AGO1 regulates major satellite transcripts and H3K9me3 distribution at pericentromeric regions in mESCs

Madlen Müller[1,2], Tara Fäh[1], Moritz Schäfer[1,2], Janina Luitz[1], Patrick Stalder[1,2], Rajika Arora[1], Richard Patryk Ngondo[1, £] and Constance Ciaudo[1, *].

[1] Swiss Federal Institute of Technology Zurich, IMHS, Chair of RNAi and Genome Integrity, Zurich, Switzerland.

[2] Life Science Zurich Graduate School, University of Zürich, Switzerland.

[£] Current address: Institut de Biologie Moléculaire des Plantes UPR-CNRS 2357, Strasbourg, France.

* Correspondence: cciaudo@ethz.ch

Running title: AGO1 regulates major satellite transcripts in mESCs.

Key words: AGO1, pericentromeres, Major satellites, H3K9me3, miR-30.

SUMMARY

In the past years, several studies reported nuclear roles for the Argonaute (AGO) proteins associating them with transcriptional activation or repression, alternative splicing and, chromatin organization. However, as most of these experiments have been conducted in human cancer cell lines, the nuclear functions of the AGO proteins in mouse early embryonic development still remains elusive. In this study, we investigated possible nuclear functions of the AGO proteins in mouse Embryonic Stem Cells (mESCs). By biochemical assays, we observed that AGO1 and AGO2 are present in a small fraction in the nucleus and even less on chromatin in mESCs. To profile the nuclear interactome of the AGO proteins, we performed immunoprecipitation followed by Mass Spectrometry and identified three novel nuclear interactors for AGO1, namely DNMT3a, HP1α, and ATRX. These interactors are well-known proteins involved in the establishment and maintenance of heterochromatin at pericentromeric regions. Indeed, upon depletion of *Ago1,* we observed a specific redistribution of the heterochromatin protein HP1α and the repressive histone mark H3K9me3, away from pericentromeric regions. Furthermore, these regions are characterized by AT-rich tandem repeats known as major satellite sequences. We demonstrated that major satellite transcripts are strongly upregulated in *Ago1_KO* mESCs. Interestingly, this phenotype was not caused by the loss of genome integrity at pericentromeres, as these could still form normally in *Ago1*_KO mESCs. Lastly, we showed that specific microRNAs loaded in AGO1, regulate the expression of the major satellite transcripts. Overall, our results demonstrate for the first time a novel role for AGO1 in regulating major satellite transcripts and localization of HP1α and H3K9me3 at pericentromeres in mESCs.

INTRODUCTION

The microRNA (miRNA) pathway is crucial in regulating early embryonic development and differentiation *in vivo* and *in vitro* (DeVeale et al., 2021). MiRNAs can fine-tune gene expression throughout early embryonic development at the post-transcriptional level. They are processed into ~22 nt long mature miRNAs by two consecutive cleavage steps conducted by the RNAse III enzyme DROSHA in the nucleus, and DICER, in the cytoplasm (Bodak et al., 2017). Mature miRNAs are loaded into Argonaute (AGO) proteins, which are key components of the RNA-Induced Silencing Complex (RISC). They guide the RISC complex to partially complementary target sequences leading to the translational inhibition of these targets (Bartel, 2018).

In mice, there are four AGO proteins (AGO1-4), but only AGO1 and AGO2 are detectably expressed during early embryonic development, with AGO2 being substantially more abundant (Boroviak et al., 2018; Lykke-Andersen et al., 2008; Müller et al., 2020). While *Ago2* deficient mice die at a post-implantation stage, due to severe developmental defects (Alisch et al., 2007; Cheloufi et al., 2010; Liu et al., 2004; Morita et al., 2007), *Ago1,3,4* deficient mice are viable (Modzelewski et al., 2012; Van Stry et al., 2012). Mouse Embryonic Stem Cells (mESCs), which are derived from the inner cell mass of the blastocyst, are a powerful tool to study early embryonic development *in vitro.* These cells are pluripotent and can differentiate into the three embryonic germ layers. As observed *in vivo*, mESCs only express AGO1 and AGO2 proteins (Boroviak et al., 2018; Lykke-Andersen et al., 2008; Müller et al., 2020). MESCs deficient for either AGO1 or AGO2 are viable, can exit from pluripotency and differentiate into the three embryonic germ layers (Ngondo et al., 2018).

In addition to their major role in the cytoplasmic miRNA pathway, several studies have reported noncanonical functions of the Argonaute proteins in the nucleus (Gagnon et al., 2014a; Li et al., 2020; Meister, 2013). AGO2 was shown to shuttle into the nucleus with the help of TNRC6A (Nishi et al., 2013). In the nucleus, guided by small RNAs (smRNAs), both AGO1 and AGO2 have been shown to localize to promoter regions and reinforce the recruitment of chromatin modifiers leading to either transcriptional activation or silencing (Cho et al., 2014; Hu et al., 2012; Janowski et al., 2006; Kim et al., 2006; Li et al., 2006; Portnoy et al., 2016). AGO1 was also found to be enriched at promoters of actively transcribed genes, where it interacts with RNA Polymerase II (RNA PolII) (Huang et al., 2013). Furthermore, AGO1 was found to localize to enhancer regions, which was dependent on a species of RNA called enhancer RNAs (eRNAs) (Alló et al., 2014; Shuaib et al., 2019). Additionally, interaction of AGO1 with enhancers was shown to be crucial for maintenance of 3D chromatin organization (Shuaib et al., 2019). Finally, AGO1 has been implicated in alternative splicing events, taking place within the nucleus (Agirre et al., 2015; Alló et al., 2014; Ameyar-Zazoua

et al., 2012). Of note, most of these chromatin-associated functions have been described mainly for AGO1, AGO2 however was reported to be involved in double strand break repair (Gao et al., 2014; Wang and Goldstein, 2016).

