2 | GS1163 | RNAseq | 50 bp PE | CGTAGA |
| HeLa_TF_siSRSF3.r2 | GS1164 | RNAseq | 50 bp PE | TCAGAG |
| HeLa_TF_siSETDB1.r2 | GS1165 | RNAseq | 50 bp PE | CACAGT |
| HeLa_siCtrl H3K9me3. r2 | GS1168 | ChIP-seq | 50 bp PE | 1AGTTCC |
| HeLa_siDIS3 H3K9me3. r2 | GS1169 | ChIP-seq | 50 bp PE | ATGTCA |
| HeLa_siSF3A3 H3K9me3. r2 | GS1170 | ChIP-seq | 50 bp PE | CCGTCC |
| HeLa_siSRSF3 H3K9me3. r2 | GS1171 | ChIP-seq | 50 bp PE | GTCCGC |
| HeLa_siSETDB1 H3K9me3. r2 | GS1172 | ChIP-seq | 50 bp PE | GTGAAA |

## 4.2 Methods

### 4.2.1 Cell culture

THP-1 and KG1a suspension cells were thawed and maintained in a medium based on RPMI-1640 or IMDM, respectively, containing 15% FCS (F7542, Sigma), 5 ml non-essential amino acids (M7145, Sigma), and 5 ml Penicillin/Streptomycin (P4333, Sigma). HeLa adherent cells were thawed and maintained in a medium based on DMEM (D6429, Sigma), containing 10% FCS (F7542, Sigma), 5 ml non-essential amino acids (M7145, Sigma), and 5 ml Penicillin/Streptomycin (P4333, Sigma). Cells were split every 3 days: for suspension cell lines – an appropriate amount of old medium with cells was added into a flask with fresh medium; for adherent cell line - trypsin was used to detach cells, then needed amount of cells were seeded into a fresh culture dish. All cell lines were cultured at 37 °C in a 5% CO2 humidified atmosphere.

### 4.2.2 Cell culture treatments

#### 4.2.2.1 5-Aza treatment

Decitabine (5-aza-2′-deoxycytidine) (617982, Sigma) was added to THP-1 and KG1a cells the next day after splitting to a final concentration of 0.5 uM in the medium. Since 10 mM 5-Aza stock was diluted in DMSO (276855, Sigma), the respective amount of DMSO was added into the control cells medium. The next day, the old medium was changed with a fresh medium with the same concentration of 5-Aza. Cells were incubated with 5-Aza in total for 72 hours. After 72 hours, cells were collected for RRBS, ChIP-seq, and histone mass spectrometry sample preparations. In addition, an aliquot of cells was kept for 4-thiouridine (T4509, Sigma) labeling.

#### 4.2.2.2 Panobinostat treatment

Panobinostat (LBH589) (S1030, Selleck Chemicals) was added to THP-1 and KG1a cells the next day after splitting to a final concentration of 10 nM in the medium. Since 1 mM Panobinostat stock was diluted in DMSO (276855, Sigma), the respective amount of DMSO was added into the control cells medium. Cells were incubated with Panobinostat in total for 24 hours. After 24 hours, cells were collected for RRBS, ChIP-seq, and histone mass spectrometry sample preparations. In addition, an aliquot of cells was kept for 4-thiouridine (T4509, Sigma) labeling.

#### 4.2.2.3 siRNA transfection

All transfection experiments in HeLa cells were performed in a 10 cm dish using a Lipofectamine RNAiMAX transfection reagent (13778150, Sigma) according to manufacture protocol. In short, cells were seeded 1.5 million cells per dish. The next day 43 ul of transfection reagent (13778150, Sigma) were mixed with 1 ml of Opti-MEM Reduced-Serum Medium (31985062, Sigma) and appropriate Silencer Select siRNA (Sigma). Then this solution was added to cells medium to a final concentration of 20 nM of siRNA in the media. Cells were incubated for 72 hours. After 72 hours, cells were collected for RRBS, ChIP-seq, RT-qPCR, and Western blot sample preparations. In addition, an aliquot of cells was kept for 4-thiouridine (T4509, Sigma) labeling.

#### 4.2.2.4 4-thiourudine metabolic labeling

4-thiourudine (T4509, Sigma) was added into the cell culture to a total concentration of 200 uM. Cells were incubated with 4-thiourudine (T4509, Sigma) for 50 min in a 10 cm dish. After labeling, cells were harvested, diluted in TRIzol reagent (15596026, Sigma), and were either stored in -80 °C or subjected directly to RNA extraction.

### 4.2.3 Protein extraction and Western blot

Whole-cell protein extracts were prepared by resuspending 1 million cell pellets in 35 ul of freshly prepared lysis buffer. Samples were vortexed for 10 s and boiled for 10 min at 95 °C. After incubation with 1 ul of Benzonase and 2.5 mM MgCl2 at 37°C for 15 min, protein extracts were mixed with 12 µl Loading buffer and boiled again for 5min at 95 °C.

Each sample (20 ul) was electrophoresed on pre-cast 4 - 20 % TG PRIME gel (43277, Serva) at a constant current of 25 mA/gel on a vertical double gel system (45-1010-I, Peqlab) in a running buffer. Triple colour protein standard III (39258, Serva) was run on every gel. Then separated proteins were electrophoretically transferred to polyvinylidene fluoride transfer membrane (T830.1, Rothi) soaked in 98 % methanol using mini trans-blot electrophoretic transfer cell (1703930, Biorad) in the transfer buffer. Blots were blocked for one hour in the blocking buffer. Blots were then incubated with one of the primary antibodies in blocking buffer for 20 hours at 4 °C in a wet chamber. After incubation with primary antibodies, blots were washed 3 times in PBST and then incubated with appropriate secondary antibodies in a blocking buffer for 1.5 hours at room temperature in a wet chamber. Finally, blots were washed 3 times in PBST and once in PBS. In the end, proteins were visualized on the Odyssey CLx imaging system (Li-Cor Biosciences).

Solutions:

- Lysis buffer: 50 mM Tris/Cl pH 7.5, 2% SDS, 1% Triton X100, 0.5x Roche Complete Protease Inhibitor Cocktail, 1 mM PMSF
- Loading buffer: 5 x Laemmli buffer (10 % SDS, 250 mM TRIS-HCl pH 6.8, 25 % (v/v) glycerol, 500 Mm DTT, 0.5 % (m/v) bromophenol blue)
- Blocking buffer: 2,5 % BSA / 2,5 % milk in PBS-T
- PBST: 1x PBS with 0.1% Tween 20
- Running buffer: 25 mM Tris, 190 mM glycine, 0.1 % SDS
- Transfer buffer: 25 mM Tris, 190 mM glycine, 15 % methanol, 0.2 % SDS

### 4.2.4 Transient transcriptome sequencing methods

#### 4.2.4.1 RNA extraction

If RNA was isolated for TT-seq, then biotinylated and unmodified RNA spike-in mix was added to the cells mixed with TRIzol: 2.4 ng of spike-in mix per 1 million cells. Total RNA was isolated from cells using the RNA Clean & Concentrator kit - 25 (R1017, Zymo Research), including digestion of remaining genomic DNA according to producer´s guidelines. Agilent 2100

Bioanalyzer was used to assess RNA quality, and only high-quality RNA (RIN > 8) was further processed.

#### 4.2.4.2 RNA spike-ins generation with and without 4-Thio-UTPs

Spike-ins used in TT-seq were derived from four selected RNAs of the ERCC RNA Spike-in Mix (4456740, Sigma). To amplify desired spike-ins from 1 ng of ERCC RNA Spike-in Mix (4456740, Sigma) cDNA synthesis was performed using the SuperScript III kit (18080093, Sigma) following the manufacturer's instructions. Desired spike-ins were amplified with 30 cycles from a cDNA template using Phusion HF DNA polymerase (M0530L, NEB) following the manufacturer's instructions. Forward and reverse primers were designed in a way that they bind only to selected spike-in sequences. PCR products were purified with QIAquick PCR purification columns (28106, Qiagen) following the manufacturer's instructions. The size and purity of PCR products were controlled by gel electrophoresis. The sequence of PCR products was verified by Sanger sequencing, performed by Eurofins.

PCR fragments were phosphorylated using T4 Polynucleotide Kinase (M0201S, NEB). pBluescript II SK (+) plasmid was cut with EcoRV-HF enzyme (R3195M, NEB) and treated with calf intestinal alkaline phosphatase (M0290, NEB) following by purification with QIAquick PCR purification columns (28106, Qiagen). Then blunt-end PCR fragments were inserted into dephosphorylated cut pBluescript II SK (+) plasmid using T4 DNA ligase (M0202M, NEB) following manufacturer's instructions.

EcoRI-HF digested pBluescript plasmids with inserted spike-ins were input for the in vitro transcription using T7 polymerase (M0251S, NEB) following manufacturer's instructions, except for IVT of labeled RNA spike-ins (referred to as #3 and #5) 10 % of UTPs were substituted with 4-Thio-UTPs (NU-1156S, Jena Biosciences). RNA spike-ins were purified with RNase-free AMPure XP beads (A63987, Beckman Coulter) following the manufacturer's instructions. Concentrations of generated RNA spike-ins were measured by Qubit fluorimeter, and quality was assessed by Agilent Bioanalyzer.

Finally, equal amounts of each spike-in were mixed into the final RNA spike-in pool with a final concentration of 1 ng/µL.

#### 4.2.4.3 Transient transcriptome sample and NGS library preparation

To 4-thiouridine labeled and extracted RNA was given a short pulse of fragmentation in a Covaris S220 device (settings: time 10 sec. temperature 4°C, duty factor 1%, peak incident power 100 Watts, cycles per burst 200). Fragmented RNA was heated for 10 min in a ThermoMixer (Eppendorf) at 60 °C and 800 rpm shaking, then it was kept on ice for 2 min. 1 ug of total RNA was put aside for library preparation as a total RNA fraction.

The rest of the RNA was biotinylated with EZ-Link biotin-HPDP (A35390, Sigma). Before the reaction, EZ-Link biotin-HPDP (A35390, Sigma) was freshly dissolved in dimethylformamide (227056, Sigma) to a concentration of 1 mg/ml. The reaction was performed in a biotinylation reaction mix. Mix was incubated for 1.5 hours in a ThermoMixer (Eppendorf) at 24 °C and 800 rpm shaking. Biotinylated RNA was cleaned using phase-lock gel tubes (10847-802, VWR). Briefly, RNA biotinylation reaction was mixed 1:1 v/v with chloroform (67-66-3, Sigma), centrifuged at 4 °C at a maximum speed. The upper phase was transferred to a new phase-lock gel tube (10847-802, VWR), and cleaning was repeated one more time. To precipitate RNA, NaCl was added to the upper phase to a total concentration of 0.5 M NaCl plus 1 volume of isopropanol (1070224000, Sigma) mixed and centrifuged for 20 min at 4 °C at a maximum speed. Then RNA pellet was washed with 75% ethanol and centrifuged for 10 min at 4 °C at a maximum speed. Then, the RNA pellet was resuspended in water. 4-thiouridine labeling and biotinylation reaction efficiency was monitored on a dot blot using 1-2 ug of biotinylated RNA.

The rest of the biotinylated RNA was used as an input for the separation of labeled from unlabeled RNA fractions using µMACS streptavidin kit (130-074-101, Miltenyi). Before separation, RNA was heated at 65 °C for 10 min and put on ice. 1 volume of biotinylated RNA was incubated with 0.5 volume of uMACS streptavidin beads was incubated for 15 min ThermoMixer (Eppendorf) at 24 °C and 800 rpm shaking. Then, RNA with beads was transferred to a separation column pre-equilibrate with nucleic acid equilibration buffer (130-074-101, Miltenyi). The column with bound beads was washed three times with heated to 65 °C freshly prepared beads-washing buffer and three times with room temperature beads-washing buffer. Flow-through after the first wash was collected for library preparation of unlabelled RNA fraction. In the end, labeled RNA was eluted two times with 100 mM DTT (10197777001, Sigma).

Separated labeled fractions and total RNA fractions were cleaned with RNA Clean & Concentrator kit - 25 (R1017, Zymo Research) according to the producer´s guidelines. Unlabeled RNA fraction was cleaned as follows. 1 volume of isopropanol (1070224000, Sigma) was added to the unlabeled RNA fraction. Mix was centrifuged for 20 min at 4 °C at a maximum speed. Then RNA pellet was washed with 75% ethanol and centrifuged for 10 min at 4 °C at a maximum speed. Then, the RNA pellet was resuspended in water. Agilent 2100 Bioanalyzer was used to assess RNA fragment distribution for all fractions.

Fraction separation efficiency was checked by RT-qPCR of low-turnover transcripts from GAPDH and fast-turnover transcripts from MYC.

Separated fractions were used as an input for strand-specific total RNA library preparation with Ovation Universal RNA-Seq System 1-16 (0343-32, Nugen) and processed according to the

manufacturer´s instruction. Libraries were quality controlled by Qubit fluorimeter and Agilent DNA Bioanalyzer.

Solutions:

- Biotinylation reaction mix per 1 ug of RNA: 1 ul 10x biotinylation buffer (100 mM Tris pH 7.5, 10 mM EDTA), 5 ul of 1 ug RNA in water, 2 ul of dimethylformamide (227056, Sigma), 2 ul of 1 mg/ml. EZ-Link biotin-HPDP (A35390, Sigma).
- Beads-washing buffer: 100 mM Tris pH 7.5, 10 mM EDTA, 1 M NaCl, 0.1 % Tween20 (P9416, Sigma)

#### 4.2.4.4 RNA Dot blot

For the dot blot RNA was denatured at 95 °C for 1 3 min, then was put on ice. 1-2 ug of biotinylated RNA and not biotinylated RNA as a negative control was applied to a nylon Hybond-N+ membrane (RPN303B, Amersham). The membrane was blocked for 30 min in the blocking buffer with rocking. Then, the blocked membrane was incubated with 1:20,2000 diluted Streptavidin-HRP (21130, Pierce) for 15 min in an incubation buffer with the rocking followed by two washes with incubation buffer, two washes with wash buffer #1, and two washes with wash buffer #2. Finally, the membrane was developed using an ECL machine (Biorad).

Solutions:

- Dot blot blocking buffer: 5 % milk in PBS-T
- Dot blot incubation buffer: 1x PBS with 10 % SDS
- Dot blot wash buffer #1: 1x PBS with 1% SDS
- Dot blot wash buffer #2: 1x PBS with 0.1% SDS

### 4.2.5 Chromatin immunoprecipitation of histone modifications and NGS library preparation

2 million cross-linked cells (1 % formaldehyde, 10 min RT) were lysed in 100 ul Buffer B and sonicated in microtubes (520045, Covaris) using a Covaris S220 machine aiming for DNA fragments distribution around 200 - 500 base pairs (settings: temperature 4 °C, duty cycle 2 %, peak incident power 105 Watts, cycles per burst 200, processing time 27 min). After shearing lysates were centrifuged for 10 min, 4 °C, 12000 g and supernatant diluted with 900ul of Buffer A. Then 150 ul of the sonicated chromatin was incubated three hours at 4 °C on a rotating wheel with H3K9me3 (39161, Active Motif) or H3K9ac (39137, Active Motif) antibodies conjugated to 10 ul of Dynabeads protein A (10001D, Sigma). Beads with bound antibodies and chromatin were washed four times with ice-cold Buffer A and once with Buffer C. Then,

beads were re-suspended in 100 µl elution buffer and incubated for 30 min at 37 °C with 3 µl RNase A (10 mg/ml) to degrade RNA. Then proteins were digested with 6 µl Proteinase K (10 mg/ml) at 55°C for one hour. Crosslink reversal of immunoprecipitated DNA was carried out overnight at 65°C. Finally, DNA was isolated and purified with SPRI AMPure XP beads (B23319, Beckman Coulter) with a sample-to-beads ratio of 1:2.

Agilent 2100 Bioanalyzer was used to assess the efficiency of DNA fragments shearing. ChIP efficiency was checked by performing qPCR for control genes - GAPDH (positive for H3K9ac mark and negative for H3K9me3) and ZNF180 (negative for H3K9ac mark and positive for H3K9me3).

Isolated DNA was used as an input for the NGS library preparation with Ultra II DNA Library Prep Kit for Illumina (E7645S, NEB) or ThruPLEX DNA-seq kit (R400675, Takara) and processed according to the manufacturer´s instruction. Libraries were quality controlled by Qubit and Agilent DNA Bioanalyzer.

Solutions:

- Buffer B (lysis buffer): 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1 % SDS, 1x protease inhibitor (Roche)
- Buffer A (IP and wash buffer): 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM EGTA, 1 % Triton X-100, 0.1 % SDS, 0.1 % Na-deoxycholate, 140 mM NaCl, 1x protease inhibitor (Roche)
- Buffer C: 10 mM Tris-HCl, pH 8.0, 10 mM EDTA
- Elution buffer: 10 mM Tris-HCl, pH 8.0, 5 mM EDTA, 0.5 % SDS, 300 mM NaCl
- TE buffer: 10 mM Tris-HCl, pH 8.0, 1 mM EDTA

### 4.2.6 RT-qPCR and qPCR

For the RT-qPCR analysis, 1 µg of total RNA was transcribed with SuperScript III kit (18080093, Sigma) into cDNA. Then cDNA was amplified with the Fast SYBR Green Master mix (4385610, Applied Biosystems) and primers using a Light-Cycler (Roche). For the qPCR, DNA fragments were directly amplified the same way using a LightCycler480 (Roche). Every PCR reaction was performed in a total volume of 10 µl in duplicates in a 384-well plate (Sarstedt). The expression level of GAPDH was used as a reference for qRT-PCR experiments, and the geometric mean of reference Ct values was used as normalization as described by others (Vandesompele et al., 2002). Ct-values were generated by the LightCycler480 software (Roche) using the 2nd derivative-max function.

### 4.2.7 Sample preparation for Reduced Representation Bisulfite sequencing

Genomic DNA was isolated from 1 million cells using the DNEasy Blood and Tissue Kit (69506, Qiagen). Isolated DNA was digested overnight with MspI enzyme (R0106S, NEB) following the manufacturer's instructions. Digested DNA was prepared for adaptor ligation by filling gaps and performing A-tailing using DNA Polymerase I Klenow fragment (M0210S, NEB) following the manufacturer's instructions. Then DNA was purified with SPRI AMPure XP beads (B23319, Beckman Coulter) with sample-to-beads ratio of 1:2. Methylated NEBNext Multiplex Oligos for Illumina, 1 µM each (E7535S, NEB) were ligated overnight to purified DNA fragments using T4 DNA ligase (M0202M, NEB) following manufacturer's instructions. Before bisulfite conversion, the same amounts of libraries were pooled together and concentrated using SPRI AMPure XP beads (B23319, Beckman Coulter) with a sample-to-beads ratio of 1:2. Then pooled libraries were subject to bisulfite conversion using the EpiTect Bisulfite Kit (59104, Qiagen) according to the manufacturer's protocol. Bisulfite converted DNA libraries were amplified with 20 cycles using PfuTurbo DNA Polymerase (600250, Agilent) and Illumina P5/P7 primers, 2.5 µM each, following manufacturer's instructions. Finally, libraries were purified with SPRI AMPure XP beads (B23319, Beckman Coulter) with a sample-to-beads ratio of 1:1.2. Libraries were checked for the correct fragment length on the Bioanalyzer 2100 (Agilent), and concentrations were measured with Qubit dsDNA HS assay kit (Q32851, Life Technologies).

### 4.2.8 High-throughput sequencing

High-throughput sequencing was performed on Illumina's HiSeq 1500 sequencer machine at the Laboratory for Functional Genome Analysis (LAFUGA) within the Gene Center (LMU Munich) using 50 bp single-end or paired-end mode.

### 4.2.9 Nuclear histone acid extraction and sample preparation for the mass spectrometry

1 million cells per sample were harvested and re-suspended in 100 ul of 0.2 M H2SO4. Re-suspended cells were vortexed and rotated overnight at 4 °C. Then, the supernatant was transferred to a new tube, and 100 % trichloroacetic acid was added to a final concentration of 26 %. The solution was incubated for 2 hours on ice to precipitate the histones. Precipitated histones were collected by centrifugation at a maximum speed 15 min at 4 °C. The pellet was washed four times with 500 uL of ice-cold acetone. Then the acetone was removed, and the pellet was subject to mass spectrometry sample preparation.

The protein pellet was resuspended in Laemmli buffer (Roti-Load 4x, Roth) and boiled at 95°C for 10 min. Then, samples were loaded on a 4-20% gradient polyacrylamide gel (V140115-1, SERVA) for SDS-PAGE separation and were run with 25 mA per gel. The gel was stained with Coomassie Blue to visualize the histone bands. Histone bands were excised, cut into smaller pieces, and incubated in 200 µl 50 % acetone in 50 mM ammonium bicarbonate 2 times for 30 min at 37 °C to de-stain gel pieces. The gel pieces were washed 2 times for 5 min with water and 2 times for 5 min with 100 % acetone to dehydrate the gel. Gel pieces were propionylated with 5 µl of Propionic anhydride, as described in (Villar-Garea et al., 2008), and dehydrated with 100% acetone. Then, histones were digested with Trypsin. Digested peptides were extracted from the gel pieces by 1st wash in 50 mM ammonium bicarbonate, 2nd wash with 50 % acetone plus 0.1 % trichloroacetic acid, and 3$^{rd}$ wash with 100 % acetone, followed by desalting using C18 Stagetips (3M Empore) and the porous carbon material (TipTop Carbon, Glygen), as described in (Rappsilber et al., 2007). Extracted peptides were resuspended in 15 µl 0.1 % trichloroacetic acid and stored at -80 °C.

### 4.2.10 Liquid chromatography-mass spectrometry analysis and quantification of histone modifications

The Q-Exactive HF mass spectrometer (Thermo Fisher) at the Protein Analysis Unit within BMC LMU was used. Q-Exactive HF operated in data-dependent mode to automatically switch between full-scan MS and MS/MS acquisition. Survey full-scan MS spectra (from m/z 250–1600) were acquired with resolution R=60,000 at m/z 400 (AGC target of 3x106). The mass spectrometric conditions were: spray voltage 1.5 kV, no sheath and auxiliary gas flow, heated capillary temperature 250ºC, ion selection threshold 33.000 counts. To identify specific post-translational modifications, the resulting spectra were analyzed with Skyline software (MacLean et al., 2010). Integrated peak values were selected manually and exported in the data table format for further analysis. The sum of the total area of integrated peak values for the same isotopically modified peptide represented the amount of a total peptide. The relative abundance of an observed modified peptide was calculated as a percentage of the overall peptide.

### 4.2.11 Bioinformatic analysis

#### 4.2.11.1 TT-seq analysis

All annotations and pipelines used in this analysis are based on the hg38 human genome. Raw reads from fastq files were mapped to human genome version hg38 with STAR (Dobin et al., 2013) with default options "--runThreadN 32 --quantMode TranscriptomeSAM GeneCounts --

outSAMtype BAM SortedByCoordinate". We used bam files generated by STAR as input for Samtools (Li et al., 2009). With Samtools we filtered out reads with low mapping score, MAPQ less than 20 using options “-b -f 2 -F 1280 -q 20”. To create a reference GTF file for LTR elements, we used a RepeatMasker based annotation table of LTR elements by merging adjacent LTR fragments with a distance between them less than 500 bp and with the same strand orientation into one LTR entity (Tarailo-Graovac and Chen, 2009). It was done to reduce excessive fragmentation of original LTR reference and to merge internal ERV loci with flanking LTRs. For coding genes reference was used protein sequence annotations from the UniProt/SwissProt database, mapped to genomic coordinates, was retrieved from the UCSC table browser (2021; Navarro Gonzalez et al., 2021) For the long non-coding genes reference, the long non-coding RNA gene annotation was used, which retrieved from GENCODE website (Frankish et al., 2019). GTF file for spike-ins annotation was created, based on spike-in sequences retrieved from NCBI GenBank, using spike-ins ERCC IDs.

To create an expression matrix for ERV loci and genes, we used HTSeq software (Anders et al., 2015), which takes into account reads direction. Expression matrixes were used as an input for R (Team, 2018). Read counts for LTR repeats and genes were normalized using the R package “DESeq2” (Love et al., 2014) separately for labeled and steady RNA fractions. Then, normalized counts were log2 transformed. LTR repeats with read-through transcription were excluded by sub-selecting for further analysis only LTR repeats containing H3K9ac peaks in any condition for each cell line, using the R package “GenomicRanges” (Lawrence et al., 2013). LTR repeats and genes with significantly changed expression within the same RNA fraction were determined through pairwise comparisons using the DESeq2 results function (adjusted p-value < 0.05). Transcript turnover for each locus was calculated by subtracting total or unlabeled RNA fraction log2 transformed normalized read counts from labeled RNA fraction log2 transformed normalized read counts. WebGestalt online tool was used for the over-representation pathway enrichment analysis against the Reactome database (Wang et al., 2017). Differentially spliced genes were identified with the R package “NBSplice” (Merino and Fernández, 2020), which takes isoform expression matrix produced by RSEM (Li and Dewey, 2011) as an input. RSEM isoform expression matrixes were produced from STAR-output bam files. ERV loci were counted as bi-directionally transcribed if reads mapped in opposite orientation to ERV loci exceeded log2 transformed and normalized expression value of 2. To plot results, R package “ggplot2” was used (Wickham, 2016).

The separation efficiency of RNA fractions was checked using the R package “rCube”(Wickham, 2016), which was developed to analyze TT-seq data. Mapped with STAR read counts from labeled and unlabeled RNA spike-ins for each sample were used as an input for a “rCube” package. Then, the ratio of unlabeled spike-ins in labeled RNA fraction versus

unlabeled spike-ins in total RNA fraction was calculated for each sample. Based on this, RNA fractions separation efficiency was estimated.

#### 4.2.11.2 ChIP-seq analysis

ChIP-seq reads were aligned to the human genome hg38 using Bowtie with options “-q -n 2 --best --chunkmbs 2000 -p 32 -S”. Then, duplicated reads were removed with the Picard software (2018). The H3K9me3 enrichment on LTR elements and coding genes was assessed using Homer (Heinz et al., 2010) through analyzeRepeats.pl with respective annotations. The total number of mapped reads was normalized to 10 million reads per experiment. Then normalized read count table was loaded into R software. To identify H3K9me3 and H3K9ac peaks, Homer software findPeaks function was used with the option “-style histone”. Then identified peaks were loaded into R software and assigned to LTRs and genes using the R package “GenomicRanges”. Input sample was used in a peak calling to correct for background.

#### 4.2.11.3 Reduced Representation Bisulfite sequencing analysis

RRBS data files were trimmed and cleaned. Bismark software (Krueger and Andrews, 2011) was used to map files to the Homo sapiens GRCh38 reference genome. Mapped data were analyzed using the R package “methylKit” (Akalin et al., 2012) to determine the methylation at individual CpG sites. P-values were generated using logistic regression and a q-value cut-off of < 0.01, and a percentage methylation difference of > 25 % was used by “methylKit” to identify differential methylation events.

# 5. Abbreviations

| | |
|---|---|
| 4sU | 4-thiouracil |
| 5-Aza | decitabine (5-aza-2′-deoxycytidine) |
| AML | acute myeloid leukemia |
| bp | base pair(s) |
| BrU | 5-bromouracil |
| cDNA | complementary DNA |
| ChIP | chromatin immunoprecipitation |
| ChIP-qPCR | chromatin immunoprecipitation followed by quantitative PCR |
| ChIP-Seq | chromatin immunoprecipitation followed by high throughput sequencing |
| CpG | cytosine nucleotide followed by a guanine in the 5’ to 3’ direction |
| CRISPR | clustered regularly interspaced short palindromic repeats |
| Dis3 | DIS3 homolog, exosome endoribonuclease and 3'-5' exoribonuclease |
| DNA | deoxyribonucleic acid |
| Dnmt | DNA methytransferase |
| DNMTi | DNA methyltransferase inhibitor |
| dNTP | Deoxynucleoside triphosphate |
| dsRNA | double-stranded RNA |
| ERV | endogeneous retroviral element |
| GAPDH | Glyceraldehyde 3-phosphate dehydrogenase |
| H3K27me3 | histone 3 Lysine 27 trimethylation |
| H3K9ac | histone 3 Lysine 9 acetylation |
| H3K9me2/3 | histone 3 Lysine 9 -, di- or tri- methylation |
| H4K20me2/3 | histone 4 Lysine 20 -, di- or tri- methylation |
| HDACi | histone deacetylase inhibitor |
| HDAC | histone deacetylase |
| HeLa | human epithelial cell line derived from cervix epithelioid carcinoma |
| HERV | human endogeneous retroviral element |
| HIV-1 | human immunodeficiency virus type 1 |

| | |
|---|---|
| IAP | internal A-type Particle |
| IGV | Integrative Genomics Viewer |
| IP | immunoprecipitation |
| IP-MS | immunoprecipitation mass spectrometry |
| kb | kilo-base pair(s) |
| KD | knockdown |
| KG1a | human macrophage cell line derived from acute myelogenous leukemia patient |
| KO | knockout |
| KRAB-ZNF | Krüppel-associated box domain zinc finger |
| lincRNA | long intergenic non-coding RNAs |
| LINEs | long interspersed elements |
| LTR | long terminal repeat (a big subclass of retrotransposons) or regulatory sequence element flanking the retrotransposon sequence |
| MBD | methyl-CpG binding domain |
| mRNA | messenger ribonucleic acid |
| MYC | MYC Proto-Oncogene, BHLH Transcription Factor |
| NEXT | the nuclear exosome targeting (complex) |
| NGS | next-generation sequencing |
| ORA | over-representation analysis |
| qPCR | quantitative polymerase chain reaction |
| RNA | ribonucleic acid |
| RNAi | RNA interference |
| RNA-Seq | RNA high throughput sequencing |
| rpm | rotations per minute |
| RRBS | reduced representation bisulfite sequencing |
| rRNA | ribosomal RNA |
| RT | room temperature |
| RT-qPCR | reverse transcription quantitative polymerase chain reaction |
| Setdb1 | SET Domain Bifurcated Histone Lysine Methyltransferase 1 |
| Sf3a3 | Splicing Factor 3a Subunit 3 |

| | |
|---|---|
| SINEs | short interspersed nuclear elements |
| siRNA | short interfering ribonucleic acid |
| Srsf3 | Serine And Arginine Rich Splicing Factor 3 |
| STAR | Spliced Transcripts Alignment to a Reference |
| TE | transposable element |
| TFBS | transcription factor binding site |
| THP-1 | human monocytic cell line derived from acute monocytic leukemia patient |
| Trim28 | Tripartite motif containing 28 |
| TSS | transcription start site |
| TT-seq | transient transcriptome sequencing |
| UTR | untranslated region |

# 6. Acknowledgements

I would like to express my deepest gratitude to my supervisor Prof. Dr. Gunnar Schotta for giving me the opportunity to work on such an amazing and interesting project. Gunnar invested a lot of time in helping me to develop the project. I am very grateful for the fruitful discussions and brainstorming we had during my Ph.D., which shaped me as a scientist.

I also want to sincerely thank the head of our department Peter Becker, BMC secretary Edith Müller, and IRTG Program coordinator Dr. Elizabeth Schroeder-Reiter for the huge support and assistance, which I received from them.

Special thanks go to my current and former lab members: Filippo Cernilogar, Zeyang Wang, Sophia Groh, Maryam Kazerani, Helena Moreno, Andrea Terrasi, Heike Bollig, Giulia Nicoletto, Angela Russo, and Eugenio Fergola for huge scientific and emotional support. They helped me to learn an enormous amount of lab techniques and shared with me a lot of fun outside the lab.

A huge portion of thankfulness goes to my peers and at the same time good friends: Daniil Pokrovskii, Andrea Lukacs, Magdalini Serefidou, Alessandro Angerilli, and others. It was worth moving to Münich just to meet you and become your friend. Thanks to you, my Ph.D. was one of the happiest times in my life.

Также хочу поблагодарить Русский ланч за поддержку и конструктивное обсуждение насущных новостей и важнейших философских вопросов.

In the end, I want to thank my parents, Elizaveta Romanova and Vladimir Shcherbakov, for believing in me, for their love and constant support.

Спасибо, у меня всё.

# 7. Appendix

### Table 8.1 Significantly deregulated genes in DIS3 KD, Hela cells

| geneID | Log2 basemean | Log2 Fold Change | padj |
|---|---|---|---|
| ACTBL2 | 6,195845 | 2,935499 | 0,002517 |
| ADAMTS1 | 10,906208 | 1,827254 | 0,012631 |
| ADM | 8,847370 | 1,889754 | 0,031962 |
| ALDH8A1 | 5,231874 | 2,187725 | 0,043946 |
| CALU | 13,805285 | 1,423829 | 0,012631 |
| CLDN1 | 9,256533 | 2,308610 | 0,011224 |
| CXCL8 | 10,952191 | 1,518616 | 0,030483 |
| CYP4F3 | 5,346227 | -0,017284 | 0,025013 |
| DIS3 | 11,487964 | -2,395067 | 0,000000 |
| DUSP1 | 11,103545 | 1,900835 | 0,030483 |
| DUSP8 | 7,680435 | 2,308234 | 0,012130 |
| F3 | 10,149839 | 1,910807 | 0,019769 |
| FST | 6,819088 | 2,362014 | 0,015772 |
| GOPC | 11,121804 | 1,199251 | 0,044973 |
| KTN1 | 13,142162 | -1,404661 | 0,011224 |
| LAMA1 | 7,715745 | 1,611503 | 0,046417 |
| MET | 11,260266 | 1,672347 | 0,012631 |
| MFAP5 | 7,529094 | 2,211431 | 0,018280 |
| MFSD6 | 8,991936 | -1,794365 | 0,025531 |
| MYBPC2 | 5,611764 | -3,418299 | 0,011224 |
| NFATC2 | 6,801137 | 2,195420 | 0,029877 |
| NR4A3 | 9,767375 | 2,002211 | 0,012631 |
| OTUD1 | 8,832919 | 1,705133 | 0,029877 |
| PTPN14 | 12,128804 | 1,368753 | 0,012631 |
| R3HDM2 | 9,306988 | -1,872698 | 0,016344 |
| RBPJ | 10,917425 | 1,260560 | 0,026077 |
| RCHY1 | 7,861696 | 2,052456 | 0,001394 |
| RHOB | 12,819420 | 1,746641 | 0,018200 |
| RNF19A | 9,929714 | 1,262398 | 0,028448 |
| RNU2-2P | 7,965721 | 1,900512 | 0,011224 |
| S100A2 | 9,864465 | -1,227112 | 0,046417 |
| STC2 | 12,565884 | 1,499507 | 0,038517 |
| STRA6 | 10,226414 | -2,013653 | 0,002517 |
| TACSTD2 | 8,833076 | 1,447448 | 0,031962 |
| TAGLN | 7,192540 | 3,115868 | 0,000007 |
| TAPT1 | 7,621472 | 1,958689 | 0,011224 |

### Table 8.2 Significantly deregulated genes (top 1000) in SF3A3 KD, Hela cells

| geneID | Log2 baseMean | log2 FoldChange | padj |
|---|---|---|---|
| AASS | 9,11638 | -3,00313 | 1,51E-12 |
| AATK | 9,39561 | -2,14746 | 5,55E-05 |
| ABCA7 | 7,10346 | -2,57155 | 1,87E-09 |
| ABCB9 | 10,62360 | -1,9442 | 0,004701 |
| ABCC3 | 6,28079 | -2,83785 | 5,00E-08 |
| ABHD15 | 10,73101 | -1,92874 | 0,000872 |
| ABL2 | 9,75957 | 2,340608 | 6,85E-08 |
| ABLIM2 | 9,57265 | -2,17775 | 0,000184 |
| ACACB | 7,62231 | -2,83846 | 3,12E-08 |
| ACBD4 | 6,87479 | -2,64687 | 6,24E-05 |
| ACP6 | 11,39145 | -1,92664 | 0,000714 |
| ACRC | 3,11891 | 2,85754 | 2,85E-05 |
| ACSF3 | 9,65091 | -2,21029 | 1,19E-05 |
| ACTG1P20 | 5,27276 | 3,28194 | 1,59E-07 |
| ACVR1B | 9,97823 | 2,123809 | 2,66E-05 |
| ACVR1C | 6,26574 | 3,0498 | 2,57E-05 |
| ADAM20 | 11,04413 | 4,061393 | 6,29E-05 |
| ADAM20P1 | 10,25787 | 2,359306 | 0,000794 |
| ADAM32 | 9,87135 | 2,786716 | 4,38E-05 |
| ADAMTS14 | 6,02597 | -3,06697 | 6,79E-05 |
| ADAMTSL4 | 9,73403 | -2,62115 | 6,68E-05 |
| ADCK1 | 9,77099 | -2,04851 | 0,005243 |
| ADCK4 | 12,20350 | -2,42785 | 9,36E-07 |
| ADGRB2 | 7,25319 | -2,39667 | 0,000307 |
| ADM | 7,70968 | 2,154373 | 9,72E-05 |
| ADRM1 | 8,02531 | 2,234216 | 2,44E-08 |
| AEN | 9,79765 | 2,204628 | 6,79E-05 |
| AGFG2 | 10,53656 | -2,29311 | 0,000307 |
| AGO4 | 7,48778 | -2,06054 | 0,000145 |
| AGRN | 11,72144 | -2,06053 | 4,28E-05 |
| AIM1L | 7,59019 | -2,66216 | 0,000141 |
| AK9 | 9,67094 | -2,13789 | 0,002646 |
| AKAP17A | 9,86197 | 2,079864 | 5,90E-07 |
| AKIRIN2 | 9,10275 | 2,038436 | 0,000244 |
| AKR7A2 | 7,65483 | -2,1739 | 3,12E-07 |
| ALDH1A3 | 7,35824 | -3,67968 | 7,78E-06 |
| ALDH3B1 | 8,27368 | -3,52117 | 8,94E-15 |
| ALDH4A1 | 9,91938 | -2,38558 | 0,000568 |
| ALDH8A1 | 10,04091 | 4,445021 | 6,47E-07 |
| ALPK1 | 6,38520 | -2,68569 | 6,77E-06 |
| ALPL | 6,97961 | -1,98525 | 0,00397 |

| | | | |
|---|---|---|---|
| ALS2CL | 11,99695 | -2,08841 | 0,000188 |
| AMMECR1L | 9,40677 | 2,04629 | 8,94E-05 |
| ANKRD22 | 6,90322 | -2,63925 | 0,000704 |
| ANP32A-IT1 | 5,43181 | 2,279165 | 0,000464 |
| AOC2 | 12,48304 | 2,901027 | 0,000149 |
| AOC3 | 9,50491 | 1,916275 | 0,001427 |
| AP5S1 | 7,57720 | -2,12989 | 0,000941 |
| AP5Z1 | 9,62060 | 2,00547 | 4,63E-05 |
| ARHGEF17 | 6,95844 | -2,21107 | 1,96E-07 |
| ARHGEF26 | 7,09656 | -2,3183 | 0,000423 |
| ARHGEF40 | 8,34924 | -2,32093 | 0,000231 |
| ARID3B | 6,51261 | 2,238906 | 3,04E-05 |
| ARL10 | 7,97413 | -2,63537 | 0,000175 |
| ARPIN | 12,31050 | -2,06084 | 0,003164 |
| ARRDC1 | 8,75004 | -2,6495 | 3,65E-05 |
| ARRDC3 | 7,69474 | 2,389261 | 4,71E-09 |
| ARSG | 10,45777 | -2,1152 | 0,000188 |
| ARV1 | 8,81466 | -3,19077 | 3,56E-11 |
| ASAH2 | 9,35489 | 3,083834 | 9,53E-05 |
| ASAP3 | 7,63371 | -2,26901 | 7,58E-06 |
| ATF3 | 8,60100 | 2,915375 | 1,62E-05 |
| ATG10 | 9,45373 | -2,22515 | 0,00057 |
| ATG12 | 5,72614 | 1,957488 | 4,48E-07 |
| ATG14 | 10,11441 | 2,638212 | 6,84E-08 |
| ATG16L2 | 5,29562 | -2,81119 | 5,32E-05 |
| ATOH8 | 6,25694 | -1,98664 | 0,00509 |
| ATP2B4 | 6,34425 | -2,19062 | 4,41E-09 |
| ATP6V0A4 | 10,90621 | -2,32517 | 0,000599 |
| ATP7B | 6,22920 | -2,20021 | 0,000469 |
| ATP8A1 | 5,96304 | -2,47652 | 0,00168 |
| AURKAPS1 | 8,93868 | 2,685765 | 8,28E-05 |
| BAAT | 5,43723 | 4,010646 | 4,78E-05 |
| BAHCC1 | 8,04257 | -2,85454 | 3,09E-07 |
| BAIAP2-AS1 | 7,99076 | -2,31405 | 0,000991 |
| BANP | 5,84989 | 2,365208 | 5,04E-07 |
| BBS2 | 8,86660 | -2,25731 | 3,57E-06 |
| BCL2L11 | 8,93275 | 1,931264 | 0,000531 |
| BCLAF1 | 5,26215 | 2,054949 | 3,13E-06 |
| BDH1 | 7,02899 | -2,47548 | 0,000845 |
| BIVM | 8,15734 | -2,73195 | 5,36E-06 |
| BMP8B | 6,84732 | -2,49399 | 0,001879 |
| BRF2 | 9,04614 | 2,305137 | 4,42E-06 |
| BTG1 | 8,51414 | 1,961161 | 1,73E-05 |
| BUD31 | 8,00719 | 2,458888 | 2,61E-08 |

| | | | |
|---|---|---|---|
| C10orf2 | 11,92704 | 2,243779 | 0,000307 |
| C10orf91 | 11,84518 | -2,18678 | 0,004421 |
| C17orf75 | 8,84737 | -2,33425 | 0,000625 |
| C18orf25 | 11,89264 | 1,96333 | 4,13E-05 |
| C19orf66 | 9,72486 | -2,88459 | 7,19E-06 |
| C1D | 7,47034 | 2,270699 | 4,83E-07 |
| C1orf27 | 8,30489 | 1,926104 | 4,71E-06 |
| C1orf52 | 11,97826 | 2,407749 | 6,39E-07 |
| C1orf56 | 9,09088 | -2,14019 | 0,000449 |
| C1QTNF6 | 10,47981 | -2,07026 | 0,001222 |
| C1RL | 7,24030 | -2,68748 | 1,04E-08 |
| C3 | 4,94485 | -2,13886 | 7,09E-07 |
| C6orf223 | 9,96040 | -2,48372 | 3,35E-05 |
| C8orf31 | 7,83779 | -2,12784 | 0,000899 |
| C8orf37 | 7,15215 | -2,06337 | 0,001309 |
| CA12 | 9,52649 | -1,95694 | 0,000331 |
| CABLES2 | 7,66294 | -2,63218 | 3,29E-07 |
| CACNA1G | 10,60926 | -3,07038 | 4,47E-05 |
| CACNA1H | 9,70921 | -2,36754 | 0,000636 |
| CADM4 | 12,87412 | -2,07496 | 0,002723 |
| CALY | 11,30044 | -2,05849 | 0,001706 |
| CAPN10-AS1 | 11,86875 | 2,309253 | 0,001441 |
| CASC15 | 13,93867 | -2,7003 | 2,70E-05 |
| CASKIN1 | 11,87580 | -2,50587 | 0,005468 |
| CBLL1 | 12,25157 | 2,562296 | 7,18E-12 |
| CBR4 | 11,96489 | -1,96218 | 0,00011 |
| CBS | 9,91255 | -3,96678 | 1,12E-07 |
| CBWD5 | 8,94740 | 2,396214 | 1,86E-08 |
| CBX7 | 6,97649 | -2,64284 | 0,000261 |
| CBY1 | 10,54005 | -2,50792 | 2,44E-05 |
| CCDC113 | 10,80991 | -2,88424 | 1,55E-05 |
| CCDC12 | 5,51429 | 1,973379 | 5,14E-08 |
| CCDC163P | 11,01170 | 2,558179 | 0,000164 |
| CCDC17 | 13,69543 | 2,235064 | 0,009314 |
| CCDC181 | 10,12068 | 4,901321 | 4,88E-06 |
| CCDC33 | 10,22872 | -2,31783 | 0,020426 |
| CCDC40 | 10,95569 | -1,99598 | 0,005078 |
| CCDC61 | 7,68731 | -2,4134 | 0,000262 |
| CCDC62 | 10,61377 | 1,958779 | 0,017622 |
| CCDC69 | 10,42492 | -3,09496 | 9,87E-09 |
| CCDC86 | 9,73560 | 2,238013 | 1,27E-06 |
| CCDC9 | 10,44974 | 1,917017 | 1,50E-05 |
| CCDC92 | 10,94572 | -2,33449 | 0,000207 |
| CCNB1IP1 | 8,15905 | 2,149914 | 0,001944 |

| | | | |
|---|---|---|---|
| CCNL1 | 10,92553 | 2,354078 | 5,68E-08 |
| CCNT1 | 5,57745 | 2,456578 | 1,35E-07 |
| CCNT2 | 9,99537 | 2,557373 | 1,87E-09 |
| CD3EAP | 6,42385 | 2,124914 | 0,000121 |
| CDC25B | 7,38557 | -2,06826 | 1,34E-06 |
| CDK11A | 8,17233 | 2,377069 | 3,14E-05 |
| CDK18 | 10,48313 | -2,81329 | 8,58E-07 |
| CDKL2 | 5,23187 | 2,663893 | 0,002535 |
| CDKL3 | 8,04739 | 2,791787 | 3,39E-05 |
| CDKN2AIP | 9,20826 | 2,349907 | 6,70E-06 |
| CEBPG | 11,83118 | 1,968918 | 6,78E-06 |
| CEBPZ | 7,39394 | 2,03608 | 1,11E-07 |
| CEP250 | 7,72859 | -1,92319 | 2,99E-06 |
| CFB | 8,52665 | -2,64558 | 5,65E-07 |
| CFP | 6,35103 | 3,114331 | 0,000109 |
| CGN | 8,07872 | -2,27087 | 0,000581 |
| CGRRF1 | 10,99105 | 2,077136 | 0,000623 |
| CHAC1 | 11,33963 | 2,731765 | 3,10E-06 |
| CHD2 | 7,44858 | 2,296936 | 6,70E-06 |
| CHDH | 9,77975 | -2,31056 | 0,002141 |
| CHST6 | 11,38198 | -2,13469 | 0,006968 |
| CHURC1 | 4,00002 | -1,95758 | 3,30E-07 |
| CIART | 7,61545 | 2,294625 | 0,007306 |
| CKLF | 5,67018 | -2,16285 | 1,94E-06 |
| CLIC3 | 3,75076 | -2,60873 | 0,000213 |
| CLIP2 | 6,84119 | -2,25717 | 7,88E-05 |
| CLK1 | 7,53128 | 2,506832 | 4,27E-08 |
| CLK3 | 10,27014 | 1,894762 | 9,16E-06 |
| CLP1 | 4,15881 | 2,543453 | 1,78E-06 |
| CLTCL1 | 8,87653 | -2,84877 | 2,09E-05 |
| CLYBL | 10,13817 | -2,09897 | 0,001269 |
| CMYA5 | 7,12102 | 5,395416 | 2,84E-20 |
| COL16A1 | 9,10628 | -2,59685 | 2,65E-08 |
| COL18A1 | 10,75848 | -2,09858 | 0,003547 |
| COL7A1 | 9,22082 | -2,1536 | 2,66E-06 |
| COL9A3 | 7,28322 | -2,09207 | 0,004903 |
| COPG2IT1 | 7,19191 | 2,17535 | 0,002542 |
| CORO6 | 8,23505 | -2,28056 | 0,00076 |
| CPNE2 | 5,12470 | -2,92384 | 4,68E-08 |
| CPT2 | 13,07423 | -2,48273 | 1,33E-08 |
| CREB3L4 | 7,19287 | -1,99583 | 0,000615 |
| CREB5 | 12,62347 | 3,072174 | 2,29E-06 |
| CREBRF | 9,88568 | 3,522009 | 1,55E-08 |
| CREG2 | 12,39420 | -2,2845 | 8,70E-05 |

| | | | |
|---|---|---|---|
| CRIP2 | 8,33148 | -2,60063 | 4,77E-06 |
| CRKL | 10,87165 | 1,951497 | 1,24E-05 |
| CRTC1 | 5,15359 | -2,90468 | 3,45E-08 |
| CRYM-AS1 | 5,46093 | 2,073245 | 0,009737 |
| CSPG4 | 7,04036 | -3,25949 | 5,35E-07 |
| CTDSP1 | 6,12865 | -2,13575 | 6,45E-06 |
| CTNNBIP1 | 10,26825 | -2,09996 | 0,001423 |
| CXCL1 | 10,57540 | -2,16387 | 6,45E-05 |
| CXXC5 | 11,50302 | -2,06065 | 2,19E-06 |
| CYB5D1 | 6,10606 | 1,897623 | 0,000699 |
| DAPK1 | 6,59460 | -2,88681 | 4,92E-07 |
| DCAF4L1 | 8,62435 | 3,900312 | 2,62E-08 |
| DDIT3 | 11,16782 | 3,843819 | 1,77E-09 |
| DDX18 | 11,67337 | 2,528994 | 2,81E-14 |
| DDX39A | 3,41360 | 1,893469 | 5,20E-05 |
| DDX46 | 9,95016 | 2,213788 | 2,48E-08 |
| DDX47 | 6,74511 | 2,040954 | 3,10E-06 |
| DDX60 | 9,47109 | -2,92435 | 2,65E-06 |
| DDX60L | 4,20169 | -2,24185 | 3,08E-05 |
| DERL3 | 7,85275 | 2,603668 | 4,54E-05 |
| DES | 9,05641 | -2,49746 | 0,000245 |
| DGKA | 8,38154 | -2,48638 | 3,99E-05 |
| DHRS12 | 9,81912 | -1,96519 | 0,004002 |
| DHRS4L2 | 6,48866 | -2,14864 | 0,000805 |
| DHTKD1 | 12,31944 | -2,16185 | 0,000716 |
| DMBT1 | 6,73099 | -2,00034 | 0,000642 |
| DNAJC2 | 9,52605 | 1,930357 | 8,62E-08 |
| DNAJC3-AS1 | 11,51280 | 2,254093 | 0,000888 |
| DNM3 | 9,73104 | -2,54859 | 2,09E-05 |
| DOHH | 8,80939 | 2,037463 | 2,22E-05 |
| DOPEY2 | 9,41301 | -2,88223 | 0,000111 |
| DPY19L2P4 | 10,01722 | 2,169882 | 0,015136 |
| DSCR9 | 7,21514 | 3,348892 | 4,17E-06 |
| DUSP10 | 9,70829 | 2,324621 | 0,000174 |
| DUSP12 | 5,97880 | 2,101581 | 3,53E-09 |
| DUSP16 | 7,20010 | 2,113354 | 2,31E-05 |
| DUSP19 | 9,75445 | -2,30407 | 0,000816 |
| DYNLT1 | 8,80154 | -2,54167 | 2,48E-08 |
| DYRK1A | 6,82611 | 1,949845 | 1,58E-05 |
| DZIP3 | 8,33550 | -2,47224 | 1,31E-08 |
| E2F6 | 10,81267 | 2,154731 | 5,39E-07 |
| EAF1 | 6,45607 | 2,33523 | 2,93E-07 |
| ECI1 | 9,21042 | -3,18768 | 2,72E-09 |
| EEFSEC | 6,91591 | -2,1764 | 9,08E-05 |