Most of the mentioned nuclear functions for the AGO proteins were described in human cancer cell lines and not studied during early embryonic development. So far only one study reported a function for AGO2 in the nucleus of mESCs (Sarshad et al., 2018). They identified by subcellular fractionation that the majority of AGO2 localizes to the nucleus of mESCs. Furthermore, they showed that nuclear AGO2 interacted with known RISC factors and expanded its post-transcriptional silencing within the nucleus not just to the 3'UTRs, but also to coding and intronic sequences (Sarshad et al., 2018).

In this study, we aimed to better understand the roles of Argonaute proteins during mouse early embryonic development. We characterized the nuclear and cytoplasmic interactomes of the two expressed AGO proteins in mESCs and identified three nuclear AGO1 interactors in our approach, known to be involved in the maintenance of heterochromatin at pericentromeric regions. Depletion of *Ago1* in mESCs led to a redistribution of HP1α and H3K9me3 away from pericentromeres. In addition, we observed an increase in major satellite transcripts upon loss of AGO1. Nevertheless, we did not observe any changes in pericentromeres integrity. These results prompted us to investigate the underlying molecular mechanism by which AGO1 regulates major satellite transcripts. Using computational analyses and molecular approaches, we found that AGO1, loaded with miR-30a, d, e-3p, regulates major satellite transcripts and mediates proper distribution of HP1α and H3K9me3 at pericentromeric regions in mESCs.

RESULTS

Nuclear interactors link AGO1 to the pericentromeres

In order to better understand the multiple functions of the Argonaute proteins during early development, we assessed the subcellular distribution of the two expressed AGO proteins in mESCs (Müller et al., 2020). Using well established biochemical assays (Gagnon et al., 2014b), we performed cytoplasmic/nucleoplasmic/chromatin fractionation of WT mESCs and analyzed the abundance of AGO1 and AGO2 in these three fractions by Western blotting (WB) (Figure 1A). In contrast to a previous report (Sarshad et al., 2018), we observed that the majority of AGO1 and AGO2 localized to the cytoplasm, whereas, only around 10-15% of both AGO proteins are present in the nuclear fraction and even less than 3% is found in the chromatin fraction (Figure 1A). Cross-contamination was controlled for by using specific subcellular markers to validate the purity of the different fractions (Figure 1A).

To characterize the interactome of AGO1 and AGO2 in the nuclear and cytoplasmic compartments, we further engineered endogenously tagged cell lines, where AGO1 and AGO2 are heterozygously BirA*/FLAG tagged N-terminally (B/F-Ago1, B/F-Ago2) (Figure S1A). This approach allowed us to circumvent the poor quality of antibodies for mouse AGO immunoprecipitations and increase the specificity of the subsequent Mass Spectrometry (MS) (Figure 1B). As a control line, we generated mESCs expressing a B/F-GFP-NLS2 construct (Figure S1B). GFP was tagged with a nuclear localization signal (NLS), to ensure its localization to the nucleus and verified by fluorescence microscopy (Figure S1B).

Using these three engineered cell lines, we performed cytoplasmic/nuclear fractionation with subsequent immunoprecipitation (IP) of the FLAG tag in the two fractions, followed by MS (Figure 1B). Of note, we tried a BioID proximity-labeling approach using the BirA* tag to pulldown AGO interactors, however this led to too many interactions for the nuclear fraction (data not shown). As the interactome of AGO2 has been previously intensively studied by several groups (Table S1) (Frohn et al., 2012; Schopp et al., 2017), we first analyzed the cytoplasmic AGO2 pulldown (see Material and Methods). As expected, we identified several known interactors of AGO2 in the cytoplasm, such as TNRC6, PABPC1, FKBP5 and FXR1 (Figures 1C and Table S1) (Chen et al., 2009; Fabian et al., 2009; Martinez et al., 2013; Vasudevan and Steitz, 2007). We also identified the known RNA binding protein PAPBC1 in the cytoplasmic AGO1 IP, however RISC components such as the TNRC6 proteins were not significantly enriched (Figure S1C). Generally, the number of significant proteins identified in the cytoplasmic AGO1 IP was lower than for AGO2 but most of the proteins significant in the AGO1 IP were also pulled down with AGO2. From these results, we concluded that our protocol successfully allows us to pull down the AGO proteins, as well as their cellular interactors.

The analysis of nuclear AGO1 and AGO2 pulldowns revealed fewer significantly enriched proteins for both baits (Figures 1D, S1D), probably due to the lower amount of protein present in this compartment (Figure 1A). In contrast to the cytoplasmic fraction, the interactomes of nuclear AGO1 and AGO2 looked fairly different. Proteins pulled down in the nuclear AGO2 IP, seemed to be mostly associated with protein synthesis or the cytoskeleton (Figure S1D). In contrast, we identified several known chromatin factors, such as CBX5 (hereafter referred to as HP1α), ATRX and DNMT3A, in the nuclear AGO1 IP (Figure 1D). Interestingly, these three proteins HP1α, ATRX and DNMT3a are known factors involved in the establishment of pericentric heterochromatin (Feldman et al., 2006; Guenatri et al., 2004; Hyun et al., 2017; Probst and Almouzni, 2008).

Ago1 depletion affects the distribution of HP1α and H3K9me3 at pericentromeric regions

The nuclear interactor of AGO1, HP1α is enriched at pericentromeric heterochromatin regions, which in mammals, is a kind of constitutive heterochromatin that is found at the centromeres. It is required for proper sister chromatid cohesion and chromosome segregation (Bernard et al., 2001; Guenatri et al., 2004; Houlard et al., 2006; Nonaka et al., 2002; Probst and Almouzni, 2008; Probst et al., 2009). Pericentromeric domains from several chromosomes are known to cluster together within interphase in order to form chromocenters (Guenatri et al., 2004; Probst and Almouzni, 2008). Chromocenters are easily visible by fluorescence microscopy with a brighter DAPI stain (Guenatri et al., 2004; Probst and Almouzni, 2008) (Figure S2A).