| | | | |
|---|---|---|---|
| EFNA4 | 10,10745 | -2,12052 | 0,002125 |
| EHD2 | 9,40238 | -2,90091 | 5,40E-12 |
| EHF | 6,80078 | -2,14985 | 0,000395 |
| EHHADH | 10,61598 | -2,23047 | 0,003222 |
| EIF1AD | 3,55498 | 2,102441 | 4,56E-06 |
| EMC3-AS1 | 9,10508 | 2,628253 | 1,36E-06 |
| ENDOG | 9,32525 | -2,12016 | 0,000182 |
| ENOX2 | 10,44344 | -1,94997 | 0,001421 |
| ENTPD5 | 10,70155 | -2,48678 | 2,14E-05 |
| EPB41L4A-AS1 | 8,06476 | 2,499114 | 5,20E-08 |
| EPM2AIP1 | 7,72810 | 1,92233 | 0,000397 |
| ETFB | 7,99543 | -2,04691 | 1,14E-06 |
| ETV3 | 8,51075 | 2,170818 | 3,39E-05 |
| EVI5L | 8,66311 | -3,78707 | 1,31E-08 |
| EVL | 11,62445 | -1,96841 | 0,00223 |
| FAHD2B | 7,39616 | -2,87868 | 4,44E-05 |
| FAM133B | 7,22463 | 2,471901 | 3,62E-09 |
| FAM149A | 8,25851 | -2,52053 | 0,000166 |
| FAM175A | 5,60530 | -3,04884 | 2,22E-08 |
| FAM20C | 9,79210 | -2,78264 | 4,63E-05 |
| FAM32A | 7,70136 | 2,150911 | 4,41E-05 |
| FAM63A | 7,94460 | -2,15927 | 0,000277 |
| FAM76A | 10,18029 | -2,45453 | 0,000336 |
| FAM83H-AS1 | 8,89629 | -2,94499 | 4,04E-05 |
| FAM86C1 | 9,40691 | -2,1194 | 0,000676 |
| FAM86JP | 9,44282 | -1,94603 | 0,006204 |
| FBF1 | 9,28026 | -2,11188 | 0,000945 |
| FBXO31 | 6,28175 | 2,028098 | 1,12E-07 |
| FBXO32 | 8,54352 | 2,514344 | 0,000168 |
| FBXW4P1 | 9,41371 | 2,943014 | 8,54E-06 |
| FEM1C | 7,71213 | 2,218593 | 1,09E-05 |
| FEZF1-AS1 | 8,36571 | -2,83099 | 0,000123 |
| FGD1 | 9,95670 | -1,93952 | 0,007842 |
| FGF18 | 9,74292 | 4,21122 | 2,42E-09 |
| FGF19 | 8,17197 | 2,533504 | 0,000872 |
| FGFR3 | 11,67566 | -2,41393 | 0,000819 |
| FIBIN | 11,74669 | -2,50584 | 0,000669 |
| FKRP | 10,48231 | -2,03701 | 4,85E-05 |
| FLJ20021 | 9,57614 | -2,20149 | 0,002843 |
| FLJ22447 | 13,56169 | -2,99252 | 7,56E-07 |
| FLYWCH2 | 11,51464 | -3,64401 | 7,63E-12 |
| FN1 | 6,19215 | -1,94052 | 4,57E-06 |
| FN3K | 10,20594 | -2,48571 | 0,000207 |
| FN3KRP | 10,14615 | -1,98366 | 0,003953 |

| | | | |
|---|---|---|---|
| FOSL1 | 10,02808 | 2,26892 | 5,64E-07 |
| FOXJ3 | 9,39639 | 2,413904 | 7,63E-12 |
| FOXN2 | 6,54017 | 2,246809 | 1,86E-08 |
| FOXO6 | 8,11643 | -2,7625 | 1,89E-07 |
| FRAT2 | 9,04849 | 2,150398 | 1,94E-05 |
| FRRS1 | 11,22031 | -2,04134 | 0,001904 |
| FRS2 | 10,20275 | 2,627392 | 7,29E-08 |
| FSIP2 | 5,49031 | -2,42431 | 4,19E-05 |
| FST | 10,57671 | 3,292949 | 2,67E-07 |
| FTH1P3 | 9,94892 | 2,986202 | 2,98E-06 |
| FUT1 | 11,01367 | 1,974361 | 0,010965 |
| FUZ | 10,10580 | -1,92744 | 0,002683 |
| FYCO1 | 10,30069 | -2,2018 | 1,46E-06 |
| GAA | 11,61538 | -2,19677 | 0,000547 |
| GABPA | 6,19377 | 1,97762 | 0,000112 |
| GAS5 | 4,82411 | 2,526088 | 1,75E-06 |
| GAS5-AS1 | 10,25070 | 2,002944 | 0,017427 |
| GAS7 | 10,05239 | 2,923743 | 7,06E-05 |
| GATA2-AS1 | 10,98049 | -2,18511 | 0,000343 |
| GBAS | 10,52812 | -2,06789 | 1,75E-05 |
| GCC1 | 9,84146 | 2,164202 | 2,08E-06 |
| GDF15 | 10,86683 | 2,574176 | 0,000331 |
| GDPD1 | 11,20553 | -2,93603 | 0,002199 |
| GEM | 9,52316 | 2,739041 | 0,000165 |
| GNAZ | 9,62912 | -2,2351 | 0,000889 |
| GNL2 | 6,74663 | 2,069317 | 2,67E-08 |
| GOLGA2P5 | 4,93035 | -3,16429 | 1,97E-06 |
| GPATCH3 | 6,88768 | 2,507336 | 1,46E-06 |
| GPBAR1 | 8,06099 | 2,000313 | 0,048341 |
| GPCPD1 | 10,80432 | 2,552185 | 7,97E-08 |
| GPR141 | 10,85685 | -3,65509 | 0,000127 |
| GPR153 | 10,23179 | -3,24937 | 1,09E-05 |
| GPR173 | 7,96600 | -2,49176 | 0,000356 |
| GPR83 | 10,48127 | 3,051169 | 0,000166 |
| GRHPR | 9,10989 | -1,9654 | 2,70E-07 |
| GRK7 | 11,45948 | 2,186835 | 0,015064 |
| GRM8 | 11,05058 | 1,982351 | 0,012378 |
| GRPEL2 | 7,49847 | 2,696636 | 7,28E-09 |
| GSTK1 | 8,68880 | -2,22273 | 8,44E-07 |
| GSTO2 | 10,74365 | -2,6296 | 0,000374 |
| GYS1 | 7,65278 | -2,81091 | 1,20E-13 |
| H6PD | 8,04321 | -3,20961 | 4,36E-09 |
| HAGH | 4,16159 | -3,56561 | 1,78E-12 |
| HAUS2 | 7,74121 | 2,653018 | 1,65E-10 |

| | | | |
|---|---|---|---|
| HAUS4 | 7,19635 | -1,95468 | 0,000453 |
| HAUS7 | 5,22856 | -2,05612 | 0,002204 |
| HCP5 | 8,40631 | -1,98649 | 0,001777 |
| HDAC6 | 10,90711 | -2,1375 | 3,81E-05 |
| HDHD2 | 9,87185 | -2,32336 | 9,36E-09 |
| HEATR5A | 10,29907 | -2,27463 | 3,32E-08 |
| HERPUD1 | 7,37158 | 1,956043 | 0,000108 |
| HEXDC | 6,77202 | -3,10232 | 2,64E-05 |
| HGSNAT | 11,25006 | -2,79606 | 9,55E-09 |
| HIC2 | 9,34624 | 2,054891 | 3,83E-05 |
| HINFP | 13,54991 | 2,158114 | 0,001046 |
| HIP1 | 7,62729 | -2,8308 | 2,44E-08 |
| HIST1H1A | 8,40681 | -1,9864 | 0,00082 |
| HIST1H2AB | 7,60702 | -1,97424 | 0,00403 |
| HIST1H2BH | 10,52266 | -2,01048 | 0,000204 |
| HIST1H2BM | 9,07408 | -2,02043 | 0,002683 |
| HIST1H3D | 11,24712 | -1,9401 | 5,78E-05 |
| HIST1H4B | 8,05006 | -1,97549 | 0,001136 |
| HIST1H4I | 12,91549 | -1,95503 | 0,000621 |
| HMGCL | 11,00921 | -1,95476 | 0,000274 |
| HMMR | 9,68035 | -2,25145 | 6,91E-05 |
| HNRNPU-AS1 | 10,30328 | 3,257911 | 2,62E-10 |
| HOOK1 | 8,36939 | -1,96173 | 0,000764 |
| HR | 10,08016 | -3,33249 | 3,55E-07 |
| HSD17B8 | 12,85770 | -2,29235 | 4,30E-05 |
| HSF4 | 4,45818 | -2,35269 | 0,000351 |
| HSPB6 | 9,24289 | -2,0686 | 0,001655 |
| HSPG2 | 5,48479 | -2,41959 | 1,66E-05 |
| HYKK | 12,06693 | -2,29271 | 0,003446 |
| IDUA | 5,37608 | -2,11985 | 0,002462 |
| IFI27L1 | 10,24528 | -2,1881 | 0,000186 |
| IFT43 | 6,55711 | -2,01527 | 0,000434 |
| IFT80 | 2,98686 | -1,9314 | 0,000111 |
| IKZF5 | 5,34681 | 2,380364 | 6,47E-08 |
| IL17RB | 8,72147 | -1,96192 | 0,008055 |
| IL1A | 11,46324 | 2,174298 | 0,001146 |
| IMPA2 | 10,38825 | -2,37809 | 6,87E-05 |
| INE1 | 12,50264 | 2,154445 | 0,011196 |
| INHBE | 8,93799 | 1,92939 | 0,024197 |
| INPP5A | 9,46126 | -1,92221 | 2,05E-05 |
| IQCA1 | 10,12911 | -2,5471 | 3,74E-05 |
| IQGAP2 | 10,27795 | -2,03209 | 0,004913 |
| IQSEC2 | 13,05132 | -2,19071 | 0,00084 |
| ISL2 | 12,95083 | 2,554548 | 8,57E-08 |

| | | | |
|---|---|---|---|
| ITGB4 | 11,50705 | -2,46893 | 3,87E-09 |
| ITGB8 | 9,31771 | -2,40462 | 8,57E-08 |
| ITPKA | 8,26796 | -1,96848 | 0,007494 |
| ITPRIP | 8,41417 | 2,163677 | 5,18E-05 |
| IVD | 12,14805 | -2,37218 | 6,26E-07 |
| JMJD6 | 9,96424 | 2,139389 | 4,91E-06 |
| JMY | 11,05780 | 1,982304 | 1,19E-07 |
| KANK2 | 11,33075 | -2,10646 | 1,83E-08 |
| KANSL1 | 10,40720 | 1,914625 | 2,44E-05 |
| KATNBL1 | 13,40570 | -2,20545 | 0,000373 |
| KBTBD8 | 10,69942 | 2,20519 | 0,005138 |
| KCNAB2 | 9,40873 | -2,18453 | 4,48E-06 |
| KCNJ14 | 9,74989 | 2,445764 | 0,000292 |
| KIAA0907 | 7,38115 | 2,209703 | 1,76E-06 |
| KIAA1841 | 8,61200 | -2,50679 | 2,13E-05 |
| KIFC2 | 9,79275 | -2,13546 | 0,00068 |
| KLF4 | 8,82822 | 2,826863 | 2,01E-12 |
| KLHL15 | 11,72453 | 2,504533 | 4,17E-06 |
| KLHL18 | 11,59751 | 2,691285 | 3,49E-08 |
| KLHL22 | 9,75784 | -3,04611 | 6,91E-05 |
| KLHL24 | 5,86964 | 1,896763 | 0,000144 |
| KLHL35 | 4,19836 | -2,70293 | 0,000216 |
| KLRC2 | 13,07417 | -3,67292 | 3,79E-06 |
| KLRC3 | 7,62316 | -2,74524 | 6,22E-06 |
| KMT2E-AS1 | 9,48773 | 2,122484 | 2,82E-06 |
| KREMEN1 | 10,66232 | -2,47952 | 0,000337 |
| KRT7 | 4,22945 | -2,55406 | 3,71E-11 |
| LAMA5 | 5,14578 | -2,09124 | 1,10E-05 |
| LAMTOR3 | 9,09937 | 1,995339 | 0,000434 |
| LATS2 | 9,58383 | 2,053393 | 2,49E-06 |
| LCAL1 | 4,71429 | 3,094572 | 0,003723 |
| LEPROT | 6,70147 | -2,27405 | 2,65E-07 |
| LGALS9B | 10,00876 | -2,44148 | 0,003606 |
| LGMN | 9,92364 | -1,96019 | 1,79E-07 |
| LHPP | 9,83038 | -2,0725 | 0,003791 |
| LINC00173 | 12,60829 | -2,09553 | 0,00116 |
| LINC00176 | 8,24266 | 2,090766 | 0,004219 |
| LINC00342 | 10,18021 | -2,38458 | 0,000161 |
| LINC00641 | 5,52317 | 3,877428 | 5,91E-13 |
| LINC01132 | 8,55602 | -2,11573 | 0,002696 |
| LINC01184 | 9,54755 | -1,93448 | 0,003492 |
| LINC01239 | 8,50212 | -2,19393 | 0,007159 |
| LINC01270 | 7,49966 | -2,29381 | 4,05E-05 |
| LINC01311 | 8,39150 | 1,923271 | 0,004239 |

| | | | |
|---|---|---|---|
| LINC01393 | 8,64262 | 2,326248 | 0,010492 |
| LMTK3 | 9,77364 | -3,1158 | 3,20E-05 |
| LOC100128288 | 7,52475 | 1,961996 | 0,035131 |
| LOC100130744 | 9,83204 | 2,087759 | 0,001116 |
| LOC100288069 | 7,53334 | 2,004668 | 0,001332 |
| LOC100506136 | 11,38160 | 2,612723 | 0,002202 |
| LOC100506314 | 11,58634 | -2,16802 | 0,00259 |
| LOC100506476 | 7,14581 | -1,96726 | 0,001939 |
| LOC100506548 | 8,53029 | 2,979778 | 6,76E-11 |
| LOC100506730 | 8,02596 | -2,27473 | 0,004155 |
| LOC100506834 | 8,84192 | -2,03647 | 0,002019 |
| LOC101060542 | 5,42878 | -2,57214 | 0,001344 |
| LOC101060553 | 9,19925 | 1,918276 | 0,017955 |
| LOC101409256 | 9,52597 | 2,079071 | 0,024514 |
| LOC101927811 | 7,87006 | -2,33846 | 0,000754 |
| LOC101928020 | 5,04176 | 2,131002 | 0,001988 |
| LOC101928034 | 8,98913 | -2,3318 | 0,00131 |
| LOC101928979 | 3,13069 | 2,119877 | 0,017071 |
| LOC150776 | 10,59616 | 1,893759 | 0,000429 |
| LOC283140 | 9,33545 | 3,250971 | 0,000616 |
| LOC283710 | 7,00827 | -2,04399 | 0,008004 |
| LOC284009 | 6,72160 | 2,059992 | 0,018291 |
| LOC388813 | 9,35832 | 2,928945 | 0,000491 |
| LOC652276 | 8,99801 | 2,363254 | 2,55E-05 |
| LOC729603 | 9,03833 | 2,428485 | 0,000658 |
| LPCAT3 | 7,74194 | -2,3775 | 9,02E-10 |
| LRRC61 | 8,33268 | -3,51727 | 5,55E-12 |
| LRRCC1 | 9,74140 | -2,42888 | 9,24E-07 |
| LRRK1 | 13,50653 | -2,05673 | 0,001081 |
| LRRK2 | 9,63237 | -2,16817 | 0,011617 |
| LSG1 | 8,62039 | 2,007475 | 8,06E-07 |
| LSM3 | 10,37214 | 1,981282 | 5,28E-07 |
| LSMEM1 | 9,36443 | 3,505094 | 2,89E-10 |
| LTB4R | 9,27486 | 2,684917 | 9,18E-06 |
| LTBP2 | 10,30517 | -2,49618 | 4,59E-08 |
| LTBP4 | 7,97154 | -2,12138 | 0,000418 |
| LUZP1 | 9,07624 | 1,984644 | 1,24E-05 |
| LVRN | 6,34255 | 2,694504 | 0,000334 |
| LYAR | 11,33601 | 2,15027 | 3,43E-05 |
| LYRM7 | 7,01055 | -3,013 | 1,47E-06 |
| LYSMD4 | 7,27554 | -2,17541 | 0,000601 |
| LZTFL1 | 8,59470 | 2,007552 | 6,45E-05 |
| MAFF | 9,31694 | 2,618271 | 1,10E-06 |
| MAFK | 8,04567 | 2,40576 | 1,46E-05 |

| | | | |
|---|---|---|---|
| MAN1A1 | 10,85608 | -2,36768 | 1,65E-09 |
| MAP1LC3B | 8,59706 | 2,30804 | 2,73E-05 |
| MAP4K2 | 8,24888 | -2,16899 | 0,000151 |
| MAPT | 10,32459 | -2,63678 | 0,000151 |
| MARK4 | 7,86904 | -2,22741 | 3,12E-07 |
| MAST3 | 5,40105 | -2,24371 | 1,09E-07 |
| MAVS | 7,75316 | -2,0827 | 1,59E-06 |
| MBD1 | 10,65609 | 1,899596 | 1,19E-05 |
| MCF2L | 7,01537 | -1,96602 | 0,004302 |
| MED10 | 9,86824 | 2,246591 | 3,77E-07 |
| MED17 | 11,40309 | 1,969748 | 8,81E-06 |
| MED26 | 11,41453 | 2,304839 | 6,39E-07 |
| MED6 | 7,05230 | 2,098672 | 1,78E-05 |
| MEIS3 | 11,73089 | -2,85609 | 1,36E-06 |
| MET | 13,80528 | 2,086254 | 2,26E-06 |
| METTL16 | 4,76314 | 2,007599 | 2,73E-05 |
| METTL7A | 9,04391 | -2,94959 | 2,61E-08 |
| MFI2 | 9,05694 | -1,9748 | 0,002119 |
| MGC16142 | 4,19276 | 2,704887 | 0,000613 |
| MGC70870 | 7,46449 | -1,96245 | 0,008022 |
| MICAL2 | 10,87408 | -1,98834 | 0,000193 |
| MIF4GD | 6,52336 | -1,99441 | 0,000473 |
| MIR1204 | 8,16186 | 1,937322 | 0,026023 |
| MIR1292 | 9,58994 | 3,407257 | 0,00086 |
| MIR1304 | 9,94697 | 2,385062 | 0,001689 |
| MIR210HG | 6,73643 | -3,38214 | 1,99E-05 |
| MIR3175 | 3,04684 | 1,976687 | 0,021944 |
| MIR5087 | 10,81935 | 2,03087 | 0,005389 |
| MIR6850 | 3,30105 | 2,46009 | 0,012517 |
| MIR8072 | 4,37036 | 2,37944 | 0,000498 |
| MKRN2 | 7,45825 | 1,967348 | 2,18E-06 |
| MKRN3 | 10,46633 | 2,28607 | 1,78E-05 |
| MLXIPL | 9,57240 | -2,90918 | 1,55E-05 |
| MMAB | 10,02976 | -1,96286 | 0,000662 |
| MMD | 10,92623 | -2,17168 | 0,000114 |
| MMP15 | 9,51871 | -1,99095 | 0,001421 |
| MORN1 | 6,65881 | -2,6557 | 5,39E-06 |
| MORN4 | 8,45159 | -2,07842 | 0,002772 |
| MPHOSPH10 | 8,57454 | 2,538414 | 4,67E-09 |
| MPI | 13,20339 | -3,17522 | 1,06E-10 |
| MPP7 | 4,85385 | -2,11095 | 0,001464 |
| MR1 | 6,71137 | -2,34298 | 0,000373 |
| MRPL32 | 9,91266 | 1,95254 | 0,000103 |
| MRPS24 | 10,89893 | 1,970498 | 0,008648 |

| | | | |
|---|---|---|---|
| MSC-AS1 | 12,50104 | -3,06114 | 5,05E-06 |
| MTF1 | 7,24531 | 2,423746 | 8,17E-07 |
| MTHFD2L | 6,99460 | 3,022282 | 4,54E-12 |
| MTMR9LP | 7,67304 | -2,17183 | 0,002766 |
| MTOR | 7,09428 | -2,06126 | 7,65E-09 |
| MTPAP | 9,80232 | 2,026722 | 1,18E-06 |
| MUC16 | 8,10702 | -4,0656 | 2,84E-20 |
| MUC5B | 10,08890 | -3,72364 | 5,22E-08 |
| MURC | 10,27677 | 1,962194 | 0,013763 |
| MVP | 8,46129 | -2,0751 | 4,92E-05 |
| MXD1 | 3,62996 | 2,446732 | 6,72E-05 |
| MYBPC2 | 6,81786 | -3,52638 | 0,000323 |
| MYC | 5,47507 | 2,030755 | 0,000192 |
| MYO16 | 4,72764 | -2,10306 | 7,09E-05 |
| MYO18A | 9,68707 | -2,04899 | 1,12E-05 |
| MYO5C | 3,22915 | -2,09935 | 0,000123 |
| MYO9A | 3,76932 | -1,94016 | 4,57E-07 |
| MYOM3 | 5,61167 | -2,69973 | 2,63E-05 |
| NAA16 | 7,91445 | 1,931655 | 3,78E-07 |
| NAPA-AS1 | 10,37941 | -2,05245 | 0,006582 |
| NAT14 | 11,64560 | -2,33514 | 3,20E-05 |
| NBEAL2 | 6,31554 | -3,05176 | 3,17E-08 |
| NCBP2 | 4,81149 | 2,09445 | 1,25E-06 |
| NDUFAF4 | 9,08980 | 2,085333 | 8,28E-05 |
| NDUFAF5 | 7,07765 | 2,525147 | 3,89E-07 |
| NECAB3 | 8,30598 | -2,44996 | 4,60E-08 |
| NFATC2 | 11,62633 | 3,088584 | 1,08E-06 |
| NFE2 | 11,22491 | -3,37938 | 6,43E-08 |
| NFYA | 6,59009 | 1,9145 | 8,23E-05 |
| NGDN | 9,34437 | 2,345328 | 1,46E-06 |
| NGFR | 7,09653 | 2,292898 | 0,030581 |
| NGRN | 9,87614 | 2,152007 | 3,36E-09 |
| NICN1 | 9,16365 | -2,66682 | 4,27E-05 |
| NKD1 | 7,56801 | -2,7206 | 0,000223 |
| NOCT | 7,94031 | 1,96486 | 0,000299 |
| NPTN-IT1 | 9,66674 | 2,389446 | 0,001773 |
| NQO2 | 11,71508 | -2,08409 | 1,86E-05 |
| NR1D1 | 11,46021 | 2,971191 | 1,94E-08 |
| NR4A2 | 10,73953 | 2,635148 | 1,58E-07 |
| NRBF2 | 11,31985 | 1,949826 | 2,22E-05 |
| NRBP2 | 10,59068 | -2,33828 | 1,16E-05 |
| NRIP3 | 7,72962 | -2,67032 | 0,000106 |
| NRSN2 | 14,53614 | -2,05687 | 2,28E-07 |
| NTS | 12,70191 | -3,20476 | 5,68E-06 |

| | | | |
|---|---|---|---|
| NUDT17 | 10,17566 | 2,331018 | 0,003193 |
| NUP50 | 2,85562 | 2,002713 | 7,80E-07 |
| NUPR1 | 11,74257 | 3,055806 | 1,63E-05 |
| OAS3 | 10,89409 | -2,39506 | 7,44E-07 |
| OBSCN | 9,63358 | -2,35358 | 2,07E-07 |
| OCEL1 | 10,83013 | -2,45212 | 3,82E-05 |
| ODC1 | 13,45463 | 2,430623 | 2,36E-06 |
| OGG1 | 7,94916 | -2,03954 | 0,000123 |
| OLFML2A | 7,50166 | -3,21033 | 1,78E-10 |
| OPLAH | 9,83476 | -2,40662 | 0,000569 |
| ORAI3 | 7,81793 | -2,5395 | 2,26E-05 |
| OSBPL5 | 10,91955 | -2,45063 | 7,67E-08 |
| OSER1 | 10,65456 | 3,114096 | 1,72E-12 |
| OTUD3 | 7,25938 | 2,256779 | 5,05E-06 |
| OVGP1 | 8,64060 | 7,173226 | 9,16E-12 |
| P2RX4 | 11,72774 | -2,48617 | 0,004256 |
| P4HTM | 5,70556 | -2,55795 | 1,50E-07 |
| PAAF1 | 11,14714 | -2,63779 | 4,12E-08 |
| PAK3 | 10,35082 | -2,37278 | 0,000847 |
| PALM | 7,04628 | -2,15106 | 0,001301 |
| PAN3-AS1 | 9,81333 | 1,967938 | 0,004873 |
| PAPLN | 10,59494 | -2,92979 | 8,92E-06 |
| PARD6A | 10,74446 | -1,96268 | 0,009355 |
| PC | 9,26811 | -2,57428 | 5,67E-08 |
| PCDHB15 | 10,58067 | 2,552342 | 5,18E-05 |
| PCED1B | 9,61950 | -2,09267 | 0,006495 |
| PCF11 | 9,04857 | 2,272871 | 8,28E-05 |
| PCGF1 | 9,93573 | 1,965274 | 4,57E-05 |
| PCYOX1L | 10,34081 | -2,28781 | 0,000531 |
| PDE8B | 4,29487 | -2,7129 | 5,53E-07 |
| PDRG1 | 6,22797 | 2,968175 | 9,36E-09 |
| PEA15 | 9,16408 | 2,094328 | 3,22E-07 |
| PECAM1 | 9,06166 | -1,92533 | 0,009364 |
| PELI3 | 8,61739 | -2,53275 | 3,81E-05 |
| PELO | 6,34971 | 1,909895 | 6,04E-06 |
| PEX1 | 9,28386 | -2,73035 | 1,01E-06 |
| PEX10 | 10,53999 | -2,12023 | 0,000192 |
| PGLS | 11,43251 | -2,42158 | 1,01E-08 |
| PHKA2 | 10,04865 | -1,99359 | 5,52E-05 |
| PHKG1 | 11,69092 | 5,383371 | 1,78E-12 |
| PI4K2B | 9,89776 | -2,14676 | 0,000179 |
| PIK3CD | 12,59911 | -2,15193 | 0,001347 |
| PIN1 | 11,26333 | -2,48377 | 9,01E-08 |
| PIR | 7,55652 | -2,09211 | 0,000902 |

| | | | |
|---|---|---|---|
| PITPNM3 | 7,95069 | -2,80171 | 7,15E-08 |
| PLAU | 9,82504 | -2,31859 | 0,00045 |
| PLCB3 | 9,09042 | -1,93043 | 0,000234 |
| PLD2 | 11,10546 | -2,32437 | 0,000199 |
| PLEKHH1 | 10,93196 | -3,70497 | 2,99E-06 |
| PLEKHJ1 | 9,72926 | -1,93077 | 7,37E-05 |
| PLGRKT | 7,12168 | -2,34997 | 0,00045 |
| PLLP | 8,21669 | -2,45365 | 0,000385 |
| PMAIP1 | 7,70552 | 2,049596 | 7,26E-05 |
| PMM1 | 9,53072 | -2,32549 | 1,62E-05 |
| PMS2CL | 9,43046 | 2,102322 | 0,000598 |
| PMS2P3 | 8,63336 | 2,092935 | 0,000751 |
| PNMA2 | 7,46216 | -3,76135 | 2,06E-11 |
| PNO1 | 8,82328 | 2,230254 | 8,85E-08 |
| POLR1C | 9,28911 | 2,46463 | 2,11E-08 |
| POLR2K | 9,79928 | 1,892724 | 8,88E-08 |
| POLR3C | 11,00056 | 1,91917 | 0,000119 |
| POLR3GL | 13,67022 | -3,013 | 2,48E-06 |
| POT1-AS1 | 4,84141 | -2,02632 | 0,00591 |
| POU2F2 | 13,54316 | 2,373118 | 0,007368 |
| POU6F1 | 7,31990 | -2,52748 | 0,000294 |
| PPFIA4 | 6,16073 | -1,96689 | 0,013893 |
| PPM1H | 8,73439 | -2,25137 | 2,34E-07 |
| PPOX | 8,56698 | -2,12807 | 0,001116 |
| PPP1R15A | 11,92156 | 2,374117 | 1,78E-06 |
| PPP1R26-AS1 | 13,13544 | -2,04728 | 0,003709 |
| PPP2R2B | 5,54081 | -2,04235 | 0,012378 |
| PPP2R3A | 9,54017 | -2,01861 | 0,002626 |
| PPP4R3A | 11,43184 | 2,296983 | 2,11E-10 |
| PPTC7 | 10,78498 | 2,268044 | 1,04E-06 |
| PREX1 | 11,18079 | -1,93434 | 0,000242 |
| PRICKLE4 | 9,79006 | -2,26006 | 0,000196 |
| PRKAB1 | 9,76058 | 2,370373 | 1,76E-06 |
| PRORSD1P | 11,64283 | -1,95127 | 0,012436 |
| PRPF3 | 8,81092 | 2,856273 | 2,16E-13 |
| PRPF38B | 5,07890 | 2,084281 | 1,25E-07 |
| PRPF39 | 11,28412 | 2,485588 | 6,92E-12 |
| PRTFDC1 | 8,17898 | -2,22669 | 0,001827 |
| PSKH1 | 9,13081 | -2,96205 | 1,74E-10 |
| PSMC1 | 5,48381 | 2,072001 | 1,46E-06 |
| PTCH2 | 8,77398 | 2,763961 | 0,000127 |
| PTP4A3 | 10,16928 | -2,14131 | 0,004349 |
| PTPN18 | 10,42625 | -2,75014 | 1,05E-07 |
| PVT1 | 10,31765 | 2,403486 | 2,64E-08 |

| | | | |
|---|---|---|---|
| PXMP4 | 10,33310 | -2,90485 | 2,83E-08 |
| PYCRL | 3,98184 | -2,20643 | 0,000872 |
| PYGL | 6,30270 | -2,17008 | 2,65E-09 |
| PYGM | 9,54253 | 1,901255 | 0,003638 |
| PYROXD2 | 9,99966 | -2,20033 | 4,97E-05 |
| R3HDM2 | 10,24342 | -2,20297 | 1,55E-05 |
| RAB26 | 10,28134 | -3,337 | 5,31E-06 |
| RAB3D | 12,12973 | -2,53499 | 6,85E-10 |
| RABGGTB | 12,85336 | 2,285483 | 1,15E-08 |
| RAD51B | 8,28533 | -2,36783 | 0,000484 |
| RASL11A | 9,58616 | 3,405873 | 2,96E-05 |
| RBAK | 10,16591 | 2,179122 | 0,00011 |
| RBM15 | 10,74956 | 1,97183 | 7,14E-07 |
| RBM22 | 10,06305 | 2,36135 | 4,23E-09 |
| RBM33 | 9,80057 | 1,987304 | 1,31E-07 |
| RBM48 | 7,96334 | 1,951524 | 0,000365 |
| RBPJ | 9,05344 | 1,940935 | 2,61E-08 |
| RCHY1 | 9,77954 | 2,891261 | 1,50E-10 |
| RECK | 6,73086 | -1,96125 | 0,00468 |
| RGL3 | 5,74542 | -2,49425 | 0,000219 |
| RGS16 | 11,67380 | 3,743069 | 2,72E-09 |
| RHPN1 | 6,39046 | -2,53643 | 0,000349 |
| RIBC2 | 7,73244 | -2,62579 | 0,000146 |
| RIOK3 | 10,42423 | 1,934124 | 0,002271 |
| RLF | 10,18604 | 2,941149 | 5,88E-15 |
| RMDN1 | 10,24878 | -2,5145 | 1,90E-08 |
| RMI2 | 8,40602 | -2,80521 | 3,54E-05 |
| RN7SK | 10,83464 | 2,009378 | 5,58E-07 |
| RNA28S5 | 12,15169 | 2,019109 | 0,000502 |
| RND3 | 9,59576 | 2,404058 | 4,98E-06 |
| RNF170 | 9,61538 | -2,186 | 0,000505 |
| RNF213 | 6,44584 | -2,05336 | 3,22E-07 |
| RNF219 | 7,50007 | 2,319071 | 2,88E-07 |
| RNF6 | 6,72506 | 2,016932 | 1,75E-05 |
| RNMT | 10,64172 | 2,335255 | 9,72E-10 |
| RNU11 | 8,99087 | 3,079103 | 7,92E-09 |
| RNU12 | 8,94321 | 3,324118 | 9,55E-11 |
| RNU2-2P | 9,49797 | 3,568267 | 5,10E-14 |
| RNU4-1 | 13,25583 | 3,260842 | 1,44E-07 |
| RNU4-2 | 6,11456 | 3,807087 | 1,12E-09 |
| RNU4ATAC | 12,05101 | 3,478941 | 1,55E-08 |
| RNU5A-1 | 10,97112 | 2,97757 | 0,000165 |
| RNU5B-1 | 11,39101 | 3,390932 | 5,47E-05 |
| RNU5D-1 | 9,99199 | 2,177691 | 0,003323 |

| | | | |
|---|---|---|---|
| RNU5E-1 | 7,19555 | 3,107408 | 3,93E-06 |
| RNU5F-1 | 3,20959 | 1,914513 | 0,006571 |
| RNU6ATAC | 9,48421 | 2,0223 | 0,013919 |
| RNVU1-14 | 9,74778 | 3,201606 | 5,11E-08 |
| RNVU1-15 | 9,69922 | 3,682032 | 1,73E-08 |
| RNVU1-6 | 8,77096 | 3,021397 | 7,05E-06 |
| RNVU1-7 | 9,27824 | 2,857263 | 0,00014 |
| RNVU1-8 | 9,21815 | 2,353742 | 0,002739 |
| RNY1 | 8,87875 | 2,224078 | 0,000172 |
| ROGDI | 10,04808 | -3,04038 | 1,75E-08 |
| RP9 | 13,32653 | 1,967298 | 0,000238 |
| RP9P | 3,59570 | 2,659653 | 2,35E-05 |
| RPA4 | 8,44044 | 2,417249 | 0,014298 |
| RPF1 | 7,80263 | 2,228039 | 1,22E-08 |
| RPP40 | 11,78369 | 2,465031 | 1,79E-05 |
| RPS16P5 | 9,25490 | 2,447072 | 0,002651 |
| RPS6KA1 | 11,48063 | -2,22261 | 3,11E-07 |
| RPS6KA2 | 11,77307 | -3,0151 | 1,33E-10 |
| RRH | 12,14526 | 4,023227 | 4,47E-05 |
| RRN3 | 14,22533 | 2,370398 | 1,79E-10 |
| RRN3P1 | 4,64116 | 3,274105 | 3,55E-07 |
| RSL1D1 | 8,69748 | 1,989065 | 6,05E-08 |
| RSRP1 | 13,44783 | 2,918381 | 5,67E-11 |
| RYBP | 3,68039 | 2,266526 | 3,12E-07 |
| S100A3 | 6,92693 | -3,04196 | 1,93E-05 |
| SAP18 | 7,58911 | 2,119788 | 5,35E-07 |
| SARS2 | 8,28479 | -2,11937 | 7,05E-06 |
| SBDSP1 | 10,24058 | 2,481644 | 1,19E-07 |
| SBNO1 | 7,83251 | 2,195117 | 2,02E-07 |
| SCAF4 | 12,07366 | 3,059528 | 1,55E-13 |
| SCARNA15 | 8,76157 | 2,479767 | 0,009856 |
| SCARNA18 | 5,36533 | 2,154167 | 0,014096 |
| SCARNA27 | 8,43972 | 1,896268 | 0,040625 |
| SCN1A | 11,79267 | -2,054 | 0,005261 |
| SCN1B | 10,00637 | -3,38273 | 1,44E-07 |
| SCN9A | 8,98389 | -2,24735 | 0,003512 |
| SCNM1 | 8,70160 | 2,283794 | 6,96E-05 |
| SCNN1A | 9,21663 | -4,42531 | 3,40E-15 |
| SDE2 | 9,75816 | 2,688412 | 1,86E-10 |
| SEC14L1P1 | 8,62953 | 3,611193 | 1,25E-06 |
| SELENBP1 | 6,17872 | -3,14612 | 2,61E-09 |
| SEMA3D | 9,26149 | -2,31571 | 3,64E-05 |
| SEMA4B | 8,58578 | -2,03997 | 9,94E-05 |
| SEPT1 | 11,26124 | -2,01081 | 0,011023 |

| | | | |
|---|---|---|---|
| SERPINB4 | 9,20942 | -2,03283 | 0,008466 |
| SERPINB9P1 | 11,96547 | -2,54997 | 7,41E-05 |
| SESN2 | 9,74456 | 2,81812 | 3,45E-06 |
| SF3A3 | 5,96199 | -2,89289 | 5,88E-15 |
| SF3B4 | 8,37459 | 1,997518 | 1,50E-06 |
| SFXN5 | 13,30927 | -1,94732 | 0,00078 |
| SH3D21 | 8,48490 | -2,97629 | 3,51E-05 |
| SH3GLB2 | 11,54232 | -1,96892 | 0,000719 |
| SH3PXD2B | 11,72573 | -2,12639 | 0,000331 |
| SHMT1 | 10,07922 | -2,33567 | 3,89E-08 |
| SIRT1 | 8,49366 | 2,004806 | 5,74E-07 |
| SLC16A10 | 9,23672 | -2,50362 | 0,005164 |
| SLC16A2 | 8,28248 | -1,93658 | 0,001295 |
| SLC16A5 | 8,72563 | -2,36251 | 8,84E-06 |
| SLC1A1 | 9,62679 | -1,92707 | 0,003077 |
| SLC22A13 | 12,97533 | 1,987445 | 0,011381 |
| SLC22A18 | 6,94456 | -2,27646 | 4,35E-05 |
| SLC23A2 | 8,50451 | -2,40487 | 0,000188 |
| SLC23A3 | 12,25394 | 2,614917 | 0,000829 |
| SLC25A10 | 9,71356 | -2,14406 | 0,000943 |
| SLC25A20 | 11,54682 | -2,34603 | 0,000116 |
| SLC25A25 | 7,88704 | 2,463423 | 1,09E-08 |
| SLC25A29 | 9,25884 | -3,3985 | 1,99E-07 |
| SLC25A33 | 9,97846 | 1,983844 | 0,003262 |
| SLC25A42 | 7,73623 | -2,67685 | 0,000615 |
| SLC25A43 | 10,99404 | -2,07881 | 0,001187 |
| SLC27A1 | 10,85136 | -2,61561 | 0,000131 |
| SLC2A12 | 8,88357 | -2,84357 | 2,71E-05 |
| SLC30A3 | 7,06376 | -1,93194 | 0,00878 |
| SLC30A4 | 6,99825 | -2,09498 | 0,005324 |
| SLC35E2B | 9,03679 | -2,0273 | 3,59E-06 |
| SLC37A1 | 10,36136 | -2,1564 | 0,000244 |
| SLC3A2 | 8,94559 | 2,344422 | 1,65E-10 |
| SLC7A5P1 | 4,35373 | 2,533266 | 0,000117 |
| SLC7A5P2 | 2,43778 | 2,421124 | 1,46E-06 |
| SLC9A8 | 3,62769 | -2,06511 | 0,000164 |
| SLCO1C1 | 5,06507 | 2,703909 | 4,00E-05 |
| SLCO2A1 | 12,42447 | -1,9622 | 0,002804 |
| SLFN5 | 11,23049 | -2,32572 | 1,97E-05 |
| SMCR8 | 10,75684 | 1,998122 | 4,85E-05 |
| SMG1P1 | 10,62696 | 2,374659 | 0,000115 |
| SMG1P2 | 10,86019 | 2,53561 | 0,000759 |
| SMG7 | 10,07598 | 1,934054 | 9,42E-06 |
| SMIM2-AS1 | 7,63804 | -2,07781 | 0,001052 |

| | | | |
|---|---|---|---|
| SNAI1 | 9,81126 | 2,405004 | 4,86E-07 |
| SNAP25 | 8,93567 | -1,98481 | 0,00013 |
| SNAPC1 | 7,64276 | 2,280461 | 3,79E-06 |
| SNHG1 | 10,57862 | 2,950777 | 3,59E-10 |
| SNHG10 | 11,18972 | 2,634654 | 1,04E-06 |
| SNHG12 | 10,35867 | 3,169403 | 2,07E-10 |
| SNHG15 | 4,51371 | 2,609815 | 1,81E-09 |
| SNHG16 | 8,82657 | 2,423369 | 2,98E-06 |
| SNHG17 | 10,67420 | 2,451438 | 5,44E-09 |
| SNHG20 | 8,78415 | 2,187495 | 0,000446 |
| SNHG21 | 11,01182 | 3,378012 | 1,25E-08 |
| SNHG3 | 6,61720 | 1,9984 | 2,39E-05 |
| SNHG4 | 8,51918 | 2,507001 | 0,000102 |
| SNHG5 | 11,68976 | 3,02112 | 1,78E-12 |
| SNHG7 | 8,48464 | 2,801219 | 2,34E-10 |
| SNIP1 | 10,33242 | 2,348149 | 2,92E-06 |
| SNORA1 | 10,53065 | 3,019296 | 1,52E-08 |
| SNORA14B | 9,80105 | 2,138523 | 0,007481 |
| SNORA16A | 8,91889 | 3,203856 | 2,08E-06 |
| SNORA18 | 9,65208 | 2,780814 | 3,22E-07 |
| SNORA27 | 6,77952 | 2,699084 | 3,10E-07 |
| SNORA32 | 8,61926 | 3,530396 | 7,18E-12 |
| SNORA4 | 7,70098 | 2,34091 | 0,00026 |
| SNORA40 | 6,80055 | 2,764569 | 1,34E-07 |
| SNORA44 | 10,24707 | 3,763682 | 4,40E-07 |
| SNORA45A | 10,55821 | 3,477396 | 1,55E-05 |
| SNORA45B | 7,90742 | 2,423975 | 0,001332 |
| SNORA52 | 6,76602 | 2,409147 | 0,000908 |
| SNORA55 | 9,06967 | 2,271167 | 0,011263 |
| SNORA5C | 7,13157 | 1,985033 | 0,002983 |
| SNORA6 | 11,15970 | 2,206924 | 0,00415 |
| SNORA61 | 8,69220 | 2,21452 | 0,001633 |
| SNORA63 | 7,56456 | 2,96494 | 9,02E-08 |
| SNORA67 | 4,81833 | 2,792487 | 9,93E-08 |
| SNORA68 | 9,73804 | 4,736848 | 5,19E-24 |
| SNORA70 | 10,23983 | 2,598552 | 1,33E-06 |
| SNORA71C | 9,91814 | 2,266073 | 0,013042 |
| SNORA74B | 7,63929 | 1,917594 | 0,020648 |
| SNORA7B | 9,39217 | 2,098253 | 0,010603 |
| SNORA8 | 9,33348 | 2,407615 | 4,21E-05 |
| SNORA80B | 9,23477 | 2,588448 | 0,001851 |
| SNORD100 | 9,98642 | 2,897952 | 0,00084 |
| SNORD101 | 8,02345 | 2,087403 | 0,010469 |
| SNORD102 | 6,90731 | 2,994475 | 5,55E-09 |
| SNORD11B | 7,44011 | 2,554283 | 0,001976 |
| SNORD13 | 10,96820 | 3,05685 | 0,000109 |
| SNORD14B | 10,34197 | 2,473846 | 0,001233 |
| SNORD18C | 8,26581 | 2,000043 | 0,009294 |
| SNORD19 | 8,04342 | 2,177489 | 0,008342 |
| SNORD2 | 10,30249 | 2,41038 | 0,000107 |
| SNORD25 | 10,57339 | 2,504628 | 8,28E-05 |
| SNORD26 | 10,59188 | 2,164043 | 4,46E-05 |
| SNORD29 | 8,40872 | 1,929911 | 0,000753 |
| SNORD30 | 7,99618 | 2,554672 | 0,000152 |
| SNORD34 | 10,10564 | 3,240768 | 2,09E-05 |
| SNORD35A | 9,45885 | 3,17746 | 1,96E-07 |
| SNORD35B | 9,62522 | 2,214929 | 0,001577 |
| SNORD38B | 8,42095 | 3,70515 | 4,53E-08 |
| SNORD45C | 8,91944 | 2,476506 | 0,000548 |
| SNORD5 | 11,21195 | 2,89996 | 1,13E-06 |
| SNORD50A | 9,92961 | 2,339569 | 0,004605 |
| SNORD50B | 6,14540 | 2,32623 | 0,006407 |
| SNORD51 | 11,56026 | 2,394799 | 0,023495 |
| SNORD55 | 10,15907 | 3,589485 | 9,11E-09 |
| SNORD56 | 11,27879 | 2,818958 | 0,002548 |
| SNORD57 | 13,11675 | 2,125497 | 0,027796 |
| SNORD6 | 11,60726 | 3,301455 | 8,82E-06 |
| SNORD63 | 11,30146 | 2,498387 | 0,004671 |
| SNORD81 | 13,09281 | 1,922277 | 0,007752 |
| SNORD82 | 10,85937 | 1,912599 | 0,048924 |
| SNORD84 | 9,56247 | 2,496462 | 0,00519 |
| SNORD99 | 9,30172 | 2,015977 | 0,036049 |
| SNRPB | 10,43755 | 2,094906 | 6,24E-07 |
| SOCS1 | 10,14422 | 1,985716 | 0,014971 |
| SOCS5 | 11,90410 | 2,03725 | 8,17E-07 |
| SPATA2 | 8,04571 | 2,206621 | 3,89E-07 |
| SPATA24 | 7,72288 | -2,60611 | 0,002819 |
| SPEG | 9,23691 | -2,07307 | 0,001587 |
| SPEN | 9,16988 | 1,904288 | 9,18E-07 |
| SPINK13 | 8,93090 | -2,39148 | 3,51E-05 |
| SPIRE2 | 9,87701 | -2,20404 | 0,000531 |
| SPPL2B | 10,88033 | -2,96613 | 4,19E-10 |
| SPTLC3 | 10,00774 | -2,19148 | 0,000401 |
| SPTY2D1 | 6,83923 | 2,014552 | 0,000303 |
| SRGAP2 | 5,17791 | -1,96184 | 8,24E-05 |
| SRRM1 | 10,18707 | 2,176176 | 1,51E-08 |
| SRRM5 | 9,68918 | 2,943155 | 1,28E-07 |
| SSBP3 | 8,76392 | -2,14105 | 3,33E-08 |

| | | | |
|---|---|---|---|
| SSPN | 8,60846 | 3,042873 | 0,000605 |
| SSSCA1-AS1 | 9,44593 | 2,663029 | 0,000356 |
| ST18 | 8,02508 | -2,00395 | 0,002378 |
| ST3GAL3 | 9,23525 | -3,97859 | 4,66E-05 |
| ST3GAL5 | 8,78369 | -2,21617 | 0,000907 |
| STC2 | 4,23199 | 2,007175 | 4,64E-06 |
| STEAP4 | 6,63424 | -2,24149 | 0,006262 |
| STRA6 | 6,62637 | -3,37496 | 1,39E-13 |
| STX11 | 6,99763 | 1,90146 | 0,01648 |
| STX3 | 12,28953 | 2,163307 | 1,24E-07 |
| STXBP5 | 8,48287 | -2,36799 | 6,00E-07 |
| STXBP5-AS1 | 9,73668 | -1,99171 | 0,005498 |
| SUOX | 11,64820 | -2,34172 | 3,46E-05 |
| SURF1 | 10,03333 | -2,69492 | 2,43E-06 |
| SYF2 | 10,40293 | 2,200385 | 5,79E-05 |
| SYNC | 12,86782 | -2,35516 | 4,73E-05 |
| SYTL2 | 9,79593 | 2,773281 | 0,00045 |
| TAB1 | 12,28739 | -2,81206 | 6,23E-10 |
| TAF1A | 9,60079 | 2,560798 | 1,77E-05 |
| TAF1D | 8,34451 | 2,288293 | 3,70E-08 |
| TAF9B | 11,90483 | -2,67016 | 1,61E-07 |
| TAGLN | 6,85616 | 2,507162 | 2,09E-05 |
| TANGO2 | 9,23473 | -2,06243 | 0,000641 |
| TARBP1 | 8,32926 | -2,20681 | 3,41E-05 |
| TAS2R30 | 6,95455 | 1,900873 | 0,032364 |
| TAZ | 10,36869 | -2,11973 | 0,001384 |
| TBC1D16 | 4,29062 | -2,46173 | 0,000132 |
| TBR1 | 6,71347 | 2,574676 | 0,003064 |
| TBX19 | 10,01718 | 2,948444 | 7,38E-05 |
| TCEA2 | 10,53682 | -2,85155 | 4,49E-10 |
| TCEAL3 | 2,56165 | -1,95605 | 0,000943 |
| TCTE3 | 7,96062 | 2,19828 | 0,006769 |
| TECPR1 | 13,03963 | -2,00424 | 7,37E-05 |
| TECPR2 | 11,03696 | -2,52852 | 5,49E-06 |
| TFAP2A | 3,05203 | -2,29536 | 5,00E-08 |
| TFB2M | 10,80482 | 2,644103 | 2,30E-09 |
| TFIP11 | 9,27222 | 2,508708 | 1,18E-07 |
| TGFBR3L | 11,62075 | -2,35602 | 0,00031 |
| TGM2 | 9,04017 | -2,15852 | 3,79E-06 |
| THAP1 | 9,64334 | 2,037373 | 0,000505 |
| THUMPD3-AS1 | 7,71171 | 1,934743 | 0,000848 |
| TK2 | 8,35475 | -3,40937 | 4,01E-08 |
| TLCD2 | 11,31753 | -2,32152 | 2,14E-05 |
| TLE2 | 11,61645 | -2,4451 | 0,001198 |

| | | | |
|---|---|---|---|
| TM7SF2 | 8,24136 | -2,08319 | 0,001988 |
| TMCC1-AS1 | 6,78922 | -1,92908 | 0,008375 |
| TMEM106B | 11,33988 | -2,36372 | 2,95E-08 |
| TMEM143 | 11,53033 | -3,19099 | 1,25E-07 |
| TMEM175 | 5,61235 | -2,20897 | 0,00023 |
| TMEM218 | 8,60171 | -1,92656 | 0,002358 |
| TMEM242 | 3,60315 | 2,019906 | 0,000101 |
| TMEM37 | 11,41458 | -2,6693 | 0,000121 |
| TMEM44 | 7,81476 | -2,75345 | 3,32E-06 |
| TMEM52B | 11,96444 | -2,28129 | 0,000476 |
| TMEM53 | 9,60315 | -3,01238 | 7,38E-05 |
| TMEM63A | 11,52597 | -1,97062 | 0,000143 |
| TMEM81 | 10,98005 | 2,373536 | 0,00018 |
| TMOD2 | 9,02557 | -2,08033 | 0,003261 |
| TMTC4 | 5,18793 | -2,42537 | 1,16E-05 |
| TNS2 | 12,03989 | -2,69005 | 3,94E-14 |
| TNS4 | 8,81948 | -2,09397 | 7,15E-07 |
| TOM1L2 | 8,90842 | -3,19917 | 5,94E-09 |
| TOMM20 | 7,28984 | 2,562767 | 3,59E-10 |
| TOX2 | 9,37737 | -1,98417 | 0,002014 |
| TP53I11 | 5,89625 | -2,07473 | 0,001618 |
| TPCN2 | 9,73432 | -2,25607 | 0,000152 |
| TRAK2 | 9,64865 | -2,01991 | 1,26E-07 |
| TRANK1 | 8,00084 | -2,12669 | 0,00106 |
| TRIM11 | 8,15561 | 1,994224 | 7,46E-06 |
| TRIM45 | 10,45690 | -1,99849 | 0,003085 |
| TRIM47 | 9,10209 | -2,02468 | 3,92E-05 |
| TRIM52 | 3,19991 | 2,042686 | 0,000641 |
| TRIM65 | 4,72760 | -2,33071 | 5,88E-05 |
| TRIM72 | 9,20701 | 3,423959 | 0,000197 |
| TRIOBP | 9,23971 | -3,3585 | 6,77E-17 |
| TRMT10C | 3,31473 | 2,00468 | 3,70E-06 |
| TRMT61A | 9,47866 | 2,060637 | 2,26E-05 |
| TSC2 | 6,01275 | -1,96442 | 2,55E-05 |
| TTC12 | 5,44543 | -2,60564 | 6,45E-05 |
| TTYH2 | 9,06522 | -2,01923 | 0,005345 |
| TTYH3 | 8,31031 | -2,01352 | 0,000161 |
| TUBGCP5 | 9,46837 | -1,94509 | 3,14E-05 |
| TXNDC2 | 9,24808 | 2,660684 | 0,000161 |
| TXNL4B | 9,18581 | 2,162974 | 7,55E-06 |
| UACA | 8,19822 | -1,99116 | 2,95E-05 |
| UBE2D4 | 9,83133 | -2,71506 | 4,62E-06 |
| UBE2G1 | 9,08422 | 1,899187 | 5,28E-07 |
| UPP1 | 5,95194 | 2,473809 | 5,05E-06 |

| | | | |
|---|---|---|---|
| USP36 | 8,53556 | 2,027389 | 1,56E-07 |
| USPL1 | 6,81777 | 3,326234 | 6,77E-17 |
| UTP23 | 9,58575 | 2,01384 | 2,14E-05 |
| VASH1 | 4,75639 | -2,21105 | 0,000554 |
| VCPIP1 | 7,68044 | 1,937554 | 0,000258 |
| VCPKMT | 11,15000 | 2,093577 | 0,000331 |
| VIPR2 | 8,84391 | -2,84655 | 8,78E-05 |
| VSIG10L | 13,49392 | -2,16259 | 0,004385 |
| VTRNA1-3 | 10,39263 | 2,653298 | 4,71E-07 |
| WASH7P | 6,79337 | 2,109553 | 0,001445 |
| WDR20 | 8,80213 | 2,264441 | 3,15E-06 |
| WDR45B | 9,65204 | 1,933405 | 0,000369 |
| WDR74 | 11,26691 | 2,011303 | 1,24E-07 |
| WNT6 | 9,80513 | -2,16231 | 0,002548 |
| XPC | 3,99152 | -2,00987 | 1,86E-07 |
| YOD1 | 8,30675 | 1,993112 | 0,000234 |
| YRDC | 10,27432 | 2,541191 | 1,14E-09 |
| YTHDC1 | 9,85780 | 1,923559 | 2,40E-05 |
| YTHDF3 | 9,10316 | 1,999768 | 5,85E-06 |
| ZBED3 | 10,06651 | -1,93181 | 0,000464 |
| ZBED9 | 9,10202 | 2,501699 | 5,18E-05 |
| ZBTB21 | 11,49397 | 2,428608 | 1,65E-10 |
| ZBTB42 | 7,94592 | -3,13499 | 9,51E-05 |
| ZBTB43 | 9,59577 | 2,339711 | 0,000174 |
| ZBTB47 | 3,95505 | -1,98502 | 0,000465 |
| ZBTB5 | 9,94873 | 2,128216 | 3,07E-05 |
| ZBTB6 | 12,89775 | 1,928589 | 0,000117 |
| ZC3H12A | 10,40974 | 2,04727 | 2,30E-05 |
| ZC3H12D | 8,31971 | -2,84683 | 0,000148 |
| ZCCHC24 | 6,87006 | -2,77138 | 9,16E-12 |
| ZCCHC3 | 9,00769 | 2,582001 | 1,18E-07 |
| ZDHHC16 | 10,28233 | -2,18025 | 7,25E-08 |
| ZFAND2A | 6,43579 | 2,885332 | 6,72E-09 |
| ZFAS1 | 10,20344 | 2,478027 | 1,48E-08 |
| ZFP36L2 | 10,72257 | 2,077279 | 6,78E-06 |
| ZGLP1 | 7,62357 | 1,91329 | 0,029909 |
| ZHX3 | 11,81274 | -1,94285 | 4,05E-06 |
| ZKSCAN8 | 11,51225 | 2,259585 | 9,78E-06 |
| ZMYM3 | 8,42992 | -2,21379 | 4,62E-05 |
| ZMYM6NB | 5,96346 | -1,99177 | 0,00415 |
| ZNF12 | 6,96992 | 2,266633 | 8,99E-06 |
| ZNF16 | 13,40964 | 1,956643 | 0,000141 |
| ZNF175 | 10,14536 | 1,967885 | 0,001301 |
| ZNF202 | 9,43850 | 2,876114 | 8,45E-11 |
| ZNF252P-AS1 | 11,40494 | 2,247819 | 0,006854 |
| ZNF26 | 8,48016 | 2,327424 | 3,45E-06 |
| ZNF263 | 11,73653 | 2,698613 | 4,45E-08 |
| ZNF280D | 14,20743 | -2,36605 | 1,65E-06 |
| ZNF292 | 12,21053 | 2,191488 | 3,23E-09 |
| ZNF296 | 13,03861 | 3,287092 | 1,10E-10 |
| ZNF304 | 13,12339 | 2,090561 | 5,53E-05 |
| ZNF317 | 10,57901 | 2,107649 | 0,000179 |
| ZNF35 | 14,48272 | 2,158487 | 1,62E-05 |
| ZNF350 | 10,67442 | 2,138307 | 0,00012 |
| ZNF384 | 11,94010 | 1,946827 | 7,17E-06 |
| ZNF394 | 11,82745 | 2,367316 | 7,83E-07 |
| ZNF408 | 12,64551 | 2,700757 | 5,09E-10 |
| ZNF410 | 13,88036 | 2,203364 | 3,39E-06 |
| ZNF432 | 13,07248 | 2,164723 | 0,000531 |
| ZNF441 | 10,89819 | 2,257817 | 0,000252 |
| ZNF467 | 10,22172 | -2,85597 | 9,68E-06 |
| ZNF473 | 7,64241 | 2,764063 | 6,96E-09 |
| ZNF488 | 10,23293 | -2,59956 | 0,000385 |
| ZNF529 | 8,89102 | 2,235245 | 2,31E-05 |
| ZNF574 | 8,49740 | 2,073536 | 1,86E-05 |
| ZNF597 | 9,62624 | 2,319507 | 7,39E-06 |
| ZNF616 | 8,32146 | 2,15388 | 6,96E-05 |
| ZNF622 | 9,38694 | 2,07455 | 4,13E-06 |
| ZNF654 | 9,17792 | 2,203634 | 0,000135 |
| ZNF669 | 11,47856 | 2,947535 | 1,31E-08 |
| ZNF672 | 11,93362 | -2,25768 | 1,89E-07 |
| ZNF707 | 9,19630 | 1,98039 | 0,000106 |
| ZNF721 | 8,48115 | 2,414509 | 1,01E-05 |
| ZNF830 | 9,22291 | 2,419516 | 5,33E-08 |
| ZNF841 | 11,94646 | 1,966798 | 0,000177 |
| ZNF92 | 8,47775 | 2,697597 | 1,39E-07 |
| ZPR1 | 8,96773 | 2,012363 | 1,76E-06 |
| ZSCAN12P1 | 7,41406 | 2,284069 | 0,005273 |

**Table 8.3 Significantly deregulated genes in SRSF3 KD, Hela cells**

| geneID | Log2 baseMean | log2 FoldChange | padj |
|---|---|---|---|
| ABCC8 | 6,02597 | 2,761138 | 0,000501 |
| ABCD1 | 9,73403 | -1,41038 | 0,019855 |

| | | | |
|---|---|---|---|
| ABHD14B | 8,75180 | -2,55975 | 3,40E-05 |
| ABHD5 | 10,53656 | 1,706557 | 0,003805 |
| ABLIM1 | 12,42842 | 1,266308 | 0,029016 |
| ACTA2 | 6,26616 | -1,98778 | 0,032493 |
| ADAM15 | 10,09538 | -1,38914 | 0,04446 |
| ADAM32 | 6,34425 | 3,510596 | 5,18E-06 |
| ADAMTS14 | 5,96304 | -2,43942 | 0,014314 |
| ADCY3 | 9,55791 | -2,53396 | 7,31E-05 |
| ADGRE2 | 9,25915 | -1,36718 | 0,010211 |
| AFF3 | 4,94485 | 2,51389 | 0,021709 |
| AGAP1 | 10,28278 | 1,03049 | 0,038013 |
| AGPAT1 | 10,60926 | -1,31701 | 0,006879 |
| AK4 | 10,10650 | -1,61667 | 0,002451 |
| AKAP6 | 5,84546 | 2,198434 | 0,030872 |
| ALDH8A1 | 5,23187 | 2,668907 | 0,002299 |
| ALOX12-AS1 | 5,22838 | 2,330599 | 0,015581 |
| AQP3 | 10,91077 | 1,730214 | 0,008897 |
| AREG | 8,70892 | 3,543708 | 1,39E-07 |
| ARFGEF3 | 6,73099 | -1,84055 | 0,021307 |
| ARID1B | 11,19380 | -1,07907 | 0,031883 |
| ASNS | 9,74292 | 1,710054 | 0,016303 |
| ATAD3B | 9,20022 | 1,368593 | 0,013109 |
| ATP2A3 | 7,09954 | -1,96785 | 0,040427 |
| ATP2B4 | 10,10580 | -1,33337 | 0,004863 |
| ATP6V0A1 | 10,33339 | 2,090116 | 0,000651 |
| ATP6V0A2 | 9,58504 | 1,317043 | 0,012048 |
| ATP6V0E2 | 7,41690 | -2,06926 | 0,017214 |
| ATP8A2 | 3,29975 | 1,741084 | 0,040427 |
| AVIL | 7,12794 | 2,241772 | 0,032013 |
| BACH2 | 6,71483 | 2,208305 | 0,019645 |
| BBIP1 | 8,91732 | 1,538135 | 0,013109 |
| BBS4 | 8,64165 | 1,341545 | 0,045113 |
| BIRC3 | 10,49843 | 2,456356 | 0,002299 |
| BRI3BP | 8,41417 | -1,29716 | 0,049371 |
| BRINP2 | 5,02125 | 0,633513 | 0,000384 |
| BZW2 | 9,96382 | -1,62378 | 0,032493 |
| C11orf86 | 5,81614 | 2,007741 | 0,03276 |
| C11orf96 | 5,16374 | 2,594127 | 0,004702 |
| C16orf71 | 3,61159 | 2,365172 | 0,030235 |
| C1D | 8,02596 | 1,510163 | 0,010418 |
| C1S | 10,77657 | 1,721686 | 0,014559 |
| C1orf233 | 6,28041 | -2,61553 | 0,002208 |
| C22orf29 | 8,33613 | -1,65297 | 0,030872 |
| C2orf74 | 9,00808 | 1,807228 | 0,017506 |
| C9orf69 | 9,31694 | -1,73272 | 0,000208 |
| CA12 | 10,85608 | -2,6912 | 5,52E-06 |
| CA2 | 8,17949 | 1,947093 | 0,00515 |
| CALCOCO1 | 10,28339 | 1,973769 | 0,012642 |
| CAP2 | 7,46449 | -2,01516 | 0,00461 |
| CASP6 | 7,45825 | -1,61387 | 0,018976 |
| CASP9 | 9,93368 | 2,037311 | 0,00898 |
| CAT | 9,57240 | -1,27804 | 0,037665 |
| CBX6 | 10,66824 | -1,30882 | 0,047349 |
| CCDC144CP | 8,93164 | 2,655019 | 1,39E-07 |
| CCDC69 | 7,07765 | -1,51229 | 0,032059 |
| CCDC85B | 10,30181 | -1,6488 | 0,035063 |
| CCDC86 | 11,62633 | -1,52024 | 0,011086 |
| CCDC91 | 9,59990 | 1,196918 | 0,035534 |
| CCL20 | 9,16307 | 2,385915 | 0,00461 |
| CCND1 | 12,91469 | -1,88145 | 0,015125 |
| CDKN2AIP | 9,16408 | 1,499256 | 0,034687 |
| CEBPB-AS1 | 5,73418 | 2,747949 | 0,005621 |
| CHMP1B | 11,18079 | 1,415962 | 0,000973 |
| CIPC | 8,65222 | -1,76699 | 0,004128 |
| CLDN1 | 9,25653 | 2,448055 | 0,000476 |
| CLK1 | 10,42423 | 1,272351 | 0,049925 |
| CLN6 | 9,36430 | -1,97514 | 0,017586 |
| CNNM3 | 9,35548 | -1,61532 | 0,018359 |
| COQ10B | 9,82335 | 1,351637 | 0,020157 |
| CREB5 | 8,50451 | 3,511164 | 1,72E-06 |
| CREBRF | 9,71356 | 2,211276 | 0,00461 |
| CRIM1 | 12,86083 | 1,306667 | 0,018333 |
| CRK | 10,99404 | 1,570173 | 0,003937 |
| CROCC | 7,87682 | -1,54623 | 0,012642 |
| CTSL | 11,59780 | 1,651533 | 0,001805 |
| CUL9 | 8,48464 | -1,34473 | 0,032311 |
| CWC25 | 9,80105 | 1,7513 | 0,005284 |
| CYB561A3 | 8,58923 | -1,31945 | 0,04642 |
| CYB5D2 | 7,69767 | 1,71212 | 0,017334 |
| CYP4V2 | 7,87806 | 3,658702 | 9,37E-06 |
| CYSRT1 | 6,30314 | 1,836209 | 0,017277 |
| DCAF16 | 9,72051 | -1,87362 | 0,00259 |
| DCLK1 | 9,38083 | 1,370647 | 0,049598 |
| DCP1B | 7,01594 | -1,45852 | 0,043732 |
| DCTD | 9,07319 | -1,35586 | 0,007495 |
| DDIT3 | 8,91944 | 1,967362 | 0,018171 |
| DDX39A | 11,30146 | 1,334381 | 0,035219 |
| DENND2D | 6,87308 | -1,7882 | 0,020512 |

| | | | |
|---|---|---|---|
| DLGAP1-AS1 | 8,32926 | 1,661953 | 0,002347 |
| DNAJB1 | 11,58644 | 1,113588 | 0,048452 |
| DNAJB4 | 10,01003 | 1,55963 | 0,035063 |
| DNAJC3-AS1 | 6,78922 | 1,816728 | 0,045822 |
| DPAGT1 | 8,64592 | -2,00806 | 0,00149 |
| DPYD | 9,23971 | -1,34759 | 0,033352 |
| DUSP7 | 9,16730 | -1,65287 | 0,029062 |
| DYNLT3 | 9,65204 | 1,573306 | 0,001592 |
| ECM2 | 3,68957 | 2,584749 | 0,01365 |
| EGR1 | 13,28935 | -1,81587 | 0,017302 |
| EIF2AK3 | 9,43100 | 1,393726 | 0,045822 |
| EIF4EBP2 | 11,37876 | -1,85405 | 0,012642 |
| EIF6 | 10,89819 | -1,21305 | 0,039322 |
| ELF3 | 7,64241 | 2,311459 | 0,013403 |
| ELFN2 | 8,77549 | -1,52763 | 0,049598 |
| EMC3 | 9,95707 | 1,464623 | 0,014186 |
| EMC7 | 11,93362 | 1,228083 | 0,015581 |
| ENTPD6 | 8,10249 | -1,24925 | 0,044988 |
| EREG | 8,40320 | 2,689791 | 8,33E-05 |
| ERV3-1 | 5,67146 | 2,247037 | 0,0253 |
| ESCO1 | 9,46361 | 1,316385 | 0,030235 |
| ESM1 | 3,47274 | 1,640723 | 0,044451 |
| ESYT1 | 10,08941 | -1,01503 | 0,045822 |
| ETAA1 | 9,76565 | 1,312549 | 0,016646 |
| ETS2 | 11,59811 | 1,244229 | 0,018976 |
| ETV4 | 10,37684 | -2,61226 | 0,000137 |
| EVC | 8,37566 | -1,62086 | 0,03096 |
| EVPL | 9,74293 | -1,23789 | 0,047253 |
| FAAH2 | 7,72067 | 1,643933 | 0,049925 |
| FAAP24 | 7,40483 | 1,430089 | 0,045775 |
| FAM134B | 8,46493 | 1,759675 | 0,022376 |
| FAM171B | 8,22648 | -2,23252 | 0,000128 |
| FAM220A | 7,34709 | -1,5412 | 0,017075 |
| FAM222B | 11,07774 | 1,253293 | 0,041937 |
| FAM46C | 4,59498 | -1,84321 | 0,041684 |
| FAM76B | 9,17901 | 1,075785 | 0,042522 |
| FAM9B | 6,22286 | 3,263056 | 0,000113 |
| FBL | 11,59168 | -1,39101 | 0,038978 |
| FBXO32 | 7,71538 | 3,15207 | 3,38E-05 |
| FBXO38 | 9,66118 | 1,458133 | 0,004863 |
| FBXW4P1 | 6,43530 | -1,86289 | 0,032493 |
| FER | 10,27139 | 1,177938 | 0,045822 |
| FLJ42969 | 3,78502 | 2,917382 | 0,033068 |
| FMO4 | 3,33517 | 2,608768 | 0,022408 |

| | | | |
|---|---|---|---|
| FNIP1 | 10,38113 | 1,558707 | 0,006879 |
| FOXJ2 | 9,68604 | -1,1955 | 0,027808 |
| FRG1 | 10,17239 | 1,295359 | 0,040415 |
| FST | 6,81909 | 2,079247 | 0,011086 |
| FZD10 | 8,01988 | -1,79433 | 0,012755 |
| GABRR2 | 3,94976 | 2,332425 | 0,009984 |
| GALNT2 | 13,92461 | 1,379184 | 0,022376 |
| GARS | 12,07013 | 1,119238 | 0,049779 |
| GAS7 | 5,11067 | 2,422309 | 0,00742 |
| GCNT3 | 3,79907 | 2,614593 | 0,029062 |
| GDPD1 | 4,86480 | -3,03476 | 0,003038 |
| GLB1L2 | 6,32772 | -1,75274 | 0,024028 |
| GPCPD1 | 9,19100 | 1,446852 | 0,025249 |
| GPR141 | 5,81720 | -2,20446 | 0,023275 |
| GPR153 | 6,58963 | -2,49971 | 0,008855 |
| GPR3 | 5,44917 | -2,24646 | 0,012174 |
| GREB1 | 6,13090 | 2,570739 | 0,005318 |
| GRM1 | 4,10386 | 2,279629 | 0,007749 |
| HDAC5 | 8,76505 | -1,3668 | 0,019543 |
| HECTD2-AS1 | 5,30716 | 2,997549 | 0,006451 |
| HEXA-AS1 | 5,53896 | 2,644907 | 0,003486 |
| HIF1A | 13,46465 | -1,13212 | 0,03403 |
| HK2 | 10,52311 | -1,24939 | 0,032493 |
| HKDC1 | 4,32206 | 2,29924 | 0,022376 |
| HMBOX1 | 9,60373 | 1,674619 | 0,002299 |
| HNRNPL | 12,09617 | -1,30478 | 0,003804 |
| HOXA13 | 6,79824 | -2,35525 | 0,000373 |
| HSPB8 | 9,03768 | 1,803997 | 0,022376 |
| IDH3B | 10,47377 | 1,249989 | 0,041684 |
| IFI44L | 8,29542 | -1,46854 | 0,013403 |
| IFRD1 | 11,00728 | 1,913062 | 9,54E-05 |
| IGSF1 | 7,59770 | 1,89201 | 0,022376 |
| IL16 | 3,28136 | 2,270507 | 0,039322 |
| IL18 | 10,08600 | 1,339065 | 0,030218 |
| IL1A | 7,65676 | 3,744986 | 3,73E-07 |
| IL1B | 5,55655 | 2,320229 | 0,040486 |
| IL20RB | 8,82875 | 2,206748 | 0,016646 |
| IQGAP1 | 13,62480 | -1,41691 | 0,004097 |
| ITCH | 11,48057 | 1,187831 | 0,031176 |
| ITGB4 | 11,42571 | -1,47028 | 0,006879 |
| ITPRIPL2 | 10,89577 | -1,35909 | 0,041467 |
| JMJD1C | 11,68426 | 1,252461 | 0,045775 |
| JMJD4 | 7,26822 | -2,02617 | 0,008942 |
| KCNK1 | 10,61034 | 1,201154 | 0,018 |

| KCTD11 | 7,22381 | -2,24539 | 8,44E-05 |
|---|---|---|---|
| KIAA0100 | 12,54923 | -1,25698 | 0,012063 |
| KITLG | 8,75202 | 1,588886 | 0,01313 |
| KLC1 | 10,98513 | 1,266636 | 0,011086 |
| KLF4 | 10,36562 | 1,791383 | 0,000197 |
| KRT17 | 13,83614 | 2,312855 | 8,44E-05 |
| LAMA1 | 7,71575 | 2,229012 | 4,57E-05 |
| LAMB3 | 11,15593 | 1,978459 | 0,013109 |
| LATS2 | 11,53220 | 1,995138 | 7,16E-05 |
| LCN2 | 5,69922 | 2,698003 | 0,008816 |
| LDOC1L | 9,86128 | -1,34901 | 0,025084 |
| LGALS9 | 6,55352 | 2,510195 | 0,000707 |
| LINC00520 | 5,03996 | 2,278088 | 0,029062 |
| LINC00958 | 9,25813 | -1,31222 | 0,020578 |
| LINC01140 | 5,98188 | 3,735369 | 2,94E-06 |
| LIPG | 7,01509 | -2,18052 | 0,002456 |
| LMBRD1 | 9,25163 | 1,385849 | 0,049925 |
| LMO7 | 13,35703 | 1,26364 | 0,045345 |
| LOC100131496 | 4,05227 | 2,612071 | 0,020132 |
| LOC100506746 | 6,17042 | 2,495592 | 0,016175 |
| LOC101929125 | 4,78701 | 2,255498 | 0,035641 |
| LOC102724467 | 4,37451 | 2,332264 | 0,039976 |
| LOC388813 | 3,79555 | 1,798719 | 0,041723 |
| LONRF1 | 6,96846 | -1,87494 | 0,022656 |
| LRP1B | 8,71327 | 1,63039 | 0,012642 |
| LVRN | 5,27207 | 2,332261 | 0,010868 |
| LZTFL1 | 8,80957 | 1,470237 | 0,029119 |
| LZTR1 | 8,78582 | -1,38471 | 0,011937 |
| MALL | 6,17040 | -2,87358 | 0,000508 |
| MAP2 | 10,31017 | 2,599997 | 0,000274 |
| MAP4K2 | 8,23478 | -2,36461 | 0,000299 |
| MARCH10 | 4,28122 | 2,667517 | 0,041723 |
| MARVELD1 | 10,02926 | -1,19202 | 0,015581 |
| MAVS | 9,87542 | -1,67333 | 0,001509 |
| MBNL2 | 11,82450 | 1,359153 | 0,005318 |
| MCM3AP-AS1 | 7,32134 | -1,56124 | 0,029765 |
| MED4 | 10,52023 | 1,423588 | 0,004691 |
| MED6 | 9,35215 | 1,415411 | 0,032493 |
| MET | 11,26027 | 2,322013 | 2,80E-06 |
| METTL8 | 8,76041 | -1,46107 | 0,001805 |
| MEX3A | 9,30694 | -2,11133 | 0,000315 |
| MFAP5 | 7,52909 | 1,924562 | 0,014559 |
| MGAT5 | 10,02638 | -1,48499 | 0,003804 |
| MGEA5 | 12,24503 | 1,385568 | 0,029062 |

| MIAT | 6,27669 | 1,874706 | 0,036209 |
|---|---|---|---|
| MICAL2 | 8,27380 | -2,14009 | 0,000502 |
| MIR100HG | 9,05293 | 1,527018 | 0,008298 |
| MMP24-AS1 | 9,16893 | -1,17035 | 0,047669 |
| MOB1B | 9,87508 | -1,17219 | 0,014661 |
| MOK | 7,04938 | 2,628568 | 4,86E-05 |
| MOSPD1 | 8,38635 | 1,602655 | 0,038978 |
| MPI | 9,21323 | -1,62185 | 0,012668 |
| MPLKIP | 8,41797 | -1,39476 | 0,020571 |
| MTR | 10,50050 | -1,5222 | 0,003857 |
| MTUS1 | 11,09618 | 1,897861 | 0,017302 |
| MUC5AC | 7,60899 | -2,10352 | 0,029156 |
| MYBPC2 | 5,61176 | -2,87616 | 0,018171 |
| MYO16 | 7,62170 | 1,710071 | 0,007541 |
| MYT1L | 6,05010 | 3,463966 | 0,000141 |
| MZT1 | 9,47878 | -1,23844 | 0,026762 |
| NCKIPSD | 9,73122 | -1,51856 | 0,000644 |
| NEU4 | 5,08117 | -2,29622 | 0,014559 |
| NEURL1 | 9,13552 | -1,90091 | 0,008442 |
| NFATC2 | 6,80114 | 2,684848 | 0,000271 |
| NFIL3 | 9,32118 | 1,666746 | 0,002816 |
| NFIX | 9,41529 | -2,07938 | 0,000208 |
| NGEF | 6,00515 | -2,19887 | 0,016487 |
| NHSL1 | 8,48614 | -1,31768 | 0,012063 |
| NNT | 10,67413 | -1,07733 | 0,042679 |
| NOTCH2 | 12,33958 | -1,30204 | 0,002503 |
| NPRL3 | 9,49848 | -1,65602 | 0,000351 |
| NR1D1 | 10,68864 | -1,44261 | 0,049371 |
| NRBF2 | 8,99494 | 1,715668 | 0,002208 |
| NTS | 7,05124 | -2,48294 | 0,005995 |
| NUP50-AS1 | 6,06975 | 1,800826 | 0,049598 |
| NUPR1 | 7,93820 | 3,447266 | 2,08E-05 |
| OAS3 | 8,53202 | -1,54571 | 0,014186 |
| OIP5-AS1 | 9,14140 | -1,56527 | 0,001219 |
| ORAI1 | 7,54141 | -3,28704 | 1,85E-07 |
| ORAI3 | 6,73490 | -1,87474 | 0,01313 |
| ORC3 | 9,01419 | -1,41713 | 0,033797 |
| OVGP1 | 7,01238 | 2,47694 | 0,004702 |
| PARVA | 10,80055 | 1,046777 | 0,041723 |
| PAXIP1-AS1 | 6,41343 | 1,725953 | 0,02966 |
| PCBP1-AS1 | 6,43700 | 1,832285 | 0,019204 |
| PCCB | 9,88569 | -1,63285 | 0,014123 |
| PCDH1 | 4,54511 | -2,50883 | 0,002373 |
| PCSK7 | 9,01755 | 1,424271 | 0,044918 |

| | | | |
|---|---|---|---|
| PDCD6 | 11,66970 | 1,596237 | 0,000184 |
| PDLIM5 | 11,71387 | -1,53078 | 0,003482 |
| PEG10 | 14,41320 | -2,13454 | 5,18E-06 |
| PFKFB3 | 9,04140 | -1,49315 | 0,01313 |
| PGAP2 | 7,72291 | -1,61495 | 0,01365 |
| PHLDB2 | 12,20021 | 1,401548 | 0,032493 |
| PIK3R2 | 10,64666 | -1,51552 | 0,039742 |
| PIP4K2A | 9,66249 | 1,914803 | 5,13E-05 |
| PLCB1 | 9,05877 | 1,337143 | 0,035837 |
| PLS3 | 11,92310 | 1,248931 | 0,048153 |
| PLSCR1 | 8,66869 | -1,36963 | 0,012755 |
| PLXNA1 | 10,41045 | -1,46024 | 0,000887 |
| PNP | 10,15425 | -1,01951 | 0,035534 |
| POLD2 | 11,22066 | -1,53626 | 0,049371 |
| PPID | 10,59541 | 1,175804 | 0,0487 |
| PPP1R12B | 9,62752 | 1,698775 | 0,000293 |
| PPP1R13L | 9,82064 | -1,57963 | 0,001307 |
| PRR16 | 4,36957 | -2,13188 | 0,030894 |
| PSMD6-AS2 | 6,96717 | 1,623929 | 0,015354 |
| PTBP1 | 12,91776 | -1,78666 | 0,000326 |
| PTCH2 | 7,46060 | 2,214741 | 0,015581 |
| PTGS2 | 13,71385 | -2,27014 | 2,08E-05 |
| PTPRR | 5,53350 | 2,923534 | 0,008653 |
| RAB30 | 10,11934 | 1,476763 | 0,037701 |
| RAB7A | 12,96877 | 1,326384 | 0,010418 |
| RALGPS1 | 7,22558 | 1,819362 | 0,036209 |
| RAP2A | 9,73556 | -1,2232 | 0,016175 |
| RASSF5 | 5,85189 | -2,79956 | 0,000833 |
| RBMS2 | 9,75122 | -1,38774 | 0,00764 |
| RCHY1 | 7,86170 | 1,544753 | 0,009467 |
| RETSAT | 8,60168 | -2,05213 | 0,004097 |
| RFK | 9,76252 | 1,95809 | 0,000239 |
| RFX3 | 7,17944 | -2,11442 | 0,001342 |
| RNF19A | 9,92971 | 1,374557 | 0,001516 |
| RNF19B | 8,35941 | 1,55119 | 0,041723 |
| RNF6 | 11,10625 | 1,747196 | 0,002299 |
| RNU11 | 9,50743 | 1,563832 | 0,032493 |
| RNU12 | 9,00112 | 1,785886 | 0,007495 |
| RNU2-2P | 7,96572 | 1,661419 | 0,008042 |
| RNU4-1 | 8,29206 | 1,767876 | 0,035953 |
| RNU4-2 | 7,93846 | 2,02148 | 0,012617 |
| RNU4ATAC | 8,01020 | 2,359829 | 0,001587 |
| RNU5B-1 | 4,99669 | 2,085815 | 0,042522 |
| RNU5D-1 | 7,30529 | 2,448033 | 0,008298 |
| RNU5F-1 | 6,84689 | 1,949413 | 0,033139 |
| RNVU1-14 | 9,15881 | 1,894153 | 0,01365 |
| RNVU1-15 | 8,74370 | 2,337178 | 0,003588 |
| RNVU1-6 | 6,55246 | 1,867453 | 0,038378 |
| RNVU1-7 | 6,70321 | 2,276743 | 0,016379 |
| RRAGD | 6,16897 | 2,174095 | 0,033768 |
| RTCA | 8,94722 | 1,236252 | 0,025626 |
| SAPCD2 | 9,76563 | -1,74268 | 0,004457 |
| SARAF | 12,65856 | 0,982829 | 0,045276 |
| SCARA3 | 11,95190 | 1,422908 | 0,010211 |
| SCNN1A | 7,92623 | -1,49354 | 0,033768 |
| SCX | 3,76358 | -1,97616 | 0,019645 |
| SEL1L | 11,26738 | 1,378442 | 0,019855 |
| SEMA5A | 6,08116 | 3,420864 | 0,000113 |
| SERINC5 | 7,63120 | -1,54947 | 0,044451 |
| SESN2 | 9,40596 | 2,126412 | 0,005098 |
| SH3BP4 | 10,77508 | 1,050315 | 0,048153 |
| SH3PXD2A | 10,32668 | -1,3175 | 0,026572 |
| SIX2 | 8,39513 | -2,03732 | 0,000208 |
| SLC16A3 | 12,67533 | -1,71335 | 0,008897 |
| SLC25A10 | 8,51746 | -2,3242 | 0,00259 |
| SLC25A22 | 8,75887 | -1,20109 | 0,032493 |
| SLC29A2 | 8,30939 | -2,22379 | 0,013831 |
| SLC2A4RG | 11,18292 | -1,77746 | 3,29E-06 |
| SLC37A3 | 9,26813 | 1,564877 | 0,013109 |
| SLC3A2 | 13,21208 | 1,772686 | 3,38E-05 |
| SLC5A3 | 9,20980 | -1,87403 | 0,002456 |
| SLC6A8 | 10,43842 | -1,76074 | 1,25E-05 |
| SLC9A2 | 8,34680 | -1,90255 | 0,030872 |
| SLCO4A1 | 9,56301 | -1,864 | 0,030223 |
| SLIT2 | 9,17121 | -1,30374 | 0,005627 |
| SMC1B | 4,50508 | 2,604923 | 0,031883 |
| SMIM19 | 6,87836 | 2,033301 | 0,000508 |
| SMIM2-AS1 | 7,15927 | 2,671874 | 0,000163 |
| SNHG3 | 11,01167 | -1,43385 | 0,020157 |
| SNORD47 | 5,97363 | -2,0162 | 0,019855 |
| SNRNP48 | 10,45310 | 1,131135 | 0,030602 |
| SOD2 | 11,85516 | 1,902161 | 0,019883 |
| SPACA6P-AS | 4,69733 | 2,162078 | 0,035534 |
| SPPL2B | 8,60297 | -1,39897 | 0,029765 |
| SREBF1 | 11,71292 | -1,30644 | 0,022376 |
| SRF | 10,37502 | -1,49875 | 0,001805 |
| SRRM5 | 7,23261 | 2,108076 | 0,002272 |
| SRSF3 | 11,96464 | -2,32545 | 1,85E-07 |

| | | | |
|---|---|---|---|
| SSPN | 4,65375 | 2,871892 | 0,008242 |
| SSSCA1 | 9,10459 | -1,36871 | 0,033797 |
| ST3GAL3 | 6,18330 | -2,32136 | 0,007541 |
| ST6GALNAC4 | 9,43521 | -1,69833 | 0,016303 |
| STAG2 | 12,23250 | 1,649567 | 0,000435 |
| STEAP3 | 9,21162 | -1,56705 | 0,037827 |
| STOM | 11,50113 | 1,38364 | 0,021995 |
| STRA6 | 10,22641 | -1,92183 | 0,000508 |
| STX3 | 9,43243 | 1,446715 | 0,00578 |
| SULT1E1 | 5,02948 | 0,61576 | 0,00079 |
| SUSD5 | 6,57719 | -2,39077 | 0,003619 |
| SYNPO2 | 8,22666 | 1,786398 | 0,016592 |
| TAPT1 | 7,62147 | 1,431458 | 0,033352 |
| TBX19 | 5,35074 | 1,998312 | 0,041684 |
| TCP11L2 | 7,49123 | 2,219036 | 0,003619 |
| TDP2 | 9,96273 | 1,972381 | 0,008666 |
| TGM2 | 13,35880 | -1,40037 | 0,027449 |
| THOC5 | 10,28148 | 1,273037 | 0,030223 |
| THRA | 9,02940 | -1,48643 | 0,018976 |
| THRB | 8,94447 | 1,366125 | 0,03096 |
| TINAGL1 | 8,66658 | -2,41704 | 3,75E-05 |
| TMEM107 | 9,28952 | 1,691762 | 0,041684 |
| TMEM109 | 10,97235 | -1,11449 | 0,029119 |
| TMEM129 | 7,77576 | -1,89128 | 0,009984 |
| TMEM158 | 7,39985 | -2,24939 | 0,000208 |
| TMEM177 | 6,56583 | -1,97357 | 0,009467 |
| TMEM208 | 10,20060 | 1,696401 | 0,013035 |
| TMEM5 | 7,04143 | -1,54585 | 0,01905 |
| TMPPE | 7,23097 | -1,46217 | 0,045358 |
| TNFAIP1 | 11,14910 | 1,285526 | 0,025785 |
| TNFAIP6 | 7,46769 | 1,737375 | 0,036209 |
| TNS2 | 9,81448 | -1,17164 | 0,013109 |
| TNS4 | 12,32183 | -2,05729 | 2,08E-05 |
| TP53 | 10,66838 | -1,33388 | 0,015125 |
| TP63 | 5,62137 | 3,402219 | 0,000842 |
| TRAF1 | 3,75393 | 2,572499 | 0,020132 |
| TRAK2 | 10,65575 | -1,15925 | 0,025785 |
| TRIM10 | 3,67407 | 2,352493 | 0,018171 |
| TRIM15 | 4,52384 | 2,784489 | 0,039689 |
| TRIM25 | 10,93661 | -1,6118 | 0,014186 |
| TRMT112 | 11,83952 | 1,069624 | 0,03096 |
| TTC13 | 8,78047 | 1,315812 | 0,032493 |
| TTC17 | 11,64139 | 1,579087 | 0,000659 |
| UBAP2L | 12,57569 | -1,64535 | 0,000113 |
| UBASH3B | 7,23404 | -1,61619 | 0,047979 |
| UBE2G1 | 10,58737 | 1,054536 | 0,0487 |
| UBIAD1 | 8,81448 | -1,66071 | 0,009984 |
| UBQLN4 | 10,07594 | -1,36218 | 0,005627 |
| UCN2 | 5,57477 | -3,07226 | 0,001219 |
| UNC13A | 9,72297 | -1,91196 | 0,019855 |
| UNC5D | 5,86323 | -2,0531 | 0,004378 |
| UPRT | 7,05020 | -1,79113 | 0,013035 |
| UVSSA | 8,15826 | 1,758418 | 0,01846 |
| VPS9D1-AS1 | 8,44661 | -1,77154 | 0,03096 |
| VTN | 5,33885 | 1,92978 | 0,030223 |
| WDR45 | 10,10812 | 1,309873 | 0,036264 |
| WSB1 | 11,36240 | 1,346718 | 0,038728 |
| ZADH2 | 9,27515 | -1,94038 | 3,38E-05 |
| ZBED6CL | 7,93164 | -1,68209 | 0,013109 |
| ZBTB20 | 11,16222 | 1,364732 | 0,035953 |
| ZC3HAV1L | 8,56212 | -2,23165 | 3,46E-05 |
| ZCCHC2 | 10,27871 | 1,929861 | 0,001113 |
| ZFAND2A | 8,49288 | 1,525881 | 0,023298 |
| ZFHX2 | 5,45853 | -2,15423 | 0,019645 |
| ZFYVE1 | 8,32087 | 1,942697 | 0,022376 |
| ZMYND10 | 3,50865 | 1,614909 | 0,029453 |
| ZNF226 | 9,00433 | 1,70947 | 0,010211 |
| ZNF396 | 5,10430 | 2,376281 | 0,033139 |
| ZNF436-AS1 | 6,64863 | 2,05114 | 0,032493 |
| ZNF524 | 5,60238 | -2,06241 | 0,018359 |
| ZNF525 | 7,81113 | -1,44561 | 0,016175 |
| ZNF556 | 6,32480 | -2,41752 | 0,001113 |
| ZNF576 | 7,38023 | -1,81382 | 0,007182 |
| ZNF580 | 9,24542 | -2,33623 | 0,000283 |
| ZNF596 | 6,32904 | 2,137245 | 0,032493 |
| ZNF888 | 7,03980 | -1,51843 | 0,0487 |

**Table 8.4 Significantly deregulated genes in SETDB1 KD, Hela cells**

| geneID | Log2 baseMean | log2 FoldChange | padj |
|---|---|---|---|
| AIM1L | 6,97649 | -2,61875 | 0,03253 |
| AOX1 | 7,17621 | -2,64588 | 0,00066 |
| CYP4F3 | 5,34623 | -0,0176 | 0,03623 |

| | | | |
|---|---|---|---|
| MET | 11,26027 | 1,994506 | 0,00185 |
| RBPJ | 10,91743 | 1,393577 | 0,025736 |

| | | | |
|---|---|---|---|
| SETDB1 | 9,51316 | -1,74201 | 0,018858 |

**Table 8.5 Significantly differentially spliced genes in DIS3 KD, Hela cells**

| geneID | padj |
|---|---|
| ALG3 | 0,03791 |
| ARHGAP5 | 9,40E-10 |
| CALD1 | 0,000759 |
| CASP8AP2 | 0,025815 |
| CHURC1 | 0,000245 |
| CKAP2 | 0,025815 |
| CLPTM1 | 0,00468 |
| CRTC3 | 0,007993 |
| CSDE1 | 7,29E-07 |
| EZR | 0,000664 |

| | |
|---|---|
| HIST1H2BD | 3,04E-06 |
| HNRNPD | 1,26E-05 |
| HNRNPDL | 0,005631 |
| HNRNPU | 7,29E-07 |
| ILF3 | 0,00349 |
| IRAK1 | 0,000107 |
| NF1 | 0,008422 |
| OSBPL8 | 4,16E-10 |
| PRKDC | 9,40E-10 |
| RIF1 | 0,010541 |
| RPL15 | 0,001031 |

| | |
|---|---|
| SEC16A | 0,006419 |
| SLC22A23 | 0,004142 |
| TAX1BP1 | 0,036277 |
| TOMM40 | 0,001249 |
| TTC3 | 0,000107 |
| UBA1 | 2,89E-05 |
| UBAP2L | 0,005531 |
| YWHAZ | 0,000867 |
| ZNF146 | 0,001031 |