In order to study whether AGO1 might be important for the integrity of mouse pericentromeric regions, we used two *Ago1*_KO mESC lines generated using a paired CRISPR-Cas9 approach (Wettstein et al., 2016). The *Ago1*_KO1 mESC line was previously published by (Ngondo et al., 2018). The *Ago1*_KO2 mESCs line was newly generated and validated for the absence of AGO1 expression (Figures S2B, C). We first analyzed the distribution of HP1α in Wild Type (WT) versus *Ago1*_KO mESCs by indirect Immunofluorescence (IF) and observed colocalization of HP1α with DAPI rich regions in WT mESCs (Figure 2A). However, this colocalization of HP1α with DAPI rich regions, was significantly reduced in *Ago1*_KO mESCs (Figure 2A). In addition, we noted a slight increase of HP1α protein expression in *Ago1*_KO mESCs compared to WT cells (Figure S2D). We also assessed the distribution of DNMT3a by IF in WT and *Ago1*_KO mESCs (Figure S2E). Unlike HP1α, no apparent colocalization with DAPI rich regions could be observed for the antibody staining for DNMT3a, but we found that DNMT3a integrated intensity was stronger in *Ago1*_KO than in WT mESCs, which was further confirmed by WB (Figures S2E, F).

HP1α is known to recruit other proteins, such as the methyltransferase SUV39H1/2 to pericentromeric heterochromatin regions, which leads to the spreading of heterochromatin and specifically to the deposition of the repressive histone mark H3K9me3 (Bannister et al., 2001; Hyun et al., 2017; Lachner et al., 2001). Therefore, we subsequently assessed the colocalization of H3K9me3 with DAPI rich regions in WT and *Ago1*_KO mESCs by IF (Figure 2B). Surprisingly, we observed a strong redistribution of H3K9me3 in *Ago1*_KO mESCs, away from the pericentromeric regions (Figure 2B). Nevertheless, we did not observe a global change of H3K9me3 protein levels in *Ago1*_KO (Figure S2G). To strengthen these results, we performed H3K9me3 Chromatin Immunoprecipitation followed by quantitative PCR (ChIP-qPCR) in WT versus *Ago1*_KO mESCs. We validated that H3K9me3 pulldown was successful by using published primer pairs for H3K9me3 control regions and assessed the enrichment at these sites over an intergenic region (Karimi et al., 2011; Ngondo et al., 2018) (Figure S2H). Indeed, when we compared the enrichment of H3K9me3 at major satellites sequences in WT over Input versus *Ago1*_KO over Input, we observed that upon *Ago1* depletion, around 50% of H3K9me3 is lost at these sites (Figure 2C).

In conclusion, we observed a redistribution of the heterochromatin protein HP1α and the repressive histone mark H3K9me3, away from pericentromeric regions in *Ago1*_KO mESCs.

AGO1 complementation rescues the distribution of H3K9me3 at pericentromeric regions

In order to determine, whether the redistribution of H3K9me3 away from pericentromeric regions is specific to the loss of AGO1, we aimed to complement our *Ago1*_KO mESCs. We transfected the *Ago1*_KO2 mESC line with a vector expressing N-terminally HA-tagged AGO1. This vector additionally contains two selection markers, a GFP and a puromycin resistance genes. Cells expressing AGO1 were selected for a week with antibiotic resistance, followed by FACS sorting on GFP (Figure S3A). Finally, we verified that the GFP sorted mixed population expresses HA-AGO1 by IF and WB (Figures 3A, S3B). We next tested whether reintroduction of AGO1 can rescue the distribution of H3K9me3 foci at DAPI rich regions and performed a co-stain for H3K9me3 and HA in WT, *Ago1*_KO2 and the *Ago1*_KO2 + HA-AGO1 mixed population (Figure 3A). We observed a rescue of the H3K9me3 distribution at the pericentromeres upon reintroduction of AGO1 in *Ago1*_KO mESC line (Figure 3A).

Major satellite transcripts are upregulated in *Ago1*_KO mESCs

In Mouse, pericentromeric heterochromatin regions are characterized by AT-rich tandem repeats, known as major satellite repeat sequences. Major satellites consist of 234 bp tandem repeat sequences that can stretch over several kilobases (Guenatri et al., 2004; Komissarov

et al., 2011). The minor satellite sequences adjacent to the major satellites are localized to the centromeric part of the chromosome (Figure 3B) (Guenatri et al., 2004; Vissel and Choo, 1989). Even though pericentromeric regions are marked by repressive heterochromatin marks, pericentromeric transcripts, such as major satellite transcripts, have been previously reported *in vivo* during mouse early development and also *in vitro* in mESCs (Lehnertz et al., 2003; Probst et al., 2010; Rudert et al., 1995).

In order to assess whether the depletion of *Ago1* also has an impact on the major satellite transcripts, we performed an RNA Fluorescent *In Situ* Hybridization (FISH) (Figure 3C). We observed RNA FISH signal for the major satellites in WT mESCs, however the signal in the *Ago1*_KO mESCs was much stronger and we also detected significantly more major satellite RNA foci compared to WT (Figures 3C, S3C). To better quantify the amount of major satellite transcripts in both cell lines, we measured their relative expression using a stringent RT-qPCR protocol in WT and *Ago1*_KO mESCs (see Material and Methods and Figure S3D). We detected a significant upregulation of major satellite mRNAs in the *Ago1*_KO mESCs with two independent primer pairs (Figure 3D).

Finally, to better understand whether the upregulation of major satellite transcripts was linked to changes at the chromatin level, we analyzed the IF images for the H3K9me3 staining and quantified the number of DAPI foci (chromocenters) in WT versus *Ago1*_KO mESCs. We observed no decrease of DAPI foci formation in *Ago1*_KO compared to WT mESCs (Figure 3E). Additionally, we also performed DNA FISH for the major satellite repeats in these two cell lines. The major satellite DNA FISH signal was similar between WT and *Ago1*_KO mESCs and we did not detect a more dispersed signal for the major satellites in *Ago1*_KO mESCs, indicating that the overall structures of chromocenters is preserved in *Ago1*_KO mESCs (Figure 3F).

In summary, we observed that upon *Ago1* depletion, major satellite transcripts are upregulated in mESCs. Furthermore, this phenotype was not caused by the loss of chromocenters structure, as they could still form normally in *Ago1*_KO mESCs.

MiR-30a, d, e-3p regulate major satellite transcripts in mESCs

We aimed to identify the molecular mechanism causing the upregulation of major satellite transcripts in *Ago1*_KO mESCs. The AGO proteins are best known for their role in post-transcriptional silencing rather than transcriptional regulation (Meister, 2013). Therefore, we sought to identify whether miRNAs might target major satellite transcripts in mESCs. Most major satellite annotations, as obtained from RepeatMasker and Dfam (Bao et al., 2015; Smit et al.; Storer et al., 2021), were not properly mapped to any of the chromosomes and therefore annotated in unmapped genomic contigs (Figure S4A). In accordance with reports of major satellite sequences being several kilobase long, we selected those annotations that mapped

to regions of at least 20 kbps of length, which are regions mapped to chromosome X, 9 and the contigs JH583204.1 and GL456383.1. We searched for miRNAs with high confidence seed matches within these sequences and identified that three miRNAs from the miR-30 family have a high number of 8mer binding sites within major satellite sequences (Figure 4A). The miR-30 family is composed of 6 pre-miRNAs, (miR-30a, miR-30b, miR-30c-1, miR-30c-2, miR-30d, miR-30e) located on three different chromosomes, which are all expressed in WT mESCs (Figures S4B, C, Table S3). While all the mature miR-30-5p share the same seed sequence, only the miR-30a-3p, miR-30d-3p and miR-30e-3p have identical seeds, which match the major satellite sequences (Figures 4A, S4D). Additionally, by analyzing published AGO1 RNA immunoprecipitation and sequencing (RIP-seq) data (Ngondo et al., 2018), we identified that these three miRNAs are preferentially loaded into AGO1 (Figure 4B, Table S3). To investigate a possible regulation of major satellite transcripts by miR-30a-3p, miR-30d-3p and miR-30e-3p, we used miRNA inhibitors against the three miRNAs WT mESCs. We transfected WT mESCs either with a negative control inhibitor or with a pool of miR-30a-3p, miR-30d-3p and miR-30e-3p inhibitors. We monitored the major satellite transcripts level 36 hours after transfection by RT-qPCR and identified an increase of around 2-fold upon transfection with the miR-30a-3p, miR-30d-3p and miR-30e-3p inhibitors compared to the negative control (Figure 4C). Taken together, these results indicate a role for the miR-30a-3p, miR-30d-3p and miR-30e-3p in fine-tuning major satellite transcript levels.

DISCUSSION

Since the discovery that AGO proteins can localize to the nucleus in mammalian cells, numerous studies have attempted to describe their nuclear functions (Meister, 2013). While in human cells, nuclear AGO's have been linked to functions in transcriptional gene regulation, splicing, chromatin organization and double-strand break repair (Agirre et al., 2015; Alló et al., 2014; Ameyar-Zazoua et al., 2012; Cho et al., 2014; Gao et al., 2014; Hu et al., 2012; Huang et al., 2013; Janowski et al., 2006; Kim et al., 2006; Li et al., 2006; Portnoy et al., 2016; Shuaib et al., 2019; Wang and Goldstein, 2016), little is known about their role during early embryonic development.

In this study, we aimed to understand the multiple roles of AGO1 in mESCs. By analyzing the subcellular distribution of the AGO proteins, we identified that only a small fraction (10-15%) localized to the nucleus. This is in contrast to a previous report of (Sarshad et al., 2018) that identified up to 70% of AGO2 in the nucleus of mESCs, despite them using the same protocol (Gagnon et al., 2014b). Furthermore, to uncover the nuclear function of AGO1 and AGO2, we studied their nuclear interactome. Considering that mouse AGO antibodies are not well suited for pulldown experiments and the low amount of AGO proteins in the nucleus, we genetically

engineered endogenously BirA*/FLAG tagged AGO1 and AGO2 cell lines. Pulldown of FLAG-AGO1 and AGO2 from the two subcellular compartments showed a lower amount of significantly enriched interactors within the nucleus compared to the cytoplasm, which is not too surprising, given the low abundance of AGO1 and AGO2 in this compartment. Additionally, we only identified known components of the RISC complex as interactors of cytoplasmic AGO2 (Table S1), but not nuclear AGO2 as previously described (Sarshad et al., 2018). However, the cell lines used in the two studies differ substantially. We used endogenously tagged AGO1 and AGO2 mESCs lines, whereas an inducible *Ago1-4*$^{-/-}$ mESC line complemented with a FLAG-HA-tagged human AGO2 was used in the other study (Sarshad et al., 2018; Su et al., 2009; Zamudio et al., 2014). Therefore, it is difficult to directly compare the two different pulldown results.

Interestingly, we identified three proteins involved in the maintenance of heterochromatin at pericentromeric regions in the nuclear AGO1 immunoprecipitation. AGO1 had previously been reported to interact with RNA Polymerase II in human cells, where AGO1 was linked to chromatin and active promoters (Alló et al., 2014; Huang et al., 2013; Shuaib et al., 2019).

We therefore decided to conduct a genetic approach by depleting *Ago1* from WT mESCs (Figure S2B) and assessed by immunofluorescence the localization of the heterochromatin protein HP1α and the repressive histone mark H3K9me3 in the mutant cell line compared to WT mESCs (Figures 2A, 2B). Surprisingly, we observed a redistribution of HP1α and the repressive histone mark H3K9me3 away from pericentromeric regions in *Ago1*_KO mESCs. The redistribution of H3K9me3 was found to be specific to the loss of AGO1, as reintroducing AGO1 could rescue the phenotype (Figure 3A). We questioned whether major satellites residing within pericentromeric regions are upregulated at the transcript level in *Ago1*_KO mESCs. Indeed, we observed an increase of pericentromeric major satellite transcripts in *Ago1*_KO mESCs by RT-qPCR and RNA FISH (Figures 3C and D). This increase was not caused by a change in the number of chromocenters, as was confirmed by DNA FISH (Figure 3F). Finally, we wondered whether AGO1 could regulate pericentromeric transcripts by a miRNA-mediated mechanism. By computational analysis, we identified that miR-30a-3p, miR-30d-3p and miR-30e-3p possibly target major satellite transcripts and that blocking these miRNAs in WT mESCs using inhibitors increased the major satellite transcripts compared to a negative control.