**Table 8.6 Significantly differentially spliced genes in SF3A3 KD, Hela cells**

| geneID | padj |
|---|---|
| ABL2 | 2,47E-10 |
| ACADVL | 6,83E-07 |
| ACOT9 | 0,02793 |
| ADARB1 | 1,29E-07 |
| ADD3 | 0,010126 |
| ADGRE2 | 0,001092 |
| ADPGK | 0,010452 |
| AK2 | 2,51E-10 |
| AKAP13 | 0,041438 |
| AKT1S1 | 3,33E-05 |
| AMD1 | 4,16E-06 |
| AMOTL1 | 0,024337 |
| ANAPC7 | 0,034381 |
| ANKHD1 | 0,010452 |
| ANKRD11 | 0,032465 |
| ANKRD17 | 1,14E-08 |
| ANLN | 9,62E-08 |
| ANTXR1 | 8,23E-14 |
| AP1B1 | 0,00208 |
| AP2B1 | 4,88E-06 |
| AP3M1 | 3,25E-12 |
| APEX1 | 0,039752 |
| APP | 0,000371 |

| | |
|---|---|
| ARAF | 0,007125 |
| ARHGAP11A | 0,027775 |
| ARHGDIA | 1,07E-07 |
| ARID5B | 0,009658 |
| ARMCX6 | 0,035884 |
| ARPC2 | 0,000157 |
| ATF7IP | 0,039519 |
| ATP5J | 0,025994 |
| ATP6V0B | 0,04889 |
| ATP6V1H | 0,005699 |
| ATPIF1 | 0,000678 |
| ATRX | 0,001412 |
| BACH1 | 0,000696 |
| BAG4 | 3,93E-05 |
| BAZ2B | 0,031673 |
| BBX | 5,34E-11 |
| BCAT1 | 8,73E-10 |
| BCLAF1 | 0,000182 |
| BCOR | 0,045597 |
| BICD2 | 0,015574 |
| BRD2 | 9,52E-12 |
| BRD9 | 7,32E-10 |
| BRMS1 | 0,018938 |
| BTBD3 | 1,94E-06 |

| | |
|---|---|
| BTF3 | 3,28E-08 |
| BTF3L4 | 0,014271 |
| C11orf49 | 0,010597 |
| C21orf59 | 0,011791 |
| C8orf59 | 0,000113 |
| CAB39 | 7,23E-05 |
| CAPRIN1 | 1,09E-08 |
| CBFB | 0,0464 |
| CBX5 | 2,81E-05 |
| CCDC14 | 0,018043 |
| CCDC50 | 5,71E-07 |
| CCNT1 | 8,31E-23 |
| CCT3 | 0,001083 |
| CCT8 | 0,002633 |
| CD44 | 0,0406 |
| CDCA7L | 0,005594 |
| CDK13 | 6,42E-08 |
| CDK2 | 0,00671 |
| CDKN1A | 0,00671 |
| CDKN1C | 0,036689 |
| CES2 | 0,001477 |
| CHM | 0,018772 |
| CHMP2B | 0,000824 |
| CHORDC1 | 0,009035 |

| | |
|---|---|
| CHRAC1 | 0,031808 |
| CKAP5 | 0,020551 |
| CLEC16A | 0,008174 |
| CLINT1 | 0,018549 |
| CLPB | 0,004429 |
| CMPK1 | 2,08E-05 |
| CNBP | 0,043823 |
| CNOT1 | 8,07E-12 |
| CNOT7 | 0,002793 |
| COG5 | 0,012078 |
| COL4A6 | 1,38E-05 |
| COX16 | 1,68E-08 |
| CPEB4 | 0,014615 |
| CREB1 | 0,023185 |
| CSDE1 | 1,57E-07 |
| CSNK1A1 | 0,049915 |
| CTBP2 | 0,007919 |
| CTDP1 | 0,000751 |
| CUL3 | 0,002515 |
| CWC27 | 0,020994 |
| CYP51A1 | 0,000484 |
| DAB2 | 2,32E-15 |
| DDB2 | 0,038924 |
| DDX39B | 0,005558 |
| DDX3X | 0,000143 |
| DENND1A | 0,000112 |
| DFFA | 0,012676 |
| DGUOK | 0,02819 |
| DHX33 | 2,96E-08 |
| DIDO1 | 2,38E-09 |
| DKC1 | 0,043486 |
| DNAJB12 | 0,04634 |
| DNM1L | 0,026935 |
| DROSHA | 0,000535 |
| DTL | 3,48E-05 |
| DYNLL1 | 6,83E-07 |
| DYRK1A | 0,000368 |
| EARS2 | 0,015574 |
| ECHDC1 | 0,010999 |
| EDIL3 | 2,84E-05 |
| EEF1B2 | 1,17E-13 |
| EEF1D | 2,32E-06 |
| EIF3B | 5,34E-11 |
| EIF3M | 0,000726 |
| EIF4B | 7,23E-07 |
| EIF4G1 | 8,73E-15 |
| EIF4G2 | 1,93E-07 |
| ELF2 | 0,005272 |
| ELF4 | 0,014462 |
| ELOVL5 | 0,001406 |
| ELP3 | 1,64E-06 |
| ENG | 0,012342 |
| ENO1 | 0,011807 |
| ERAP1 | 0,046517 |
| ERCC1 | 0,0058 |
| ERLIN1 | 5,91E-05 |
| ETF1 | 2,08E-06 |
| ETFA | 0,007372 |
| EXOSC2 | 0,010452 |
| EYA4 | 0,007188 |
| EZR | 1,45E-05 |
| FAM127B | 0,002053 |
| FAM133B | 3,48E-05 |
| FAM21A | 0,010798 |
| FAM222B | 7,90E-09 |
| FAM96A | 0,000564 |
| FBXW11 | 0,039428 |
| FGD5-AS1 | 0,036865 |
| FIP1L1 | 0,039428 |
| FLNA | 0,0152 |
| FNTA | 0,02897 |
| FOLR1 | 0,010452 |
| FOXJ3 | 0,046404 |
| FTSJ1 | 0,036804 |
| FXR1 | 0,03456 |
| GART | 0,015895 |
| GGT1 | 0,016026 |
| GIT2 | 0,000157 |
| GNAS | 0,000371 |
| GNB1 | 2,13E-15 |
| GNL3 | 0,032176 |
| GPBP1 | 0,002244 |
| GUK1 | 0,014815 |
| HAT1 | 0,00599 |
| HAUS2 | 0,005779 |
| HBS1L | 0,000321 |
| HDLBP | 2,38E-09 |
| HIF1A | 1,58E-25 |
| HIPK3 | 0,027003 |
| HIST1H2BD | 0,001361 |
| HJURP | 0,020433 |
| HMGA1 | 0,002432 |
| HMGCS1 | 1,14E-05 |
| HMGN3 | 0,003904 |
| HNRNPA1 | 3,13E-16 |
| HNRNPA2B1 | 1,93E-11 |
| HNRNPAB | 7,74E-10 |
| HNRNPDL | 6,12E-08 |
| HNRNPF | 0,002466 |
| HNRNPH1 | 0,011791 |
| HNRNPK | 9,63E-12 |
| HNRNPL | 0,000263 |
| HNRNPM | 0,005084 |
| HNRNPU | 2,52E-05 |
| HNRNPUL1 | 0,000531 |
| HTATSF1 | 0,0026 |
| ICE2 | 0,011018 |
| IL18 | 0,017067 |
| ILF3 | 0,008781 |
| INF2 | 0,000335 |
| ING1 | 0,048995 |
| JMJD1C | 1,16E-11 |
| KCTD2 | 0,01238 |
| KDM2A | 0,00312 |
| KHDRBS1 | 3,51E-11 |
| KIF1B | 5,71E-07 |
| KTN1 | 5,41E-07 |
| LDB1 | 0,025192 |
| LINC00473 | 0,02793 |
| LSM1 | 0,008226 |
| LSS | 6,64E-05 |
| MAEA | 0,007264 |
| MAP3K13 | 0,001917 |
| MAPK7 | 0,012078 |
| MAPKAP1 | 0,04889 |
| MARCH6 | 0,020569 |
| MARCH7 | 0,017238 |
| MATR3 | 1,54E-28 |
| MCM7 | 7,16E-05 |
| MDM2 | 1,11E-05 |
| MEF2A | 0,002757 |
| MFAP3 | 0,000788 |

| | |
|---|---|
| MGAT4B | 0,004809 |
| MGST1 | 0,000223 |
| MICAL3 | 0,000285 |
| MINOS1 | 0,000186 |
| MIS12 | 0,004409 |
| MKI67 | 0,000796 |
| MKRN1 | 3,86E-05 |
| MOCS2 | 0,025587 |
| MORF4L1 | 0,000202 |
| MOV10 | 0,004576 |
| MRPL33 | 0,041438 |
| MRPL4 | 5,17E-10 |
| MRPS2 | 0,029538 |
| MRPS21 | 1,46E-05 |
| MTCH1 | 0,000435 |
| MTHFS | 0,043 |
| MTMR2 | 0,002793 |
| MTRF1L | 0,010332 |
| MXI1 | 2,86E-05 |
| MYO19 | 3,16E-05 |
| MYO1B | 0,048499 |
| N4BP2L2 | 0,024873 |
| NAP1L1 | 1,91E-05 |
| NBR1 | 0,017571 |
| NDRG1 | 5,00E-07 |
| NEDD4L | 0,002948 |
| NEMP1 | 0,010744 |
| NET1 | 0,026298 |
| NF1 | 8,34E-10 |
| NFIA | 2,85E-06 |
| NFKB2 | 0,014573 |
| NGFRAP1 | 1,13E-05 |
| NGRN | 0,011893 |
| NOB1 | 0,01249 |
| NONO | 1,29E-07 |
| NOP16 | 0,011219 |
| NPM1 | 3,96E-106 |
| NQO1 | 2,37E-06 |
| NRP1 | 0,004037 |
| NUDT4 | 0,00238 |
| NUP54 | 0,000696 |
| NUP98 | 0,00056 |
| NUPL1 | 0,000288 |
| NUSAP1 | 0,000568 |
| NXF1 | 0,00608 |
| NXPE3 | 0,019358 |
| ODC1 | 1,76E-05 |
| OLA1 | 0,000309 |
| OSMR | 1,67E-05 |
| OTUB1 | 3,54E-07 |
| PAPOLA | 0,024799 |
| PCDH7 | 7,77E-06 |
| PDCD2 | 0,000684 |
| PDHA1 | 4,86E-08 |
| PFN2 | 0,020508 |
| PHF20L1 | 0,001328 |
| PHLDB2 | 5,18E-05 |
| PITPNB | 1,97E-06 |
| PLA2G4A | 3,17E-06 |
| PLEC | 0,035588 |
| PLK4 | 0,027667 |
| PNISR | 3,52E-24 |
| POLDIP3 | 2,20E-11 |
| POLR1D | 0,027999 |
| POLR2H | 0,027796 |
| PPP1R12A | 0,001079 |
| PPP6R3 | 0,000702 |
| PQBP1 | 0,023368 |
| PRC1 | 0,000273 |
| PRRC2A | 0,004654 |
| PSMA4 | 8,82E-06 |
| PSMA5 | 0,017477 |
| PSMD11 | 0,008782 |
| PSMD12 | 0,033156 |
| PSMD13 | 0,005659 |
| PSPC1 | 0,006775 |
| PTBP1 | 2,69E-16 |
| PTEN | 0,001774 |
| PTP4A2 | 1,15E-06 |
| PTTG1IP | 0,04634 |
| PVRL2 | 0,028289 |
| PVRL3 | 0,002147 |
| QKI | 0,025573 |
| QTRTD1 | 0,035884 |
| RAB1A | 2,42E-07 |
| RABGGTB | 0,000433 |
| RAD1 | 8,05E-08 |
| RAD23A | 0,027796 |
| RAN | 4,66E-39 |
| RBM19 | 0,012419 |
| RBM34 | 0,012752 |
| RBM38 | 0,007919 |
| RCC2 | 0,011853 |
| RDX | 0,021824 |
| RECQL | 0,04634 |
| RIF1 | 1,29E-07 |
| RLIM | 0,001328 |
| RNPS1 | 0,00498 |
| RPAIN | 0,037877 |
| RPAP3 | 7,74E-10 |
| RPF2 | 4,42E-07 |
| RPL13A | 0,015211 |
| RPL15 | 0,002239 |
| RPL17 | 2,90E-05 |
| RPL31 | 1,49E-05 |
| RPL41 | 7,27E-11 |
| RPL9 | 0,000134 |
| RPLP1 | 1,97E-06 |
| RPRD2 | 0,023922 |
| RPS10 | 0,027949 |
| RPS17 | 8,43E-41 |
| RPS24 | 8,93E-33 |
| RPS27A | 2,38E-09 |
| RQCD1 | 0,00027 |
| RREB1 | 0,000588 |
| RRN3 | 5,37E-05 |
| RSRC2 | 0,014462 |
| RTFDC1 | 6,05E-06 |
| SACS | 0,014815 |
| SEC11A | 0,004409 |
| SEC16A | 0,000568 |
| Sep 15 | 5,95E-05 |
| Sep 07 | 0,029692 |
| Sep 09 | 8,23E-10 |
| SETD5 | 0,023185 |
| SETDB1 | 0,004429 |
| SF1 | 1,10E-08 |
| SHISA5 | 5,29E-08 |
| SHMT2 | 0,0013 |
| SLBP | 0,001497 |
| SLC22A23 | 0,027775 |
| SLC25A3 | 7,68E-12 |

| | |
|---|---|
| SLC38A2 | 3,61E-10 |
| SLC39A7 | 6,64E-05 |
| SLC39A8 | 0,014573 |
| SLC3A2 | 2,55E-14 |
| SLMO2 | 6,24E-47 |
| SMAD5 | 3,92E-07 |
| SMG7 | 0,000881 |
| SMPD4 | 0,000881 |
| SMS | 4,88E-06 |
| SNHG16 | 1,75E-08 |
| SNRNP70 | 0,00082 |
| SNX5 | 0,000286 |
| SOD2 | 0,008344 |
| SON | 3,68E-10 |
| SPATS2 | 0,048499 |
| SPECC1L | 0,024254 |
| SPTBN1 | 2,15E-24 |
| SRP19 | 5,55E-07 |
| SRPR | 0,002147 |
| SRRM1 | 1,13E-10 |
| SRSF3 | 0,000202 |
| SRSF5 | 1,17E-11 |
| SRSF6 | 4,07E-07 |
| SRSF7 | 1,91E-05 |
| SSSCA1 | 0,019229 |
| STAG2 | 0,022531 |
| STAU1 | 0,012752 |
| STRBP | 0,027667 |
| STX16 | 0,04353 |
| SUGT1 | 0,001502 |
| SUMO1 | 0,040962 |
| SURF4 | 0,019333 |
| SWAP70 | 0,001181 |
| TAF12 | 0,011872 |
| TAF6 | 0,042663 |
| TAX1BP1 | 1,17E-05 |
| TBC1D14 | 1,84E-05 |
| TBRG4 | 0,00346 |
| TCEB1 | 0,020678 |
| TCOF1 | 0,027796 |
| TCP1 | 5,44E-08 |
| TFAM | 1,96E-11 |
| TFIP11 | 0,011893 |
| THUMPD1 | 4,22E-07 |
| TIMM23 | 4,24E-06 |
| TMEM209 | 1,91E-05 |
| TMPO | 3,91E-10 |
| TNIP1 | 3,28E-08 |
| TNPO1 | 0,001768 |
| TNPO3 | 0,000195 |
| TOR1AIP2 | 1,15E-06 |
| TPD52L2 | 0,011018 |
| TPT1 | 6,62E-11 |
| TRIP12 | 0,005411 |
| TRMT112 | 1,70E-09 |
| TROAP | 0,003293 |
| TROVE2 | 0,002216 |
| TSPAN3 | 0,000239 |
| TTC19 | 0,012752 |
| TTC3 | 1,60E-25 |
| TUBGCP3 | 0,029884 |
| TXN | 4,88E-06 |
| TXNL4A | 0,04889 |
| TXNRD1 | 2,69E-16 |
| TYW3 | 0,000519 |
| UBAP1 | 7,83E-05 |
| UBAP2L | 0,006988 |
| UBE2D3 | 9,96E-07 |
| UBE2E1 | 0,000288 |
| UBE2H | 2,85E-07 |
| UBE2I | 0,011893 |
| UBE2K | 7,09E-05 |
| UBE2V1 | 5,19E-07 |
| UBQLN1 | 7,68E-07 |
| UBTF | 0,03193 |
| UFM1 | 4,73E-06 |
| UGGT1 | 6,34E-09 |
| UQCRH | 0,001454 |
| USB1 | 0,018573 |
| USP10 | 7,68E-07 |
| USP14 | 0,000739 |
| USP47 | 5,66E-12 |
| UXT | 0,011791 |
| VAPB | 1,97E-07 |
| VPS13C | 3,30E-07 |
| VPS28 | 0,032683 |
| VPS37A | 0,004522 |
| VTA1 | 0,015655 |
| WAC | 0,00491 |
| WBP5 | 3,14E-06 |
| WBSCR22 | 2,00E-08 |
| WDR74 | 1,77E-07 |
| WNK1 | 0,00238 |
| WWP2 | 0,000431 |
| XYLT2 | 0,037222 |
| YAP1 | 0,001454 |
| YBX3 | 8,77E-08 |
| YTHDF3 | 2,63E-19 |
| YWHAB | 4,61E-06 |
| ZC3H14 | 5,38E-05 |
| ZCCHC6 | 0,000149 |
| ZEB2 | 0,03575 |
| ZFAND5 | 1,07E-05 |
| ZFAS1 | 3,08E-08 |
| ZFP91 | 0,00243 |
| ZNF121 | 0,004985 |
| ZNF146 | 1,61E-28 |
| ZNF317 | 1,49E-16 |
| ZNF384 | 1,17E-08 |
| ZNF451 | 0,011327 |
| ZNF480 | 0,009035 |
| ZNF692 | 0,021467 |
| ZRANB2 | 2,53E-12 |
| ZYX | 0,013253 |

**Table 8.7 Significantly differentially spliced genes in SRSF3 KD, Hela cells**

| geneID | padj |
|---|---|
| ACIN1 | 0,026184 |
| ACSL3 | 0,041914 |
| ADARB1 | 4,60E-07 |
| ADGRE2 | 0,001396 |
| ADGRE5 | 1,04E-05 |
| AK2 | 0,001262 |
| ALG3 | 0,019384 |
| ALS2 | 0,017955 |

| | |
|---|---|
| AMOTL1 | 0,049452 |
| ANKHD1 | 0,04108 |
| ANKRD11 | 0,036729 |
| ANKRD17 | 5,72E-24 |
| AP1G1 | 0,001201 |
| AP2B1 | 9,59E-09 |
| API5 | 0,046423 |
| ARAP1 | 0,024047 |
| ARHGEF12 | 0,01432 |
| ARPC2 | 2,79E-06 |
| ASCC3 | 8,35E-06 |
| ASPH | 0,028162 |
| ATF7 | 0,044029 |
| ATF7IP | 0,023859 |
| ATG13 | 0,007861 |
| ATP6V0C | 1,04E-05 |
| ATXN2 | 0,00465 |
| BBX | 5,11E-05 |
| BCCIP | 0,000141 |
| BRMS1 | 0,026184 |
| BTF3 | 0,018698 |
| BZW1 | 1,65E-17 |
| C11orf24 | 6,85E-07 |
| C5orf24 | 0,004098 |
| CAMTA2 | 0,02196 |
| CBX5 | 0,040002 |
| CD63 | 7,50E-12 |
| CDK11B | 0,01432 |
| CDK12 | 0,000762 |
| CDV3 | 0,001107 |
| CEP152 | 0,007054 |
| CEP170 | 5,40E-05 |
| CHCHD10 | 0,022028 |
| CHD4 | 0,009366 |
| CHPF | 0,044717 |
| CLIP1 | 0,00192 |
| CNN2 | 0,017121 |
| CNOT4 | 0,01405 |
| CNOT7 | 0,037409 |
| COL4A6 | 0,00465 |
| CPEB4 | 0,022028 |
| CRK | 0,000211 |
| CSNK1A1 | 0,025987 |
| CSNK1E | 0,002685 |
| CYP51A1 | 0,006891 |
| DCAF8 | 0,031006 |
| DDB2 | 0,01405 |
| DDX46 | 0,000225 |
| DICER1 | 0,028005 |
| DIS3 | 0,00321 |
| DLGAP4 | 0,000298 |
| DYRK2 | 0,001645 |
| EDIL3 | 0,016884 |
| EEF1B2 | 0,013374 |
| EIF3J | 0,028743 |
| EIF4G2 | 2,48E-07 |
| EIF4G3 | 0,034936 |
| EIF4H | 8,52E-05 |
| EIF6 | 0,030406 |
| ELF4 | 6,93E-10 |
| ELK4 | 0,035763 |
| EPB41L2 | 1,98E-06 |
| ERBB2IP | 5,78E-06 |
| FAM127B | 0,0001 |
| FAM214A | 0,00371 |
| FAM222B | 5,32E-05 |
| FBXO38 | 0,029366 |
| FERMT2 | 0,010111 |
| FEZ2 | 0,034674 |
| FNDC3B | 0,043772 |
| GAPDH | 0,003366 |
| GART | 5,78E-06 |
| GLG1 | 0,037116 |
| GMPR2 | 0,029178 |
| GOLIM4 | 7,51E-09 |
| GTF3C1 | 2,77E-11 |
| HAT1 | 0,004869 |
| HERPUD1 | 0,003782 |
| HIF1A | 5,46E-10 |
| HNRNPA1 | 0,001968 |
| HNRNPA2B1 | 1,04E-39 |
| HNRNPDL | 0,040002 |
| HSPA14 | 0,001417 |
| HTATIP2 | 0,014134 |
| HTRA3 | 0,017121 |
| IARS | 0,004621 |
| INTS6 | 0,048368 |
| JMJD1C | 6,59E-11 |
| JUP | 9,73E-06 |
| KCTD2 | 0,018802 |
| KDM2A | 0,000879 |
| KIF1B | 0,001004 |
| KIF23 | 0,004058 |
| KTN1 | 0,000673 |
| LARP4 | 1,38E-05 |
| LIN54 | 4,60E-07 |
| LINC00473 | 0,001745 |
| LRIF1 | 9,33E-10 |
| MADD | 0,018942 |
| MAMLD1 | 0,044029 |
| MAP4K4 | 1,63E-10 |
| MAP7D1 | 0,005427 |
| MAPKAPK2 | 0,005312 |
| MATR3 | 0,006219 |
| METTL13 | 0,004309 |
| MFAP3 | 0,007944 |
| MFSD1 | 0,033152 |
| MGA | 1,50E-27 |
| MIR22HG | 0,018265 |
| MKI67 | 0,01405 |
| MKRN1 | 0,000196 |
| MOV10 | 0,018842 |
| MPZL1 | 0,005392 |
| MRPL33 | 0,049452 |
| MTHFD1L | 0,030624 |
| MTMR1 | 0,018802 |
| MYO1B | 0,000342 |
| NASP | 0,001004 |
| NCOA4 | 7,78E-10 |
| NCOA6 | 4,32E-05 |
| NCOA7 | 0,016377 |
| NDUFV3 | 0,000403 |
| NEDD4L | 0,004415 |
| NGFRAP1 | 0,006778 |
| NIPA1 | 0,000292 |
| NKIRAS2 | 0,017331 |
| NR1D2 | 0,000757 |
| NUP98 | 0,045123 |
| NUPL1 | 0,026248 |
| ORMDL1 | 0,000105 |
| OSMR | 0,00148 |
| P4HA1 | 0,036371 |

| | |
|---|---|
| PALLD | 9,87E-18 |
| PDLIM7 | 1,41E-05 |
| PEX26 | 0,040002 |
| PFN2 | 9,33E-10 |
| PHF20L1 | 0,000105 |
| PHF3 | 3,11E-06 |
| PHF8 | 0,000995 |
| PIP5K1C | 0,007657 |
| PKM | 4,48E-18 |
| PLEC | 0,002929 |
| PPP1R12A | 4,16E-12 |
| PPP2R5C | 0,00401 |
| PRRC2A | 5,46E-10 |
| PRSS23 | 0,005392 |
| PTBP1 | 0,017628 |
| PXN | 1,46E-05 |
| QRICH1 | 0,01529 |
| RBAK | 5,25E-08 |
| RBBP6 | 0,0003 |
| RIMKLB | 0,000404 |
| RNF182 | 0,006014 |
| RNF212 | 0,026597 |
| RNF216 | 3,40E-05 |
| RNPS1 | 0,017955 |
| RPL13 | 0,004299 |
| RPL15 | 0,036371 |
| RPL41 | 0,019766 |
| RPL8 | 0,017874 |
| RPRD2 | 0,012366 |
| RPS15A | 0,020398 |
| SACS | 0,00451 |

| | |
|---|---|
| SCAF4 | 0,028183 |
| SEC14L1 | 0,001745 |
| SEC16A | 8,35E-06 |
| SEC24C | 0,00148 |
| Sep 15 | 0,000689 |
| SKP2 | 1,16E-08 |
| SLC22A23 | 0,00401 |
| SLC3A2 | 0,001481 |
| SLCO3A1 | 0,02526 |
| SLMO2 | 0,018942 |
| SMARCA1 | 0,002706 |
| SMARCA4 | 0,033756 |
| SPAG9 | 0,012699 |
| SPATS2 | 0,006014 |
| SPECC1 | 0,000462 |
| SRP9 | 0,037116 |
| SRSF1 | 1,38E-05 |
| SRSF3 | 0,001004 |
| SSFA2 | 0,018802 |
| STX16 | 0,000111 |
| SYNPO | 0,005883 |
| TBL1X | 1,11E-08 |
| TFDP1 | 0,007944 |
| TFE3 | 9,02E-07 |
| TGM2 | 8,94E-12 |
| TMEM164 | 0,009203 |
| TMEM87A | 0,044029 |
| TMX2 | 0,034464 |
| TNRC6B | 0,000333 |
| TOMM40 | 0,006891 |
| TP53BP1 | 0,040981 |

| | |
|---|---|
| TPM2 | 1,10E-07 |
| TRA2A | 0,00371 |
| TRIOBP | 0,002083 |
| TTC3 | 0,002506 |
| TXNRD1 | 0,00401 |
| UBA1 | 7,59E-05 |
| UBE2D3 | 0,005489 |
| UFM1 | 0,000147 |
| UQCC1 | 0,031366 |
| URI1 | 6,49E-05 |
| USP15 | 0,001889 |
| USP47 | 0,003455 |
| USP8 | 0,032712 |
| VEGFB | 0,015047 |
| WAC | 0,002978 |
| WARS | 0,018325 |
| WIPI2 | 0,006778 |
| WNK1 | 5,61E-08 |
| WNT5A | 5,05E-13 |
| XPR1 | 1,38E-05 |
| YAP1 | 0,000102 |
| YWHAZ | 0,001616 |
| ZAK | 0,000225 |
| ZBTB20 | 0,001266 |
| ZFP36L1 | 0,028005 |
| ZNF106 | 0,049452 |
| ZNF146 | 2,20E-06 |
| ZNF207 | 0,030624 |
| ZNF226 | 0,008498 |
| ZNF385A | 0,01103 |
| ZNF664 | 5,03E-06 |

**Table 8.8 Significantly differentially spliced genes in SETDB1 KD, Hela cells**

| geneID | padj |
|---|---|
| ALG3 | 0,012753 |
| ANKRD12 | 0,013964 |
| ATF7IP | 0,018013 |

| | |
|---|---|
| MATR3 | 0,008153 |
| NAA50 | 0,013964 |
| OSBPL8 | 3,84E-06 |
| RAP1GAP2 | 1,69E-05 |

| | |
|---|---|
| RPL15 | 0,000544 |
| SPTAN1 | 0,011168 |

**Table 8.9 Significantly deregulated ERVs with H3K9ac peak in DIS3 KD, Hela cells**

| RepeatID | chr | start | end | Log2 base Mean | log2 Fold Change | p-value |
|---|---|---|---|---|---|---|

| | | | | | | |
|---|---|---|---|---|---|---|
| 102872_LTR40c | chr13 | 79484963 | 79485180 | 2,0464 | 2,2476 | 0,024475873 |
| 106708_MER4A1,MER4-int,MER4-int,MER4-int,MER4-int,MER4-int,MER4-int,MER4A1 | chr13 | 99180737 | 99187756 | 2,2493 | -3,0800 | 0,021773382 |
| 108409_LTR9A1,LTR7Y,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7Y,LTR9A1,LTR9 | chr13 | 109264649 | 109272374 | 14,2478 | 1,4280 | 0,006845062 |
| 108993_LTR73,MSTB1,MSTA1 | chr13 | 113188387 | 113189461 | 2,6017 | 2,2913 | 0,008895736 |
| 114326_MLT1F2,MLT1F2,MER41B,LTR56 | chr14 | 52070629 | 52072212 | 4,2565 | 1,6061 | 0,016904551 |
| 114334_MER21B,MER21B | chr14 | 52110793 | 52111338 | 2,3064 | 2,6815 | 0,008906234 |
| 117114_MER41B | chr14 | 71130672 | 71131303 | 0,0000 | 3,2560 | 0,022390513 |
| 117517_MamGyp-int | chr14 | 73712618 | 73713668 | 9,9633 | -1,1101 | 0,030298611 |
| 122888_MER54A | chr14 | 104748811 | 104749690 | 3,5486 | -2,6721 | 0,043144196 |
| 127636_MLT1L,MLT1L | chr15 | 51884686 | 51885098 | 3,0375 | 1,8238 | 0,031486613 |
| 128281_MER65C | chr15 | 56734076 | 56734537 | 0,0000 | 3,1240 | 0,014013592 |
| 13009_LTR8B | chr1 | 89952737 | 89953657 | 1,1678 | 2,9407 | 0,005665741 |
| 130354_MSTB | chr15 | 74205225 | 74205636 | 4,7363 | -2,8683 | 0,009349059 |
| 130355_MLT1J | chr15 | 74210732 | 74210968 | 5,7248 | -2,4166 | 0,003305641 |
| 133645_THE1B | chr15 | 97690688 | 97691045 | 3,8364 | 2,0156 | 0,019538294 |
| 135957_MER4D1,LTR29,LTR29 | chr16 | 11973184 | 11974203 | 4,1072 | -2,0294 | 0,013265442 |
| 136667_MLT1K | chr16 | 15489167 | 15489300 | 3,4838 | 1,8059 | 0,025318796 |
| 137476_LTR14C,HERVK14C-int,LTR14C | chr16 | 19393581 | 19402152 | 1,6370 | 2,4816 | 0,022268747 |
| 139550_MER90a,MER90a | chr16 | 29927421 | 29928300 | 3,7808 | -2,2072 | 0,027115774 |
| 139597_LTR2,LTR21A,HERVFH21-int,HERVFH21-int,HERVFH21-int,HERVFH21-int,LTR2,LTR2,HERVFH21-int,HERVFH21-int,LTR21A | chr16 | 30541068 | 30548758 | 3,9344 | 1,5484 | 0,017663139 |
| 143827_MER51C | chr16 | 67481515 | 67482239 | 5,3532 | -1,9445 | 0,015505942 |
| 146925_MLT1C,MLT1C | chr16 | 87551576 | 87552100 | 0,4278 | 3,2738 | 0,019823782 |
| 149747_MER4B | chr17 | 18721345 | 18721943 | 2,6017 | 2,6744 | 0,002686261 |
| 151864_MER21A | chr17 | 42552016 | 42552560 | 5,6903 | -1,6371 | 0,035971397 |
| 165954_MLT1C,LTR16E1 | chr18 | 62438346 | 62438841 | 2,0024 | 1,8904 | 0,04019927 |
| 169589_LTR38A1,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7C,MER92B | chr19 | 6574424 | 6577385 | 0,0000 | 2,5490 | 0,039105075 |
| 170720_MLT1C | chr19 | 14506165 | 14506489 | 0,8888 | 2,9230 | 0,016410144 |
| 171986_LTR7C,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7C,HERVK22-int,HERVK22-int,HERVK22-int,LTR22B1 | chr19 | 23549492 | 23558908 | 9,7149 | 1,5282 | 0,000911756 |
| 172986_LTR12,HERV9-int,HERV9N-int,HERV9N-int,LTR12 | chr19 | 32693105 | 32700522 | 5,2171 | -1,5674 | 0,036636342 |
| 174419_LTR8B,LTR101_Mam | chr19 | 42851137 | 42852046 | 3,7969 | 1,8300 | 0,016897595 |
| 174424_LTR8B | chr19 | 42875094 | 42875873 | 2,5094 | 2,4995 | 0,003436217 |
| 174545_LTR8B | chr19 | 43265000 | 43265754 | 1,1678 | 2,2680 | 0,018105341 |
| 174565_MSTA,LTR7Y,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7Y,THE1D-int,MSTA | chr19 | 43323083 | 43330524 | 8,6620 | 2,2309 | 0,037539846 |
| 174907_MLT1D,MLT1D,MER57A-int | chr19 | 45376452 | 45377187 | 0,5953 | 2,6085 | 0,047249214 |
| 179012_MLT1O | chr2 | 11389467 | 11389950 | 2,3060 | 2,0940 | 0,030964902 |
| 180667_LTR5A | chr2 | 20501251 | 20502198 | 3,7708 | 3,2573 | 1,76E-07 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 181697_MER31-int,MER31-int,MER31-int,MER31-int | chr2 | 29126581 | 29129623 | 4,3737 | 1,4735 | 0,033416336 |
| 18798_MER61C,LTR76,LTR12C,LTR76,MER61-int,MER61-int | chr1 | 151905123 | 151909112 | 4,6251 | -1,7236 | 0,045261248 |
| 189766_HERVP71A-int,LTR71A | chr2 | 82778296 | 82780489 | 4,9944 | 1,4504 | 0,031313513 |
| 189768_LTR1D | chr2 | 82790255 | 82791056 | 0,5953 | 2,8389 | 0,02700155 |
| 190180_MLT2B3,MLT2B3 | chr2 | 84919720 | 84920696 | 5,6149 | 1,1941 | 0,029610912 |
| 191933_MER11C,HERVK11-int,MER11A | chr2 | 100361987 | 100365971 | 1,8057 | 2,7877 | 0,004048739 |
| 193508_MER4A1_ | chr2 | 109701267 | 109701835 | 1,2898 | 2,8425 | 0,01706303 |
| 197878_MLT1K | chr2 | 136236799 | 136237352 | 4,3900 | 1,2649 | 0,041903665 |
| 198214_MLT1C | chr2 | 138035738 | 138036195 | 4,2486 | 1,7536 | 0,042791649 |
| 199036_MER72 | chr2 | 142875416 | 142876090 | 0,0000 | 4,2427 | 0,000753033 |
| 207723_MLT1J2 | chr2 | 201150528 | 201150886 | 3,9508 | 1,6933 | 0,020473398 |
| 216666_LTR101_Mam,LTR101_Mam | chr20 | 10353417 | 10354360 | 1,6114 | 2,6965 | 0,010240958 |
| 216971_LTR35 | chr20 | 11791706 | 11792313 | 4,5982 | 1,5713 | 0,013304706 |
| 216996_MER65C,MLT1A0,LTR16C | chr20 | 11849612 | 11850903 | 5,8968 | 1,4495 | 0,013469865 |
| 223433_MLT1G1 | chr20 | 51347649 | 51347806 | 1,5664 | 2,3174 | 0,049513158 |
| 223434_MLT1L | chr20 | 51351034 | 51351189 | 0,6945 | 3,0377 | 0,014962254 |
| 223915_LTR33 | chr20 | 54131541 | 54131974 | 1,0231 | 2,6848 | 0,027104846 |
| 224663_LTR12C | chr20 | 57348213 | 57349645 | 0,0000 | 2,1817 | 0,045324697 |
| 227993_MER21C | chr21 | 20720794 | 20721574 | 0,5953 | 2,8879 | 0,019843045 |
| 231802_MLT1D | chr21 | 38840820 | 38841271 | 0,0000 | 2,5503 | 0,030935314 |
| 231819_LTR29 | chr21 | 38908060 | 38908632 | 1,3447 | 2,9551 | 0,004711898 |
| 231858_THE1B | chr21 | 39047561 | 39047875 | 0,0000 | 2,7213 | 0,043761808 |
| 233235_ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,LTR16A | chr21 | 46456086 | 46458103 | 7,1285 | -1,2161 | 0,026803091 |
| 23339_MLT2F,MLT2F | chr1 | 186621027 | 186621930 | 0,0000 | 3,4760 | 0,013692269 |
| 234928_MLT1G1,MSTA | chr22 | 27715992 | 27717421 | 2,5070 | 2,0790 | 0,028658239 |
| 237269_MER54A | chr22 | 45583568 | 45584193 | 4,6390 | -1,5646 | 0,044747378 |
| 242004_LTR33 | chr3 | 23042038 | 23042500 | 0,6945 | 3,4490 | 0,004280598 |
| 245428_MER4D1,MER4D1,LTR18B,MER4D1,LTR7B,MER4D1 | chr3 | 41962833 | 41965130 | 5,0890 | 1,4569 | 0,028129239 |
| 245761_LTR16A1 | chr3 | 43734082 | 43734486 | 2,9004 | 1,7071 | 0,032292116 |
| 254229_LTR16A2 | chr3 | 99503437 | 99503786 | 0,0000 | 3,2569 | 0,008398355 |
| 254500_LTR6A,HERVS71-int,LTR6A,MSTA | chr3 | 101632044 | 101642824 | 3,1201 | -2,0774 | 0,020919406 |
| 256914_LTR12C,LTR12C,MLT2A1 | chr3 | 115791379 | 115793165 | 4,4545 | 2,6935 | 6,26E-05 |
| 256915_LTR7,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7,LTR2 | chr3 | 115793482 | 115800126 | 0,6945 | 3,0687 | 0,013317307 |
| 256916_MER77,MER77B | chr3 | 115800142 | 115801155 | 1,1678 | 2,2993 | 0,046004567 |
| 261606_MLT2C1,MLT1D-int,MLT1D,ERVL-B4-int | chr3 | 146967031 | 146970434 | 0,0000 | 3,3803 | 0,006188256 |
| 261962_MER74A,MLT1F2,MLT1F2 | chr3 | 149388603 | 149389300 | 0,0000 | 3,3809 | 0,007752889 |
| 263134_MLT1L | chr3 | 156819003 | 156819149 | 1,6114 | 2,5335 | 0,018560311 |
| 263136_MER67C | chr3 | 156827170 | 156827706 | 4,7588 | 1,8439 | 0,002705403 |
| 26316_MER21C | chr1 | 206889720 | 206890488 | 5,8311 | 1,9537 | 0,005984569 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 267556_THE1C | chr3 | 185960874 | 185961009 | 3,7218 | 1,7021 | 0,046558272 |
| 26866_ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,LTR16A,ERVL-B4-int,MLT1N2 | chr1 | 210436678 | 210441114 | 2,3749 | 3,0886 | 0,000140189 |
| 27181_MLT1K,MLT1K | chr1 | 212595202 | 212595647 | 1,0417 | 2,8866 | 0,014963992 |
| 277706_MER21C | chr4 | 41506728 | 41507652 | 4,9099 | 1,6653 | 0,006283213 |
| 279326_LTR21A,MSTA,MSTA | chr4 | 52713762 | 52715242 | 9,6539 | -1,4584 | 0,020538998 |
| 281911_LTR25 | chr4 | 66574695 | 66575495 | 0,4278 | 3,1081 | 0,010785204 |
| 285250_MLT1E | chr4 | 86929098 | 86929646 | 1,7631 | 2,8041 | 0,004705837 |
| 29384_LTR16E2,LTR16E2,LTR16E2,LTR16,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,LTR16 | chr1 | 227356153 | 227360947 | 0,0000 | 3,1852 | 0,010900696 |
| 294566_MLT1L,MLT1G | chr4 | 142092587 | 142093449 | 0,4278 | 2,9525 | 0,014305258 |
| 29612_MER101-int,MER101-int | chr1 | 228648617 | 228650136 | 5,6401 | -1,6932 | 0,016970814 |
| 297371_LTR16A2,ERV3-16A3_I-int,MSTB1,ERVL-E-int,ERVL-E-int,ERVL-E-int,ERVL-E-int,LTR16A2 | chr4 | 159545279 | 159549677 | 3,0853 | 2,2605 | 0,009079677 |
| 301412_LTR54,LTR54,MER57C2 | chr4 | 183410827 | 183411525 | 1,2898 | 2,7660 | 0,010869263 |
| 302569_LTR41B,LTR41B | chr5 | 337615 | 338050 | 4,3665 | -2,1031 | 0,041098154 |
| 302573_MLT1A,MLT1A-int,MLT1A-int,MLT1A | chr5 | 360002 | 362364 | 7,7561 | -1,0820 | 0,033218474 |
| 303619_MLT1A1 | chr5 | 6409005 | 6409403 | 2,3060 | 2,2208 | 0,01241208 |
| 304500_MLT1E1A,MLT1E1A,LTR41 | chr5 | 10474219 | 10475470 | 0,0000 | 2,5503 | 0,022386651 |
| 305302_MLT1C | chr5 | 14038719 | 14039172 | 0,9169 | 2,9405 | 0,017597082 |
| 30559_LTR7,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7 | chr1 | 233683448 | 233689204 | 6,9396 | 1,8880 | 0,000283071 |
| 305792_HERV16-int,HERV16-int,LTR16 | chr5 | 16439974 | 16441746 | 3,9129 | 1,5224 | 0,030566098 |
| 305939_MamGypLTR2b | chr5 | 17375279 | 17375517 | 2,4757 | 1,7464 | 0,030976527 |
| 310336_MER21A,MER21A,MER21A,MER21-int,MER4B-int,MER21-int,MER21-int,MER21-int,MER21-int,MER21A,MER21A | chr5 | 38631831 | 38638329 | 0,0000 | 2,5686 | 0,048855217 |
| 311108_MER31-int,MER31-int,MER39 | chr5 | 43051693 | 43053018 | 6,5734 | -1,6026 | 0,01744089 |
| 311203_LTR49-int,LTR49-int,LTR49-int,LTR49-int | chr5 | 43571512 | 43575275 | 8,3728 | -1,2140 | 0,031763869 |
| 322253_LTR10F,LTR10F,LTR10F,LTR10F,MER41A,LTR10F,LTR10F,MER41A,MLT1D | chr5 | 115928854 | 115931428 | 2,6639 | 3,8603 | 8,21E-06 |
| 323546_MLT1A1 | chr5 | 123820312 | 123820716 | 0,8888 | 3,6518 | 0,000778118 |
| 327126_LTR62 | chr5 | 150702658 | 150702964 | 1,4841 | 2,2713 | 0,035334116 |
| 330466_MLT1B | chr5 | 172802673 | 172803064 | 0,0000 | 3,7484 | 0,010681767 |
| 330561_MamGypsy2-I,MLT1L,MamGypsy2-I | chr5 | 173452137 | 173453120 | 1,5925 | 3,1584 | 0,001213333 |
| 330562_LTR16C,MLT1A-int,MLT1A,MLT1F1,MLT1F1,MLT1F1,MLT1F1 | chr5 | 173453476 | 173455503 | 6,7414 | 2,5287 | 8,17E-06 |
| 330563_MLT1B | chr5 | 173456048 | 173456471 | 0,8888 | 3,9622 | 0,000149341 |
| 330564_MamGypsy2-I | chr5 | 173456725 | 173456843 | 1,2751 | 2,7502 | 0,020085408 |
| 330838_MER77B | chr5 | 174829605 | 174830167 | 1,8581 | 2,6481 | 0,00626227 |
| 332093_MLT1B | chr6 | 971506 | 971875 | 4,9260 | 2,0507 | 0,001319433 |
| 332747_MLT1A0 | chr6 | 4932153 | 4932521 | 0,0000 | 4,3315 | 0,000603287 |
| 334339_MLT1B,MLT1C,MLT1C | chr6 | 14731204 | 14732784 | 1,7631 | 2,8978 | 0,008324531 |
| 336076_LTR7,MER34 | chr6 | 26028813 | 26029778 | 5,7056 | 2,5839 | 0,000686974 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 336099_MER101 | chr6 | 26122928 | 26123343 | 4,9148 | 2,0641 | 0,004434013 |
| 336114_LOR1b | chr6 | 26205550 | 26205846 | 4,7709 | 2,5526 | 0,000336309 |
| 336134_MER68,LTR3,HERVK3-int,HERVK3-int,LTR3 | chr6 | 26287804 | 26296266 | 0,0000 | 3,0718 | 0,01836215 |
| 336318_MER67A,MER67A | chr6 | 27126648 | 27127002 | 1,9637 | 2,8701 | 0,002965536 |
| 336319_ERVL-E-int,ERVL-E-int | chr6 | 27128639 | 27129713 | 4,1258 | 2,4726 | 0,005817853 |
| 336321_LTR48B | chr6 | 27138223 | 27138815 | 6,0371 | 1,5570 | 0,015945183 |
| 336332_LTR10B2,HERVIP10B3-int,HERVIP10B3-int,HERVIP10B3-int,LTR10B2 | chr6 | 27178167 | 27186538 | 5,0651 | -2,7803 | 0,000280262 |
| 336333_LTR14,HERVKC4-int | chr6 | 27187520 | 27190816 | 2,1920 | -2,6565 | 0,017794008 |
| 336502_MER39B,MER34,LTR48 | chr6 | 27897622 | 27898881 | 5,6940 | 2,3489 | 0,000221922 |
| 336504_MER4E1 | chr6 | 27899681 | 27900456 | 5,1339 | 2,2113 | 0,00092288 |
| 33679_MER51B,MER51-int,MER51-int,MER51-int,MER51-int,MER41-int,MER51-int,MER51-int,MER51-int,MER51B,MSTD | chr10 | 5065471 | 5071306 | 2,5710 | 1,9419 | 0,02313785 |
| 336983_MLT1G3,MLT1G3 | chr6 | 30767566 | 30768367 | 0,0000 | 2,7899 | 0,010623622 |
| 33728_MSTD,MSTB1 | chr10 | 5291580 | 5292813 | 4,1531 | 1,4891 | 0,038698681 |
| 33729_MLT1A0,MLT1A0 | chr10 | 5292175 | 5292941 | 0,0000 | 3,1766 | 0,018197125 |
| 33968_MLT2A2 | chr10 | 6779419 | 6779969 | 4,9216 | 1,5193 | 0,032320354 |
| 340497_MSTD,MER89 | chr6 | 56334668 | 56335782 | 0,0000 | 4,8346 | 5,62E-05 |
| 344641_MLT1A0 | chr6 | 84974690 | 84975015 | 4,8882 | -1,8767 | 0,03503529 |
| 352234_LTR64,LTR64,LTR64 | chr6 | 134286803 | 134287963 | 4,8892 | 1,2563 | 0,040635493 |
| 352494_MER4D1,MER4D1 | chr6 | 136068053 | 136069183 | 0,0000 | 2,7033 | 0,032367129 |
| 353949_LTR7Y,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7Y | chr6 | 144923174 | 144928858 | 9,0288 | 1,3751 | 0,043177995 |
| 354424_MLT1C | chr6 | 147137790 | 147138276 | 7,2391 | -2,2544 | 0,001595741 |
| 354996_LTR12_,HERV9-int,HERV9-int,LTR12_ | chr6 | 150849126 | 150856692 | 3,2572 | 1,5435 | 0,049035409 |
| 354997_LTR10E | chr6 | 150856969 | 150857593 | 3,5876 | 2,7244 | 0,000309452 |
| 354998_LTR47B4 | chr6 | 150857629 | 150858046 | 2,8525 | 3,1317 | 9,47E-05 |
| 356114_MER4C | chr6 | 158854265 | 158854723 | 1,1678 | 2,4888 | 0,045162821 |
| 356264_MLT1K | chr6 | 159578009 | 159578187 | 0,4278 | 3,2260 | 0,010560918 |
| 359105_MLT1A0 | chr7 | 7911209 | 7911546 | 1,4841 | 2,3371 | 0,033034814 |
| 359855_LTR7B,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7B | chr7 | 12659787 | 12665594 | 4,5674 | 3,1608 | 5,13E-06 |
| 360605_MLT1D | chr7 | 17372108 | 17372580 | 0,0000 | 3,7866 | 0,002348694 |
| 360616_MLT2B3 | chr7 | 17406566 | 17407073 | 0,0000 | 4,3301 | 0,000360772 |
| 360617_MLT1E1A | chr7 | 17408187 | 17408704 | 0,4278 | 4,5281 | 0,000239583 |
| 361329_LTR45 | chr7 | 22582602 | 22583135 | 6,4945 | 1,2421 | 0,033823425 |
| 361330_MLT1G3 | chr7 | 22583268 | 22583703 | 3,2572 | 1,6649 | 0,047139906 |
| 361335_LTR38A1 | chr7 | 22598761 | 22598946 | 1,0231 | 2,8814 | 0,0184557 |
| 363471_LTR7C | chr7 | 36659793 | 36660261 | 0,0000 | 2,9043 | 0,031989607 |
| 3635_MSTC,MSTC,MSTC | chr1 | 26349611 | 26350303 | 0,6945 | 2,8043 | 0,048210615 |
| 365657_LTR22A | chr7 | 50533873 | 50534328 | 0,5953 | 2,6652 | 0,034622663 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 373984_MLT1J1,LTR7,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7,MSTD | chr7 | 109641330 | 109648236 | 5,0827 | 2,7464 | 0,000280453 |
| 375662_LTR4,HERV3-int,HERV3-int | chr7 | 120052770 | 120061151 | 2,1803 | 2,7748 | 0,005832087 |
| 376985_MLT1C | chr7 | 128112087 | 128112329 | 0,5953 | 2,9198 | 0,011151652 |
| 377329_MLT1H | chr7 | 131188268 | 131188771 | 1,4841 | 3,2183 | 0,001638418 |
| 379341_MER92C,MER92C | chr7 | 144842031 | 144842865 | 0,0000 | 3,3172 | 0,007963303 |
| 380070_MER34A1 | chr7 | 148939147 | 148939707 | 4,7871 | -1,6769 | 0,035039362 |
| 380282_LTR67B,MLT1N2 | chr7 | 150313854 | 150314406 | 3,8166 | 2,1136 | 0,013282194 |
| 38047_MLT1L,MLT1L | chr10 | 32983298 | 32984137 | 3,4550 | 1,7585 | 0,024236225 |
| 381187_MLT1B,MSTB,MLT1B | chr7 | 155206324 | 155207101 | 0,0000 | 3,2781 | 0,015690832 |
| 384429_PRIMA4_LTR,LTR19A,PRIMA4_LTR | chr8 | 16716812 | 16717922 | 1,2751 | 3,0744 | 0,005549477 |
| 386254_LTR56,LTR23 | chr8 | 27773106 | 27773577 | 2,3496 | 2,1929 | 0,010753143 |
| 386562_LTR85c,LTR104_Mam,LTR104_Mam | chr8 | 29847620 | 29848168 | 4,0275 | 1,6735 | 0,011558268 |
| 388297_THE1D | chr8 | 40159843 | 40160206 | 2,6639 | 2,0474 | 0,019868411 |
| 388301_MER4B | chr8 | 40176104 | 40176745 | 2,5094 | 2,4529 | 0,006275453 |
| 390728_MLT2B2,MSTD | chr8 | 58842232 | 58842528 | 2,7926 | 1,9426 | 0,041621538 |
| 390729_MER74B | chr8 | 58850173 | 58850725 | 3,6824 | 2,1556 | 0,005540463 |
| 391112_LTR32 | chr8 | 61722412 | 61722865 | 3,3868 | 2,7513 | 0,000466611 |
| 391203_MLT2B3,LTR9B,MLT2B3 | chr8 | 62162880 | 62164245 | 1,2751 | 2,8112 | 0,004777316 |
| 395815_MLT1F1 | chr8 | 91180080 | 91180616 | 0,0000 | 3,2503 | 0,008483028 |
| 39678_LTR12_,LTR12_ | chr10 | 47166757 | 47167766 | 0,0000 | 2,4462 | 0,039247284 |
| 398233_LTR26,LTR18A,ERVL-E-int,MLT1C,MLT1C,LTR16E2 | chr8 | 106059648 | 106063263 | 3,6863 | 1,8142 | 0,016615408 |
| 400369_MLT1C | chr8 | 118018108 | 118018589 | 0,8888 | 2,7340 | 0,036985469 |
| 405013_MER65D,MER65D | chr8 | 140465770 | 140466426 | 2,8265 | 2,1361 | 0,011283475 |
| 406606_LTR10D,HERVI-int,HERVI-int,LTR10D,MLT1J | chr9 | 6078707 | 6088395 | 0,4278 | 2,8248 | 0,031500349 |
| 410894_ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,MLT1A1,ERV3-16A3_I-int,ERV3-16A3_I-int | chr9 | 34158412 | 34161501 | 2,0847 | 2,5771 | 0,013832567 |
| 410895_MLT1B | chr9 | 34158616 | 34159001 | 2,1206 | 2,7375 | 0,008510381 |
| 421792_MLT1C | chr9 | 118337004 | 118337415 | 0,4278 | 3,0409 | 0,024724544 |
| 421840_MLT1D | chr9 | 118608790 | 118609212 | 0,0000 | 2,9504 | 0,024792122 |
| 422295_LTR86A1 | chr9 | 121325898 | 121326350 | 1,2898 | 3,0838 | 0,006210792 |
| 422579_LTR33 | chr9 | 124143487 | 124143581 | 0,0000 | 3,0101 | 0,023370091 |
| 422973_LTR42,LTR42 | chr9 | 128503574 | 128504231 | 0,0000 | 2,7724 | 0,048093083 |
| 430223_MamGypLTR1a,MamGypLTR1a | chrX | 30283642 | 30284239 | 1,0231 | 3,5784 | 0,001740012 |
| 433876_MER68B,MER68-int,MER68-int,LTR47A,MER68-int,MER68-int | chrX | 47499228 | 47503560 | 5,7505 | 1,4888 | 0,013948033 |
| 445565_MER61B,MER61-int,MER61-int,MER61-int | chrX | 141934261 | 141939406 | 7,8032 | -1,2081 | 0,031195538 |
| 46848_MER31A | chr10 | 93437917 | 93438368 | 3,7516 | 1,7999 | 0,024669361 |
| 48149_LTR79,LTR106_Mam,LTR106_Mam | chr10 | 105586599 | 105587547 | 0,0000 | 2,9673 | 0,035321335 |
| 53647_MER4C | chr11 | 8160917 | 8161342 | 1,1678 | 2,3902 | 0,039876186 |
| 57066_LTR4,HERV3-int,LTR4 | chr11 | 29518505 | 29527987 | 0,0000 | 2,6112 | 0,023706564 |
| 57763_MLT1A0 | chr11 | 33693282 | 33693613 | 0,8888 | 2,8074 | 0,02218126 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 61431_MSTD | chr11 | 60918249 | 60918594 | 2,4075 | -2,7296 | 0,04134651 |
| 61553_LTR3A,HERVK3-int,HERVK3-int,HERVK3-int,LTR3A | chr11 | 62049778 | 62056355 | 4,6331 | 1,2766 | 0,040775262 |
| 61939_LTR8 | chr11 | 65383794 | 65383913 | 0,0000 | 2,7967 | 0,03327574 |
| 62140_LTR23,LTR23,LTR23,ERV24_Prim-int,ERV24_Prim-int,ERV24_Prim-int,LTR49-int,LTR49-int,MER41A,LTR49-int,MER57A-int | chr11 | 67560725 | 67564653 | 0,0000 | 2,8166 | 0,022889378 |
| 63117_MER52D,MER52D | chr11 | 73796027 | 73797877 | 5,6645 | -1,2927 | 0,044718789 |
| 64913_MLT1J1 | chr11 | 86735412 | 86735518 | 0,0000 | 2,8075 | 0,038201291 |
| 66293_MLT1C | chr11 | 95877507 | 95877969 | 1,2898 | 2,7145 | 0,021141913 |
| 69226_LTR35A | chr11 | 114294850 | 114295386 | 4,3705 | 1,4785 | 0,036345316 |
| 70874_ERVL-B4-int,MLT2B1,MLT2B1 | chr11 | 128093965 | 128097207 | 4,8032 | 1,2825 | 0,048025437 |
| 72661_MER21C | chr12 | 8681348 | 8682152 | 9,8283 | 1,1980 | 0,035146551 |
| 72920_MLT1C,MER74B | chr12 | 10204729 | 10205892 | 3,4612 | 2,1256 | 0,010206803 |
| 74418_MLT1J2 | chr12 | 19946540 | 19946934 | 3,2091 | 1,6896 | 0,043696616 |
| 74423_LTR35A,LTR35A | chr12 | 19970993 | 19971796 | 7,6980 | 1,3309 | 0,024861415 |
| 76002_LTR1A2 | chr12 | 30696466 | 30697255 | 3,3496 | -2,2694 | 0,044462477 |
| 84191_MLT1F2 | chr12 | 88298777 | 88299346 | 0,0000 | 3,0297 | 0,029730605 |
| 85580_MLT1A0 | chr12 | 95806434 | 95806773 | 0,0000 | 3,2647 | 0,010537786 |
| 86370_LTR22B1,HERVK22-int,HERVK22-int,HERVK22-int,HERVK22-int,LTR22B1 | chr12 | 100159364 | 100165461 | 7,6938 | -1,1773 | 0,046448927 |
| 88362_MLT1J1 | chr12 | 111839719 | 111840110 | 6,8325 | -1,9190 | 0,007129046 |
| 90476_MER66B | chr12 | 124696652 | 124697106 | 1,8704 | 3,6706 | 0,000421598 |
| 92137_MLT1C | chr12 | 130133799 | 130134228 | 7,5923 | -2,1727 | 0,002544136 |
| 95363_MER21C | chr13 | 35927589 | 35928289 | 0,0000 | 3,0203 | 0,018960648 |