An involvement of smRNA loaded AGO proteins in the regulation of pericentromeric regions has previously been reported in *s.pombe*, where Ago1 loaded with siRNAs is guided to pericentromeres (Verdel et al., 2004). Ago1 together with Tas3 and Chp1 forms the RNA-induced transcriptional silencing complex (RITS). The RITS is guided to centromeric repeats by siRNAs, which are derived from this region. Targeting the RITS complex to centromeric

repeats is needed for the localization of the HP1α homolog Swi6 and the nucleation of heterochromatin H3K9me at these sites (Bühler et al., 2006; Goto and Nakayama, 2012; Motamedi et al., 2004; Verdel et al., 2004). However, our findings differ from the ones in yeast as we did not identify smRNAs derived from pericentromeric regions to be loaded in AGO1, rather miRNAs. Though, there has been reports suggesting that smRNAs from pericentromeric regions might also be present in mammalian cells (Hsieh et al., 2011; Kanellopoulou et al., 2005), we identified that AGO1 in mESCs is probably guided to major satellite transcripts by specific miRNAs, miR-30a-3p, miR-30d-3p and miR-30e-3p.

Importantly, even though we identified a decrease of H3K9me3 at pericentromeric regions upon the depletion of *Ago1* in mESCs, these cells are viable and can differentiate *in vitro*. Ngondo et al., have reported that the depletion of *Ago1* does not affect the cell cycle nor their potential to differentiate (Ngondo et al., 2018; Van Stry et al., 2012). It appears that the loss of HP1α and H3K9me3 disturbs the environment more locally without affecting overall cell viability. While we do not currently know how mESCs cope with this loss, it is possible that the plasticity of stem cells or the reestablishment of heterochromatin at pericentromeric regions upon differentiation may be required in order to survive. As several studies in human and cancer cells have already described a nuclear role for the AGO proteins, especially also for AGO1 (Agirre et al., 2015; Alló et al., 2014; Ameyar-Zazoua et al., 2012; Cho et al., 2014; Hu et al., 2012; Huang et al., 2013; Janowski et al., 2006; Kim et al., 2006; Li et al., 2006; Portnoy et al., 2016; Shuaib et al., 2019), it will be interesting to study whether the decrease of H3K9me3 at pericentromeric regions also occurs in these cell types. Interestingly, the upregulation of major satellite transcripts in several cancer lines has already been described (Hall et al., 2012), however, we do not know whether this might be linked to a nuclear AGO1 function.

While there are still open questions and further experiments required to identify the complete underlying molecular mechanism, how AGO1 regulates heterochromatin at pericentromeric regions and major satellite transcripts. Our results, allow us to propose a working model in which AGO1 regulates major satellite transcripts via miR-30a, d, e-3p. From our data we hypothesize that AGO1 might not necessarily interact directly with HP1α, but due to its targeting to major satellite transcripts by miRNAs, be in close proximity thereby aiding to maintain a specific environment at these regions (Figure 4D).

MATAERIALS AND METHODS

Mouse ESC lines

WT E14 (129/Ola background) and *Dnmt* TKO (Tsumura et al., 2006) mESCs were cultured in Dulbecco's Modified Eagle Media (DMEM, Sigma-Aldrich), supplemented with 15% fetal bovine serum (FBS, Life Technologies), 100 U/ml LIF (Millipore), 0.1 mM 2-β-mercaptoethanol (Life Technologies), and 1% Penicillin/Streptomycin (Sigma-Aldrich). MESCs were cultured on 0.2% gelatin-coated culture flasks in the absence of feeder cells. The culture medium was changed daily and all cells were grown at 37°C in 8% CO_2.

CRISPR/Cas9 mediated gene knockout

The generation of the *Ago1*_KO1 cell line was previously published by (Ngondo et al., 2018). The *Ago1*_KO2 cell line was generated using a paired CRISPR/Cas9 approach, as described by (Wettstein et al., 2016). E14 mESCs were transfected with lipofectamine 2000 (Invitrogen) and the pX458-sgRNA_*Ago1*_5/6 (addgene #172470, #172471). After 48h, GFP-positive cells were single sorted into 96-well plates (TPP). In order to confirm the deletion, genotyping at DNA level was performed, with the primers PS_Ago1_FW/RW_1 listed in the Table S2. MESC clones were then amplified and the absence of AGO1 protein and RNA was additionally verified by Western Blotting and RT-qPCR, respectively.

Generation of BirA-FLAG tagged cell lines

Endogenous tagging of *Ago1* and *Ago2* with BirA*-FLAG was achieved by transfection of two plasmids. The first encoding CRISPR-Cas9 and a single-guide RNA targeting the N-terminus of either *Ago1* or *Ago2*. The second encoding the repair template, which consisted of two around 500 bp long sequences homologous to the left and right side of the cut and flanking a BirA*-FLAG tag. Single-guide RNAs were cloned into the pX458 (addgene #48138) by restriction digest and subsequent ligation (Wettstein et al., 2016). The repair template plasmid was generated by PCR amplifying the homology arms from genomic DNA (Ago1_HR_left_F/R, Ago1_HR_right_F/R, Ago2_HR_left_F/R, Ago2_HR_right_F/R, Table S2). The outer primers contained 15 bp sequences complementary to the pUC19 vector (addgene #50005). The BirA*-FLAG sequence was amplified from pMSCV_Hyg_FLAG-BirA*-eGFP-NLS2 (addgene #170917) and flanked with 15 bp sequences complementary to the homology arms (Ago1_Flag-BirA_F/R, Ago2_Flag-BirA_F/R). The three fragments were cloned into the linearized pUC19 plasmid (addgene #50005) using In Fusion cloning (Takara). The two plasmids (addgene #170920, addgene