### Table 8.10 Significantly deregulated ERVs with H3K9ac peak in SF3A3 KD, Hela cells

| RepeatID | chr | start | end | Log2 base Mean | log2 Fold Change | p-value |
|---|---|---|---|---|---|---|
| 340497_MSTD,MER89 | chr6 | 56334668 | 56335782 | 0,0000 | 7,2397 | 2,48E-08 |
| 311159_HERVH-int,HERVH-int,HERVH-int,LTR7C,HERVIP10F-int,LTR10F | chr5 | 43329036 | 43336643 | 0,9169 | 6,6361 | 2,07E-06 |
| 65973_LTR76,MER90 | chr11 | 93716498 | 93717881 | 0,4278 | 6,5829 | 1,56E-06 |
| 57066_LTR4,HERV3-int,LTR4 | chr11 | 29518505 | 29527987 | 0,0000 | 6,2768 | 4,21E-06 |
| 62142_LTR49-int,LTR49-int,LTR49-int,LTR49-int,MER51E,LTR49-int,LTR13_,LTR49-int,LTR49-int,LTR49-int,ERV24B_Prim-int,LTR23 | chr11 | 67569605 | 67575312 | 0,0000 | 5,9141 | 1,13E-05 |
| 175394_LTR13 | chr19 | 48876221 | 48877217 | 0,0000 | 5,9131 | 1,71E-05 |
| 310721_MLT1A0 | chr5 | 40812184 | 40812545 | 0,0000 | 5,8508 | 9,70E-05 |
| 360617_MLT1E1A | chr7 | 17408187 | 17408704 | 0,4278 | 5,5928 | 1,16E-05 |
| 322253_LTR10F,LTR10F,LTR10F,LTR10F,MER41A,LTR10F,LTR10F,MER41A,MLT1D | chr5 | 115928854 | 115931428 | 2,6639 | 5,4980 | 3,21E-10 |
| 360616_MLT2B3 | chr7 | 17406566 | 17407073 | 0,0000 | 5,4358 | 1,94E-05 |
| 2452_LTR4 | chr1 | 16905463 | 16905857 | 5,6725 | 5,4112 | 4,64E-07 |
| 261606_MLT2C1,MLT1D-int,MLT1D,ERVL-B4-int | chr3 | 146967031 | 146970434 | 0,0000 | 5,2877 | 9,58E-05 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 76032_HUERS-P3-int,LTR9B,LTR9B,LTR13_,LTR9B | chr12 | 30850122 | 30852149 | 0,0000 | 5,1859 | 0,000208 |
| 365235_LTR56,LTR54,MER41B,LTR54,MER51B,MER51B | chr7 | 48085504 | 48088059 | 0,6945 | 5,1796 | 1,79E-05 |
| 62140_LTR23,LTR23,LTR23,ERV24_Prim-int,ERV24_Prim-int,ERV24_Prim-int,LTR49-int,LTR49-int,MER41A,LTR49-int,MER57A-int | chr11 | 67560725 | 67564653 | 0,0000 | 5,0914 | 0,00024 |
| 360605_MLT1D | chr7 | 17372108 | 17372580 | 0,0000 | 5,0340 | 0,000135 |
| 304500_MLT1E1A,MLT1E1A,LTR41 | chr5 | 10474219 | 10475470 | 0,0000 | 5,0246 | 0,000134 |
| 376985_MLT1C | chr7 | 128112087 | 128112329 | 0,5953 | 4,9773 | 8,39E-05 |
| 107172_MLT1A,MLT1A-int,MLT1A | chr13 | 102191290 | 102193373 | 0,0000 | 4,9438 | 0,001068 |
| 174809_MER41B | chr19 | 44587758 | 44588155 | 0,4278 | 4,8609 | 0,000515 |
| 335868_MER51A,MER4A1,MER4-int | chr6 | 24668520 | 24670953 | 0,9169 | 4,8377 | 0,000115 |
| 18275_LTR75B | chr1 | 147083988 | 147084093 | 2,1920 | 4,7452 | 1,65E-06 |
| 395815_MLT1F1 | chr8 | 91180080 | 91180616 | 0,0000 | 4,7067 | 0,000428 |
| 309230_MER31A | chr5 | 32653512 | 32653978 | 1,4511 | 4,7066 | 1,23E-06 |
| 169589_LTR38A1,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7C,MER92B | chr19 | 6574424 | 6577385 | 0,0000 | 4,5825 | 0,001029 |
| 76186_LTR8B,PrimLTR79 | chr12 | 31749068 | 31749678 | 0,8888 | 4,5561 | 9,77E-05 |
| 377329_MLT1H | chr7 | 131188268 | 131188771 | 1,4841 | 4,5483 | 1,24E-05 |
| 354253_MLT1D-int,MLT1D | chr6 | 146289966 | 146291094 | 1,0231 | 4,5466 | 0,000832 |
| 73309_MLT1N2 | chr12 | 12803118 | 12803632 | 1,0231 | 4,5430 | 0,000394 |
| 165954_MLT1C,LTR16E1 | chr18 | 62438346 | 62438841 | 2,0024 | 4,5284 | 1,63E-06 |
| 214802_MER65C | chr20 | 1393576 | 1394007 | 0,5953 | 4,5033 | 0,001304 |
| 207188_MLT1J1,MLT1J2 | chr2 | 197461148 | 197461427 | 0,0000 | 4,4515 | 0,00426 |
| 342969_LTR33B | chr6 | 74137805 | 74138271 | 0,0000 | 4,4488 | 0,026893 |
| 199835_MER74A,LTR8 | chr2 | 148813726 | 148815177 | 0,4278 | 4,4418 | 0,002145 |
| 61241_MER41C,MER41C,MER41C | chr11 | 59813317 | 59814385 | 0,0000 | 4,4180 | 0,001783 |
| 26859_LTR82B,LTR82B,LTR82A | chr1 | 210372339 | 210373878 | 7,3678 | 4,3641 | 2,26E-11 |
| 211390_LTR48B,LTR29,MER34 | chr2 | 224146256 | 224146903 | 0,0000 | 4,3612 | 0,002286 |
| 16077_LOR1b,MER51A,MER110-int | chr1 | 109648877 | 109651358 | 0,0000 | 4,3495 | 0,00873 |
| 332747_MLT1A0 | chr6 | 4932153 | 4932521 | 0,0000 | 4,3390 | 0,000604 |
| 174545_LTR8B | chr19 | 43265000 | 43265754 | 1,1678 | 4,3137 | 9,51E-06 |
| 28495_MER39,MER21A,MER21A,LTR48B | chr1 | 221838607 | 221839445 | 0,0000 | 4,2820 | 0,003596 |
| 224663_LTR12C | chr20 | 57348213 | 57349645 | 0,0000 | 4,2720 | 0,000715 |
| 64458_LTR30,HERV30-int,HERV30-int,LTR30 | chr11 | 83216718 | 83220622 | 0,5953 | 4,2607 | 0,010366 |
| 164702_THE1D | chr18 | 54967703 | 54968072 | 0,4278 | 4,2393 | 0,005306 |
| 137867_MER67D,THE1B,MER67D | chr16 | 21302435 | 21303327 | 0,0000 | 4,2315 | 0,008813 |
| 423318_LTR13 | chr9 | 132252083 | 132253091 | 0,0000 | 4,2312 | 0,021144 |
| 2448_LTR75B | chr1 | 16896666 | 16896830 | 6,1673 | 4,2069 | 5,69E-13 |
| 352494_MER4D1,MER4D1 | chr6 | 136068053 | 136069183 | 0,0000 | 4,1977 | 0,002389 |
| 237208_MER41C,LTR46,LTR46,MER41C,LTR7Y,MER41C | chr22 | 45160741 | 45162810 | 2,7291 | 4,1939 | 3,18E-06 |
| 302627_MLT2A1 | chr5 | 918576 | 919018 | 0,8888 | 4,1868 | 0,000295 |
| 174723_MLT1A0 | chr19 | 44164478 | 44164746 | 1,5664 | 4,1661 | 9,57E-05 |
| 223242_MLT1K,MLT1L | chr20 | 50439615 | 50440547 | 0,8888 | 4,1347 | 0,000253 |
| 119060_LTR12_,MLT1I | chr14 | 83556406 | 83557464 | 0,4278 | 4,1273 | 0,002477 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 397337_MER52A | chr8 | 100907235 | 100908779 | 1,8057 | 4,1060 | 3,29E-05 |
| 81998_LTR16D | chr12 | 75483410 | 75483538 | 1,0417 | 4,0640 | 0,000418 |
| 356264_MLT1K | chr6 | 159578009 | 159578187 | 0,4278 | 4,0597 | 0,001988 |
| 336983_MLT1G3,MLT1G3 | chr6 | 30767566 | 30768367 | 0,0000 | 4,0309 | 0,000776 |
| 66291_MLT1E2,MLT1E2 | chr11 | 95800020 | 95800796 | 1,2898 | 4,0273 | 0,001656 |
| 254500_LTR6A,HERVS71-int,LTR6A,MSTA | chr3 | 101632044 | 101642824 | 3,1201 | 3,9888 | 7,04E-08 |
| 420158_MLT2C1,MLT2C1 | chr9 | 107755729 | 107756199 | 0,6945 | 3,9818 | 0,003418 |
| 296967_MER34A | chr4 | 156950066 | 156950681 | 1,1678 | 3,9816 | 0,008904 |
| 47956_LTR36,MER4CL34 | chr10 | 104252956 | 104254461 | 0,0000 | 3,9662 | 0,002194 |
| 72088_LTR10F,LTR10F,MER31B | chr12 | 4319491 | 4320728 | 1,9586 | 3,9578 | 2,39E-05 |
| 305939_MamGypLTR2b | chr5 | 17375279 | 17375517 | 2,4757 | 3,9352 | 1,41E-06 |
| 421840_MLT1D | chr9 | 118608790 | 118609212 | 0,0000 | 3,9328 | 0,004972 |
| 129003_MLT1J | chr15 | 63134292 | 63134755 | 0,0000 | 3,9276 | 0,004917 |
| 78402_LTR2,HERVE_a-int,HERVE_a-int,Harlequin-int,Harlequin-int,Harlequin-int,LTR2,MamGypLTR1d,MamGypLTR1d | chr12 | 49733989 | 49740949 | 0,0000 | 3,9211 | 0,014343 |
| 149790_MLT1B | chr17 | 19111648 | 19111720 | 2,3064 | 3,9204 | 7,51E-05 |
| 406606_LTR10D,HERVI-int,HERVI-int,LTR10D,MLT1J | chr9 | 6078707 | 6088395 | 0,4278 | 3,9170 | 0,004667 |
| 422580_MER77B,LTR8,LTR8,MER77B | chr9 | 124144785 | 124146556 | 1,0231 | 3,9066 | 0,00077 |
| 421253_MER50,MER50 | chr9 | 114884685 | 114885308 | 0,0000 | 3,9060 | 0,01475 |
| 2397_LTR75 | chr1 | 16513454 | 16513604 | 5,2406 | 3,9040 | 6,28E-09 |
| 431166_LTR47A2,LTR22A | chrX | 35622090 | 35623063 | 2,5710 | 3,8865 | 1,77E-05 |
| 6901_LTR8,LTR8 | chr1 | 51557827 | 51558840 | 0,4278 | 3,8848 | 0,006644 |
| 159549_MLT1A1 | chr18 | 22468268 | 22468645 | 1,7631 | 3,8761 | 0,00035 |
| 54217_MER41B | chr11 | 12043904 | 12043975 | 0,0000 | 3,8710 | 0,012107 |
| 386254_LTR56,LTR23 | chr8 | 27773106 | 27773577 | 2,3496 | 3,8569 | 9,10E-06 |
| 289252_MLT1J2 | chr4 | 112084112 | 112084414 | 0,0000 | 3,8511 | 0,010118 |
| 108993_LTR73,MSTB1,MSTA1 | chr13 | 113188387 | 113189461 | 2,6017 | 3,8484 | 1,36E-05 |
| 246756_MER4B,MER4-int,MER4-int,MER4-int,MER4-int | chr3 | 52155573 | 52158226 | 1,8331 | 3,8363 | 0,000761 |
| 423802_MER31A | chr9 | 135962011 | 135962421 | 1,0231 | 3,8314 | 0,004954 |
| 233200_LTR13 | chr21 | 46143384 | 46144361 | 0,0000 | 3,8112 | 0,01549 |
| 373984_MLT1J1,LTR7,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7,MSTD | chr7 | 109641330 | 109648236 | 5,0827 | 3,8097 | 4,90E-07 |
| 362113_MER68 | chr7 | 27677894 | 27678417 | 0,4278 | 3,8057 | 0,005489 |
| 150361_LTR8,LTR8,LTR8 | chr17 | 27376102 | 27377435 | 0,0000 | 3,7894 | 0,010773 |
| 369010_MER21C | chr7 | 74097730 | 74098330 | 0,0000 | 3,7860 | 0,009235 |
| 146925_MLT1C,MLT1C | chr16 | 87551576 | 87552100 | 0,4278 | 3,7600 | 0,008801 |
| 150331_MLT1N2 | chr17 | 27244355 | 27244874 | 0,5953 | 3,7510 | 0,009094 |
| 29384_LTR16E2,LTR16E2,LTR16E2,LTR16,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,LTR16 | chr1 | 227356153 | 227360947 | 0,0000 | 3,7394 | 0,003972 |
| 310724_MLT1A1,MSTB,MLT1A1 | chr5 | 40821647 | 40822500 | 3,1555 | 3,7373 | 5,50E-05 |
| 6255_MER66-int,MER66-int,MER66C,MER66C,MER65A,MER65A | chr1 | 46332950 | 46335031 | 0,9169 | 3,7271 | 0,006142 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 126599_LTR13 | chr15 | 44784188 | 44785182 | 0,0000 | 3,7195 | 0,008323 |
| 372948_MER90,LTR64,LTR64,LTR64 | chr7 | 102394288 | 102395842 | 0,4278 | 3,7150 | 0,005109 |
| 74619_MLT2F,ERVL-E-int,ERVL-E-int,ERVL-E-int,ERVL-E-int | chr12 | 21492563 | 21499139 | 2,5830 | 3,7137 | 0,000139 |
| 207689_LTR7B | chr2 | 200862213 | 200862677 | 0,0000 | 3,7059 | 0,012318 |
| 338667_LTR7B | chr6 | 43370197 | 43370576 | 1,5664 | 3,7023 | 0,000327 |
| 185552_MLT1K | chr2 | 55622744 | 55623179 | 0,6945 | 3,7013 | 0,014204 |
| 214_MER65A,MER65A,MER65-int,MER65-int,MER65-int | chr1 | 1575524 | 1577276 | 0,0000 | 3,6974 | 0,014024 |
| 422545_MLT1J | chr9 | 123933309 | 123933813 | 1,0231 | 3,6896 | 0,002229 |
| 261962_MER74A,MLT1F2,MLT1F2 | chr3 | 149388603 | 149389300 | 0,0000 | 3,6863 | 0,004486 |
| 20071_MER4E1,MER4E1,MER4D1,MER4D1,MER4D1,MER41B,MER4D1,MER4D1,MER4B | chr1 | 162493563 | 162497426 | 0,5953 | 3,6832 | 0,012407 |
| 116963_LTR12C,LTR12C | chr14 | 70254543 | 70256182 | 0,0000 | 3,6749 | 0,013274 |
| 145703_MSTB1,MSTB1,MSTB1 | chr16 | 81181316 | 81182333 | 1,1386 | 3,6657 | 0,002685 |
| 94501_LTR1A2,ERV3-16A3_LTR,LTR78B,LTR78B | chr13 | 30689734 | 30691658 | 0,6945 | 3,6583 | 0,00654 |
| 191403_MER21B | chr2 | 96302454 | 96303153 | 3,8839 | 3,6533 | 8,51E-07 |
| 334825_LTR12C | chr6 | 17580547 | 17581971 | 0,4278 | 3,6450 | 0,008379 |
| 181448_LTR38B,LTR38B,LTR15,HERV15-int,HERV15-int,HERV15-int,HERV15-int,HERV15-int,LTR15 | chr2 | 26296968 | 26306395 | 3,5733 | 3,6305 | 1,09E-06 |
| 360654_MLT1G3,MLT1F2 | chr7 | 17598305 | 17599376 | 0,0000 | 3,6166 | 0,01246 |
| 180667_LTR5A | chr2 | 20501251 | 20502198 | 3,7708 | 3,6028 | 8,01E-09 |
| 360615_MamRep1527 | chr7 | 17405800 | 17405886 | 0,0000 | 3,5915 | 0,008634 |
| 174565_MSTA,LTR7Y,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7Y,THE1D-int,MSTA | chr19 | 43323083 | 43330524 | 8,6620 | 3,5862 | 0,000943 |
| 174810_MLT1F2 | chr19 | 44588555 | 44589012 | 0,0000 | 3,5817 | 0,018956 |
| 358216_MLT1G3,THE1B,THE1B | chr7 | 1520721 | 1521796 | 0,0000 | 3,5674 | 0,010865 |
| 311153_LTR16E2,ERVL-E-int,ERVL-E-int | chr5 | 43281268 | 43284094 | 4,0957 | 3,5491 | 3,66E-06 |
| 170407_LTR26E | chr19 | 12293899 | 12294377 | 0,0000 | 3,5485 | 0,026565 |
| 6163_LTR45C | chr1 | 45501583 | 45501680 | 1,0231 | 3,5404 | 0,00973 |
| 150332_LTR79 | chr17 | 27245527 | 27245775 | 0,0000 | 3,5319 | 0,013965 |
| 310110_MLT2A1,HERVL-int,HERVL-int,HERVL-int,HERVL-int,MLT2A1,MLT2A1,LTR40b | chr5 | 37264508 | 37272840 | 3,6397 | 3,5234 | 2,02E-05 |
| 303601_MLT1D | chr5 | 6355873 | 6356389 | 1,0231 | 3,5074 | 0,003626 |
| 239753_MER31A | chr3 | 8964026 | 8964477 | 0,4278 | 3,4899 | 0,01136 |
| 54337_ERVL-E-int,LTR50 | chr11 | 12985677 | 12986127 | 0,6945 | 3,4852 | 0,008358 |
| 45789_MLT1L | chr10 | 86972539 | 86972659 | 1,6370 | 3,4571 | 0,001165 |
| 1255_MER84,MER84 | chr1 | 7960454 | 7960847 | 3,3457 | 3,4503 | 8,51E-06 |
| 270431_LTR37B | chr4 | 6985402 | 6985819 | 1,0417 | 3,4445 | 0,004104 |
| 132288_LTR2B | chr15 | 88622388 | 88622894 | 0,5953 | 3,4387 | 0,011625 |
| 4954_MER67C | chr1 | 36465009 | 36465694 | 2,8973 | 3,4317 | 2,67E-05 |
| 309409_MER89,MER89 | chr5 | 33511325 | 33512159 | 0,0000 | 3,4245 | 0,048898 |
| 227993_MER21C | chr21 | 20720794 | 20721574 | 0,5953 | 3,4194 | 0,007574 |
| 219882_LTR46 | chr20 | 26010570 | 26010894 | 0,0000 | 3,4145 | 0,019753 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 305202_MLT1F2,MSTA1 | chr5 | 13567869 | 13568936 | 1,3447 | 3,4093 | 0,004041 |
| 398233_LTR26,LTR18A,ERVL-E-int,MLT1C,MLT1C,LTR16E2 | chr8 | 106059648 | 106063263 | 3,6863 | 3,4040 | 7,27E-06 |
| 117114_MER41B | chr14 | 71130672 | 71131303 | 0,0000 | 3,3826 | 0,018617 |
| 337160_MLT1C | chr6 | 31941161 | 31941422 | 0,0000 | 3,3794 | 0,030096 |
| 125474_THE1D | chr15 | 34588272 | 34588649 | 1,4841 | 3,3631 | 0,004912 |
| 258707_LTR8 | chr3 | 128878209 | 128878882 | 0,0000 | 3,3556 | 0,015311 |
| 93506_MLT1C-int,MLT1C | chr13 | 23889834 | 23890581 | 2,9285 | 3,3520 | 9,89E-05 |
| 294078_MER66B | chr4 | 139014288 | 139014343 | 1,0417 | 3,3514 | 0,003576 |
| 181493_MER41A | chr2 | 26757783 | 26758302 | 0,4278 | 3,3361 | 0,028806 |
| 18435_LTR75 | chr1 | 148264110 | 148264219 | 2,6017 | 3,3314 | 0,006598 |
| 233019_MER41C,MER41C | chr21 | 44844722 | 44845537 | 1,6114 | 3,3298 | 0,002721 |
| 114846_LTR39,MLT1C,LTR85c,MLT1I | chr14 | 55457023 | 55459401 | 0,0000 | 3,3091 | 0,02708 |
| 170296_MER57E3,MER4D,MER49,MER49,MER57E3,LTR24C,MER41D,MER41D,LTR24C,LTR24C,LTR24C,MER57E3 | chr19 | 11925630 | 11929740 | 0,0000 | 3,2998 | 0,049269 |
| 379341_MER92C,MER92C | chr7 | 144842031 | 144842865 | 0,0000 | 3,2972 | 0,008454 |
| 354942_THE1B | chr6 | 150419520 | 150419883 | 0,4278 | 3,2965 | 0,021721 |
| 375737_MLT1B | chr7 | 120583424 | 120583812 | 0,5953 | 3,2926 | 0,023383 |
| 196556_MER92C | chr2 | 127527074 | 127527554 | 3,7526 | 3,2874 | 8,71E-05 |
| 209010_LTR12C,LTR12C,MER39B | chr2 | 209999639 | 210001961 | 2,0464 | 3,2863 | 0,000411 |
| 259931_LTR33 | chr3 | 136818116 | 136818501 | 0,8888 | 3,2810 | 0,009497 |
| 3517_MER90a,MER90a,MER90a | chr1 | 25230593 | 25231344 | 4,4247 | 3,2774 | 4,94E-06 |
| 353214_LTR37-int,LTR37-int | chr6 | 140622038 | 140623451 | 0,5953 | 3,2728 | 0,049988 |
| 207698_MER21A | chr2 | 200906050 | 200906212 | 0,0000 | 3,2627 | 0,025122 |
| 60111_THE1D | chr11 | 47138055 | 47138417 | 2,1967 | 3,2531 | 0,004602 |
| 365657_LTR22A | chr7 | 50533873 | 50534328 | 0,5953 | 3,2520 | 0,013072 |
| 18706_LTR78,LTR78 | chr1 | 150604928 | 150605456 | 1,0231 | 3,2434 | 0,013083 |
| 311158_LTR10F,HERVIP10F-int,HERVIP10F-int,LTR7C,HERVH-int | chr5 | 43315210 | 43322830 | 1,3079 | 3,2365 | 0,01232 |
| 311427_LTR78,MLT1J | chr5 | 44875718 | 44876407 | 1,2751 | 3,2332 | 0,007734 |
| 337805_MLT1J | chr6 | 37243221 | 37243424 | 0,0000 | 3,2214 | 0,029803 |
| 19207_MLT1A | chr1 | 156590639 | 156590950 | 1,9637 | 3,2178 | 0,001021 |
| 435911_MER77B,MER77B | chrX | 68503623 | 68504625 | 0,0000 | 3,1973 | 0,034709 |
| 126305_LTR12C | chr15 | 40961432 | 40962899 | 3,2395 | 3,1950 | 9,48E-05 |
| 92666_LTR8A | chr12 | 132985524 | 132985862 | 0,4278 | 3,1944 | 0,036626 |
| 33729_MLT1A0,MLT1A0 | chr10 | 5292175 | 5292941 | 0,0000 | 3,1702 | 0,018556 |
| 158995_MLT1L | chr18 | 12598742 | 12598867 | 0,0000 | 3,1640 | 0,035158 |
| 236760_MER51B | chr22 | 41827412 | 41827761 | 0,0000 | 3,1572 | 0,027409 |
| 62001_MamGypLTR1b | chr11 | 66014825 | 66014875 | 0,4278 | 3,1529 | 0,039533 |
| 365234_MER4E1,MER4E1 | chr7 | 48083792 | 48084888 | 0,0000 | 3,1493 | 0,02856 |
| 39678_LTR12_,LTR12_ | chr10 | 47166757 | 47167766 | 0,0000 | 3,1469 | 0,011932 |
| 279039_LTR88b | chr4 | 48781227 | 48781390 | 0,0000 | 3,1415 | 0,028065 |
| 281911_LTR25 | chr4 | 66574695 | 66575495 | 0,4278 | 3,1401 | 0,010398 |
| 34930_MER51C | chr10 | 13587325 | 13587915 | 2,7882 | 3,1228 | 0,002772 |
| 301711_MLT1D | chr4 | 185204548 | 185204975 | 1,1678 | 3,1220 | 0,010628 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 175372_LTR13A,LTR13A | chr19 | 48753519 | 48754784 | 1,0231 | 3,1133 | 0,049374 |
| 207621_MLT1J2,MLT1J2,MLT1I,MLT1I-int | chr2 | 200302301 | 200303390 | 1,0417 | 3,1060 | 0,011167 |
| 30798_LTR8A | chr1 | 235366295 | 235366990 | 1,0417 | 3,1025 | 0,010239 |
| 310748_THE1A,THE1A-int,THE1A | chr5 | 40881493 | 40883717 | 0,0000 | 3,0912 | 0,049854 |
| 208583_LTR44 | chr2 | 207394694 | 207395169 | 0,4278 | 3,0798 | 0,04409 |
| 445760_MLT1J2 | chrX | 142998605 | 142998826 | 1,8057 | 3,0743 | 0,002715 |
| 193827_MER21C | chr2 | 111659358 | 111660135 | 1,2751 | 3,0713 | 0,005868 |
| 176815_MER89 | chr19 | 57241086 | 57241457 | 3,5605 | 3,0697 | 0,000188 |
| 184089_MLT1F | chr2 | 46224753 | 46225256 | 1,4841 | 3,0680 | 0,004625 |
| 361335_LTR38A1 | chr7 | 22598761 | 22598946 | 1,0231 | 3,0580 | 0,013117 |
| 430223_MamGypLTR1a,MamGypLTR1a | chrX | 30283642 | 30284239 | 1,0231 | 3,0578 | 0,006751 |
| 122249_MamGypLTR3,MamGypLTR2b | chr14 | 99682655 | 99683290 | 1,6114 | 3,0211 | 0,006844 |
| 309220_ERVL-E-int,LTR33 | chr5 | 32610578 | 32611620 | 5,1944 | 3,0138 | 2,84E-07 |
| 399819_MLT1D | chr8 | 114936383 | 114936917 | 0,0000 | 3,0093 | 0,031684 |
| 158947_MER51B,MER76 | chr18 | 11955271 | 11955923 | 1,0417 | 3,0072 | 0,021321 |
| 431333_MSTD | chrX | 36501278 | 36501521 | 0,0000 | 3,0041 | 0,044952 |
| 153843_LTR71A | chr17 | 58632587 | 58633010 | 1,3079 | 2,9805 | 0,008894 |
| 108985_MSTD,MSTD | chr13 | 113154795 | 113155189 | 2,0024 | 2,9762 | 0,00457 |
| 309136_LTR83 | chr5 | 32072929 | 32073157 | 0,8888 | 2,9732 | 0,019076 |
| 70874_ERVL-B4-int,MLT2B1,MLT2B1 | chr11 | 128093965 | 128097207 | 4,8032 | 2,9697 | 4,41E-06 |
| 48333_MER92B | chr10 | 106500482 | 106501156 | 0,9169 | 2,9645 | 0,022367 |
| 233822_MER4D1,MER4D1 | chr22 | 19123202 | 19124341 | 1,5664 | 2,9584 | 0,018233 |
| 242004_LTR33 | chr3 | 23042038 | 23042500 | 0,6945 | 2,9576 | 0,013128 |
| 405633_MER39B,MER39B | chr8 | 144755893 | 144756393 | 2,4488 | 2,9352 | 0,002066 |
| 207723_MLT1J2 | chr2 | 201150528 | 201150886 | 3,9508 | 2,9304 | 6,05E-05 |
| 149502_LTR48B | chr17 | 16463067 | 16463145 | 1,3447 | 2,9297 | 0,027651 |
| 208402_MER34C2,MER34C2 | chr2 | 206211783 | 206212065 | 0,0000 | 2,9285 | 0,046708 |
| 61210_MER34A1,MER34A1 | chr11 | 59561187 | 59562092 | 1,0231 | 2,9143 | 0,031545 |
| 190424_LTR3B,LTR3B_,MLT1F2-int,MLT1F2 | chr2 | 86703756 | 86704999 | 0,0000 | 2,9137 | 0,044497 |
| 369246_MER31A,MER31A | chr7 | 76293518 | 76294272 | 1,4841 | 2,8992 | 0,008434 |
| 76182_LTR33,LTR8,LTR8,LTR33 | chr12 | 31733069 | 31734503 | 0,8888 | 2,8970 | 0,043206 |
| 210495_MLT1O,MER39,MER39 | chr2 | 219595280 | 219596817 | 4,1348 | 2,8821 | 0,000322 |
| 85580_MLT1A0 | chr12 | 95806434 | 95806773 | 0,0000 | 2,8796 | 0,020495 |
| 149675_MER21B | chr17 | 17958669 | 17959343 | 0,6945 | 2,8785 | 0,044252 |
| 89933_LTR37B | chr12 | 120290225 | 120290296 | 3,0398 | 2,8594 | 0,002741 |
| 126730_MLT1M | chr15 | 45614980 | 45615181 | 0,0000 | 2,8562 | 0,047188 |
| 76190_LTR8B,LTR8B | chr12 | 31756862 | 31757926 | 0,0000 | 2,8441 | 0,022108 |
| 371535_HERV1_LTRb | chr7 | 92135482 | 92135990 | 2,1920 | 2,8386 | 0,005732 |
| 113975_LTR23,LTR23 | chr14 | 49619341 | 49620016 | 6,5269 | 2,8379 | 2,74E-07 |
| 6204_MLT1B,LTR37A,LTR37A,MER92-int,MER92-int,MER92-int,MER92-int,MER92-int | chr1 | 45890280 | 45894714 | 0,5953 | 2,8351 | 0,036102 |
| 336134_MER68,LTR3,HERVK3-int,HERVK3-int,LTR3 | chr6 | 26287804 | 26296266 | 0,0000 | 2,8337 | 0,027014 |
| 405517_MER57B2 | chr8 | 143684040 | 143684157 | 2,4488 | 2,8314 | 0,001847 |
| 176442_MLT1D,MLT1D,MLT1D,MLT1D | chr19 | 55060856 | 55062242 | 3,1928 | 2,8296 | 0,000787 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 206016_LTR16A | chr2 | 190607881 | 190607928 | 0,6945 | 2,8287 | 0,034404 |
| 223434_MLT1L | chr20 | 51351034 | 51351189 | 0,6945 | 2,8276 | 0,022748 |
| 128726_LTR37B,LTR37B | chr15 | 60372423 | 60373112 | 0,9169 | 2,8272 | 0,020225 |
| 220616_MLT1O,MLT1K | chr20 | 35560810 | 35561593 | 0,6945 | 2,8146 | 0,046145 |
| 107272_MER4A1 | chr13 | 102844793 | 102845223 | 3,2630 | 2,8066 | 0,000506 |
| 66195_MSTD | chr11 | 95283790 | 95284174 | 1,0231 | 2,8050 | 0,02484 |
| 199036_MER72 | chr2 | 142875416 | 142876090 | 0,0000 | 2,8034 | 0,014237 |
| 189768_LTR1D | chr2 | 82790255 | 82791056 | 0,5953 | 2,8000 | 0,029077 |
| 247571_LTR12C | chr3 | 59050675 | 59052063 | 0,0000 | 2,7815 | 0,043542 |
| 219339_MER67C,MER110A | chr20 | 23360499 | 23361459 | 2,5897 | 2,7736 | 0,003782 |
| 256915_LTR7,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7,LTR2 | chr3 | 115793482 | 115800126 | 0,6945 | 2,7704 | 0,024104 |
| 84191_MLT1F2 | chr12 | 88298777 | 88299346 | 0,0000 | 2,7665 | 0,043428 |
| 61939_LTR8 | chr11 | 65383794 | 65383913 | 0,0000 | 2,7631 | 0,035452 |
| 233263_MER4-int,MER4-int,MER4A1,MER4B | chr21 | 46631792 | 46635087 | 3,4907 | 2,7631 | 0,000755 |
| 125527_MER76 | chr15 | 34986472 | 34987062 | 5,3606 | 2,7595 | 4,16E-06 |
| 422579_LTR33 | chr9 | 124143487 | 124143581 | 0,0000 | 2,7555 | 0,035203 |
| 53257_LTR75_1,LTR12C,LTR75_1 | chr11 | 5682462 | 5684428 | 0,0000 | 2,7550 | 0,029648 |
| 48149_LTR79,LTR106_Mam,LTR106_Mam | chr10 | 105586599 | 105587547 | 0,0000 | 2,7441 | 0,047894 |
| 126575_LTR22B1 | chr15 | 44588523 | 44588976 | 2,4137 | 2,7428 | 0,011366 |
| 151862_MER41B | chr17 | 42535067 | 42535524 | 1,4449 | 2,7418 | 0,012278 |
| 61968_MER90a | chr11 | 65722145 | 65722623 | 0,8888 | 2,7411 | 0,025463 |
| 175172_MER66B,MER66B | chr19 | 47289205 | 47289947 | 1,5601 | 2,7181 | 0,018327 |
| 12835_MER51A,MER51A,MER51A | chr1 | 88892399 | 88893266 | 0,0000 | 2,7161 | 0,039325 |
| 234710_MLT2A1 | chr22 | 26511033 | 26511483 | 4,0718 | 2,7050 | 8,70E-05 |
| 362853_MER76 | chr7 | 32669162 | 32669692 | 0,6945 | 2,7042 | 0,030777 |
| 269292_MER4A1 | chr4 | 498234 | 498501 | 5,4122 | 2,7021 | 2,89E-06 |
| 369193_LTR26 | chr7 | 75878190 | 75878797 | 2,3496 | 2,6890 | 0,004229 |
| 290373_MER65A | chr4 | 118590733 | 118591172 | 0,0000 | 2,6886 | 0,047894 |
| 175003_HERVH-int,HERVH-int,LTR7C | chr19 | 46201640 | 46203386 | 10,6610 | 2,6880 | 1,69E-05 |
| 76187_PrimLTR79,LTR29,LOR1a,LTR29 | chr12 | 31750190 | 31751223 | 0,0000 | 2,6878 | 0,027881 |
| 262126_MER90a,MER90a,HERVH-int,HERVH-int,LTR7B | chr3 | 150600113 | 150602459 | 3,6206 | 2,6837 | 0,004119 |
| 130113_ERV3-16A3_I-int,ERV3-16A3_I-int | chr15 | 72618776 | 72620666 | 5,0493 | 2,6811 | 1,39E-05 |
| 78659_MLT1C,MLT1C | chr12 | 53056040 | 53056776 | 2,5830 | 2,6787 | 0,00528 |
| 245924_MER70C | chr3 | 44509537 | 44509934 | 2,0847 | 2,6729 | 0,01044 |
| 445311_MamGyp-int,MamGyp-int | chrX | 140712043 | 140712665 | 0,4278 | 2,6627 | 0,036323 |
| 235981_THE1D | chr22 | 35230880 | 35231246 | 4,4786 | 2,6600 | 0,000142 |
| 294566_MLT1L,MLT1G | chr4 | 142092587 | 142093449 | 0,4278 | 2,6584 | 0,02549 |
| 151945_MER11C | chr17 | 43365611 | 43366679 | 3,4619 | 2,6531 | 0,002585 |
| 170016_LTR7C,HERVH-int,HERVH-int | chr19 | 9539342 | 9540733 | 3,5920 | 2,6522 | 0,004923 |
| 33728_MSTD,MSTB1 | chr10 | 5291580 | 5292813 | 4,1531 | 2,6483 | 0,000237 |
| 156910_MLT1B,MLT1B,MLT1B | chr17 | 82287795 | 82288442 | 0,8888 | 2,6243 | 0,039796 |
| 9901_LTR46 | chr1 | 71081717 | 71082102 | 2,8961 | 2,6239 | 0,002807 |
| 352345_MLT1A0 | chr6 | 135049088 | 135049438 | 2,4757 | 2,6197 | 0,00619 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 26866_ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,LTR16A,ERVL-B4-int,MLT1N2 | chr1 | 210436678 | 210441114 | 2,3749 | 2,6185 | 0,001254 |
| 384436_THE1B | chr8 | 16749475 | 16749828 | 5,7223 | 2,6167 | 7,93E-06 |
| 285250_MLT1E | chr4 | 86929098 | 86929646 | 1,7631 | 2,6138 | 0,008492 |
| 54336_LTR33,ERVL-E-int | chr11 | 12983723 | 12985098 | 1,8331 | 2,6082 | 0,01482 |
| 180571_LTR37A | chr2 | 20052516 | 20052735 | 1,2898 | 2,5924 | 0,038012 |
| 111945_MER4C,MER4-int,MER4-int,MER4-int,MER4-int,MER4-int,MER4C | chr14 | 39376616 | 39382803 | 1,6370 | 2,5902 | 0,023774 |
| 232916_LTR8A,LTR8,LTR8A | chr21 | 44339862 | 44341132 | 4,7817 | 2,5887 | 0,001818 |
| 330563_MLT1B | chr5 | 173456048 | 173456471 | 0,8888 | 2,5860 | 0,011282 |
| 417931_LTR36,LTR36,MER50 | chr9 | 94138049 | 94139311 | 0,5953 | 2,5833 | 0,046334 |
| 259086_MLT1F | chr3 | 131503400 | 131503801 | 1,3447 | 2,5818 | 0,049175 |
| 128281_MER65C | chr15 | 56734076 | 56734537 | 0,0000 | 2,5798 | 0,035622 |
| 3316_MLT1F,LTR2C,LTR2C,MLT1F,MLT1F1,MLT1F1,MLT1F1 | chr1 | 23774095 | 23775687 | 3,4006 | 2,5702 | 0,001909 |
| 313328_MER34D | chr5 | 60700947 | 60701269 | 2,5481 | 2,5685 | 0,010747 |
| 61999_LTR3A | chr11 | 66010024 | 66010487 | 1,6370 | 2,5628 | 0,026944 |
| 114113_LTR90B | chr14 | 50669136 | 50669234 | 1,9586 | 2,5626 | 0,018568 |
| 390728_MLT2B2,MSTD | chr8 | 58842232 | 58842528 | 2,7926 | 2,5602 | 0,007538 |
| 433318_MER31B,MER31B | chrX | 44543239 | 44543999 | 2,1743 | 2,5584 | 0,011511 |
| 353948_MER110A | chr6 | 144922959 | 144923148 | 0,8888 | 2,5576 | 0,047671 |
| 233034_MSTB | chr21 | 45004930 | 45005334 | 2,4757 | 2,5531 | 0,004634 |
| 147462_LTR79 | chr17 | 2975524 | 2975946 | 6,5238 | 2,5441 | 1,27E-06 |
| 189516_LTR2752 | chr2 | 81481147 | 81482372 | 2,6276 | 2,5417 | 0,005356 |
| 207537_ERV3-16A3_LTR,ERV3-16A3_LTR | chr2 | 199941703 | 199942424 | 0,9169 | 2,5367 | 0,041263 |
| 52858_MER4D1,MER4D1 | chr11 | 2397539 | 2398606 | 4,4710 | 2,5362 | 0,001301 |
| 65875_LTR2,Harlequin-int,Harlequin-int,LTR2 | chr11 | 93103523 | 93110857 | 3,0186 | 2,5312 | 0,017477 |
| 326114_MER4A1 | chr5 | 143263724 | 143264164 | 1,6540 | 2,5297 | 0,018555 |
| 303333_MER21C,MER21C | chr5 | 5351314 | 5352309 | 1,2898 | 2,5232 | 0,031842 |
| 236741_LTR9,LTR9 | chr22 | 41683804 | 41684287 | 4,4536 | 2,5202 | 0,000278 |
| 19171_MLT1A | chr1 | 155932071 | 155932116 | 3,6046 | 2,5177 | 0,001162 |
| 54333_LTR33A,LTR79,LTR79 | chr11 | 12962647 | 12963475 | 1,6370 | 2,5093 | 0,02703 |
| 414587_MER41B | chr9 | 76414364 | 76415007 | 1,6114 | 2,5057 | 0,038836 |
| 310343_LTR9D | chr5 | 38660104 | 38660746 | 2,6127 | 2,5025 | 0,010942 |
| 185243_MLT1D | chr2 | 53403376 | 53403849 | 0,0000 | 2,4987 | 0,046807 |
| 364707_MER67C,MER67C,MER67C,MER67C | chr7 | 44987085 | 44987850 | 3,8839 | 2,4983 | 0,005799 |
| 381107_ERV24B_Prim-int,ERV24B_Prim-int,LTR48 | chr7 | 154922857 | 154929139 | 1,6370 | 2,4929 | 0,031334 |
| 229896_LTR50,LTR33A | chr21 | 28993609 | 28994214 | 1,7631 | 2,4914 | 0,014537 |
| 170045_LTR10F | chr19 | 9673925 | 9674425 | 6,7400 | 2,4909 | 2,56E-07 |
| 303619_MLT1A1 | chr5 | 6409005 | 6409403 | 2,3060 | 2,4830 | 0,00532 |
| 29183_LTR65 | chr1 | 225996608 | 225997251 | 6,0931 | 2,4813 | 9,02E-05 |
| 309216_MLT1A0,LTR85c,MamGypLTR2b,MSTA,MSTA | chr5 | 32579594 | 32581319 | 1,1678 | 2,4690 | 0,039808 |
| 422905_MLT1C,MER21A,MER21A,MER21A | chr9 | 127437514 | 127439241 | 2,7926 | 2,4642 | 0,009009 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 32569_LTR2 | chr1 | 247127448 | 247127909 | 4,5186 | 2,4597 | 0,000328 |
| 423452_LTR107_Mam | chr9 | 133337105 | 133337232 | 2,3496 | 2,4591 | 0,021869 |
| 384429_PRIMA4_LTR,LTR19A,PRIMA4_LTR | chr8 | 16716812 | 16717922 | 1,2751 | 2,4513 | 0,026439 |
| 202801_MER57B1 | chr2 | 169802207 | 169802278 | 2,0464 | 2,4493 | 0,026206 |
| 257557_MER48,LTR9B,LTR10F,HERVIP10FH-int,HERVIP10FH-int,HERVIP10FH-int,LTR10F | chr3 | 120615854 | 120623770 | 5,2938 | 2,4443 | 0,001576 |
| 88119_LTR6A | chr12 | 109901363 | 109901927 | 3,0278 | 2,4363 | 0,020832 |
| 152503_LTR33,LTR33 | chr17 | 50376277 | 50376888 | 2,7311 | 2,4290 | 0,022782 |
| 87450_LTR85b,LTR85b,LTR85b,LTR85b | chr12 | 106030919 | 106031829 | 1,8331 | 2,4190 | 0,037462 |
| 223304_MLT1O | chr20 | 50958125 | 50958171 | 3,8855 | 2,4185 | 0,001664 |
| 116616_MER67B | chr14 | 67357605 | 67358230 | 6,1853 | 2,4171 | 3,12E-05 |
| 386311_MER4D1,MER4D1,MER4D1,MER4CL34,MER4CL34 | chr8 | 28090664 | 28092598 | 3,3133 | 2,4166 | 0,004584 |
| 204_LTR10A,LTR10A | chr1 | 1470518 | 1471375 | 3,7835 | 2,3991 | 0,006941 |
| 224734_MER41C | chr20 | 57749591 | 57750106 | 1,4841 | 2,3977 | 0,049104 |
| 134074_MSTC | chr15 | 100601187 | 100601552 | 5,0902 | 2,3975 | 0,000354 |
| 358921_LTR37A,LTR37A,LTR37-int,LTR37-int,LTR37-int | chr7 | 6707620 | 6709689 | 1,8727 | 2,3894 | 0,03364 |
| 302626_MER9a2 | chr5 | 909535 | 910048 | 2,8660 | 2,3875 | 0,003794 |
| 6165_LTR34,LTR34,LTR34 | chr1 | 45522400 | 45523599 | 5,1251 | 2,3837 | 0,001141 |
| 28498_MLT1A1 | chr1 | 221856528 | 221856873 | 2,4757 | 2,3784 | 0,03023 |
| 6256_MER41A | chr1 | 46339833 | 46340059 | 1,9586 | 2,3710 | 0,016222 |
| 134343_LTR8,LTR8,LTR8 | chr16 | 1421273 | 1422556 | 4,7565 | 2,3645 | 0,000955 |
| 52781_MER4D0,MER4D0 | chr11 | 1764351 | 1764718 | 1,6370 | 2,3609 | 0,030382 |
| 168896_MER65C,MER65C,MER65C | chr18 | 79988964 | 79989672 | 3,6046 | 2,3591 | 0,001403 |
| 224466_MER21C | chr20 | 56470524 | 56471246 | 6,1030 | 2,3452 | 9,82E-06 |
| 190180_MLT2B3,MLT2B3 | chr2 | 84919720 | 84920696 | 5,6149 | 2,3352 | 2,00E-05 |
| 82681_MER41A | chr12 | 79546312 | 79546839 | 3,1190 | 2,3343 | 0,009551 |
| 117975_MER66B,MER21B | chr14 | 77758400 | 77759142 | 5,4701 | 2,3322 | 0,000197 |
| 305792_HERV16-int,HERV16-int,LTR16 | chr5 | 16439974 | 16441746 | 3,9129 | 2,3321 | 0,000924 |
| 144250_LTR33,LTR33 | chr16 | 72091104 | 72091526 | 3,0278 | 2,3306 | 0,018379 |
| 254229_LTR16A2 | chr3 | 99503437 | 99503786 | 0,0000 | 2,3284 | 0,0452 |
| 417084_MER41A | chr9 | 89308866 | 89309412 | 3,9344 | 2,3225 | 0,003169 |
| 61648_LTR19A,LTR19A,LTR19A | chr11 | 62846498 | 62847374 | 3,3596 | 2,3086 | 0,004818 |
| 64434_LTR54 | chr11 | 82908539 | 82908994 | 3,7068 | 2,3024 | 0,004581 |
| 336590_LTR25 | chr6 | 28335119 | 28335915 | 3,7022 | 2,3021 | 0,003034 |
| 107084_MLT1A | chr13 | 101490036 | 101490411 | 3,4907 | 2,3002 | 0,014609 |
| 128150_MLT1A1 | chr15 | 55791381 | 55791725 | 1,7631 | 2,2995 | 0,048918 |
| 61552_MER4E,LTR7Y,MER4E | chr11 | 62044620 | 62045792 | 5,0130 | 2,2986 | 0,002096 |
| 78948_LTR24B | chr12 | 56725875 | 56726175 | 2,3538 | 2,2854 | 0,026681 |
| 330838_MER77B | chr5 | 174829605 | 174830167 | 1,8581 | 2,2831 | 0,018387 |
| 147335_LTR13A | chr17 | 708055 | 708969 | 5,0922 | 2,2786 | 0,000126 |
| 159123_ERVL-E-int,MLT1C,ERVL-E-int,ERVL-E-int,ERVL-E-int,MLT2C1,MLT2C1 | chr18 | 13720053 | 13723766 | 7,1330 | 2,2569 | 1,28E-05 |
| 395416_THE1D | chr8 | 88959291 | 88959654 | 1,6370 | 2,2523 | 0,030257 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 154650_MLT1H2,MLT2B4,MLT2B4,MLT1H2 | chr17 | 66509019 | 66509784 | 3,1846 | 2,2461 | 0,006206 |
| 149501_MLT1A0 | chr17 | 16459972 | 16460309 | 5,0915 | 2,2442 | 0,001607 |
| 134349_LTR56 | chr16 | 1475641 | 1475812 | 4,4457 | 2,2430 | 0,000392 |
| 5179_MER34C2 | chr1 | 38011666 | 38012045 | 2,7278 | 2,2394 | 0,011546 |
| 53707_LTR9B,LTR9B | chr11 | 8700745 | 8701663 | 6,9983 | 2,2289 | 0,001468 |
| 131492_MLT1C | chr15 | 82749521 | 82749958 | 3,0931 | 2,2272 | 0,020265 |
| 53768_MER4A1 | chr11 | 9364993 | 9365144 | 3,0375 | 2,2209 | 0,012972 |
| 359105_MLT1A0 | chr7 | 7911209 | 7911546 | 1,4841 | 2,2175 | 0,043562 |
| 156711_LTR37B | chr17 | 80148136 | 80148210 | 2,1206 | 2,2009 | 0,045193 |
| 62316_LTR48B,LTR48B,MER57A-int,MER57A-int,MER57A-int,MER57A-int,MER57A-int,MER101-int,MER101-int,MER54B | chr11 | 68884446 | 68888758 | 4,5301 | 2,1974 | 0,001429 |
| 29338_MER4C | chr1 | 226938812 | 226939189 | 2,1803 | 2,1936 | 0,029883 |
| 214326_MSTA1,MSTA1 | chr2 | 240584103 | 240584856 | 5,1585 | 2,1772 | 0,000568 |
| 358299_MER21C | chr7 | 2359154 | 2359905 | 5,8282 | 2,1713 | 5,96E-05 |
| 309395_MLT1D,MLT1D | chr5 | 33480481 | 33480805 | 6,0078 | 2,1446 | 0,000344 |
| 237335_MER4B,MER4B | chr22 | 46334637 | 46335459 | 4,1498 | 2,1421 | 0,005742 |
| 309392_LTR8B | chr5 | 33476702 | 33477479 | 6,7500 | 2,1377 | 8,99E-06 |
| 157053_MER49,MER49 | chr18 | 268925 | 270113 | 5,3492 | 2,1287 | 0,000771 |
| 245428_MER4D1,MER4D1,LTR18B,MER4D1,LTR7B,MER4D1 | chr3 | 41962833 | 41965130 | 5,0890 | 2,1210 | 0,001402 |
| 169358_MER68 | chr19 | 4248941 | 4249144 | 3,3461 | 2,1167 | 0,009778 |
| 1712_MER21A | chr1 | 11984495 | 11985279 | 3,5174 | 2,1149 | 0,005088 |
| 354997_LTR10E | chr6 | 150856969 | 150857593 | 3,5876 | 2,1128 | 0,005278 |
| 256914_LTR12C,LTR12C,MLT2A1 | chr3 | 115791379 | 115793165 | 4,4545 | 2,1009 | 0,00185 |
| 72920_MLT1C,MER74B | chr12 | 10204729 | 10205892 | 3,4612 | 2,0990 | 0,01141 |
| 43901_LTR54,LTR54 | chr10 | 75209150 | 75209895 | 3,7280 | 2,0989 | 0,014652 |
| 236641_MLT1K | chr22 | 40651455 | 40651837 | 2,1920 | 2,0962 | 0,039507 |
| 405695_HERVK13-int,HERVK13-int,HERVK13-int,HERVK13-int,LTR13 | chr8 | 145058310 | 145066800 | 8,7070 | 2,0906 | 0,000232 |
| 76181_MER65D,MER65D | chr12 | 31730202 | 31731056 | 5,7414 | 2,0805 | 0,007723 |
| 330562_LTR16C,MLT1A-int,MLT1A,MLT1F1,MLT1F1,MLT1F1,MLT1F1 | chr5 | 173453476 | 173455503 | 6,7414 | 2,0752 | 0,000254 |
| 29512_MER21B,MER21B | chr1 | 228085857 | 228087054 | 7,4290 | 2,0647 | 0,000391 |
| 245761_LTR16A1 | chr3 | 43734082 | 43734486 | 2,9004 | 2,0582 | 0,01013 |
| 5958_ERVL-B4-int,LTR33,LTR33 | chr1 | 43472109 | 43473182 | 3,6477 | 2,0561 | 0,012905 |
| 177146_MER77B | chr2 | 270619 | 271027 | 3,6824 | 2,0559 | 0,008711 |
| 224667_MLT1B | chr20 | 57384943 | 57385325 | 2,6639 | 2,0506 | 0,043887 |
| 218431_MLT2F | chr20 | 18793953 | 18794451 | 3,3723 | 2,0454 | 0,013039 |
| 220598_LTR5_Hs | chr20 | 35262123 | 35263091 | 2,8167 | 2,0305 | 0,024866 |
| 76276_LTR24B | chr12 | 31958464 | 31958860 | 3,7257 | 2,0224 | 0,011061 |
| 109846_LTR16E1,LTR42,LTR16E1 | chr14 | 24308661 | 24309370 | 6,3950 | 2,0142 | 0,000322 |
| 336487_MER51D | chr6 | 27837034 | 27837666 | 7,6363 | 2,0136 | 0,000535 |
| 174927_MLT1D | chr19 | 45576557 | 45576946 | 3,7806 | 2,0072 | 0,00694 |
| 339401_MER54B | chr6 | 48827340 | 48827491 | 3,0112 | 2,0041 | 0,04221 |
| 130354_MSTB | chr15 | 74205225 | 74205636 | 4,7363 | -5,0145 | 0,000105 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 354424_MLT1C | chr6 | 147137790 | 147138276 | 7,2391 | -4,5513 | 1,21E-08 |
| 130042_MamRep1527,MamRep1527 | chr15 | 71742553 | 71743434 | 4,0532 | -4,3620 | 0,000284 |
| 371572_MLT2F | chr7 | 92461138 | 92461486 | 4,0521 | -4,2268 | 0,001129 |
| 354388_MamGypsy2-I,MamGypsy2-I,MamGypsy2-I,MamGypsy2-I | chr6 | 146968034 | 146969888 | 4,6199 | -4,1848 | 0,000772 |
| 43975_LTR101_Mam | chr10 | 75533060 | 75533796 | 5,1542 | -4,1520 | 8,95E-05 |
| 311099_MER4B | chr5 | 43011374 | 43011953 | 4,5886 | -4,0998 | 0,001445 |
| 427907_LTR16C | chrX | 17559556 | 17560007 | 3,6212 | -3,9337 | 0,001193 |
| 225655_MSTA | chr20 | 63763318 | 63763712 | 3,6012 | -3,8181 | 0,001809 |
| 154372_MER41-int,MER41-int,MER41-int,MER41-int,MER41-int,MER41-int,LTR8,LTR8,MER41-int | chr17 | 64421421 | 64424385 | 3,8203 | -3,7664 | 0,010449 |
| 130355_MLT1J | chr15 | 74210732 | 74210968 | 5,7248 | -3,6840 | 6,01E-05 |
| 86370_LTR22B1,HERVK22-int,HERVK22-int,HERVK22-int,LTR22B1 | chr12 | 100159364 | 100165461 | 7,6938 | -3,6678 | 7,13E-09 |
| 207952_MER92-int,MER92-int,MER92-int | chr2 | 203221784 | 203222618 | 5,1728 | -3,5747 | 0,000109 |
| 72351_MER57B2 | chr12 | 6368130 | 6368540 | 3,2755 | -3,4881 | 0,007794 |
| 270080_MER31B | chr4 | 5371060 | 5371517 | 2,9508 | -3,4595 | 0,00513 |
| 365153_LTR1A2,LTR1A2,LTR1A2,HUERS-P2-int,LTR1A2 | chr7 | 47595689 | 47600882 | 4,2868 | -3,4147 | 0,002443 |
| 205157_MLT1E1A,MLT1E1A-int | chr2 | 185689133 | 185690788 | 3,1928 | -3,3863 | 0,004361 |
| 146520_MLT1D | chr16 | 85363956 | 85364451 | 2,9921 | -3,3593 | 0,006049 |
| 64931_LTR33 | chr11 | 86792065 | 86792537 | 3,9459 | -3,2961 | 0,003893 |
| 271856_MER4B,LTR41B,LTR41B | chr4 | 13653432 | 13654981 | 3,0278 | -3,2593 | 0,008671 |
| 115766_MLT1J | chr14 | 61588236 | 61588642 | 6,1387 | -3,2369 | 5,35E-05 |
| 311108_MER31-int,MER31-int,MER39 | chr5 | 43051693 | 43053018 | 6,5734 | -3,2348 | 8,00E-06 |
| 223770_MLT1F2,MLT1F2,MLT1F2 | chr20 | 53163185 | 53163915 | 4,1936 | -3,2179 | 0,003724 |
| 344641_MLT1A0 | chr6 | 84974690 | 84975015 | 4,8882 | -3,2055 | 0,001114 |
| 129998_LTR101_Mam | chr15 | 71450119 | 71450426 | 2,9285 | -3,1818 | 0,011929 |
| 216969_MLT1H | chr20 | 11790025 | 11790539 | 3,0112 | -3,1294 | 0,01672 |
| 231671_MSTA | chr21 | 38141066 | 38141444 | 5,4836 | -3,1125 | 0,000188 |
| 396207_MLT1J2,MLT1J2 | chr8 | 93845794 | 93846165 | 2,7278 | -3,1091 | 0,011885 |
| 184188_MLT1D | chr2 | 47032808 | 47032859 | 3,6126 | -3,0972 | 0,019272 |
| 311105_LTR49-int,LTR29,LTR49 | chr5 | 43047721 | 43048472 | 2,6276 | -3,0869 | 0,008045 |
| 122504_LTR81B,THE1B | chr14 | 101729530 | 101730580 | 2,0717 | -3,0428 | 0,033491 |
| 375206_MSTD | chr7 | 117675772 | 117676061 | 4,0521 | -3,0238 | 0,013685 |
| 148150_LTR56 | chr17 | 8584348 | 8584757 | 2,9285 | -3,0190 | 0,017629 |
| 311106_LTR61 | chr5 | 43049250 | 43049811 | 6,4505 | -2,9890 | 7,55E-05 |
| 220553_MLT1D,MLT1D | chr20 | 34576335 | 34577038 | 2,7200 | -2,9716 | 0,016507 |
| 231424_LTR22C0 | chr21 | 36883808 | 36884293 | 2,6639 | -2,9424 | 0,019204 |
| 445565_MER61B,MER61-int,MER61-int,MER61-int | chrX | 141934261 | 141939406 | 7,8032 | -2,9378 | 4,65E-07 |
| 298989_MER9a3 | chr4 | 168993974 | 168994479 | 5,6522 | -2,9097 | 4,86E-05 |
| 305300_THE1B | chr5 | 14034375 | 14034718 | 2,3060 | -2,8767 | 0,007675 |
| 92137_MLT1C | chr12 | 130133799 | 130134228 | 7,5923 | -2,8547 | 9,55E-05 |
| 122888_MER54A | chr14 | 104748811 | 104749690 | 3,5486 | -2,8547 | 0,039126 |
| 174996_MLT2B3 | chr19 | 46180519 | 46181084 | 7,6618 | -2,8248 | 2,35E-06 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 161760_MLT1J,MLT1K | chr18 | 36179946 | 36180799 | 8,6186 | -2,8217 | 2,16E-05 |
| 78705_MLT1D,MLT1D | chr12 | 54273054 | 54273871 | 2,4422 | -2,8058 | 0,022397 |
| 305383_MLT1D | chr5 | 14611117 | 14611644 | 6,5975 | -2,7714 | 0,000154 |
| 79912_LTR7 | chr12 | 62729866 | 62730257 | 3,2141 | -2,7681 | 0,031287 |
| 374288_MLT1H2,MLT1H2,MLT1H2-int | chr7 | 111498977 | 111499803 | 2,5094 | -2,7398 | 0,042803 |
| 233216_MER65A,MER65A,MER65A,MER65A,MER65A | chr21 | 46289308 | 46290205 | 5,4911 | -2,7331 | 0,012956 |
| 231670_LTR37A | chr21 | 38138980 | 38139309 | 3,1201 | -2,7302 | 0,02499 |
| 269666_MLT1B | chr4 | 3643871 | 3644247 | 2,3224 | -2,7122 | 0,031416 |
| 13303_MLT1D,LTR87,LTR87,LTR87 | chr1 | 91779752 | 91781165 | 2,9967 | -2,7107 | 0,03595 |
| 133836_MLT1J | chr15 | 98888740 | 98889204 | 7,4337 | -2,6957 | 5,13E-06 |
| 303693_MLT1A0,MLT1A0,MLT1C | chr5 | 6778572 | 6779983 | 6,0043 | -2,6883 | 0,000768 |
| 245154_MLT1F2 | chr3 | 40446499 | 40447096 | 3,0994 | -2,6828 | 0,034812 |
| 313137_MER21A,MER39B,MER21A | chr5 | 59300586 | 59301352 | 7,2530 | -2,6388 | 0,000118 |
| 96735_LTR27 | chr13 | 44304686 | 44305278 | 5,8677 | -2,6366 | 0,001305 |
| 328158_MER92B,MER67C | chr5 | 157477736 | 157478162 | 5,0423 | -2,6327 | 0,013341 |
| 185456_MER21C,MER21C,MER21C | chr2 | 54530737 | 54531360 | 4,0994 | -2,6210 | 0,009738 |
| 182884_MLT1K,MLT1K | chr2 | 37378220 | 37378691 | 5,3388 | -2,6163 | 0,001591 |
| 52613_LTR30 | chr11 | 309444 | 310164 | 6,1268 | -2,5928 | 0,000293 |
| 69227_MLT1F2 | chr11 | 114300870 | 114301047 | 6,4306 | -2,5899 | 6,83E-05 |
| 134480_LTR53B | chr16 | 3149105 | 3149611 | 4,0957 | -2,5820 | 0,014655 |
| 330764_MLT1G1,MLT1G1 | chr5 | 174382061 | 174382989 | 5,2106 | -2,5723 | 0,001907 |
| 96654_LTR84b | chr13 | 43874862 | 43875321 | 4,1348 | -2,5618 | 0,008999 |
| 227928_LTR16D2,MLT2A2,LTR16D2 | chr21 | 20451066 | 20451920 | 1,9637 | -2,5570 | 0,035617 |
| 134376_LTR8,LTR8 | chr16 | 1943651 | 1944737 | 3,3231 | -2,5359 | 0,031341 |
| 235140_MLT1D | chr22 | 30235166 | 30235604 | 5,3699 | -2,5241 | 0,000518 |
| 311104_MER34C2 | chr5 | 43047004 | 43047092 | 3,8761 | -2,4864 | 0,03231 |
| 178854_MLT2B1,MLT1D | chr2 | 10328351 | 10329665 | 4,8168 | -2,4688 | 0,011214 |
| 416552_MLT1D | chr9 | 86982300 | 86982744 | 3,4006 | -2,4566 | 0,026587 |
| 12162_MLT1G-int,MLT1G-int,MLT1G,MLT1G | chr1 | 84102015 | 84103764 | 4,8052 | -2,4466 | 0,002747 |
| 74424_MamGypLTR1d | chr12 | 19972339 | 19972561 | 3,2091 | -2,4412 | 0,019998 |
| 61931_MLT1K | chr11 | 65307405 | 65307653 | 4,4188 | -2,4380 | 0,007358 |
| 2812_LTR88a | chr1 | 19010649 | 19010983 | 3,7442 | -2,4236 | 0,048858 |
| 345333_LTR23 | chr6 | 89079375 | 89079667 | 1,5253 | -2,4132 | 0,049313 |
| 25458_MLT1K,MLT1K | chr1 | 198117948 | 198118393 | 3,9228 | -2,4034 | 0,017121 |
| 133615_ERVL-E-int,ERVL-E-int,ERVL-E-int,LTR33C | chr15 | 97496529 | 97501541 | 5,0461 | -2,3919 | 0,004543 |
| 58182_LTR37A,LTR37A,MamGypLTR1c | chr11 | 36479097 | 36479916 | 4,6527 | -2,3854 | 0,017878 |
| 414603_MLT1D | chr9 | 76442076 | 76442393 | 5,5430 | -2,3838 | 0,006006 |
| 327191_LTR37B | chr5 | 151082715 | 151082796 | 2,4137 | -2,3752 | 0,049191 |
| 223211_MLT1K | chr20 | 50313276 | 50313754 | 4,6689 | -2,3653 | 0,005141 |
| 309115_MLT1L,MLT1L | chr5 | 31860976 | 31861329 | 4,5774 | -2,3375 | 0,01631 |
| 311101_LTR8A | chr5 | 43021199 | 43021861 | 4,5522 | -2,3340 | 0,012009 |
| 265613_MLT1A0 | chr3 | 171800648 | 171801002 | 3,9185 | -2,3088 | 0,025108 |
| 149954_MLT1G,MLT1G | chr17 | 20185163 | 20185645 | 5,2831 | -2,3028 | 0,004138 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 115485_LTR16,ERV3-16A3_I-int | chr14 | 59439651 | 59440492 | 4,8912 | -2,2963 | 0,019028 |
| 311203_LTR49-int,LTR49-int,LTR49-int,LTR49-int | chr5 | 43571512 | 43575275 | 8,3728 | -2,2776 | 7,03E-05 |
| 231699_MLT1F1 | chr21 | 38323589 | 38324106 | 5,7287 | -2,2751 | 0,004866 |
| 277700_LTR81B,LTR81B,MLT1D,MLT1D,LTR81C | chr4 | 41434348 | 41435591 | 6,3308 | -2,2597 | 0,002749 |
| 302701_HERVE-int,LTR5B | chr5 | 1566721 | 1568346 | 4,2336 | -2,2530 | 0,014387 |
| 151682_LTR26B | chr17 | 40853212 | 40853741 | 6,0868 | -2,2343 | 0,000958 |
| 178862_MLT1A,MER66B,MLT1A | chr2 | 10388820 | 10389634 | 3,7620 | -2,2312 | 0,039453 |
| 380287_MER41B | chr7 | 150322211 | 150322830 | 4,6793 | -2,2250 | 0,01589 |
| 310411_MER4C | chr5 | 38929776 | 38930205 | 5,6213 | -2,2226 | 0,001915 |
| 149596_MSTB1 | chr17 | 17109946 | 17110376 | 5,0975 | -2,2135 | 0,00471 |
| 133643_MER57B2 | chr15 | 97688823 | 97689222 | 4,8562 | -2,1841 | 0,02555 |
| 302555_LTR5A | chr5 | 196783 | 197781 | 4,7817 | -2,1787 | 0,010685 |
| 311107_MER31-int | chr5 | 43050583 | 43051100 | 4,2746 | -2,1624 | 0,036083 |
| 126497_MER21B | chr15 | 43353444 | 43354238 | 4,2158 | -2,1559 | 0,027273 |
| 6833_MLT1C | chr1 | 50985077 | 50985466 | 4,2034 | -2,1330 | 0,03962 |
| 388928_MSTD,MSTA,MSTA,MSTD | chr8 | 47304110 | 47305134 | 4,2057 | -2,1151 | 0,030051 |
| 405302_MLT1H,LTR9,LTR8A,LTR8A,LTR8A,LTR8A,LTR8A,LTR8A,LTR8A,LTR8A,LTR8A | chr8 | 142672339 | 142674643 | 5,7440 | -2,0938 | 0,004849 |
| 388306_LTR72B,LTR72B,MER66C,MER66C | chr8 | 40204982 | 40206661 | 6,0722 | -2,0914 | 0,005839 |
| 237334_MER4A,MER65A,MER65A | chr22 | 46333166 | 46334422 | 6,9133 | -2,0890 | 0,002817 |
| 302705_LTR5_Hs | chr5 | 1595976 | 1596944 | 3,6693 | -2,0878 | 0,036943 |
| 318609_MER39,LTR79 | chr5 | 94422091 | 94423331 | 3,7147 | -2,0828 | 0,047771 |
| 175007_THE1D | chr19 | 46208548 | 46208910 | 4,2725 | -2,0699 | 0,02323 |
| 298520_ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,LTR16A | chr4 | 165971221 | 165975281 | 4,7773 | -2,0618 | 0,021809 |
| 302559_MER4B | chr5 | 269274 | 269837 | 4,5819 | -2,0557 | 0,021824 |
| 405655_LTR46,LTR46-int,HERV17-int,LTR46-int,LTR46-int,LTR46-int,LTR46-int,LTR46-int,LTR46-int,LTR46,LTR46 | chr8 | 144858830 | 144868579 | 5,9906 | -2,0440 | 0,002634 |
| 237371_MLT1B,MLT1B | chr22 | 46739898 | 46740454 | 4,7730 | -2,0059 | 0,032348 |
| 409617_LTR26,LTR19B,LTR26 | chr9 | 26261245 | 26262406 | 4,2158 | -2,0025 | 0,036658 |