#170918 for AGO2 and addgene # 170921, addgene #170919 for AGO1) were simultaneously transfected into E14 mESCs with lipofectamine 2000 (Invitrogen) and single-cell sorted on GFP after 48h. Clones were screened at the genomic level by PCR (FLAG-BirA-AGO1_F/R, FLAG-BirA-AGO2_F/R, Table S2) and by testing AGO1 (CST #5053, 1:2000) and AGO2 (CST #2897, 1:2000) expression on Western blot (Figure S1A). Heterozygous clones were selected for further experiments. The control cell line was generated by transfecting E14 mESCs with pMSCV_Hyg_FLAG-BirA*-eGFP-NLS2 (addgene #170917) and single-cell sorted on GFP after 48h of transfection with lipofectamine 2000 (Invitrogen). The plasmid was generated by cloning EGFP from pMSCV-PIG (addgene #21654) and FLAG-BirA* from pDEST5-BirA*-FLAG-pcDNA5-FRT-TO (Couzens et al., 2013) into pMSCV_hyg (Takara) by In Fusion cloning. The NLS2 sequence was also added by In Fusion cloning using primers eGFP_NLS2 F/R, which contain the NLS sequence obtained from (Takahashi et al., 2013). All primers are listed in Table S2.

Generation of *Ago1*_KO complemented cell lines

For the rescue experiments, the AGO1 complemented *Ago1*_KO2 cells were obtained by stably transfecting the pMSCV_PIG_3xHA-AGO1 plasmid (addgene #170916) with lipofectamine 3000 (Invitrogen). Cells were grown for one week under puromycin selection and then sorted by FACS to select only GFP expressing cells. We sorted two mixed population into separate dishes of around 10000-20000 cells. The mixed populations were expanded and the expression of HA-AGO1 was tested by Western blot and Immunofluorescence (Figures 3A, S3B).

Cytoplasmic/ Nucleoplasmic/ Chromatin fractionation

Cytoplasmic/nucleoplasmic/chromatin fractionation was performed following the protocol of (Gagnon et al., 2014b). Cells were grown to near confluency in two 75 cm^2 (T75) flasks (TPP). 10 million cells per condition were used. Freshly harvested cells were incubated for 10 min in ice-cold Hypotonic lysis buffer (HLB) complemented with EDTA-free protease inhibitor cocktail (Roche) and Phosphatase inhibitor cocktail (Roche). After centrifugation (800xg for 8 min at 4°C) the cytoplasmic fraction was transferred to a new tube containing 5M NaCl. Pellets were washed four times with HLB (200xg for 2 min). After the last wash ice-cold modified Wuarin-Schipler buffer (MWS) (10 mM Tris-HCl (pH 7.0), 4 mM EDTA,

0.3 M NaCl, 1 M urea, and 1% (vol/vol) IGEPAL-C630), complemented with EDTA-free protease inhibitor cocktail and Phosphatase inhibitor cocktail, was added and after vortexing, incubated for 15 min on ice. After centrifugation (1000xg for 5 min at 4°C) the nucleoplasmic fraction was transferred to a new tube. The chromatin pellet was washed twice with MWS buffer, vortexed, incubated on ice for 5 min and centrifuged at 500xg for 3 min at 4°C. Ice-cold NLB was added to the chromatin pellet, which was sonicated twice at 20% for 15 sec with 2 min incubations on ice in between. The three fractions were centrifuged for 15 min at 18000xg and the supernatant was transferred to a new tube.

The fractions were then analyzed by Western blot. To ensure proper representation of all the fractions, more of the nuclear (x4) and the chromatin (x8) fraction were loaded (Figure 1A).

Analysis of the Western blot signal was performed using ImageLab (Bio-Rad Laboratories). The intensity of the bands was calculated relative to the WT band. The intensities of the nuclear and chromatin fractions were adjusted according to the additional loading and the fact that they were resuspended in half the amount of buffer compared to the cytoplasm.

Immunoprecipitation followed by Mass spectrometry analysis

Cytoplasmic/nuclear fractionation was performed following the protocol of (Gagnon et al., 2014b). Cells were grown to near confluency in two 150 cm^2 (T150) flasks (TPP) per condition. 40 million cells were used per condition. Freshly harvested cells were incubated for 10 min in ice-cold Hypotonic Lysis Buffer (HLB) (10 mM Tris (pH 7.5), 10 mM NaCl, 3 mM MgCl2, 0.3% (vol/vol) IGEPAL-C630 and 10% (vol/vol) glycerol) complemented with EDTA-free protease inhibitor cocktail (Roche) and Phosphatase inhibitor cocktail (Roche). After centrifugation (800xg for 8 min at 4°C) the cytoplasmic fraction was transferred to a new tube containing 5M NaCl. Nuclear pellets were washed four times with HLB (200xg for 2 min) and then, lysed using the Nuclear Lysis Buffer (NLB) (20 mM Tris (pH 7.5), 150 mM KCl, 3 mM MgCl2, 0.3% (vol/vol) IGEPAL-C630 and 10% (vol/vol) glycerol) complemented with EDTA-free protease inhibitor cocktail (Roche) and Phosphatase inhibitor cocktail (Roche). Sonication was performed twice at 20% for 15 sec with 2 min incubations on ice in between. Both fractions were centrifuged for 15 min at 18000xg and the supernatant was transferred to a new tube.