### Table 8.11 Significantly deregulated ERVs with H3K9ac peak in SRSF3 KD, Hela cells

| RepeatID | chr | start | end | Log2 base Mean | log2 Fold Change | p-value |
|---|---|---|---|---|---|---|
| 231802_MLT1D | chr21 | 38840820 | 38841271 | 0,0000 | 6,4012 | 4,08E-06 |
| 191933_MER11C,HERVK11-int,MER11A | chr2 | 100361987 | 100365971 | 1,8057 | 6,2217 | 6,34E-10 |
| 336983_MLT1G3,MLT1G3 | chr6 | 30767566 | 30768367 | 0,0000 | 6,1892 | 2,61E-06 |
| 283124_MER110,LTR27E | chr4 | 74111330 | 74112735 | 0,4278 | 6,1890 | 1,74E-05 |
| 404366_MER77B | chr8 | 137149571 | 137150246 | 0,0000 | 6,1659 | 8,96E-05 |
| 119060_LTR12_,MLT1I | chr14 | 83556406 | 83557464 | 0,4278 | 6,1492 | 3,97E-05 |
| 231819_LTR29 | chr21 | 38908060 | 38908632 | 1,3447 | 5,9701 | 3,98E-08 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 365235_LTR56,LTR54,MER41B,LTR54,MER51B,MER51B | chr7 | 48085504 | 48088059 | 0,6945 | 5,9268 | 1,22E-06 |
| 360617_MLT1E1A | chr7 | 17408187 | 17408704 | 0,4278 | 5,5413 | 1,36E-05 |
| 95363_MER21C | chr13 | 35927589 | 35928289 | 0,0000 | 5,5209 | 0,000145 |
| 39678_LTR12_,LTR12_ | chr10 | 47166757 | 47167766 | 0,0000 | 5,4915 | 7,36E-05 |
| 391203_MLT2B3,LTR9B,MLT2B3 | chr8 | 62162880 | 62164245 | 1,2751 | 5,4785 | 7,92E-08 |
| 199036_MER72 | chr2 | 142875416 | 142876090 | 0,0000 | 5,4375 | 4,48E-05 |
| 360616_MLT2B3 | chr7 | 17406566 | 17407073 | 0,0000 | 5,3960 | 2,16E-05 |
| 363471_LTR7C | chr7 | 36659793 | 36660261 | 0,0000 | 5,3386 | 0,000405 |
| 281911_LTR25 | chr4 | 66574695 | 66575495 | 0,4278 | 5,3283 | 5,99E-05 |
| 372948_MER90,LTR64,LTR64,LTR64 | chr7 | 102394288 | 102395842 | 0,4278 | 5,2352 | 0,000175 |
| 445311_MamGyp-int,MamGyp-int | chrX | 140712043 | 140712665 | 0,4278 | 5,1990 | 0,000182 |
| 254229_LTR16A2 | chr3 | 99503437 | 99503786 | 0,0000 | 5,1648 | 0,000148 |
| 47956_LTR36,MER4CL34 | chr10 | 104252956 | 104254461 | 0,0000 | 5,1261 | 0,000169 |
| 268561_MLT1J1,LTR68,MLT1J1-int,MLT1J1-int | chr3 | 193208256 | 193210275 | 0,0000 | 5,1217 | 0,002038 |
| 328477_MLT1F2 | chr5 | 159583365 | 159583953 | 0,0000 | 5,0532 | 0,001063 |
| 349886_MER89 | chr6 | 119147502 | 119148054 | 0,0000 | 5,0439 | 0,000806 |
| 328915_MER21C | chr5 | 162326720 | 162327076 | 0,0000 | 5,0392 | 0,001459 |
| 294566_MLT1L,MLT1G | chr4 | 142092587 | 142093449 | 0,4278 | 5,0088 | 0,000137 |
| 304500_MLT1E1A,MLT1E1A,LTR41 | chr5 | 10474219 | 10475470 | 0,0000 | 4,9978 | 0,000143 |
| 422580_MER77B,LTR8,LTR8,MER77B | chr9 | 124144785 | 124146556 | 1,0231 | 4,9783 | 2,48E-05 |
| 122645_MER21C,MER21C,MER4D,MER49,MER21C | chr14 | 103197604 | 103199290 | 0,0000 | 4,9652 | 0,00076 |
| 114918_MER4A1_ | chr14 | 55806058 | 55806626 | 0,5953 | 4,9603 | 0,001023 |
| 379341_MER92C,MER92C | chr7 | 144842031 | 144842865 | 0,0000 | 4,9110 | 0,000307 |
| 363629_MER51C,LTR36,LTR36 | chr7 | 37664572 | 37665167 | 0,0000 | 4,8938 | 0,000866 |
| 223242_MLT1K,MLT1L | chr20 | 50439615 | 50440547 | 0,8888 | 4,8231 | 2,30E-05 |
| 376985_MLT1C | chr7 | 128112087 | 128112329 | 0,5953 | 4,7921 | 0,000139 |
| 314619_MER34-int,MER31-int,MER31-int,MSTD,LTR8,MSTD,MER31-int,MER31-int,MER31-int,MER34A,MER34A | chr5 | 68451544 | 68456623 | 0,0000 | 4,7708 | 0,006918 |
| 310604_LTR1A1,MER11C,HERVK11-int,HERVK11-int,MER11C,LTR36 | chr5 | 40102031 | 40113106 | 0,0000 | 4,7156 | 0,003074 |
| 222156_LTR3B_,LTR3B_ | chr20 | 44730512 | 44731043 | 0,0000 | 4,7003 | 0,001671 |
| 62090_LTR47A,LTR47A | chr11 | 66969092 | 66969809 | 0,0000 | 4,6986 | 0,003994 |
| 430223_MamGypLTR1a,MamGypLTR1a | chrX | 30283642 | 30284239 | 1,0231 | 4,6943 | 6,26E-05 |
| 301412_LTR54,LTR54,MER57C2 | chr4 | 183410827 | 183411525 | 1,2898 | 4,6387 | 3,05E-05 |
| 330563_MLT1B | chr5 | 173456048 | 173456471 | 0,8888 | 4,6299 | 1,17E-05 |
| 396218_LTR26D | chr8 | 93879287 | 93879904 | 0,0000 | 4,6049 | 0,002804 |
| 216666_LTR101_Mam,LTR101_Mam | chr20 | 10353417 | 10354360 | 1,6114 | 4,6004 | 2,01E-05 |
| 234928_MLT1G1,MSTA | chr22 | 27715992 | 27717421 | 2,5070 | 4,5993 | 1,73E-06 |
| 90536_LTR8 | chr12 | 124919339 | 124920129 | 2,0024 | 4,5790 | 5,73E-06 |
| 347809_THE1C,MLT1J1 | chr6 | 105427196 | 105428170 | 0,0000 | 4,5657 | 0,002708 |
| 360615_MamRep1527 | chr7 | 17405800 | 17405886 | 0,0000 | 4,5200 | 0,001574 |
| 360605_MLT1D | chr7 | 17372108 | 17372580 | 0,0000 | 4,5102 | 0,000476 |
| 247605_LTR90A,LTR90A,LTR90A | chr3 | 59245841 | 59246485 | 1,6370 | 4,5009 | 9,01E-05 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 76187_PrimLTR79,LTR29,LOR1a,LTR29 | chr12 | 31750190 | 31751223 | 0,0000 | 4,4740 | 0,002244 |
| 363605_MER41B,LTR84b,LTR84b | chr7 | 37549995 | 37551316 | 0,0000 | 4,4729 | 0,004421 |
| 265615_MLT2B3 | chr3 | 171824370 | 171824953 | 0,5953 | 4,4418 | 0,002538 |
| 227993_MER21C | chr21 | 20720794 | 20721574 | 0,5953 | 4,4299 | 0,000911 |
| 261606_MLT2C1,MLT1D-int,MLT1D,ERVL-B4-int | chr3 | 146967031 | 146970434 | 0,0000 | 4,4265 | 0,000717 |
| 277721_LTR85c | chr4 | 41708348 | 41708843 | 1,7632 | 4,4084 | 2,86E-07 |
| 330561_MamGypsy2-I,MLT1L,MamGypsy2-I | chr5 | 173452137 | 173453120 | 1,5925 | 4,4056 | 7,72E-06 |
| 40065_MLT1E3-int,LTR8B,MLT1E3 | chr10 | 49568154 | 49570415 | 1,1678 | 4,3671 | 0,001872 |
| 285250_MLT1E | chr4 | 86929098 | 86929646 | 1,7631 | 4,3522 | 1,61E-05 |
| 216968_LTR49 | chr20 | 11788711 | 11789303 | 0,5953 | 4,3488 | 0,001791 |
| 210672_MLT1H1 | chr2 | 220228331 | 220228909 | 0,0000 | 4,3279 | 0,007942 |
| 64045_LTR17 | chr11 | 80699210 | 80699970 | 0,5953 | 4,3217 | 0,003457 |
| 60180_LTR104_Mam | chr11 | 47944052 | 47944267 | 0,5953 | 4,3101 | 0,002267 |
| 370464_MLT2B5 | chr7 | 84701856 | 84702482 | 0,0000 | 4,2659 | 0,004061 |
| 375662_LTR4,HERV3-int,HERV3-int | chr7 | 120052770 | 120061151 | 2,1803 | 4,2642 | 3,02E-05 |
| 45701_MLT1K | chr10 | 86347371 | 86347790 | 0,0000 | 4,2615 | 0,003902 |
| 336987_MER83,MER83,MER83B-int,MER83 | chr6 | 30778923 | 30781026 | 3,5404 | 4,2611 | 4,71E-06 |
| 247468_MER34B,MER34B | chr3 | 58469982 | 58470742 | 0,5953 | 4,2610 | 0,004481 |
| 29384_LTR16E2,LTR16E2,LTR16E2,LTR16,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,ERV3-16A3_I-int,LTR16 | chr1 | 227356153 | 227360947 | 0,0000 | 4,2447 | 0,001449 |
| 399819_MLT1D | chr8 | 114936383 | 114936917 | 0,0000 | 4,2244 | 0,004684 |
| 33679_MER51B,MER51-int,MER51-int,MER51-int,MER41-int,MER51-int,MER51-int,MER51-int,MER51B,MSTD | chr10 | 5065471 | 5071306 | 2,5710 | 4,2116 | 9,60E-07 |
| 76186_LTR8B,PrimLTR79 | chr12 | 31749068 | 31749678 | 0,8888 | 4,1922 | 0,000311 |
| 126590_LTR16A2,LTR16A2 | chr15 | 44726308 | 44726827 | 0,0000 | 4,1874 | 0,006632 |
| 46149_MER4B,MER4B | chr10 | 89326830 | 89327747 | 0,6945 | 4,1781 | 0,002099 |
| 386742_LTR2C,HERVE-int | chr8 | 30728580 | 30733679 | 1,9878 | 4,1751 | 7,38E-06 |
| 223769_MLT1N2,THE1D | chr20 | 53161089 | 53161806 | 0,4278 | 4,1741 | 0,005272 |
| 18275_LTR75B | chr1 | 147083988 | 147084093 | 2,1920 | 4,1726 | 2,31E-05 |
| 323546_MLT1A1 | chr5 | 123820312 | 123820716 | 0,8888 | 4,1706 | 0,00014 |
| 239753_MER31A | chr3 | 8964026 | 8964477 | 0,4278 | 4,1522 | 0,00341 |
| 37546_MLT1I,MLT1I | chr10 | 29807089 | 29807629 | 1,1678 | 4,1517 | 0,000429 |
| 76190_LTR8B,LTR8B | chr12 | 31756862 | 31757926 | 0,0000 | 4,1429 | 0,003631 |
| 365234_MER4E1,MER4E1 | chr7 | 48083792 | 48084888 | 0,0000 | 4,1399 | 0,006499 |
| 332747_MLT1A0 | chr6 | 4932153 | 4932521 | 0,0000 | 4,1366 | 0,000952 |
| 364728_MER21C,MER21C,MER21C,MLT1A0,MER21C,MSTC,MER68 | chr7 | 45114607 | 45117141 | 0,4278 | 4,1299 | 0,019119 |
| 214529_MER4CL34,MER4CL34 | chr2 | 242086910 | 242087363 | 0,0000 | 4,1197 | 0,010391 |
| 356264_MLT1K | chr6 | 159578009 | 159578187 | 0,4278 | 4,1015 | 0,00181 |
| 223434_MLT1L | chr20 | 51351034 | 51351189 | 0,6945 | 4,0965 | 0,001505 |
| 209592_MER39,MER39 | chr2 | 213833836 | 213834775 | 1,5664 | 4,0820 | 0,00104 |
| 369246_MER31A,MER31A | chr7 | 76293518 | 76294272 | 1,4841 | 4,0813 | 0,000256 |
| 2452_LTR4 | chr1 | 16905463 | 16905857 | 5,6725 | 4,0682 | 0,000126 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 358426_MER41B,MER41B,LTR61,MER74A,MER74A | chr7 | 3193353 | 3194959 | 0,0000 | 4,0626 | 0,004565 |
| 261530_LTR54,LTR101_Mam,LTR24C,MLT1A-int,MLT1A-int,MLT1A,MLT1A-int | chr3 | 146506652 | 146508716 | 2,8961 | 4,0384 | 6,89E-05 |
| 422299_MER49,MER49,MER49 | chr9 | 121370822 | 121372225 | 0,0000 | 4,0357 | 0,016345 |
| 130455_LTR16A | chr15 | 75109020 | 75109441 | 0,4278 | 4,0163 | 0,006881 |
| 231858_THE1B | chr21 | 39047561 | 39047875 | 0,0000 | 4,0114 | 0,006227 |
| 354253_MLT1D-int,MLT1D | chr6 | 146289966 | 146291094 | 1,0231 | 4,0070 | 0,002903 |
| 28495_MER39,MER21A,MER21A,LTR48B | chr1 | 221838607 | 221839445 | 0,0000 | 3,9936 | 0,005971 |
| 149884_LTR33A,LTR50,MER4-int,MER4-int,MER4C | chr17 | 19645723 | 19647663 | 3,4612 | 3,9769 | 1,04E-05 |
| 322253_LTR10F,LTR10F,LTR10F,LTR10F,MER41A,LTR10F,LTR10F,MER41A,MLT1D | chr5 | 115928854 | 115931428 | 2,6639 | 3,9749 | 4,52E-06 |
| 211390_LTR48B,LTR29,MER34 | chr2 | 224146256 | 224146903 | 0,0000 | 3,9688 | 0,004702 |
| 117585_LTR33 | chr14 | 74573324 | 74573759 | 1,0231 | 3,9638 | 0,00505 |
| 385372_LTR67B | chr8 | 21918159 | 21918599 | 1,1678 | 3,9432 | 0,000571 |
| 62142_LTR49-int,LTR49-int,LTR49-int,LTR49-int,MER51E,LTR49-int,LTR13_,LTR49-int,LTR49-int,LTR49-int,ERV24B_Prim-int,LTR23 | chr11 | 67569605 | 67575312 | 0,0000 | 3,9368 | 0,0015 |
| 18435_LTR75 | chr1 | 148264110 | 148264219 | 2,6017 | 3,9335 | 0,001553 |
| 182938_LTR16B1,LTR16B1 | chr2 | 37585545 | 37586128 | 0,0000 | 3,9295 | 0,017312 |
| 303333_MER21C,MER21C | chr5 | 5351314 | 5352309 | 1,2898 | 3,9277 | 0,001169 |
| 310336_MER21A,MER21A,MER21A,MER21-int,MER4B-int,MER21-int,MER21-int,MER21-int,MER21-int,MER21A,MER21A | chr5 | 38631831 | 38638329 | 0,0000 | 3,9246 | 0,00614 |
| 193508_MER4A1_ | chr2 | 109701267 | 109701835 | 1,2898 | 3,9205 | 0,00126 |
| 126599_LTR13 | chr15 | 44784188 | 44785182 | 0,0000 | 3,9118 | 0,006005 |
| 217926_MLT1F2,MLT1F2 | chr20 | 16530206 | 16531067 | 0,6945 | 3,9041 | 0,003604 |
| 85580_MLT1A0 | chr12 | 95806434 | 95806773 | 0,0000 | 3,9007 | 0,0034 |
| 335864_LTR26C,MER4B,MLT1H1 | chr6 | 24658846 | 24660011 | 1,3685 | 3,8905 | 0,004635 |
| 210468_MLT1N2 | chr2 | 219277555 | 219278123 | 1,6114 | 3,8870 | 0,000612 |
| 388301_MER4B | chr8 | 40176104 | 40176745 | 2,5094 | 3,8827 | 1,61E-05 |
| 70892_MLT1H | chr11 | 128208282 | 128208749 | 0,4278 | 3,8814 | 0,0083 |
| 381107_ERV24B_Prim-int,ERV24B_Prim-int,LTR48 | chr7 | 154922857 | 154929139 | 1,6370 | 3,8803 | 0,000979 |
| 305939_MamGypLTR2b | chr5 | 17375279 | 17375517 | 2,4757 | 3,8798 | 1,94E-06 |
| 276362_LTR67B,LTR67B | chr4 | 34182283 | 34182687 | 1,6114 | 3,8337 | 0,001837 |
| 225596_LTR26C,LTR1A2,LTR26C,MER41A | chr20 | 63230278 | 63232402 | 1,8704 | 3,8266 | 0,000291 |
| 211318_LTR9 | chr2 | 223723885 | 223724514 | 1,7357 | 3,8164 | 0,00055 |
| 100077_LTR8B | chr13 | 63435583 | 63436318 | 0,4278 | 3,8131 | 0,00947 |
| 362853_MER76 | chr7 | 32669162 | 32669692 | 0,6945 | 3,8125 | 0,003136 |
| 420158_MLT2C1,MLT2C1 | chr9 | 107755729 | 107756199 | 0,6945 | 3,8107 | 0,004898 |
| 33681_MER67A,LTR84b,MER67C,MER67C,LTR84b,MER31B,LTR49,MER4A1,LTR49,PRIMA4_LTR,LTR49 | chr10 | 5071504 | 5075361 | 1,7631 | 3,8082 | 0,00066 |
| 61939_LTR8 | chr11 | 65383794 | 65383913 | 0,0000 | 3,8044 | 0,006779 |
| 150332_LTR79 | chr17 | 27245527 | 27245775 | 0,0000 | 3,7989 | 0,009182 |
| 159549_MLT1A1 | chr18 | 22468268 | 22468645 | 1,7631 | 3,7988 | 0,000451 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 258707_LTR8 | chr3 | 128878209 | 128878882 | 0,0000 | 3,7950 | 0,007358 |
| 402002_MamGypLTR3a,LTR8B,LTR8B,MLT2C1,MLT1E2 | chr8 | 126432991 | 126435155 | 0,8888 | 3,7778 | 0,002956 |
| 169914_MER67C,MER67C,MER67C | chr19 | 9059268 | 9059907 | 1,6370 | 3,7754 | 0,00124 |
| 142111_MLT1K | chr16 | 56605624 | 56606045 | 1,0231 | 3,7752 | 0,011822 |
| 330838_MER77B | chr5 | 174829605 | 174830167 | 1,8581 | 3,7653 | 0,000115 |
| 305202_MLT1F2,MSTA1 | chr5 | 13567869 | 13568936 | 1,3447 | 3,7403 | 0,001714 |
| 223851_MSTC | chr20 | 53827048 | 53827453 | 0,4278 | 3,7386 | 0,011879 |
| 223243_MLT1L | chr20 | 50441553 | 50442010 | 0,4278 | 3,7385 | 0,015876 |
| 364987_LTR38A1,LTR38-int,LTR57-int,LTR38-int,MER4-int,MER101-int,LTR38-int,LOR1-int,LOR1-int,MER49,PABL_B-int | chr7 | 46539582 | 46544757 | 0,0000 | 3,7378 | 0,025149 |
| 26866_ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,ERVL-B4-int,LTR16A,ERVL-B4-int,MLT1N2 | chr1 | 210436678 | 210441114 | 2,3749 | 3,7370 | 4,24E-06 |
| 62227_LTR54B | chr11 | 68011103 | 68011560 | 1,2898 | 3,7359 | 0,002792 |
| 297371_LTR16A2,ERV3-16A3_I-int,MSTB1,ERVL-E-int,ERVL-E-int,ERVL-E-int,ERVL-E-int,LTR16A2 | chr4 | 159545279 | 159549677 | 3,0853 | 3,7337 | 1,80E-05 |
| 137476_LTR14C,HERVK14C-int,LTR14C | chr16 | 19393581 | 19402152 | 1,6370 | 3,7290 | 0,000738 |
| 53257_LTR75_1,LTR12C,LTR75_1 | chr11 | 5682462 | 5684428 | 0,0000 | 3,7205 | 0,005761 |
| 335638_THE1D,MLT1J2 | chr6 | 23174352 | 23175036 | 0,0000 | 3,6698 | 0,009142 |
| 307274_LTR8B,MLT1E,MLT1E | chr5 | 23466346 | 23468395 | 0,4278 | 3,6629 | 0,035143 |
| 244891_MLT1G,MLT1G | chr3 | 39179673 | 39180013 | 0,6945 | 3,6605 | 0,007183 |
| 326109_THE1B | chr5 | 143245080 | 143245449 | 0,8888 | 3,6546 | 0,004067 |
| 3296_MER74B | chr1 | 23588913 | 23589424 | 0,0000 | 3,6523 | 0,027 |
| 302627_MLT2A1 | chr5 | 918576 | 919018 | 0,8888 | 3,6351 | 0,001512 |
| 309146_MSTC,MLT1E1,MLT1E1,MSTB1,MSTB1,MLT1E1 | chr5 | 32181342 | 32183498 | 0,0000 | 3,6336 | 0,024989 |
| 213730_MER72,LTR8A,LTR9B,LTR8A | chr2 | 236973614 | 236975320 | 1,0231 | 3,6324 | 0,019467 |
| 108940_MER4B,MER4D,MER49,MER4D,MER49,MER21C,MER21C | chr13 | 112722053 | 112724847 | 0,6945 | 3,6270 | 0,008312 |
| 38047_MLT1L,MLT1L | chr10 | 32983298 | 32984137 | 3,4550 | 3,6249 | 3,44E-06 |
| 184089_MLT1F | chr2 | 46224753 | 46225256 | 1,4841 | 3,6198 | 0,000887 |
| 436586_MER4A,MER4B-int,MER4B-int,PABL_B,MER4B-int | chrX | 76145962 | 76148830 | 2,8134 | 3,6190 | 0,000168 |
| 277720_MSTB1 | chr4 | 41702697 | 41703118 | 1,8057 | 3,6144 | 0,001284 |
| 152064_LTR12C | chr17 | 44568043 | 44569479 | 2,0024 | 3,6099 | 0,000855 |
| 188108_MER110-int,MER110-int,ERV24_Prim-int | chr2 | 73383752 | 73385020 | 0,6945 | 3,6014 | 0,010109 |
| 149848_LTR47B4,MLT1C | chr17 | 19448804 | 19449858 | 1,0417 | 3,5951 | 0,010747 |
| 422579_LTR33 | chr9 | 124143487 | 124143581 | 0,0000 | 3,5785 | 0,009316 |
| 395374_MER11C,HERVK11-int,HERVK11-int,HERVK11-int,HERVK11-int,HERVK11-int,HERVK11-int | chr8 | 88670469 | 88676597 | 0,0000 | 3,5745 | 0,026129 |
| 306230_HERVL-int | chr5 | 18746583 | 18751737 | 1,0231 | 3,5709 | 0,013225 |
| 308381_MSTD-int,MSTD | chr5 | 28374407 | 28375937 | 0,4278 | 3,5658 | 0,022094 |
| 58162_MER31-int,MER31-int,MER34A,MLT2B1 | chr11 | 36328877 | 36331898 | 0,0000 | 3,5629 | 0,014262 |
| 2308_THE1C | chr1 | 15460869 | 15461247 | 0,0000 | 3,5601 | 0,03234 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 209010_LTR12C,LTR12C,MER39B | chr2 | 209999639 | 210001961 | 2,0464 | 3,5565 | 0,000133 |
| 258511_MLT1I,MER48,MER48 | chr3 | 127592117 | 127593007 | 0,9169 | 3,5522 | 0,00284 |
| 374802_MLT1E1A,MSTB,MLT1E1A,LTR40c,MLT2A2 | chr7 | 114900114 | 114902199 | 0,4278 | 3,5496 | 0,014323 |
| 61241_MER41C,MER41C,MER41C | chr11 | 59813317 | 59814385 | 0,0000 | 3,5396 | 0,009013 |
| 128281_MER65C | chr15 | 56734076 | 56734537 | 0,0000 | 3,5385 | 0,006801 |
| 365066_LTR105_Mam,MER66B,LTR105_Mam | chr7 | 46969711 | 46970402 | 0,6945 | 3,5310 | 0,00818 |
| 310110_MLT2A1,HERVL-int,HERVL-int,HERVL-int,HERVL-int,MLT2A1,MLT2A1,LTR40b | chr5 | 37264508 | 37272840 | 3,6397 | 3,5300 | 1,94E-05 |
| 356290_LTR78,LTR78 | chr6 | 159701115 | 159701703 | 0,0000 | 3,5167 | 0,03251 |
| 391009_MLT1K | chr8 | 61053565 | 61054057 | 0,6945 | 3,5064 | 0,021533 |
| 82095_LTR37B,LTR37B | chr12 | 75951819 | 75952542 | 0,0000 | 3,5042 | 0,018815 |
| 130465_MER41-int,LTR8,MER41-int,LTR8 | chr15 | 75149446 | 75151009 | 1,2751 | 3,5031 | 0,016817 |
| 149747_MER4B | chr17 | 18721345 | 18721943 | 2,6017 | 3,5014 | 9,14E-05 |
| 33674_MST-int,LTR14A,HERVK14-int | chr10 | 5034451 | 5038306 | 1,9637 | 3,4892 | 0,002252 |
| 410806_MER77B | chr9 | 33328899 | 33329590 | 0,6945 | 3,4863 | 0,014294 |
| 108993_LTR73,MSTB1,MSTA1 | chr13 | 113188387 | 113189461 | 2,6017 | 3,4821 | 7,98E-05 |
| 395416_THE1D | chr8 | 88959291 | 88959654 | 1,6370 | 3,4809 | 0,000913 |
| 223435_MLT1J | chr20 | 51359049 | 51359386 | 0,8888 | 3,4606 | 0,007312 |
| 213549_MLT1K | chr2 | 235726924 | 235727166 | 0,0000 | 3,4582 | 0,027609 |
| 167345_MER74A | chr18 | 70423190 | 70423726 | 0,4278 | 3,4566 | 0,026006 |
| 354726_MER74A | chr6 | 149168837 | 149169382 | 0,4278 | 3,4551 | 0,038129 |
| 328914_MLT1B | chr5 | 162324014 | 162324381 | 1,0231 | 3,4511 | 0,006885 |
| 13009_LTR8B | chr1 | 89952737 | 89953657 | 1,1678 | 3,4423 | 0,001296 |
| 336134_MER68,LTR3,HERVK3-int,HERVK3-int,LTR3 | chr6 | 26287804 | 26296266 | 0,0000 | 3,4414 | 0,010142 |
| 323763_MER61A,MER61-int,MER61-int,MER61-int | chr5 | 124962834 | 124967101 | 0,0000 | 3,4380 | 0,026045 |
| 247571_LTR12C | chr3 | 59050675 | 59052063 | 0,0000 | 3,4326 | 0,017019 |
| 214802_MER65C | chr20 | 1393576 | 1394007 | 0,5953 | 3,4308 | 0,010753 |
| 305300_THE1B | chr5 | 14034375 | 14034718 | 2,3060 | 3,4283 | 0,000856 |
| 423217_MER50 | chr9 | 131361255 | 131361850 | 0,0000 | 3,4281 | 0,021717 |
| 207621_MLT1J2,MLT1J2,MLT1I,MLT1I-int | chr2 | 200302301 | 200303390 | 1,0417 | 3,4092 | 0,005618 |
| 397710_LTR16,LTR88b | chr8 | 102781298 | 102781810 | 0,6945 | 3,3956 | 0,021254 |
| 134431_LTR8B,LTR8B | chr16 | 2795578 | 2796769 | 0,0000 | 3,3954 | 0,025859 |
| 208625_MLT1J2 | chr2 | 207632892 | 207633051 | 0,0000 | 3,3949 | 0,025497 |
| 263182_MLT1D | chr3 | 157127528 | 157127915 | 1,6114 | 3,3831 | 0,004043 |
| 158617_LTR45C,LTR45C | chr18 | 9877073 | 9877678 | 0,9169 | 3,3795 | 0,009695 |
| 122890_LTR12C | chr14 | 104750447 | 104752026 | 0,0000 | 3,3795 | 0,038655 |
| 256915_LTR7,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7,LTR2 | chr3 | 115793482 | 115800126 | 0,6945 | 3,3786 | 0,007082 |
| 377988_ERVL-B4-int,MER21C | chr7 | 135853041 | 135854687 | 1,4841 | 3,3783 | 0,006641 |
| 373451_LTR88c | chr7 | 106424428 | 106424689 | 0,8888 | 3,3772 | 0,010064 |
| 64913_MLT1J1 | chr11 | 86735412 | 86735518 | 0,0000 | 3,3756 | 0,016603 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 373489_LTR10F | chr7 | 106569509 | 106569937 | 0,0000 | 3,3727 | 0,015185 |
| 6901_LTR8,LTR8 | chr1 | 51557827 | 51558840 | 0,4278 | 3,3662 | 0,016257 |
| 167346_MSTA | chr18 | 70424407 | 70424799 | 1,2898 | 3,3636 | 0,008895 |
| 432509_LTR33C | chrX | 41010228 | 41010837 | 0,8888 | 3,3635 | 0,008916 |
| 199034_LTR27E,LTR27E | chr2 | 142868234 | 142869554 | 2,1743 | 3,3607 | 0,001161 |
| 242004_LTR33 | chr3 | 23042038 | 23042500 | 0,6945 | 3,3590 | 0,005331 |
| 140641_LTR79,LTR38C,LTR38C | chr16 | 48613208 | 48614829 | 2,5830 | 3,3460 | 0,000365 |
| 406606_LTR10D,HERVI-int,HERVI-int,LTR10D,MLT1J | chr9 | 6078707 | 6088395 | 0,4278 | 3,3431 | 0,013342 |
| 76032_HUERS-P3-int,LTR9B,LTR9B,LTR13_,LTR9B | chr12 | 30850122 | 30852149 | 0,0000 | 3,3413 | 0,009705 |
| 389861_MER61-int,MER61A | chr8 | 53344423 | 53348184 | 3,3596 | 3,3360 | 0,000189 |
| 97570_ERVL-E-int,ERVL-E-int,ERVL-E-int | chr13 | 48784453 | 48785752 | 0,6945 | 3,3230 | 0,01599 |
| 199033_MLT1A1 | chr2 | 142865355 | 142865823 | 1,8331 | 3,3158 | 0,002655 |
| 354996_LTR12_,HERV9-int,HERV9-int,LTR12_ | chr6 | 150849126 | 150856692 | 3,2572 | 3,3131 | 2,46E-05 |
| 245314_MSTA,LTR16C | chr3 | 41118364 | 41119221 | 0,0000 | 3,3121 | 0,045657 |
| 90468_LTR16A2,LTR16A2,MLT1L | chr12 | 124679371 | 124680369 | 2,5830 | 3,2965 | 0,000514 |
| 114846_LTR39,MLT1C,LTR85c,MLT1I | chr14 | 55457023 | 55459401 | 0,0000 | 3,2894 | 0,027882 |
| 174545_LTR8B | chr19 | 43265000 | 43265754 | 1,1678 | 3,2885 | 0,000682 |
| 149603_MER4B | chr17 | 17281998 | 17282592 | 1,6114 | 3,2742 | 0,00692 |
| 332636_MER21C | chr6 | 4317632 | 4318540 | 1,0417 | 3,2705 | 0,02293 |
| 303601_MLT1D | chr5 | 6355873 | 6356389 | 1,0231 | 3,2701 | 0,006319 |
| 107172_MLT1A,MLT1A-int,MLT1A | chr13 | 102191290 | 102193373 | 0,0000 | 3,2698 | 0,014047 |
| 29338_MER4C | chr1 | 226938812 | 226939189 | 2,1803 | 3,2693 | 0,00129 |
| 433493_LTR33 | chrX | 45478534 | 45478984 | 1,0231 | 3,2681 | 0,017101 |
| 192142_LTR1F2 | chr2 | 101355655 | 101356399 | 6,2010 | 3,2595 | 2,10E-06 |
| 6204_MLT1B,LTR37A,LTR37A,MER92-int,MER92-int,MER92-int,MER92-int,MER92-int | chr1 | 45890280 | 45894714 | 0,5953 | 3,2514 | 0,018995 |
| 290373_MER65A | chr4 | 118590733 | 118591172 | 0,0000 | 3,2503 | 0,021148 |
| 289252_MLT1J2 | chr4 | 112084112 | 112084414 | 0,0000 | 3,2489 | 0,024565 |
| 180571_LTR37A | chr2 | 20052516 | 20052735 | 1,2898 | 3,2436 | 0,0107 |
| 364393_HERVL74-int,HERVL74-int,HERVL74-int,HERVL74-int | chr7 | 42889718 | 42892557 | 2,4422 | 3,2426 | 0,002634 |
| 369010_MER21C | chr7 | 74097730 | 74098330 | 0,0000 | 3,2411 | 0,021716 |
| 12835_MER51A,MER51A,MER51A | chr1 | 88892399 | 88893266 | 0,0000 | 3,2364 | 0,018484 |
| 174441_LTR8B,LTR8B,LTR101_Mam | chr19 | 42932564 | 42933610 | 2,0517 | 3,2345 | 0,000204 |
| 180667_LTR5A | chr2 | 20501251 | 20502198 | 3,7708 | 3,2263 | 2,37E-07 |
| 397257_MamGypsy2-I,MamGypsy2-LTR,THE1C | chr8 | 100483237 | 100483987 | 0,8888 | 3,2212 | 0,013173 |
| 422545_MLT1J | chr9 | 123933309 | 123933813 | 1,0231 | 3,2212 | 0,006934 |
| 350969_LTR14B,HERVK14-int,MER41C,HERVK14-int,HERVK14-int,HERV17-int,HERV17-int,LTR14B,MLT2D | chr6 | 125910948 | 125913531 | 3,5098 | 3,2163 | 3,54E-05 |
| 41750_MLT1K,MLT1B,MLT1B | chr10 | 59387678 | 59388512 | 0,5953 | 3,2105 | 0,023268 |
| 380297_LTR39,LTR39 | chr7 | 150365304 | 150366354 | 1,8331 | 3,2028 | 0,007741 |
| 307307_MLT1G | chr5 | 23631749 | 23632090 | 0,4278 | 3,1954 | 0,031334 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 72088_LTR10F,LTR10F,MER31B | chr12 | 4319491 | 4320728 | 1,9586 | 3,1938 | 0,000606 |
| 76182_LTR33,LTR8,LTR8,LTR33 | chr12 | 31733069 | 31734503 | 0,8888 | 3,1874 | 0,027126 |
| 27989_MLT2C1 | chr1 | 218705212 | 218705520 | 0,4278 | 3,1871 | 0,046843 |
| 150361_LTR8,LTR8,LTR8 | chr17 | 27376102 | 27377435 | 0,0000 | 3,1844 | 0,026213 |
| 386311_MER4D1,MER4D1,MER4D1,MER4CL34,MER4CL34 | chr8 | 28090664 | 28092598 | 3,3133 | 3,1815 | 0,000189 |
| 277264_MLT1E1,MLT1E1 | chr4 | 38374430 | 38375217 | 0,4278 | 3,1735 | 0,024849 |
| 73309_MLT1N2 | chr12 | 12803118 | 12803632 | 1,0231 | 3,1718 | 0,010755 |
| 90541_HERVIP10F-int,HERVIP10FH-int,LTR10E | chr12 | 124936821 | 124938971 | 0,0000 | 3,1704 | 0,034793 |
| 305290_MLT1O,LTR16E1 | chr5 | 13986305 | 13987278 | 2,3064 | 3,1704 | 0,001874 |
| 274672_LTR1D1,LTR79,LTR79,LTR79 | chr4 | 26530206 | 26531459 | 1,8331 | 3,1677 | 0,010012 |
| 18467_LTR75 | chr1 | 148387824 | 148387992 | 2,9285 | 3,1648 | 0,002228 |
| 421792_MLT1C | chr9 | 118337004 | 118337415 | 0,4278 | 3,1631 | 0,020361 |
| 57066_LTR4,HERV3-int,LTR4 | chr11 | 29518505 | 29527987 | 0,0000 | 3,1607 | 0,008586 |
| 207736_LTR41C,LTR41C | chr2 | 201232109 | 201232551 | 0,5953 | 3,1601 | 0,028194 |
| 223915_LTR33 | chr20 | 54131541 | 54131974 | 1,0231 | 3,1520 | 0,01079 |
| 53422_MLT1C,MLT1C | chr11 | 6684040 | 6684904 | 0,5953 | 3,1515 | 0,023927 |
| 190424_LTR3B,LTR3B_,MLT1F2-int,MLT1F2 | chr2 | 86703756 | 86704999 | 0,0000 | 3,1501 | 0,032771 |
| 112529_LTR16C,ERV3-16A3_I-int,LTR16C | chr14 | 42193074 | 42194546 | 0,0000 | 3,1397 | 0,030312 |
| 150331_MLT1N2 | chr17 | 27244355 | 27244874 | 0,5953 | 3,1356 | 0,024701 |
| 334684_MLT1C | chr6 | 16770362 | 16770821 | 0,6945 | 3,1350 | 0,038692 |
| 233263_MER4-int,MER4-int,MER4A1,MER4B | chr21 | 46631792 | 46635087 | 3,4907 | 3,1347 | 0,000133 |
| 236173_MLT1B | chr22 | 36371142 | 36371526 | 0,0000 | 3,1271 | 0,026155 |
| 28414_MER39B | chr1 | 221391423 | 221392078 | 1,0231 | 3,1257 | 0,029693 |
| 222152_MLT1A0 | chr20 | 44695924 | 44696275 | 0,9169 | 3,1229 | 0,033614 |
| 219338_MER49 | chr20 | 23358905 | 23359635 | 0,4278 | 3,1206 | 0,039654 |
| 408933_MLT1C | chr9 | 22102351 | 22102627 | 1,8057 | 3,1114 | 0,013056 |
| 129177_LTR33 | chr15 | 65185928 | 65186027 | 0,4278 | 3,1099 | 0,030745 |
| 84191_MLT1F2 | chr12 | 88298777 | 88299346 | 0,0000 | 3,1048 | 0,026799 |
| 132288_LTR2B | chr15 | 88622388 | 88622894 | 0,5953 | 3,1026 | 0,020263 |
| 417931_LTR36,LTR36,MER50 | chr9 | 94138049 | 94139311 | 0,5953 | 3,0988 | 0,020709 |
| 46158_MER72B,MER4A,MER50,MER50,MER50,MER4E | chr10 | 89351916 | 89354482 | 1,3685 | 3,0986 | 0,010353 |
| 302643_LTR8,LTR8 | chr5 | 1117287 | 1118335 | 0,0000 | 3,0953 | 0,047847 |
| 337802_MLT1H,MLT1F2 | chr6 | 37239908 | 37240685 | 1,8704 | 3,0930 | 0,007296 |
| 410796_MLT1K,LTR16C,MLT1K | chr9 | 33230882 | 33231965 | 2,3060 | 3,0906 | 0,002107 |
| 422295_LTR86A1 | chr9 | 121325898 | 121326350 | 1,2898 | 3,0903 | 0,006175 |
| 53964_MER65A | chr11 | 10862485 | 10862946 | 0,5953 | 3,0848 | 0,032333 |
| 88122_LTR2B,LTR2B,MER41A | chr12 | 110049151 | 110050896 | 2,5830 | 3,0827 | 0,002673 |
| 23339_MLT2F,MLT2F | chr1 | 186621027 | 186621930 | 0,0000 | 3,0805 | 0,025174 |
| 369193_LTR26 | chr7 | 75878190 | 75878797 | 2,3496 | 3,0791 | 0,001058 |
| 391112_LTR32 | chr8 | 61722412 | 61722865 | 3,3868 | 3,0695 | 9,62E-05 |
| 422578_MLT1J,LTR7B,MLT1J,LTR7B | chr9 | 124142902 | 124144014 | 3,0216 | 3,0674 | 0,000818 |
| 362113_MER68 | chr7 | 27677894 | 27678417 | 0,4278 | 3,0667 | 0,020703 |