Immunoprecipitation was performed using Anti-Flag M2 Magnetic beads (Sigma-Aldrich). Beads were incubated in HLB or NLB plus 1% BSA for 2-3 hours on a rotating wheel at 4°C. Cytoplasmic and nuclear fractions were added to the beads and incubated overnight at 4°C. Beads were washed three

times with HLB or NLB respectively and 6 times with 1X TBS. Half the beads were removed to test success of IP by Western blot, while the rest of the beads were sent to be analyzed by the Functional Genomics Center Zurich (FGCZ), where samples were further processed as follows: beads were washed twice with digestion buffer (10 mM Tris/2 mM CaCl2, pH 8.2). Then, 45 μl 10 mM Tris/2 mM CaCl2, pH 8.2 buffer was added plus 5 μl of trypsin (100 ng/ul in 10 mM HCl). Samples were subjected to microwave assisted digestion for 30 min at 60°C. Supernatants were collected and the peptides were extracted from beads. The digested samples were dried and dissolved in 20 μl 0.1% formic acid, then diluted 1:10 in 0.1% formic acid and transferred to the autosampler vials for Liquid chromatography-mass spectrometry analysis (LC-MS/MS). Two μl were injected on a nanoAcquity UPLC coupled to a Q-Exactive mass spectrometer (Thermo).

The MS data were processed for identification using the Mascot search engine. The spectra were searched against the Swissprot (mouse) protein database. The protein identification results were then imported into Scaffold for visualization. In Scaffold the following parameters were set; Protein threshold: 95%, Minimum number of peptides per protein: 2, Peptide threshold: 80%. Triplicates were combined in Scaffold and spectrum counts were exported into Excel for further processing. Analysis of the MS data was performed in Perseus (Tyanova et al., 2016) as follows: spectrum counts were logarithmised (base 2) and the samples were grouped according to the replicates, with GFP as the control. The analysis was performed separately for AGO1 and AGO2. Proteins were filtered to include at least two valid values across the three replicates for each protein, as well as at least two values greater than 0. Invalid values were imputed following a normal distribution. For each condition, an unpaired two-sample t-test was performed to identify significantly enriched proteins over the control. Proteins were considered significant if Log2(Fold-change) >2 and -Log(p-value) >1.5 for the cytoplasm and Log2(Fold-change) >1 and -Log(p-value) >1 for the nucleus. Results were visualized as Volcano plots generated in Prism and significant proteins were displayed in Cytoscape (Shannon et al., 2003).

Western Blot Analysis

Total cellular proteins were extracted using RIPA lysing buffer (50 mM Tris-HCL pH 8.0, 150 mM NaCl, 1% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.5% sodium dodecyl sulfate supplemented with EDTA-free protease inhibitor cocktail (Roche)). Protein concentration was determined by a Bradford assay (Bio-Rad Laboratories). Proteins were separated on an SDS-PAGE gel and transferred to a

PVDF membrane (Sigma-Aldrich). Membranes were blocked for at least 30 min in blocking solution (5% milk in 1X TBS-T: TBS pH 7.6: 50 mM Tris-HCL, 150 mM NaCl and 0.1% Tween-20) and incubated overnight with primary antibodies diluted in blocking solution at 4°C. Primary antibodies used were: HP1α (CST #2616, 1:2000), DNMT3a (abcam ab188470, 1:2000), AGO1 (CST #5053, 1:2000), AGO2 (CST #2897, 1:1500), LAMIN B1 (abcam ab16048, 1:10000), TUBULIN (Sigma-Aldrich T6199, 1:10000), H3K9me3 (abcam ab8898, 1:2000), HA (Roche 3F10, 1:2000).

After washing 3 times in 1X TBS-T for 10 min, membranes were incubated with the secondary antibody for 1h at room temperature (rabbit-IgG HRP-linked 1:10000, Cell Signaling Technology (#7074), mouse-IgG HRP-linked 1:10000, Cell Signaling Technology (#7076), rat-IgG HRP-linked 1:10000 (#7077)). After incubation, membranes were washed again 3 times 10min in 1X TBS-T and developed using the Clarify Western ECL substrate kit (Bio-Rad) or SuperSignal West Femto (ThermoFisher Scientific). Membranes were imaged using the ChemiDoc MP imaging system (Bio-Rad Laboratories).

Analysis of the Western blot signal was performed using ImageLab (Bio-Rad Laboratories). Coomassie or TUBULIN was used as normalizer. Intensities of the bands were calculated relative to the WT band.

Immunofluorescence and analysis

Approximately 100000 cells were plated the night before into 6-well plates (TPP), containing coverslips coated with fibronectin (1:100 in 1X PBS, Merck). The next day, cells were washed once with 1X PBS.

For the H3K9me3 and HA staining, cells were fixed with ice-cold Methanol for 10 min at -20°C. After fixation they were washed 3 times with 1X PBS and blocked for 20 min in blocking solution (1% BSA in 1X PBS-Tween 20 (0.1%)).

For the HP1α and Dnmt3a staining, a nuclear pre-extraction was performed. Cells were washed once with ice-cold 1X PBS for 3 min on ice and then incubated in CSK buffer (0.1% Triton X-100, 10 mM PIPES, 100 mM NaCl, 3 mM $MgCl_2$, 300 mM Sucrose) for 3 min, also on ice. Afterwards cells were washed once with 1X PBS and fixed with 3.7% formaldehyde (Sigma-Aldrich) for 10 min at room temperature. After fixation, cells were washed twice with 1X PBS for 5 min at room temperature and then permeabilized with CSK buffer (same as above) for 4 min on ice. After two additional wash steps with 1X PBS at room temperature, cells were blocked in blocking solution (1% BSA in 1X PBS-Tween 20 (0.1%)) for 20 min at room temperature. After blocking, cells were incubated with the primary

antibodies diluted in blocking solution (H3K9me3: ab8898, 1:500, HP1α: CST #2616, 1:200, Dnmt3a: ab188470, 1:200, HA: Roche 3F10, 1:250) for 1h at room temperature. Coverslips were washed three times for 5 min at room temperature with 1X PBS-Tween 20 (0.1%). Then, cells were incubated with secondary antibodies diluted in blocking solution (1:2000, Invitrogen) for 1 h at room temperature in the dark. Again, coverslips were washed three times for 5 min at room temperature with 1X PBS-Tween 20 (0.1%) and once with 1X PBS. Counterstain with DAPI (0.1 μl/ml) in 1X PBS was performed for 4 min at room temperature. Cells were washed once with 1X PBS and mounted on microscopy slides on a drop of antifade medium (Vectashield, Vector Laboratories). Slides were imaged on a DeltaVision Multiplexed system with an Olympus IX71 inverse microscope equipped with a 60x 1.4NA DIC Oil PlanApoN objective and a pco.edge 5.5 camera, provided by the ScopeM facility of ETH.