| 391123_LTR16E1 | chr8 | 61759745 | 61759978 | 1,7357 | 3,0660 | 0,00721 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 219339_MER67C,MER110A | chr20 | 23360499 | 23361459 | 2,5897 | 3,0617 | 0,001401 |
| 359855_LTR7B,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7B | chr7 | 12659787 | 12665594 | 4,5674 | 3,0600 | 1,02E-05 |
| 405013_MER65D,MER65D | chr8 | 140465770 | 140466426 | 2,8265 | 3,0593 | 0,000294 |
| 254228_LTR90A,LTR90A | chr3 | 99503207 | 99503869 | 0,4278 | 3,0515 | 0,029134 |
| 333168_LTR51,LTR51 | chr6 | 7699067 | 7699990 | 1,6114 | 3,0463 | 0,006213 |
| 94888_LTR32 | chr13 | 33218109 | 33218549 | 1,3447 | 3,0376 | 0,014945 |
| 354997_LTR10E | chr6 | 150856969 | 150857593 | 3,5876 | 3,0363 | 5,94E-05 |
| 128963_MSTA,MLT1C,MLT1C,LTR16E2,LTR16E2 | chr15 | 62893810 | 62895404 | 0,5953 | 3,0343 | 0,025562 |
| 57761_LTR44 | chr11 | 33690157 | 33690624 | 0,8888 | 3,0328 | 0,017373 |
| 426151_MLT1H-int,MSTB1,MSTB1,MLT1H-int | chrX | 8867247 | 8867987 | 0,5953 | 3,0267 | 0,038913 |
| 61553_LTR3A,HERVK3-int,HERVK3-int,HERVK3-int,LTR3A | chr11 | 62049778 | 62056355 | 4,6331 | 3,0237 | 1,14E-06 |
| 60159_LTR8A,LTR8A | chr11 | 47776009 | 47776969 | 0,0000 | 3,0227 | 0,040547 |
| 28415_MLT1G,MLT1G,LTR85a | chr1 | 221395224 | 221395889 | 3,3270 | 3,0197 | 0,001059 |
| 176485_MER74A,MER74A | chr19 | 55495558 | 55496389 | 0,0000 | 3,0195 | 0,037208 |
| 189768_LTR1D | chr2 | 82790255 | 82791056 | 0,5953 | 3,0192 | 0,020221 |
| 208402_MER34C2,MER34C2 | chr2 | 206211783 | 206212065 | 0,0000 | 3,0170 | 0,041905 |
| 57762_LTR33,LTR33 | chr11 | 33693083 | 33693680 | 0,0000 | 3,0148 | 0,049028 |
| 68873_MamGypLTR3a,MamGypLTR3 | chr11 | 112121530 | 112122088 | 0,0000 | 3,0044 | 0,028043 |
| 233021_LTR24B,ERV24B_Prim-int,ERV24B_Prim-int,ERV24B_Prim-int,ERV24B_Prim-int,ERV24B_Prim-int | chr21 | 44874842 | 44877939 | 2,7311 | 3,0029 | 0,001281 |
| 246756_MER4B,MER4-int,MER4-int,MER4-int,MER4-int | chr3 | 52155573 | 52158226 | 1,8331 | 2,9987 | 0,007809 |
| 53386_LTR56,LTR29 | chr11 | 6388617 | 6389735 | 0,9169 | 2,9972 | 0,026651 |
| 90476_MER66B | chr12 | 124696652 | 124697106 | 1,8704 | 2,9950 | 0,003831 |
| 212989_LTR19A | chr2 | 233250718 | 233251134 | 0,4278 | 2,9947 | 0,034661 |
| 360654_MLT1G3,MLT1F2 | chr7 | 17598305 | 17599376 | 0,0000 | 2,9867 | 0,031884 |
| 354806_PABL_B | chr6 | 149855094 | 149855610 | 0,4278 | 2,9860 | 0,045373 |
| 148102_MLT1K,MLT1K | chr17 | 8170989 | 8171563 | 3,0019 | 2,9843 | 0,02147 |
| 336332_LTR10B2,HERVIP10B3-int,HERVIP10B3-int,HERVIP10B3-int,LTR10B2 | chr6 | 27178167 | 27186538 | 5,0651 | 2,9800 | 1,59E-05 |
| 13703_MSTA,ERVL-B4-int,ERVL-B4-int | chr1 | 94784675 | 94786547 | 2,8961 | 2,9799 | 0,001715 |
| 369881_MamGypLTR1a | chr7 | 80722421 | 80722752 | 2,1967 | 2,9749 | 0,002962 |
| 169589_LTR38A1,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7C,MER92B | chr19 | 6574424 | 6577385 | 0,0000 | 2,9747 | 0,019983 |
| 3646_LTR103b_Mam | chr1 | 26502461 | 26502559 | 0,4278 | 2,9746 | 0,043917 |
| 302626_MER9a2 | chr5 | 909535 | 910048 | 2,8660 | 2,9671 | 0,000319 |
| 308592_MLT1H | chr5 | 29263966 | 29264456 | 2,7882 | 2,9523 | 0,028836 |
| 375737_MLT1B | chr7 | 120583424 | 120583812 | 0,5953 | 2,9514 | 0,038961 |
| 170000_MER9a3,HERVK9-int | chr19 | 9453903 | 9459482 | 1,0231 | 2,9375 | 0,037432 |
| 410774_MLT1G1,MLT1G1,MLT1G1-int,LTR45C,MLT1G1-int,MLT1G1-int | chr9 | 33171116 | 33172806 | 1,1386 | 2,9348 | 0,030073 |
| 261962_MER74A,MLT1F2,MLT1F2 | chr3 | 149388603 | 149389300 | 0,0000 | 2,9247 | 0,017376 |
| 223210_MSTD | chr20 | 50310962 | 50311280 | 3,2708 | 2,9196 | 0,001174 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 89933_LTR37B | chr12 | 120290225 | 120290296 | 3,0398 | 2,9167 | 0,002243 |
| 2453_LTR75B | chr1 | 16906658 | 16906798 | 2,8332 | 2,9127 | 0,003858 |
| 157432_LTR10A | chr18 | 2302842 | 2303392 | 3,1846 | 2,9050 | 0,001529 |
| 353703_MLT1B,MLT1B-int | chr6 | 143336453 | 143337436 | 1,4511 | 2,9043 | 0,011106 |
| 410794_MLT1D | chr9 | 33227916 | 33228347 | 0,8888 | 2,9009 | 0,023392 |
| 165954_MLT1C,LTR16E1 | chr18 | 62438346 | 62438841 | 2,0024 | 2,8995 | 0,001893 |
| 7689_HERVL-int,MLT2A2 | chr1 | 57522592 | 57523176 | 1,3447 | 2,8881 | 0,019091 |
| 61551_MER4A,MER4A,LTR78B | chr11 | 62042968 | 62044321 | 1,3447 | 2,8828 | 0,028053 |
| 381187_MLT1B,MSTB,MLT1B | chr7 | 155206324 | 155207101 | 0,0000 | 2,8815 | 0,029488 |
| 167332_LTR8 | chr18 | 70384448 | 70385140 | 1,6114 | 2,8761 | 0,026741 |
| 146728_LTR12C | chr16 | 86553261 | 86554406 | 3,0216 | 2,8725 | 0,0019 |
| 261961_MER51A | chr3 | 149385643 | 149386323 | 1,5253 | 2,8692 | 0,005202 |
| 263134_MLT1L | chr3 | 156819003 | 156819149 | 1,6114 | 2,8689 | 0,008078 |
| 175394_LTR13 | chr19 | 48876221 | 48877217 | 0,0000 | 2,8654 | 0,01705 |
| 62091_MLT1F2,MSTD | chr11 | 66969809 | 66970790 | 0,6945 | 2,8632 | 0,042557 |
| 421840_MLT1D | chr9 | 118608790 | 118609212 | 0,0000 | 2,8434 | 0,029477 |
| 169274_LTR3B_ | chr19 | 3175887 | 3176369 | 1,2898 | 2,8423 | 0,048242 |
| 328584_MLT1K,MER34A | chr5 | 160165662 | 160166646 | 3,1526 | 2,8349 | 0,001724 |
| 107938_MLT1A0,MLT1A1,MLT1A0,MLT1A0,MLT1C | chr13 | 106569373 | 106571480 | 0,0000 | 2,8346 | 0,031001 |
| 359105_MLT1A0 | chr7 | 7911209 | 7911546 | 1,4841 | 2,8329 | 0,010543 |
| 377329_MLT1H | chr7 | 131188268 | 131188771 | 1,4841 | 2,8305 | 0,005403 |
| 354942_THE1B | chr6 | 150419520 | 150419883 | 0,4278 | 2,8276 | 0,043832 |
| 224663_LTR12C | chr20 | 57348213 | 57349645 | 0,0000 | 2,8092 | 0,015049 |
| 151862_MER41B | chr17 | 42535067 | 42535524 | 1,4449 | 2,8090 | 0,010317 |
| 353502_LTR30,LTR7,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7 | chr6 | 142014658 | 142021782 | 2,5070 | 2,8085 | 0,011212 |
| 28033_LTR16C | chr1 | 218929541 | 218929846 | 2,3060 | 2,8070 | 0,006864 |
| 183371_MER67C | chr2 | 40718463 | 40718999 | 2,7926 | 2,8061 | 0,00618 |
| 386562_LTR85c,LTR104_Mam,LTR104_Mam | chr8 | 29847620 | 29848168 | 4,0275 | 2,8060 | 2,16E-05 |
| 283121_MER41B | chr4 | 74096174 | 74096724 | 4,2887 | 2,8024 | 0,000327 |
| 377340_MLT1A0 | chr7 | 131225116 | 131225458 | 3,3723 | 2,7950 | 0,001755 |
| 232916_LTR8A,LTR8,LTR8A | chr21 | 44339862 | 44341132 | 4,7817 | 2,7930 | 0,000768 |
| 414587_MER41B | chr9 | 76414364 | 76415007 | 1,6114 | 2,7873 | 0,022103 |
| 170720_MLT1C | chr19 | 14506165 | 14506489 | 0,8888 | 2,7869 | 0,021901 |
| 116963_LTR12C,LTR12C | chr14 | 70254543 | 70256182 | 0,0000 | 2,7851 | 0,047015 |
| 216663_MLT1C,LTR16E2,MER21C | chr20 | 10349639 | 10351578 | 3,6475 | 2,7770 | 0,002274 |
| 334825_LTR12C | chr6 | 17580547 | 17581971 | 0,4278 | 2,7688 | 0,036991 |
| 6256_MER41A | chr1 | 46339833 | 46340059 | 1,9586 | 2,7686 | 0,005054 |
| 294078_MER66B | chr4 | 139014288 | 139014343 | 1,0417 | 2,7625 | 0,015451 |
| 311102_MER4-int,MER4-int,MER4-int,MER4-int,MER4-int,MER11C,MER4-int | chr5 | 43021861 | 43026332 | 3,1104 | 2,7604 | 0,004612 |
| 61949_MER4A,MER4A | chr11 | 65485053 | 65485854 | 1,0231 | 2,7590 | 0,042901 |
| 64916_LTR78,LTR78 | chr11 | 86742007 | 86742917 | 5,2057 | 2,7580 | 0,000597 |
| 245428_MER4D1,MER4D1,LTR18B,MER4D1,LTR7B,MER4D1 | chr3 | 41962833 | 41965130 | 5,0890 | 2,7547 | 3,26E-05 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 62140_LTR23,LTR23,LTR23,ERV24_Prim-int,ERV24_Prim-int,ERV24_Prim-int,LTR49-int,LTR49-int,MER41A,LTR49-int,MER57A-int | chr11 | 67560725 | 67564653 | 0,0000 | 2,7478 | 0,026223 |
| 148328_LTR88a,LTR88a,LTR88b,MLT1K | chr17 | 9894458 | 9895604 | 1,3447 | 2,7427 | 0,043081 |
| 423653_MLT1D | chr9 | 135136898 | 135137143 | 1,2898 | 2,7342 | 0,048536 |
| 245278_LTR16E1 | chr3 | 40970488 | 40971010 | 2,8332 | 2,7307 | 0,004648 |
| 132791_LTR33 | chr15 | 91943836 | 91944054 | 2,0847 | 2,7193 | 0,017014 |
| 179012_MLT1O | chr2 | 11389467 | 11389950 | 2,3060 | 2,7162 | 0,005249 |
| 396196_LTR64 | chr8 | 93753915 | 93754432 | 1,2898 | 2,7025 | 0,036197 |
| 150347_LTR8A,LTR8A | chr17 | 27339953 | 27340913 | 2,1920 | 2,6847 | 0,028582 |
| 426220_MSTA | chrX | 9164452 | 9164853 | 2,1920 | 2,6841 | 0,022525 |
| 388297_THE1D | chr8 | 40159843 | 40160206 | 2,6639 | 2,6828 | 0,002311 |
| 32565_HERVK14C-int,HERVK14C-int,HERVK14C-int,HERVK14C-int,LTR14C | chr1 | 247106027 | 247111180 | 1,3447 | 2,6771 | 0,0339 |
| 234929_MLT1H | chr22 | 27719591 | 27719907 | 3,6477 | 2,6689 | 0,000773 |
| 52781_MER4D0,MER4D0 | chr11 | 1764351 | 1764718 | 1,6370 | 2,6433 | 0,015542 |
| 340497_MSTD,MER89 | chr6 | 56334668 | 56335782 | 0,0000 | 2,6350 | 0,012086 |
| 33729_MLT1A0,MLT1A0 | chr10 | 5292175 | 5292941 | 0,0000 | 2,6332 | 0,041975 |
| 151945_MER11C | chr17 | 43365611 | 43366679 | 3,4619 | 2,6293 | 0,00281 |
| 156738_LTR7,MER4E1,LTR38A1,LTR38A1,LTR38A1 | chr17 | 80258156 | 80260057 | 4,9463 | 2,6273 | 0,000407 |
| 310467_LTR33A_ | chr5 | 39458734 | 39459239 | 3,2082 | 2,6249 | 0,002554 |
| 156902_MER21A,MER21A,LTR8,LTR8,MER21A | chr17 | 82206284 | 82208047 | 2,1394 | 2,6237 | 0,023756 |
| 348592_MER34C2,MER34C2 | chr6 | 110816383 | 110817068 | 1,7357 | 2,6217 | 0,010103 |
| 61552_MER4E,LTR7Y,MER4E | chr11 | 62044620 | 62045792 | 5,0130 | 2,6163 | 0,00046 |
| 107272_MER4A1 | chr13 | 102844793 | 102845223 | 3,2630 | 2,6163 | 0,001175 |
| 305302_MLT1C | chr5 | 14038719 | 14039172 | 0,9169 | 2,6126 | 0,032995 |
| 174723_MLT1A0 | chr19 | 44164478 | 44164746 | 1,5664 | 2,6058 | 0,012806 |
| 365657_LTR22A | chr7 | 50533873 | 50534328 | 0,5953 | 2,5976 | 0,038795 |
| 111363_LTR7B,MLT1J2,MLT1J2,HERVH-int,HERVH-int,HERVH-int,LTR7B,MLT1H | chr14 | 34650295 | 34655394 | 2,0464 | 2,5890 | 0,042442 |
| 175822_LTR3A,HERVK3-int,HERVK3-int,HERVK3-int,HERVK3-int,LTR3A,HERVK3-int,LTR3A,LTR3A | chr19 | 51804695 | 51811939 | 2,7010 | 2,5888 | 0,020685 |
| 398233_LTR26,LTR18A,ERVL-E-int,MLT1C,MLT1C,LTR16E2 | chr8 | 106059648 | 106063263 | 3,6863 | 2,5853 | 0,00065 |
| 381685_LTR56,LTR12 | chr7 | 158862275 | 158863424 | 2,1967 | 2,5846 | 0,016416 |
| 386254_LTR56,LTR23 | chr8 | 27773106 | 27773577 | 2,3496 | 2,5824 | 0,00279 |
| 231779_MER4A1,MER4A1 | chr21 | 38735057 | 38735794 | 2,2493 | 2,5745 | 0,030474 |
| 30798_LTR8A | chr1 | 235366295 | 235366990 | 1,0417 | 2,5721 | 0,031254 |
| 210495_MLT1O,MER39,MER39 | chr2 | 219595280 | 219596817 | 4,1348 | 2,5663 | 0,001352 |
| 245761_LTR16A1 | chr3 | 43734082 | 43734486 | 2,9004 | 2,5657 | 0,001293 |
| 237208_MER41C,LTR46,LTR46,MER41C,LTR7Y,MER41C | chr22 | 45160741 | 45162810 | 2,7291 | 2,5607 | 0,004241 |
| 193827_MER21C | chr2 | 111659358 | 111660135 | 1,2751 | 2,5530 | 0,021331 |
| 72660_MER41A,MER41A | chr12 | 8679819 | 8680678 | 1,0417 | 2,5352 | 0,033451 |
| 270431_LTR37B | chr4 | 6985402 | 6985819 | 1,0417 | 2,5316 | 0,032294 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 74418_MLT1J2 | chr12 | 19946540 | 19946934 | 3,2091 | 2,5281 | 0,00257 |
| 384368_LTR8B,LTR8B | chr8 | 16409119 | 16409756 | 1,2751 | 2,5240 | 0,030346 |
| 102872_LTR40c | chr13 | 79484963 | 79485180 | 2,0464 | 2,5226 | 0,011847 |
| 64151_MLT2B4,LTR71A | chr11 | 81330990 | 81332135 | 2,1967 | 2,5224 | 0,023577 |
| 197878_MLT1K | chr2 | 136236799 | 136237352 | 4,3900 | 2,5205 | 4,57E-05 |
| 417084_MER41A | chr9 | 89308866 | 89309412 | 3,9344 | 2,5103 | 0,001412 |
| 46156_MER68-int,LTR15,LTR15,MER68-int | chr10 | 89349524 | 89351909 | 2,9344 | 2,4996 | 0,010498 |
| 447216_MLT1L,MLT1A,MLT1L | chrX | 150691822 | 150692455 | 1,2898 | 2,4945 | 0,043048 |
| 2454_LOR1b | chr1 | 16907530 | 16907942 | 3,9344 | 2,4887 | 0,005328 |
| 216665_LTR35 | chr20 | 10352220 | 10352804 | 2,4410 | 2,4861 | 0,009547 |
| 333824_MLT1B | chr6 | 11771634 | 11772010 | 3,3461 | 2,4839 | 0,007975 |
| 52858_MER4D1,MER4D1 | chr11 | 2397539 | 2398606 | 4,4710 | 2,4802 | 0,001654 |
| 423736_LTR8,MER4CL34,MER4CL34 | chr9 | 135468601 | 135470603 | 3,7701 | 2,4774 | 0,012219 |
| 53647_MER4C | chr11 | 8160917 | 8161342 | 1,1678 | 2,4741 | 0,033911 |
| 275773_MER65-int,LTR27E,MER101-int,MER101-int,MER101-int,MER65-int,MER34,MER65-int,MER65-int,MER65D | chr4 | 31612896 | 31618631 | 2,7278 | 2,4689 | 0,01473 |
| 318610_MER39 | chr5 | 94424057 | 94424230 | 1,9878 | 2,4642 | 0,047899 |
| 202686_MER74B | chr2 | 168801999 | 168802275 | 2,3060 | 2,4594 | 0,00853 |
| 396207_MLT1J2,MLT1J2 | chr8 | 93845794 | 93846165 | 2,7278 | 2,4483 | 0,036847 |
| 149790_MLT1B | chr17 | 19111648 | 19111720 | 2,3064 | 2,4482 | 0,012576 |
| 390729_MER74B | chr8 | 58850173 | 58850725 | 3,6824 | 2,4467 | 0,001662 |
| 364707_MER67C,MER67C,MER67C,MER67C | chr7 | 44987085 | 44987850 | 3,8839 | 2,4381 | 0,007057 |
| 57763_MLT1A0 | chr11 | 33693282 | 33693613 | 0,8888 | 2,4362 | 0,045209 |
| 6165_LTR34,LTR34,LTR34 | chr1 | 45522400 | 45523599 | 5,1251 | 2,4294 | 0,00091 |
| 246219_LTR71A,HERVP71A-int,HERVP71A-int,LTR71A | chr3 | 46111060 | 46118672 | 3,6762 | 2,4222 | 0,013907 |
| 114334_MER21B,MER21B | chr14 | 52110793 | 52111338 | 2,3064 | 2,4128 | 0,018423 |
| 238008_MER57F,MER41-int | chr22 | 50539605 | 50540368 | 1,7631 | 2,4079 | 0,033302 |
| 153843_LTR71A | chr17 | 58632587 | 58633010 | 1,3079 | 2,3989 | 0,033371 |
| 4954_MER67C | chr1 | 36465009 | 36465694 | 2,8973 | 2,3914 | 0,003376 |
| 364726_MER49,LTR39,MER4D,MER21C | chr7 | 45112600 | 45113830 | 4,1067 | 2,3839 | 0,004169 |
| 334339_MLT1B,MLT1C,MLT1C | chr6 | 14731204 | 14732784 | 1,7631 | 2,3838 | 0,029052 |
| 114113_LTR90B | chr14 | 50669136 | 50669234 | 1,9586 | 2,3792 | 0,028449 |
| 183370_MLT2B3 | chr2 | 40716955 | 40717505 | 2,4757 | 2,3764 | 0,022948 |
| 128178_MER34C2,MER34C2,MLT2A2 | chr15 | 56018071 | 56019660 | 2,0847 | 2,3756 | 0,032036 |
| 203054_MER49 | chr2 | 171521096 | 171521986 | 1,4841 | 2,3666 | 0,047456 |
| 237955_ERV3-16A3_I-int | chr22 | 49933204 | 49933439 | 3,4198 | 2,3652 | 0,004909 |
| 223211_MLT1K | chr20 | 50313276 | 50313754 | 4,6689 | 2,3579 | 0,001892 |
| 333035_MLT1B | chr6 | 6686244 | 6686631 | 3,4097 | 2,3559 | 0,009685 |
| 405303_LTR8 | chr8 | 142676693 | 142676841 | 3,3928 | 2,3455 | 0,004026 |
| 223433_MLT1G1 | chr20 | 51347649 | 51347806 | 1,5664 | 2,3387 | 0,047771 |
| 320974_MER67C | chr5 | 108746877 | 108747564 | 2,3680 | 2,3360 | 0,016433 |
| 390728_MLT2B2,MSTD | chr8 | 58842232 | 58842528 | 2,7926 | 2,3325 | 0,014746 |
| 309159_LTR26B,LTR26B,LTR26B | chr5 | 32315144 | 32316045 | 3,7461 | 2,3167 | 0,011006 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 53963_THE1D | chr11 | 10860037 | 10860402 | 3,5426 | 2,3137 | 0,009568 |
| 90445_MLT1I,MLT1B,MLT1B | chr12 | 124624440 | 124625709 | 4,4713 | 2,3121 | 0,011667 |
| 96193_MER89-int | chr13 | 41133777 | 41134462 | 2,5094 | 2,3121 | 0,01626 |
| 115761_MLT1L,MLT1L | chr14 | 61559677 | 61560109 | 2,7311 | 2,3089 | 0,017544 |
| 369052_LTR8B | chr7 | 74452254 | 74452948 | 3,4169 | 2,3077 | 0,002452 |
| 199035_MamGypLTR1b | chr2 | 142874498 | 142874653 | 2,3064 | 2,2992 | 0,028726 |
| 174424_LTR8B | chr19 | 42875094 | 42875873 | 2,5094 | 2,2936 | 0,007366 |
| 233034_MSTB | chr21 | 45004930 | 45005334 | 2,4757 | 2,2902 | 0,010921 |
| 337804_MLT1L,LTR8A,LTR8A,MLT1H,MLT1H,MLT1N2 | chr6 | 37242880 | 37245653 | 3,7824 | 2,2831 | 0,012701 |
| 354998_LTR47B4 | chr6 | 150857629 | 150858046 | 2,8525 | 2,2739 | 0,004605 |
| 76181_MER65D,MER65D | chr12 | 31730202 | 31731056 | 5,7414 | 2,2736 | 0,0036 |
| 229896_LTR50,LTR33A | chr21 | 28993609 | 28994214 | 1,7631 | 2,2703 | 0,025483 |
| 128283_MLT1K,MER34A,MER34A,MLT1K,MLT1J2,MLT1J2-int | chr15 | 56737063 | 56738545 | 2,5897 | 2,2682 | 0,034379 |
| 63476_MLT1B | chr11 | 76770087 | 76770368 | 3,6714 | 2,2658 | 0,005107 |
| 302625_LTR13 | chr5 | 902585 | 903579 | 3,0112 | 2,2643 | 0,01396 |
| 255474_LTR18A,LTR101_Mam,MLT1D | chr3 | 106765348 | 106767559 | 3,4181 | 2,2560 | 0,014125 |
| 445760_MLT1J2 | chrX | 142998605 | 142998826 | 1,8057 | 2,2535 | 0,026429 |
| 90449_MLT1L,MLT1L,MLT1G,MLT1G,MLT1L | chr12 | 124633596 | 124634867 | 4,1936 | 2,2377 | 0,005639 |
| 440483_LTR33 | chrX | 103405721 | 103405967 | 1,1678 | 2,2345 | 0,046775 |
| 357287_HERV15-int | chr6 | 166004974 | 166005081 | 3,7662 | 2,2255 | 0,003386 |
| 267556_THE1C | chr3 | 185960874 | 185961009 | 3,7218 | 2,2180 | 0,009631 |
| 384429_PRIMA4_LTR,LTR19A,PRIMA4_LTR | chr8 | 16716812 | 16717922 | 1,2751 | 2,2066 | 0,044961 |
| 337769_MER57A1,PABL_A-int,PABL_A-int,PABL_A-int,PABL_A-int,PABL_A-int,MER57-int,MER57-int,MER57-int,MER57A1 | chr6 | 36878575 | 36885493 | 2,9004 | 2,1978 | 0,031143 |
| 74423_LTR35A,LTR35A | chr12 | 19970993 | 19971796 | 7,6980 | 2,1933 | 0,000218 |
| 349445_MLT1D | chr6 | 116280940 | 116281424 | 4,6980 | 2,1922 | 0,003082 |
| 327126_LTR62 | chr5 | 150702658 | 150702964 | 1,4841 | 2,1896 | 0,042668 |
| 52635_LTR13A,LTR13A,LTR13A,MER41A,LTR40b | chr11 | 353858 | 355901 | 5,5212 | 2,1842 | 0,001427 |
| 154650_MLT1H2,MLT2B4,MLT2B4,MLT1H2 | chr17 | 66509019 | 66509784 | 3,1846 | 2,1792 | 0,007847 |
| 258315_MLT1N2,MLT1K | chr3 | 126782496 | 126782900 | 3,4225 | 2,1789 | 0,034571 |
| 196556_MER92C | chr2 | 127527074 | 127527554 | 3,7526 | 2,1787 | 0,009193 |
| 174419_LTR8B,LTR101_Mam | chr19 | 42851137 | 42852046 | 3,7969 | 2,1671 | 0,004697 |
| 383005_MSTB1 | chr8 | 8297515 | 8297928 | 4,2414 | 2,1316 | 0,014159 |
| 72359_LTR83,LTR83,LTR83 | chr12 | 6431060 | 6431907 | 1,9586 | 2,1273 | 0,042674 |
| 232816_MER74C,MER74C | chr21 | 43710739 | 43710969 | 4,0886 | 2,1238 | 0,010373 |
| 236641_MLT1K | chr22 | 40651455 | 40651837 | 2,1920 | 2,1232 | 0,03678 |
| 151943_PRIMA41-int,PRIMA41-int,PRIMA41-int,PRIMA41-int,MER41C | chr17 | 43357143 | 43360454 | 8,3388 | 2,1222 | 0,042898 |
| 216998_MLT1H | chr20 | 11852388 | 11852948 | 4,1340 | 2,1205 | 0,015222 |
| 74424_MamGypLTR1d | chr12 | 19972339 | 19972561 | 3,2091 | 2,1072 | 0,012989 |
| 318609_MER39,LTR79 | chr5 | 94422091 | 94423331 | 3,7147 | 2,1054 | 0,03454 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 245893_MER84,MER84-int,MER84-int,MER84-int,MER84-int,MER84 | chr3 | 44333731 | 44337778 | 2,8588 | 2,1007 | 0,036512 |
| 70874_ERVL-B4-int,MLT2B1,MLT2B1 | chr11 | 128093965 | 128097207 | 4,8032 | 2,1000 | 0,001184 |
| 133645_THE1B | chr15 | 97690688 | 97691045 | 3,8364 | 2,0936 | 0,015382 |
| 173263_MLT1A0,MER41D | chr19 | 34672788 | 34673925 | 4,0641 | 2,0936 | 0,008997 |
| 318806_LTR6B,HERVS71-int,HERVS71-int,HERVS71-int,HERVS71-int | chr5 | 95707336 | 95712075 | 5,0924 | 2,0930 | 0,007643 |
| 90469_MLT1L,MLT1L,MLT1C | chr12 | 124680927 | 124682225 | 4,4288 | 2,0923 | 0,006281 |
| 397337_MER52A | chr8 | 100907235 | 100908779 | 1,8057 | 2,0879 | 0,03243 |
| 423082_MER110A,MLT1D,MER110A,MER110A | chr9 | 129753488 | 129754525 | 3,0112 | 2,0857 | 0,038168 |
| 43975_LTR101_Mam | chr10 | 75533060 | 75533796 | 5,1542 | 2,0840 | 0,008356 |
| 156905_MLT1C | chr17 | 82216111 | 82216271 | 2,3060 | 2,0827 | 0,02915 |
| 237044_MLT1B | chr22 | 44361546 | 44361921 | 2,7200 | 2,0821 | 0,042933 |
| 33728_MSTD,MSTB1 | chr10 | 5291580 | 5292813 | 4,1531 | 2,0617 | 0,00422 |
| 197879_MLT1K | chr2 | 136238081 | 136238436 | 3,0110 | 2,0612 | 0,041977 |
| 330562_LTR16C,MLT1A-int,MLT1A,MLT1F1,MLT1F1,MLT1F1,MLT1F1 | chr5 | 173453476 | 173455503 | 6,7414 | 2,0591 | 0,000283 |
| 61540_MER41B,MER41B | chr11 | 62014070 | 62014954 | 3,6693 | 2,0491 | 0,017544 |
| 277706_MER21C | chr4 | 41506728 | 41507652 | 4,9099 | 2,0485 | 0,000774 |
| 136667_MLT1K | chr16 | 15489167 | 15489300 | 3,4838 | 2,0406 | 0,011604 |
| 303619_MLT1A1 | chr5 | 6409005 | 6409403 | 2,3060 | 2,0317 | 0,022433 |
| 149956_MER41A,MER41-int,MER41-int,MER51-int,MER51-int,MER51-int,MER57-int | chr17 | 20192870 | 20198728 | 5,3097 | 2,0094 | 0,001021 |
| 18703_LTR7B | chr1 | 150580685 | 150581114 | 2,9344 | -3,2146 | 0,017308 |
| 170944_LTR7C | chr19 | 15835690 | 15836170 | 3,7257 | -2,8707 | 0,020967 |
| 328158_MER92B,MER67C | chr5 | 157477736 | 157478162 | 5,0423 | -2,8695 | 0,007033 |
| 60111_THE1D | chr11 | 47138055 | 47138417 | 2,1967 | -2,7404 | 0,021171 |
| 53398_MLT2B1 | chr11 | 6495170 | 6495556 | 2,9619 | -2,6390 | 0,039 |
| 130355_MLT1J | chr15 | 74210732 | 74210968 | 5,7248 | -2,4066 | 0,003727 |
| 20112_MLT1B | chr1 | 162855777 | 162856190 | 3,6145 | -2,3585 | 0,03179 |
| 29563_MLT1A0 | chr1 | 228392003 | 228392129 | 3,5876 | -2,3354 | 0,045141 |
| 380287_MER41B | chr7 | 150322211 | 150322830 | 4,6793 | -2,2743 | 0,013033 |
| 6833_MLT1C | chr1 | 50985077 | 50985466 | 4,2034 | -2,0985 | 0,040666 |

### Table 8.12 Significantly deregulated ERVs with H3K9ac peak in SETDB1 KD, Hela cell

| RepeatID | chr | start | end | Log2 base Mean | log2 Fold Change | p-value |
|---|---|---|---|---|---|---|
| 108349_MER21C | chr13 | 108828500 | 108829198 | 4,5439 | -1,8049 | 0,018998 |
| 114326_MLT1F2,MLT1F2,MER41B,LTR56 | chr14 | 52070629 | 52072212 | 4,2565 | 1,4666 | 0,029856 |
| 128281_MER65C | chr15 | 56734076 | 56734537 | 0,0000 | 2,4981 | 0,040708 |
| 131807_MLT1D | chr15 | 85460499 | 85461017 | 5,0269 | -2,0055 | 0,041225 |
| 148105_MER50C,MER50C,MER50C | chr17 | 8199887 | 8201250 | 2,7882 | -3,0005 | 0,01832 |
| 152064_LTR12C | chr17 | 44568043 | 44569479 | 2,0024 | 2,3954 | 0,025733 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 170944_LTR7C | chr19 | 15835690 | 15836170 | 3,7257 | -2,4475 | 0,045042 |
| 172935_MSTB2,MLT1A0,MLT1G3 | chr19 | 32372474 | 32373383 | 2,9285 | -2,4297 | 0,0365 |
| 172986_LTR12,HERV9-int,HERV9N-int,HERV9N-int,LTR12 | chr19 | 32693105 | 32700522 | 5,2171 | 2,0942 | 0,004001 |
| 174545_LTR8B | chr19 | 43265000 | 43265754 | 1,1678 | -1,9050 | 0,046815 |
| 191933_MER11C,HERVK11-int,MER11A | chr2 | 100361987 | 100365971 | 1,8057 | 2,4343 | 0,012008 |
| 192142_LTR1F2 | chr2 | 101355655 | 101356399 | 6,2010 | -1,6142 | 0,022389 |
| 199036_MER72 | chr2 | 142875416 | 142876090 | 0,0000 | 2,2025 | 0,042543 |
| 216973_MER67A | chr20 | 11797043 | 11797580 | 2,8473 | -3,3153 | 0,016011 |
| 224663_LTR12C | chr20 | 57348213 | 57349645 | 0,0000 | 4,6998 | 0,000261 |
| 231802_MLT1D | chr21 | 38840820 | 38841271 | 0,0000 | 2,7548 | 0,022673 |
| 236173_MLT1B | chr22 | 36371142 | 36371526 | 0,0000 | 2,9701 | 0,03282 |
| 236760_MER51B | chr22 | 41827412 | 41827761 | 0,0000 | 3,0318 | 0,032674 |
| 2450_ERV24_Prim-int | chr1 | 16901742 | 16902026 | 2,2521 | -2,7805 | 0,043463 |
| 247425_MER83 | chr3 | 57972370 | 57972789 | 2,4488 | -2,7014 | 0,022494 |
| 256914_LTR12C,LTR12C,MLT2A1 | chr3 | 115791379 | 115793165 | 4,4545 | 1,4159 | 0,036541 |
| 261962_MER74A,MLT1F2,MLT1F2 | chr3 | 149388603 | 149389300 | 0,0000 | 2,3721 | 0,04404 |
| 268469_MLT1C,MLT1C | chr3 | 192567498 | 192568309 | 3,1150 | -3,3646 | 0,009914 |
| 270080_MER31B | chr4 | 5371060 | 5371517 | 2,9508 | -2,7289 | 0,038471 |
| 270389_LTR12C | chr4 | 6672540 | 6673747 | 6,0203 | 1,3152 | 0,017529 |
| 294566_MLT1L,MLT1G | chr4 | 142092587 | 142093449 | 0,4278 | 2,9541 | 0,01471 |
| 306229_HERV17-int,LTR12F | chr5 | 18745563 | 18746356 | 4,2984 | 1,4733 | 0,04672 |
| 307381_LTR12C | chr5 | 23951094 | 23952511 | 6,7392 | 1,3667 | 0,040613 |
| 308184_MER61C,MER61-int,MER61-int,MER61C | chr5 | 27472073 | 27477973 | 9,9176 | 1,2927 | 0,011086 |
| 322253_LTR10F,LTR10F,LTR10F,LTR10F,MER41A,LTR10F,LTR10F,MER41A,MLT1D | chr5 | 115928854 | 115931428 | 2,6639 | 2,2111 | 0,009957 |
| 330563_MLT1B | chr5 | 173456048 | 173456471 | 0,8888 | 1,9852 | 0,048879 |
| 336983_MLT1G3,MLT1G3 | chr6 | 30767566 | 30768367 | 0,0000 | 3,1182 | 0,005775 |
| 340070_MLT1A | chr6 | 53315031 | 53315409 | 3,4225 | -2,9784 | 0,014191 |
| 353703_MLT1B,MLT1B-int | chr6 | 143336453 | 143337436 | 1,4511 | -2,2584 | 0,033997 |
| 353949_LTR7Y,HERVH-int,HERVH-int,HERVH-int,HERVH-int,HERVH-int,LTR7Y | chr6 | 144923174 | 144928858 | 9,0288 | -1,5695 | 0,021127 |
| 354997_LTR10E | chr6 | 150856969 | 150857593 | 3,5876 | 1,5039 | 0,047637 |
| 358426_MER41B,MER41B,LTR61,MER74A,MER74A | chr7 | 3193353 | 3194959 | 0,0000 | 2,6318 | 0,045298 |
| 377937_PABL_B-int,PABL_B-int,PABL_B-int,PABL_B-int | chr7 | 135320638 | 135325703 | 1,8727 | -2,5373 | 0,048642 |
| 381187_MLT1B,MSTB,MLT1B | chr7 | 155206324 | 155207101 | 0,0000 | 3,8681 | 0,005904 |
| 385462_LTR12C | chr8 | 23069936 | 23071352 | 6,8348 | 1,6764 | 0,01076 |
| 395815_MLT1F1 | chr8 | 91180080 | 91180616 | 0,0000 | 3,1515 | 0,010521 |
| 401802_LTR84a | chr8 | 125686735 | 125687396 | 2,5094 | -2,6767 | 0,049005 |
| 410765_MER31A | chr9 | 33002296 | 33002777 | 2,1743 | -2,7848 | 0,020775 |
| 422578_MLT1J,LTR7B,MLT1J,LTR7B | chr9 | 124142902 | 124144014 | 3,0216 | -3,2689 | 0,002637 |
| 433659_MER39,THE1B,MER39,MER39,LTR8A,LTR8A,MER34C2 | chrX | 46322510 | 46325275 | 8,9790 | -1,6965 | 0,001297 |
| 46156_MER68-int,LTR15,LTR15,MER68-int | chr10 | 89349524 | 89351909 | 2,9344 | -3,2755 | 0,004093 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 50459_MER34A1,MER34 | chr10 | 119875009 | 119875927 | 1,3685 | 2,0628 | 0,046081 |
| 53257_LTR75_1,LTR12C,LTR75_1 | chr11 | 5682462 | 5684428 | 0,0000 | 2,6051 | 0,037674 |
| 6123_LTR18B,LTR67B,LTR67B,LTR67B | chr1 | 45266082 | 45267979 | 3,2630 | -2,7625 | 0,032796 |
| 62142_LTR49-int,LTR49-int,LTR49-int,LTR49-int,MER51E,LTR49-int,LTR13_,LTR49-int,LTR49-int,LTR49-int,ERV24B_Prim-int,LTR23 | chr11 | 67569605 | 67575312 | 0,0000 | 2,3308 | 0,036952 |
| 66293_MLT1C | chr11 | 95877507 | 95877969 | 1,2898 | 2,4122 | 0,03995 |
| 68873_MamGypLTR3a,MamGypLTR3 | chr11 | 112121530 | 112122088 | 0,0000 | 2,7642 | 0,039872 |
| 78307_MLT1E2 | chr12 | 48650846 | 48651162 | 3,9463 | -2,2789 | 0,041663 |
| 92580_MER66A | chr12 | 132393604 | 132394031 | 2,7882 | -2,8343 | 0,027666 |

# 8. References

(2018). Picard toolkit. Broad Institute, GitHub repository.

(2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res *49*, D480-d489.

Ahringer, J., and Gasser, S.M. (2018). Repressive Chromatin in Caenorhabditis elegans: Establishment, Composition, and Function. Genetics *208*, 491-511.

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A., and Mason, C.E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biology *13*, R87.

Alpert, T., Herzel, L., and Neugebauer, K.M. (2017). Perfect timing: splicing and transcription rates in living cells. Wiley interdisciplinary reviews RNA *8*.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Andersen, P.R., Domanski, M., Kristiansen, M.S., Storvall, H., Ntini, E., Verheggen, C., Schein, A., Bunkenborg, J., Poser, I., Hallais, M.*, et al.* (2013). The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat Struct Mol Biol *20*, 1367-1376.

Änkö, M.L. (2014). Regulation of gene expression programmes by serine-arginine rich splicing factors. Semin Cell Dev Biol *32*, 11-21.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Molecular cell *31*, 785-799.

Ast, G. (2004). How did alternative splicing evolve? Nature reviews Genetics *5*, 773-782.

Attermann, A.S., Bjerregaard, A.M., Saini, S.K., Grønbæk, K., and Hadrup, S.R. (2018). Human endogenous retroviruses and their implication for immunotherapeutics of cancer. Annals of Oncology *29*, 2183-2191.

Bannert, N., and Kurth, R. (2006). The evolutionary dynamics of human endogenous retroviral families. Annual review of genomics and human genetics *7*, 149-173.

Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? Nucleic Acids Res *33*, 7138-7150.

Bayne, E.H., Portoso, M., Kagansky, A., Kos-Braun, I.C., Urano, T., Ekwall, K., Alves, F., Rappsilber, J., and Allshire, R.C. (2008). Splicing factors facilitate RNAi-directed silencing in fission yeast. Science *322*, 602-606.

Bellak, J.M., Hanson, K., Brockman-Schneider, R.A., Dagher, H., and Gern, J.E. (2005). RV1A and dsRNA highly induce a biphasic expression of MIP-3α/CCL20 in human bronchial epithelial cells (hBECs). Journal of Allergy and Clinical Immunology *115*, S51.

Bénit, L., Calteau, A., and Heidmann, T. (2003). Characterization of the low-copy HERV-Fc family: evidence for recent integrations in primates of elements with coding envelope genes. Virology *312*, 159-168.

Berrens, R.V., Andrews, S., Spensberger, D., Santos, F., Dean, W., Gould, P., Sharif, J., Olova, N., Chandra, T., Koseki, H.*, et al.* (2017). An endosiRNA-Based Repression Mechanism Counteracts Transposon Activation during Global DNA Demethylation in Embryonic Stem Cells. Cell Stem Cell *21*, 694-703.e697.

Blahnik, K.R., Dou, L., Echipare, L., Iyengar, S., O'Geen, H., Sanchez, E., Zhao, Y., Marra, M.A., Hirst, M., Costello, J.F.*, et al.* (2011). Characterization of the contradictory chromatin signatures at the 3' exons of zinc finger genes. PLoS One *6*, e17121.

Bourque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H.L., Macfarlan, T.S.*, et al.* (2018). Ten things you should know about transposable elements. Genome Biology *19*, 199.

Brocks, D., Schmidt, C.R., Daskalakis, M., Jang, H.S., Shah, N.M., Li, D., Li, J., Zhang, B., Hou, Y., Laudato, S.*, et al.* (2017). DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. Nature genetics *49*, 1052-1060.

Bühler, M., and Moazed, D. (2007). Transcription and RNAi in heterochromatic gene silencing. Nat Struct Mol Biol *14*, 1041-1048.

Carter, A.C., Xu, J., Nakamoto, M.Y., Wei, Y., Zarnegar, B.J., Shi, Q., Broughton, J.P., Ransom, R.C., Salhotra, A., Nagaraja, S.D.*, et al.* (2020a). Spen links RNA-mediated endogenous retrovirus silencing and X chromosome inactivation. eLife *9*, e54508.

Carter, A.C., Xu, J., Nakamoto, M.Y., Wei, Y., Zarnegar, B.J., Shi, Q., Broughton, J.P., Ransom, R.C., Salhotra, A., Nagaraja, S.D.*, et al.* (2020b). Spen links RNA-mediated endogenous retrovirus silencing and X chromosome inactivation. eLife *9*.

Chang, Y.L., Hsu, Y.J., Chen, Y., Wang, Y.W., and Huang, S.M. (2017). Theophylline exhibits anti-cancer activity via suppressing SRSF3 in cervical and breast cancer cell lines. Oncotarget *8*, 101461-101474.

Chang, Y.L., Liu, S.T., Wang, Y.W., Lin, W.S., and Huang, S.M. (2018). Amiodarone promotes cancer cell death through elevated truncated SRSF3 and downregulation of miR-224. Oncotarget *9*, 13390-13406.

Chelmicki, T., Roger, E., Teissandier, A., Dura, M., Bonneville, L., Rucli, S., Dossin, F., Fouassier, C., Lameiras, S., and Bourc'his, D. (2021). m(6)A RNA methylation regulates the fate of endogenous retroviruses. Nature *591*, 312-316.

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A.*, et al.* (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell *162*, 974-986.

Chinen, M., Morita, M., Fukumura, K., and Tani, T. (2010). Involvement of the spliceosomal U4 small nuclear RNA in heterochromatic gene silencing at fission yeast centromeres. The Journal of biological chemistry *285*, 5630-5638.

Chuong, E.B., Elde, N.C., and Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science *351*, 1083-1087.

Clark, M.B., Johnston, R.L., Inostroza-Ponta, M., Fox, A.H., Fortini, E., Moscato, P., Dinger, M.E., and Mattick, J.S. (2012). Genome-wide analysis of long noncoding RNA stability. Genome Res *22*, 885-898.

Conti, A., Rota, F., Ragni, E., Favero, C., Motta, V., Lazzari, L., Bollati, V., Fustinoni, S., and Dieci, G. (2016). Hydroquinone induces DNA hypomethylation-independent overexpression of retroelements in human leukemia and hematopoietic stem cells. Biochemical and Biophysical Research Communications *474*, 691-695.

Coppin, L., Leclerc, J., Vincent, A., Porchet, N., and Pigny, P. (2018). Messenger RNA Life-Cycle in Cancer Cells: Emerging Role of Conventional and Non-Conventional RNA-Binding Proteins? International Journal of Molecular Sciences *19*, 650.

Corbo, C., Orrù, S., and Salvatore, F. (2013). SRp20: an overview of its role in human diseases. Biochem Biophys Res Commun *436*, 1-5.

Cuellar, T.L., Herzner, A.-M., Zhang, X., Goyal, Y., Watanabe, C., Friedman, B.A., Janakiraman, V., Durinck, S., Stinson, J., Arnott, D.*, et al.* (2017). ---Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia--. The Journal of Cell Biology *216*, 3535.

Daskalakis, M., Brocks, D., Sheng, Y.-H., Islam, M.S., Ressnerova, A., Assenov, Y., Milde, T., Oehme, I., Witt, O., Goyal, A.*, et al.* (2018). Reactivation of endogenous retroviral elements via treatment with DNMT- and HDAC-inhibitors. Cell Cycle *17*, 811-822.

Davidson, L., Kerr, A., and West, S. (2012). Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. The EMBO journal *31*, 2566-2578.

Day, D.S., Luquette, L.J., Park, P.J., and Kharchenko, P.V. (2010). Estimating enrichment of repetitive elements from high-throughput sequence data. Genome Biol *11*, R69.

Derti, A., Garrett-Engele, P., Macisaac, K.D., Stevens, R.C., Sriram, S., Chen, R., Rohl, C.A., Johnson, J.M., and Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. Genome Res *22*, 1173-1183.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F.*, et al.* (2012). Landscape of transcription in human cells. Nature *489*, 101-108.

Do, D.V., Strauss, B., Cukuroglu, E., Macaulay, I., Wee, K.B., Hu, T.X., Igor, R.D.L.M., Lee, C., Harrison, A., Butler, R.*, et al.* (2018). SRSF3 maintains transcriptome integrity in oocytes by regulation of alternative splicing and transposable elements. Cell Discovery *4*, 33.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat Commun *7*, 12626.

Dumesic, P.A., and Madhani, H.D. (2013). The spliceosome as a transposon sensor. RNA biology *10*, 1653-1660.

Dumesic, P.A., Natarajan, P., Chen, C., Drinnenberg, I.A., Schiller, B.J., Thompson, J., Moresco, J.J., Yates, J.R., 3rd, Bartel, D.P., and Madhani, H.D. (2013). Stalled spliceosomes are a signal for RNAi-mediated genome defense. Cell *152*, 957-968.

Ebelt, N.D., Zuniga, E., Johnson, B.L., Diamond, D.J., and Manuel, E.R. (2020). 5-Azacytidine Potentiates Anti-tumor Immunity in a Model of Pancreatic Ductal Adenocarcinoma. Frontiers in Immunology *11*.

Eberle, A.B., Jordan-Pla, A., Ganez-Zapater, A., Hessle, V., Silberberg, G., von Euler, A., Silverstein, R.A., and Visa, N. (2015). An Interaction between RRP6 and SU(VAR)3-9 Targets RRP6 to Heterochromatin and Contributes to Heterochromatin Maintenance in Drosophila melanogaster. PLoS Genet *11*, e1005523.

Emerson, R.O., and Thomas, J.H. (2011). Gypsy and the Birth of the SCAN Domain. Journal of Virology *85*, 12043-12052.

Espada, J., Ballestar, E., Fraga, M.F., Villar-Garea, A., Juarranz, A., Stockert, J.C., Robertson, K.D., Fuks, F., and Esteller, M. (2004a). Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern. The Journal of biological chemistry *279*, 37175-37184.

Espada, J., Ballestar, E., Fraga, M.F., Villar-Garea, A., Juarranz, A., Stockert, J.C., Robertson, K.D., Fuks, F., and Esteller, M. (2004b). Human DNA Methyltransferase 1 Is Required for Maintenance of the Histone H3 Modification Pattern*. Journal of Biological Chemistry *279*, 37175-37184.

Fabian, M.R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem *79*, 351-379.

Falk, S., Finogenova, K., Melko, M., Benda, C., Lykke-Andersen, S., Jensen, T.H., and Conti, E. (2016). Structure of the RBM7–ZCCHC8 core of the NEXT complex reveals connections to splicing factors. Nature communications *7*, 13573-13573.

Fischer, S.E.J., Pan, Q., Breen, P.C., Qi, Y., Shi, Z., Zhang, C., and Ruvkun, G. (2013). Multiple small RNA pathways regulate the silencing of repeated and foreign genes in C. elegans. Genes & Development *27*, 2678-2695.

Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J.*, et al.* (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res *47*, D766-d773.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., and Nakao, M. (2003). Methyl-CpG Binding Domain 1 (MBD1) Interacts with the Suv39h1-HP1 Heterochromatic Complex for DNA Methylation-based Transcriptional Repression*. Journal of Biological Chemistry *278*, 24132-24138.

Fukuda, K., Okuda, A., Yusa, K., and Shinkai, Y. (2018). A CRISPR knockout screen identifies SETDB1-target retroelement silencing factors in embryonic stem cells. Genome Res *28*, 846-858.

Gallego-Paez, L.M., Bordone, M.C., Leote, A.C., Saraiva-Agostinho, N., Ascensão-Ferreira, M., and Barbosa-Morais, N.L. (2017). Alternative splicing: the pledge, the turn, and the prestige : The key role of alternative splicing in human biological systems. Human genetics *136*, 1015-1042.

Garland, W., Comet, I., Wu, M., Radzisheuskaya, A., Rib, L., Vitting-Seerup, K., Lloret-Llinares, M., Sandelin, A., Helin, K., and Jensen, T.H. (2019). A Functional Link between Nuclear RNA Decay and Transcriptional Control Mediated by the Polycomb Repressive Complex 2. Cell Rep *29*, 1800-1811.e1806.

Gautam, P., Yu, T., and Loh, Y.-H. (2017). Regulation of ERVs in pluripotent stem cells and reprogramming. Current Opinion in Genetics & Development *46*, 194-201.

Geis, F.K., and Goff, S.P. (2020). Silencing and Transcriptional Regulation of Endogenous Retroviruses: An Overview. Viruses *12*.

Gemmell, P., Hein, J., and Katzourakis, A. (2019). The Exaptation of HERV-H: Evolutionary Analyses Reveal the Genomic Features of Highly Transcribed Elements. Frontiers in Immunology *10*.

Ghosh, M., Shen, Z., Schaefer, T.M., Fahey, J.V., Gupta, P., and Wira, C.R. (2009). CCL20/MIP3alpha is a novel anti-HIV-1 molecule of the human female reproductive tract. American journal of reproductive immunology (New York, NY : 1989) *62*, 60-71.

Glaunsinger, B.A., and Lee, Y.J. (2010). How tails define the ending: divergent roles for polyadenylation in RNA stability and gene expression. RNA biology *7*, 13-17.

Grabski, D.F., Hu, Y., Sharma, M., and Rasmussen, S.K. (2019). Close to the Bedside: A Systematic Review of Endogenous Retroviruses and Their Impact in Oncology. The Journal of surgical research *240*, 145-155.

Grandi, N., and Tramontano, E. (2018). Human Endogenous Retroviruses Are Ancient Acquired Elements Still Shaping Innate Immune Responses. Front Immunol *9*, 2039.

Groh, S., Milton, A.V., Marinelli, L.K., Sickinger, C.V., Russo, A., Bollig, H., de Almeida, G.P., Schmidt, A., Forné, I., Imhof, A.*, et al.* (2021). Morc3 silences endogenous retroviruses by enabling Daxx-mediated histone H3.3 incorporation. Nature Communications *12*, 5996.

Guffanti, G., Bartlett, A., DeCrescenzo, P., Macciardi, F., and Hunter, R. (2019). Transposable Elements. In Behavioral Neurogenomics, E.B. Binder, and T. Klengel, eds. (Cham: Springer International Publishing), pp. 221-246.

Guo, J., Wang, X., Jia, J., and Jia, R. (2020). Underexpression of SRSF3 and its target gene RBMX predicts good prognosis in patients with head and neck cancer. Journal of oral science *62*, 175-179.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millán-Zambrano, G., Barkai, O., Chávez, S., Pérez-Ortín, J.E., Darzacq, X., and Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000-1011.

Hargous, Y., Hautbergue, G.M., Tintaru, A.M., Skrisovska, L., Golovanov, A.P., Stevenin, J., Lian, L.Y., Wilson, S.A., and Allain, F.H. (2006). Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. The EMBO journal *25*, 5126-5137.

Haselbach, D., Komarov, I., Agafonov, D.E., Hartmuth, K., Graf, B., Dybkov, O., Urlaub, H., Kastner, B., Lührmann, R., and Stark, H. (2018). Structure and Conformational Dynamics of the Human Spliceosomal B(act) Complex. Cell *172*, 454-464.e411.

He, X., Ee, P.L., Coon, J.S., and Beck, W.T. (2004). Alternative splicing of the multidrug resistance protein 1/ATP binding cassette transporter subfamily gene in ovarian cancer creates functional splice variants and is associated with increased expression of the splicing factors PTB and SRp20. Clinical cancer research : an official journal of the American Association for Cancer Research *10*, 4652-4660.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell *38*, 576-589.

Helleboid, P.-Y., Heusel, M., Duc, J., Piot, C., Thorball, C.W., Coluccio, A., Pontis, J., Imbeault, M., Turelli, P., Aebersold, R.*, et al.* (2019). The interactome of KRAB zinc finger proteins reveals the evolutionary history of their functional diversification. The EMBO journal *38*, e101220.

Hilleren, P., McCarthy, T., Rosbash, M., Parker, R., and Jensen, T.H. (2001). Quality control of mRNA 3'-end processing is linked to the nuclear exosome. Nature *413*, 538-542.

Hohn, O., Hanke, K., and Bannert, N. (2013). HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease. Front Oncol *3*, 246.

Ishak, C.A., Classon, M., and De Carvalho, D.D. (2018). Deregulation of Retroelements as an Emerging Therapeutic Opportunity in Cancer. Trends in cancer *4*, 583-597.

Ivanov, D., Stone, J.R., Maki, J.L., Collins, T., and Wagner, G. (2005). Mammalian SCAN domain dimer is a domain-swapped homolog of the HIV capsid C-terminal domain. Molecular cell *17*, 137-143.

Jern, P., and Coffin, J.M. (2008). Effects of retroviruses on host genome function. Annual review of genetics *42*, 709-732.

Jia, R., Li, C., McCoy, J.P., Deng, C.X., and Zheng, Z.M. (2010). SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance. International journal of biological sciences *6*, 806-826.

Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. Nature *467*, 608-611.

Jimenez-Vacas, J.M., Herrero-Aguayo, V., Montero-Hidalgo, A.J., Gomez-Gomez, E., Fuentes-Fayos, A.C., Leon-Gonzalez, A.J., Saez-Martinez, P., Alors-Perez, E., Pedraza-Arevalo, S., Gonzalez-Serrano, T.*, et al.* (2020). Dysregulation of the splicing machinery is directly associated to aggressiveness of prostate cancer. EBioMedicine *51*, 102547.

Jin, B., Li, Y., and Robertson, K.D. (2011). DNA methylation: superior or subordinate in the epigenetic hierarchy? Genes & cancer *2*, 607-617.

Johnson, W.E. (2019). Origins and evolutionary consequences of ancient endogenous retroviruses. Nature reviews Microbiology *17*, 355-370.

Jurica, M.S., and Moore, M.J. (2003). Pre-mRNA splicing: awash in a sea of proteins. Molecular cell *12*, 5-14.

Kallgren, S.P., Andrews, S., Tadeo, X., Hou, H., Moresco, J.J., Tu, P.G., Yates, J.R., 3rd, Nagy, P.L., and Jia, S. (2014). The proper splicing of RNAi factors is critical for pericentric heterochromatin assembly in fission yeast. PLoS Genet *10*, e1004334.

Karimi, M.M., Goyal, P., Maksakova, I.A., Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M.*, et al.* (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell *8*, 676-687.

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

Kato, M., Takemoto, K., and Shinkai, Y. (2018). A somatic role for the histone methyltransferase Setdb1 in endogenous retrovirus silencing. Nature communications *9*, 1683-1683.

Kim, J., Park, R.Y., Chen, J.K., Kim, J., Jeong, S., and Ohn, T. (2014). Splicing factor SRSF3 represses the translation of programmed cell death 4 mRNA by associating with the 5'-UTR region. Cell Death Differ *21*, 481-490.

Kim, K., Nguyen, T.D., Li, S., and Nguyen, T.A. (2018). SRSF3 recruits DROSHA to the basal junction of primary microRNAs. RNA (New York, NY) *24*, 892-898.

Kim, T.K., Hemberg, M., and Gray, J.M. (2015). Enhancer RNAs: a class of long noncoding RNAs synthesized at enhancers. Cold Spring Harbor perspectives in biology *7*, a018622.

Komashko, V.M., and Farnham, P.J. (2010). 5-azacytidine treatment reorganizes genomic histone modification patterns. Epigenetics *5*, 229-240.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.

Krishnamurthy, J., Rabinovich, B.A., Mi, T., Switzer, K.C., Olivares, S., Maiti, S.N., Plummer, J.B., Singh, H., Kumaresan, P.R., Huls, H.M.*, et al.* (2015). Genetic Engineering of T Cells to Target HERV-K, an Ancient Retrovirus on Melanoma. Clinical cancer research : an official journal of the American Association for Cancer Research *21*, 3241-3251.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571-1572.

Krug, B., De Jay, N., Harutyunyan, A.S., Deshmukh, S., Marchione, D.M., Guilhamon, P., Bertrand, K.C., Mikael, L.G., McConechy, M.K., Chen, C.C.L.*, et al.* (2019). Pervasive H3K27 Acetylation Leads to ERV Expression and a Therapeutic Vulnerability in H3K27M Gliomas. Cancer cell *35*, 782-797.e788.

Kupfer, D.M., Drabenstot, S.D., Buchanan, K.L., Lai, H., Zhu, H., Dyer, D.W., Roe, B.A., and Murphy, J.W. (2004). Introns and splicing elements of five diverse fungi. Eukaryotic cell *3*, 1088-1100.

Kuranaga, Y., Sugito, N., Shinohara, H., Tsujino, T., Taniguchi, K., Komura, K., Ito, Y., Soga, T., and Akao, Y. (2018). SRSF3, a Splicer of the PKM Gene, Regulates Cell Growth and Maintenance of Cancer-Specific Energy Metabolism in Colon Cancer Cells. Int J Mol Sci *19*.

Kurosaki, T., Popp, M.W., and Maquat, L.E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. Nature reviews Molecular cell biology *20*, 406-420.

Kury, P., Nath, A., Creange, A., Dolei, A., Marche, P., Gold, J., Giovannoni, G., Hartung, H.P., and Perron, H. (2018). Human Endogenous Retroviruses in Neurological Diseases. Trends Mol Med *24*, 379-394.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Laska, M.J., Nissen, K.K., and Nexø, B.A. (2013). (Some) Cellular Mechanisms Influencing the Transcription of Human Endogenous Retrovirus, HERV-Fc1. PLOS ONE *8*, e53895.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for Computing and Annotating Genomic Ranges. PLOS Computational Biology *9*, e1003118.

Lee, A.K., Pan, D., Bao, X., Hu, M., Li, F., and Li, C.-Y. (2020a). Endogenous Retrovirus Activation as a Key Mechanism of Anti-Tumor Immune Response in Radiotherapy. Radiation research *193*, 305-317.

Lee, A.K., Pan, D., Bao, X., Hu, M., Li, F., and Li, C.Y. (2020b). Endogenous Retrovirus Activation as a Key Mechanism of Anti-Tumor Immune Response in Radiotherapy. Radiation research *193*, 305-317.

Lee, E., Wang, J., Jung, Y., Cackowski, F.C., and Taichman, R.S. (2018). Reduction of two histone marks, H3k9me3 and H3k27me3 by epidrug induces neuroendocrine differentiation in prostate cancer. Journal of cellular biochemistry *119*, 3697-3705.

Lee, E., Wang, J., Yumoto, K., Jung, Y., Cackowski, F.C., Decker, A.M., Li, Y., Franceschi, R.T., Pienta, K.J., and Taichman, R.S. (2016). DNMT1 Regulates Epithelial-Mesenchymal Transition and Cancer Stem Cells, Which Promotes Prostate Cancer Metastasis. Neoplasia *18*, 553-566.

Lee, S.Y., Hung, S., Esnault, C., Pathak, R., Johnson, K.R., Bankole, O., Yamashita, A., Zhang, H., and Levin, H.L. (2020c). Dense Transposon Integration Reveals Essential Cleavage and Polyadenylation Factors Promote Heterochromatin Formation. Cell Rep *30*, 2686-2698.e2688.

Lekas, P., Tin, K.L., Lee, C., and Prokipcak, R.D. (2000). The human cytochrome P450 1A1 mRNA is rapidly degraded in HepG2 cells. Archives of biochemistry and biophysics *384*, 311-318.

Lemaire, R., Prasad, J., Kashima, T., Gustafson, J., Manley, J.L., and Lafyatis, R. (2002). Stability of a PKCI-1-related mRNA is controlled by the splicing factor ASF/SF2: a novel function for SR proteins. Genes & development *16*, 594-607.

Lemieux, C., Marguerat, S., Lafontaine, J., Barbezier, N., Bähler, J., and Bachand, F. (2011). A Pre-mRNA degradation pathway that selectively targets intron-containing genes requires the nuclear poly(A)-binding protein. Molecular cell *44*, 108-119.

Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *100*, 189-192.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics *12*, 323.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, J., Kannan, M., Trivett, A.L., Liao, H., Wu, X., Akagi, K., and Symer, D.E. (2014a). An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. Nucleic acids research *42*, 4546-4562.

Li, J., Kannan, M., Trivett, A.L., Liao, H., Wu, X., Akagi, K., and Symer, D.E. (2014b). An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. Nucleic Acids Res *42*, 4546-4562.

Liu, S., Wu, F., Wu, Z., Li, Y., Zhang, S., and Yu, N. (2019). IL-17A synergistically enhances TLR3-mediated IL-36γ production by keratinocytes: A potential role in injury-amplified psoriatic inflammation. Experimental Dermatology *28*, 233-239.

Liu, X., Gao, Q., Li, P., Zhao, Q., Zhang, J., Li, J., Koseki, H., and Wong, J. (2013). UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. Nature Communications *4*, 1563.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. bioRxiv, 002832.

Lu, G.Y., Huang, S.M., Liu, S.T., Liu, P.Y., Chou, W.Y., and Lin, W.S. (2014). Caffeine induces tumor cytotoxicity via the regulation of alternative splicing in subsets of cancer-associated genes. The international journal of biochemistry & cell biology *47*, 83-92.

Lynch, Vincent J., Nnamani, Mauris C., Kapusta, A., Brayer, K., Plaza, Silvia L., Mazur, Erik C., Emera, D., Sheikh, Shehzad Z., Grützner, F., Bauersachs, S.*, et al.* (2015). Ancient Transposable Elements Transformed the Uterine Regulatory Landscape and Transcriptome during the Evolution of Mammalian Pregnancy. Cell Reports *10*, 551-561.

Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J.-i., Wolf, M., and Kurumizaka, H. (2018). Structural Basis of Heterochromatin Formation by Human HP1. Molecular cell *69*, 385-397.e388.

MacLean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and MacCoss, M.J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics *26*, 966-968.

Maekawa, S., Imamachi, N., Irie, T., Tani, H., Matsumoto, K., Mizutani, R., Imamura, K., Kakeda, M., Yada, T., Sugano, S.*, et al.* (2015). Analysis of RNA decay factor mediated RNA stability contributions on RNA abundance. BMC genomics *16*, 154.

Mallet, P.L., Larochelle, M., and Bachand, F. (2017). Multiple Transcriptional and Post-transcriptional Pathways Collaborate to Control Sense and Antisense RNAs of Tf2 Retroelements in Fission Yeast. Genetics *205*, 621-632.

Mangus, D.A., Evans, M.C., and Jacobson, A. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol *4*, 223.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I.A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature *464*, 927-931.

Mazzone, R., Zwergel, C., Mai, A., and Valente, S. (2017a). Epi-drugs in combination with immunotherapy: a new avenue to improve anticancer efficacy. Clinical Epigenetics *9*, 59.

Mazzone, R., Zwergel, C., Mai, A., and Valente, S. (2017b). Epi-drugs in combination with immunotherapy: a new avenue to improve anticancer efficacy. Clin Epigenetics *9*, 59.

Medon, M., Vidacs, E., Vervoort, S.J., Li, J., Jenkins, M.R., Ramsbottom, K.M., Trapani, J.A., Smyth, M.J., Darcy, P.K., Atadja, P.W.*, et al.* (2017). HDAC Inhibitor Panobinostat Engages Host Innate Immune Defenses to Promote the Tumoricidal Effects of Trastuzumab in HER2(+) Tumors. Cancer Res *77*, 2594-2606.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nature genetics *36*, 1073-1078.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A.*, et al.* (2016). Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. Molecular cell *64*, 520-533.

Merino, G.A., and Fernández, E.A. (2020). Differential splicing analysis based on isoforms expression with NBSplice. Journal of biomedical informatics *103*, 103378.

Michlewski, G., and Cáceres, J.F. (2019). Post-transcriptional control of miRNA biogenesis. RNA (New York, NY) *25*, 1-16.

Milligan, L., Torchet, C., Allmang, C., Shipman, T., and Tollervey, D. (2005). A nuclear surveillance pathway for mRNAs with defective polyadenylation. Molecular and cellular biology *25*, 9996-10004.

Mitsuhashi, S., Nakagawa, S., Sasaki-Honda, M., Sakurai, H., Frith, M.C., and Mitsuhashi, H. (2021). Nanopore direct RNA sequencing detects DUX4-activated repeats and isoforms in human muscle cells. Human Molecular Genetics.

Muller-McNicoll, M., Botti, V., de Jesus Domingues, A.M., Brandl, H., Schwich, O.D., Steiner, M.C., Curk, T., Poser, I., Zarnack, K., and Neugebauer, K.M. (2016). SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. Genes Dev *30*, 553-566.

Murchison, E.P. (2013). In Brenner's Encyclopedia of Genetics (Second Edition), K.H. Stanley Maloy, ed. (Academic Press).

Mure, F., Corbin, A., Benbahouche, N.E.H., Bertrand, E., Manet, E., and Gruffat, H. (2018). The splicing factor SRSF3 is functionally connected to the nuclear RNA exosome for intronless mRNA decay. Scientific reports *8*, 12901-12901.

Navarro Gonzalez, J., Zweig, A.S., Speir, M.L., Schmelter, D., Rosenbloom, K.R., Raney, B.J., Powell, C.C., Nassar, L.R., Maulding, N.D., Lee, C.M.*, et al.* (2021). The UCSC Genome Browser database: 2021 update. Nucleic Acids Res *49*, D1046-d1057.

Nelson, P.N., Carnegie, P.R., Martin, J., Davari Ejtehadi, H., Hooley, P., Roden, D., Rowland-Jones, S., Warren, P., Astley, J., and Murray, P.G. (2003). Demystified. Human endogenous retroviruses. Molecular pathology : MP *56*, 11-18.

Nicetto, D., and Zaret, K.S. (2019). Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. Curr Opin Genet Dev *55*, 1-10.

Niemelä, E.H., Oghabian, A., Staals, R.H., Greco, D., Pruijn, G.J., and Frilander, M.J. (2014). Global analysis of the nuclear processing of transcripts with unspliced U12-type introns by the exosome. Nucleic Acids Res *42*, 7358-7369.

O'Brien, J., Hayder, H., Zayed, Y., and Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. Frontiers in endocrinology *9*, 402.

Ohtani, H., Liu, M., Zhou, W., Liang, G., and Jones, P.A. (2018). Switching roles for DNA and histone methylation depend on evolutionary ages of human endogenous retroviruses. Genome Res *28*, 1147-1157.

Payer, L.M., and Burns, K.H. (2019). Transposable elements in human genetic disease. Nature reviews Genetics *20*, 760-772.

Peaston, A.E., Evsikov, A.V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D., and Knowles, B.B. (2004). Retrotransposons Regulate Host Genes in Mouse Oocytes and Preimplantation Embryos. Developmental Cell *7*, 597-606.

Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation, A., Chao, J., Elliott, O., Liu, Z.P.*, et al.* (2015). RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. Cell *161*, 774-789.

Pelechano, V., and Pérez-Ortín, J.E. (2008). The transcriptional inhibitor thiolutin blocks mRNA degradation in yeast. Yeast *25*, 85-92.

Peltz, S.W., Brown, A.H., and Jacobson, A. (1993). mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes Dev *7*, 1737-1754.

Penke, T.J., McKay, D.J., Strahl, B.D., Matera, A.G., and Duronio, R.J. (2016). Direct interrogation of the role of H3K9 in metazoan heterochromatin function. Genes Dev *30*, 1866-1880.

Piir, K., Paier, A., Liiv, A., Tenson, T., and Maiväli, U. (2011). Ribosome degradation in growing bacteria. EMBO Rep *12*, 458-462.

Porrua, O., and Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. Nature reviews Molecular cell biology *16*, 190-202.

Rädle, B., Rutkowski, A.J., Ruzsics, Z., Friedel, C.C., Koszinowski, U.H., and Dölken, L. (2013). Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. Journal of visualized experiments : JoVE, 50195.

Randall, R.E., and Goodbourn, S. (2008). Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. The Journal of general virology *89*, 1-47.

Rao, Suhas S.P., Huntley, Miriam H., Durand, Neva C., Stamenova, Elena K., Bochkov, Ivan D., Robinson, James T., Sanborn, Adrian L., Machol, I., Omer, Arina D., Lander, Eric S.*, et al.* (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665-1680.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nature protocols *2*, 1896-1906.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D.*, et al.* (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Reyes-Turcu, F.E., Zhang, K., Zofall, M., Chen, E., and Grewal, S.I. (2011). Defects in RNA quality control factors reveal RNAi-independent nucleation of heterochromatin. Nat Struct Mol Biol *18*, 1132-1138.

Richard, P., and Manley, J.L. (2013). How bidirectional becomes unidirectional. Nature Structural & Molecular Biology *20*, 1022-1024.

Robert, V.J., Sijen, T., van Wolfswinkel, J., and Plasterk, R.H. (2005). Chromatin and RNAi factors protect the C. elegans germline against repetitive sequences. Genes Dev *19*, 782-787.

Rose, N.R., and Klose, R.J. (2014). Understanding the relationship between DNA methylation and histone lysine methylation. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1839*, 1362-1372.

Roulois, D., Loo Yau, H., Singhania, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J.*, et al.* (2015). DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell *162*, 961-973.

Ruprecht, K., and Mayer, J. (2019). On the origin of a pathogenic HERV-W envelope protein present in multiple sclerosis lesions. Proceedings of the National Academy of Sciences *116*, 19791-19792.

Russo, J., Heck, A.M., Wilusz, J., and Wilusz, C.J. (2017). Metabolic labeling and recovery of nascent RNA to accurately quantify mRNA stability. Methods (San Diego, Calif) *120*, 39-48.

Sadeq, S., Al-Hashimi, S., Cusack, C.M., and Werner, A. (2021). Endogenous Double-Stranded RNA. Non-coding RNA *7*.

Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., Theis, F.J., and Schotta, G. (2015). Atrx promotes heterochromatin formation at retrotransposons. EMBO Rep *16*, 836-850.

Saitoh, Y., Bureta, C., Sasaki, H., Nagano, S., Maeda, S., Furukawa, T., Taniguchi, N., and Setoguchi, T. (2019). The histone deacetylase inhibitor LBH589 inhibits undifferentiated pleomorphic sarcoma growth via downregulation of FOS-like antigen 1. Molecular Carcinogenesis *58*, 234-246.

Saksouk, N., Simboeck, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin *8*, 3.

Schlackow, M., Nojima, T., Gomes, T., Dhir, A., Carmo-Fonseca, M., and Proudfoot, N.J. (2017). Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. Molecular cell *65*, 25-38.

Schmid, M., and Jensen, T.H. (2018). Controlling nuclear RNA levels. Nature Reviews Genetics *19*, 518-529.

Schneider, C., Kudla, G., Wlotzka, W., Tuck, A., and Tollervey, D. (2012). Transcriptome-wide analysis of exosome targets. Molecular cell *48*, 422-433.

Schuch, B., Feigenbutz, M., Makino, D.L., Falk, S., Basquin, C., Mitchell, P., and Conti, E. (2014). The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. The EMBO journal *33*, 2829-2846.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev *16*, 919-932.

Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., Gagneur, J., and Cramer, P. (2016). TT-seq maps the human transient transcriptome. Science *352*, 1225-1228.

Scuto, A., Kirschbaum, M., Kowolik, C., Kretzner, L., Juhasz, A., Atadja, P., Pullarkat, V., Bhatia, R., Forman, S., Yen, Y.*, et al.* (2008). The novel histone deacetylase inhibitor, LBH589, induces expression of DNA damage response genes and apoptosis in Ph- acute lymphoblastic leukemia cells. Blood *111*, 5093-5100.

Sen, S., Jumaa, H., and Webster, N.J. (2013). Splicing factor SRSF3 is crucial for hepatocyte differentiation and metabolic function. Nat Commun *4*, 1336.

Sexton, C.E., Han, M.V. (2019). Paired-end mappability of transposable elements in the human genome. Mobile DNA *10*.

Sgarbanti, M., Remoli, A.L., Marsili, G., Ridolfi, B., Borsetti, A., Perrotti, E., Orsatti, R., Ilari, R., Sernicola, L., Stellacci, E.*, et al.* (2008). IRF-1 is required for full NF-kappaB transcriptional activity at the human immunodeficiency virus type 1 long terminal repeat enhancer. J Virol *82*, 3632-3641.

Sharif, J., Endo, T.A., Nakayama, M., Karimi, M.M., Shimada, M., Katsuyama, K., Goyal, P., Brind'Amour, J., Sun, M.A., Sun, Z.*, et al.* (2016). Activation of Endogenous Retroviruses in Dnmt1(-/-) ESCs Involves Disruption of SETDB1-Mediated Repression by NP95 Binding to Hemimethylated DNA. Cell Stem Cell *19*, 81-94.

Shen, T., Li, H., Song, Y., Li, L., Lin, J., Wei, G., and Ni, T. (2019a). Alternative polyadenylation dependent function of splicing factor SRSF3 contributes to cellular senescence. Aging (Albany NY) *11*, 1356-1388.

Shen, T., Li, H., Song, Y., Li, L., Lin, J., Wei, G., and Ni, T. (2019b). Alternative polyadenylation dependent function of splicing factor SRSF3 contributes to cellular senescence. Aging (Albany NY) *11*, 1356-1388.

Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., 3rd, Frank, J., and Manley, J.L. (2009). Molecular architecture of the human pre-mRNA 3' processing complex. Molecular cell *33*, 365-376.

Shi, Y., and Manley, J.L. (2015). The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. Genes Dev *29*, 889-897.

Siebring-van Olst, E., Blijlevens, M., de Menezes, R.X., van der Meulen-Muileman, I.H., Smit, E.F., and van Beusechem, V.W. (2017). A genome-wide siRNA screen for regulators of tumor suppressor p53 activity in human non-small cell lung cancer cells identifies components of the RNA splicing machinery as targets for anticancer treatment. Molecular oncology *11*, 534-551.

Silla, T., Schmid, M., Dou, Y., Garland, W., Milek, M., Imami, K., Johnsen, D., Polak, P., Andersen, J.S., Selbach, M.*, et al.* (2020). The human ZC3H3 and RBM26/27 proteins are critical for PAXT-mediated nuclear RNA decay. Nucleic Acids Research *48*, 2518-2530.

Singh, A.K., Choudhury, S.R., De, S., Zhang, J., Kissane, S., Dwivedi, V., Ramanathan, P., Petric, M., Orsini, L., Hebenstreit, D.*, et al.* (2019). The RNA helicase UPF1 associates with mRNAs co-transcriptionally and is required for the release of mRNAs from gene loci. eLife *8*, e41444.

Slobodin, B., Bahat, A., Sehrawat, U., Becker-Herman, S., Zuckerman, B., Weiss, A.N., Han, R., Elkon, R., Agami, R., Ulitsky, I.*, et al.* (2020). Transcription Dynamics Regulate Poly(A) Tails and Expression of the RNA Degradation Machinery to Balance mRNA Levels. Molecular cell *78*, 434-444.e435.

Snyder, R.C., and Miller, D.M. (1992). Regulation of c-myc transcription initiation and elongation. Critical reviews in oncogenesis *3*, 283-291.

Song, X., Wan, X., Huang, T., Zeng, C., Sastry, N., Wu, B., James, C.D., Horbinski, C., Nakano, I., Zhang, W.*, et al.* (2019). SRSF3-Regulated RNA Alternative Splicing Promotes Glioblastoma Tumorigenicity by Affecting Multiple Cellular Processes. Cancer Res *79*, 5288-5301.

Srinivasachar Badarinarayan, S., Shcherbakova, I., Langer, S., Koepke, L., Preising, A., Hotter, D., Kirchhoff, F., Sparrer, K.M.J., Schotta, G., and Sauter, D. (2020). HIV-1 infection

activates endogenous retroviral promoters regulating antiviral gene expression. Nucleic Acids Res *48*, 10890-10908.

Sugiyama, T., Thillainadesan, G., Chalamcharla, V.R., Meng, Z., Balachandran, V., Dhakshnamoorthy, J., Zhou, M., and Grewal, S.I.S. (2016). Enhancer of Rudimentary Cooperates with Conserved RNA-Processing Factors to Promote Meiotic mRNA Decay and Facultative Heterochromatin Assembly. Molecular cell *61*, 747-759.

Sundaram, V., Cheng, Y., Ma, Z., Li, D., Xing, X., Edge, P., Snyder, M.P., and Wang, T. (2014). Widespread contribution of transposable elements to the innovation of gene regulatory networks. Genome Research *24*, 1963-1976.

Svoboda, P. (2014). Renaissance of mammalian endogenous RNAi. FEBS letters *588*, 2550-2556.

Tanackovic, G., and Krämer, A. (2005). Human splicing factor SF3a, but not SF1, is essential for pre-mRNA splicing in vivo. Molecular biology of the cell *16*, 1366-1377.

Tang, Y., Horikawa, I., Ajiro, M., Robles, A.I., Fujita, K., Mondal, A.M., Stauffer, J.K., Zheng, Z.M., and Harris, C.C. (2013). Downregulation of splicing factor SRSF3 induces p53β, an alternatively spliced isoform of p53 that promotes cellular senescence. Oncogene *32*, 2792-2798.

Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. Current protocols in bioinformatics *Chapter 4*, Unit 4.10.

Team, R.C. (2018). R: A language and environment for statistical computing, R.F.f.S. Computing, ed. (Vienna, Austria.).

Teissandier, S., Barillot, Bourc'his (2019). Tools and best practices for retrotransposon analysis using high-throughput sequencing data. Mobile DNA *10*.

Thomas, J.H., and Schneider, S. (2011). Coevolution of retroelements and tandem zinc finger genes. Genome Research *21*, 1800-1812.

Tian, B., Pan, Z., and Lee, J.Y. (2007). Widespread mRNA polyadenylation events in introns indicate dynamic interplay between polyadenylation and splicing. Genome Res *17*, 156-165.

Timms, R.T., Tchasovnikarova, I.A., Antrobus, R., Dougan, G., and Lehner, P.J. (2016). ATF7IP-Mediated Stabilization of the Histone Methyltransferase SETDB1 Is Essential for Heterochromatin Formation by the HUSH Complex. Cell Rep *17*, 653-659.

Tsusaka, T., Shimura, C., and Shinkai, Y. (2019). ATF7IP regulates SETDB1 nuclear localization and increases its ubiquitination. EMBO reports *20*, e48297.

Tudek, A., Schmid, M., Makaras, M., Barrass, J.D., Beggs, J.D., and Jensen, T.H. (2018). A Nuclear Export Block Triggers the Decay of Newly Synthesized Polyadenylated RNA. Cell Rep *24*, 2457-2467.e2457.

Ugolini, I., and Halic, M. (2018). Fidelity in RNA-based recognition of transposable elements. Philosophical transactions of the Royal Society of London Series B, Biological sciences *373*.

Uhlitz, F., Sieber, A., Wyler, E., Fritsche-Guenther, R., Meisig, J., Landthaler, M., Klinger, B., and Blüthgen, N. (2017). An immediate–late gene expression module decodes ERK signal duration. Molecular Systems Biology *13*, 928.

Vakoc, C.R., Mandat, S.A., Olenchock, B.A., and Blobel, G.A. (2005). Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Molecular cell *19*, 381-391.

Valle-García, D., Qadeer, Z.A., McHugh, D.S., Ghiraldini, F.G., Chowdhury, A.H., Hasson, D., Dyer, M.A., Recillas-Targa, F., and Bernstein, E. (2016). ATRX binds to atypical chromatin domains at the 3' exons of zinc finger genes to preserve H3K9me3 enrichment. Epigenetics *11*, 398-414.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology *3*, research0034.0031.

Vargiu, L., Rodriguez-Tomé, P., Sperber, G.O., Cadeddu, M., Grandi, N., Blikstad, V., Tramontano, E., and Blomberg, J. (2016). Classification and characterization of human endogenous retroviruses; mosaic forms are common. Retrovirology *13*, 7.

Vasiljeva, L., Kim, M., Terzi, N., Soares, L.M., and Buratowski, S. (2008). Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. Molecular cell *29*, 313-323.

Vergara Bermejo, A., Ragonnaud, E., Daradoumis, J., and Holst, P. (2020). Cancer Associated Endogenous Retroviruses: Ideal Immune Targets for Adenovirus-Based Immunotherapy. Int J Mol Sci *21*.

Villar-Garea, A., Israel, L., and Imhof, A. (2008). Analysis of histone modifications by mass spectrometry. Current protocols in protein science *Chapter 14*, Unit 14.10.

Vogel, M.J., Guelen, L., de Wit, E., Peric-Hupkes, D., Lodén, M., Talhout, W., Feenstra, M., Abbas, B., Classen, A.K., and van Steensel, B. (2006). Human heterochromatin proteins form large domains containing KRAB-ZNF genes. Genome Res *16*, 1493-1504.

Volkman, H.E., and Stetson, D.B. (2014). The enemy within: endogenous retroelements and autoimmune disease. Nat Immunol *15*, 415-422.

Wada, T., and Becskei, A. (2017). Impact of Methods on the Measurement of mRNA Turnover. Int J Mol Sci *18*.

Wang, G.Z., Wang, Y., and Goff, S.P. (2016). Histones Are Rapidly Loaded onto Unintegrated Retroviral DNAs Soon after Nuclear Entry. Cell host & microbe *20*, 798-809.

Wang, J., Vasaikar, S., Shi, Z., Greer, M., and Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. Nucleic Acids Research *45*, W130-W137.

Wang, T., Zeng, J., Lowe, C.B., Sellers, R.G., Salama, S.R., Yang, M., Burgess, S.M., Brachmann, R.K., and Haussler, D. (2007). Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. Proceedings of the National Academy of Sciences *104*, 18613-18618.

Warf, M.B., Shepherd, B.A., Johnson, W.E., and Bass, B.L. (2012). Effects of ADARs on small RNA processing pathways in C. elegans. Genome Res *22*, 1488-1498.

Warnecke, T., Parmley, J.L., and Hurst, L.D. (2008). Finding exonic islands in a sea of non-coding sequence: splicing related constraints on protein composition and evolution are common in intron-rich genomes. Genome Biol *9*, R29.

Weigt, M., Gao, Q., Ban, H., He, H., Mastrobuoni, G., Kempa, S., Chen, W., and Li, F. (2021). Rbm10 facilitates heterochromatin assembly via the Clr6 HDAC complex. Epigenetics & Chromatin *14*, 8.

Westholm, J.O., and Lai, E.C. (2011). Mirtrons: microRNA biogenesis via splicing. Biochimie *93*, 1897-1904.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).

Will, C.L., and Lührmann, R. (2011). Spliceosome structure and function. Cold Spring Harbor perspectives in biology *3*.

Wu, Y., Liu, W., Chen, J., Liu, S., Wang, M., Yang, L., Chen, C., Qi, M., Xu, Y., Qiao, Z.*, et al.* (2019). Nuclear Exosome Targeting Complex Core Factor Zcchc8 Regulates the Degradation of LINE1 RNA in Early Embryos and Embryonic Stem Cells. Cell Rep *29*, 2461-2472.e2466.

Yang, B.X., El Farran, C.A., Guo, H.C., Yu, T., Fang, H.T., Wang, H.F., Schlesinger, S., Seah, Y.F., Goh, G.Y., Neo, S.P.*, et al.* (2015). Systematic identification of factors for provirus silencing in embryonic stem cells. Cell *163*, 230-245.

Yang, X.J., and Seto, E. (2007). HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. Oncogene *26*, 5310-5318.

Zhang, F., Wang, J., Xu, J., Zhang, Z., Koppetsch, B.S., Schultz, N., Vreven, T., Meignin, C., Davis, I., Zamore, P.D.*, et al.* (2012). UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell *151*, 871-884.

Zhang, X., Yan, C., Zhan, X., Li, L., Lei, J., and Shi, Y. (2018). Structure of the human activated spliceosome in three conformational states. Cell research *28*, 307-322.

Zhao, Q., Zhang, J., Chen, R., Wang, L., Li, B., Cheng, H., Duan, X., Zhu, H., Wei, W., Li, J.*, et al.* (2016). Dissecting the precise role of H3K9 methylation in crosstalk with DNA maintenance methylation in mammals. Nature Communications *7*, 12464.

Zhou, F., Krishnamurthy, J., Wei, Y., Li, M., Hunt, K., Johanning, G.L., Cooper, L.J., and Wang-Johanning, F. (2015). Chimeric antigen receptor T cells targeting HERV-K inhibit breast cancer and its metastasis through downregulation of Ras. Oncoimmunology *4*, e1047582.

Zhou, L., Guo, J., and Jia, R. (2019). Oncogene SRSF3 suppresses autophagy via inhibiting BECN1 expression. Biochem Biophys Res Commun *509*, 966-972.

Zhou, Z., Gong, Q., Lin, Z., Wang, Y., Li, M., Wang, L., Ding, H., and Li, P. (2020). Emerging Roles of SRSF3 as a Therapeutic Target for Cancer. Frontiers in Oncology *10*.

Zhu, Y., Wang, G.Z., Cingöz, O., and Goff, S.P. (2018). NP220 mediates silencing of unintegrated retroviral DNA. Nature *564*, 278-282.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. *31*, 88-100.