For GFP expressing FLAG-BirA-GFP-NLS2 cells, 100000 cells were plated the night before into 6-well plates (TPP), containing coverslips coated with fibronectin (1:100 in 1X PBS, Merck). The next day, cells were washed once with 1X PBS and fixed with 3.7% formaldehyde (Sigma-Aldrich) for 10 min at room temperature. After fixation, cells were washed twice with 1X PBS for 5 min at room temperature and then permeabilized with CSK buffer (0.5% Triton X-100, 10 mM PIPES, 100 mM NaCl, 3 mM $MgCl_2$, 300 mM Sucrose) for 4 min on ice. After two additional wash steps with 1X PBS, cells were counterstained with DAPI (0.1 μl/ml) in 1X PBS for 4 min at room temperature. Cells were washed once with 1X PBS and mounted on microscopy slides on a drop of antifade medium (Vectashield, Vector Laboratories). Slides were imaged on a DeltaVision Multiplexed system as above provided by the ScopeM facility of ETH.

For image analysis, deconvolved images were processed with Fiji (Schindelin et al., 2012). A Z-projection of the Max intensity has been performed for each image and used for further analysis. Foci count and intensity analysis was performed on the Z-projected images, with the help of CellProfiler (Mcquin et al., 2018). In CellProfiler, nuclei were identified by using the IdentifyPrimaryObjects module and the Otsu thresholding method. Nuclei were edited manually and the DAPI and H3K9me3 was enhanced with the EnhanceOrSupressFeatures module to detect speckles. Foci were identified by using IdentifyPrimaryObject. For the DAPI foci, the RobustBackground was used as a thresholding method and the threshold strategy was set to Global. Typical diameter of objects, in pixel units was set to 5-35. For the H3K9me3 foci the same thresholding method and strategy was used, but the typical diameter of objects, in pixel units was set to 7-35. Foci were related to the edited nuclei and the

H3K9me3 foci were related to the DAPI foci. The MeasureObjectIntensity module was used to measure object intensity and foci count and results were exported to a csv file. Overlayed objects were saved as png.

RNA extraction and quantitative RT-qPCR Analysis

RNA extraction and RT PCR analysis has been performed as previously described by (Bodak and Ciaudo, 2016). Briefly, total RNA from mESC pellets was extracted using Trizol (Life Technologies) according to standard protocols. RNA quality was checked, by running 1 μg on a 1% agarose gel (Sigma). 2 μg of RNA was treated with DNase (RQ1 RNase-free DNase Kit, Promega), according to manufacturer's protocol. Reverse transcription was performed using the GoScript Reverse Transcriptase Kit (Promega), according to the manufacturer's protocol. A positive (+RT) and negative reaction (-RT) were performed in parallel. cDNA of positive and negative RT reactions was checked for the absence of genomic contamination, by performing a PCR for β-actin. RT-qPCR was performed on a LightCycler 480 instrument (Roche), using 2 μl of diluted cDNA (1:5) and the KAPA SYBR FAST qPCR kit (Kapa Biosystems). Samples were run in a technical triplicate and expression levels were calculated using the $2^{-\Delta\Delta CT}$ method and normalized to the housekeeping gene Gapdh.

For qPCR on major satellite transcripts, 20 μg of RNA was treated twice with 1U of DNaseI (Qiagen) per μg of RNA. RiboLock was added to reduce RNA degradation. DNase treated RNA was purified using Direct-zol RNA mini prep kit (Zymo Research). Reverse transcription and qPCR were performed as described above.

Primers are listed in the supplementary Table S2.

RNA FISH

The plasmid pCR4 Maj9-2 (a kind gift from the Almouzni lab, originally from (Lehnertz et al., 2003)) was used to generate the RNA FISH probe by nick translation (Abott). In brief, 2 μg of plasmid, 3 μl of nick translation enzyme, 2.5 μl 0.2 mM red-dUTP, 5 μl 0.1 mM dTTP, 10 μl 0.1 mM dNTP mix and 5 μl 10X nick translation buffer were incubated for 15 h at 15°C. Nick translation efficiency was checked for by running 3 μl probe on a 1% agarose gel (Sigma). The rest of the probe was cleaned-up with the

Zymoclean Gel DNA Recovery Kit (Zymo Research). The probe was dried down to 5 μl using a speed vac and then resuspended in hybridization solution (50 μl Deionized Formamide, 10 μl 20X SSC, 2 μl 100mg/ml BSA, 20 μl 50% Dextran Sulfate, 3 μl Salmon Sperm, 10 μl RVC). Prior to use, the probe was diluted 1:2 in hybridization solution.

Approximately 150000 cells were plated the night before into 6-well plates, containing coverslips coated with fibronectin (1:100 in 1X PBS, Merck). A nuclear pre-extraction was performed. Cells were washed once with ice-cold 1X PBS for 3 min on ice and then incubated in CSK buffer (0.1% Triton X-100, 10 mM PIPES, 100 mM NaCl, 3 mM $MgCl_2$, 300 mM Sucrose) for 3 min, also on ice. Afterwards cells were washed once with 1X PBS and fixed with 3.7% formaldehyde (Sigma-Aldrich) for 10 min at room temperature. After fixation, cells were washed twice with 1X PBS for 5 min at room temperature and then permeabilized with CSK buffer (same as above) for 4 min on ice. After two additional wash steps with 1X PBS at room temperature, cells were blocked in blocking solution (1% BSA in 1X PBS and RVC (1 mM)) for 30 min at room temperature. Then coverslips were washed once with 2X SSC and dehydrated with ethanol (70% EtOH for 3 min, 90% EtOH for 3 min, 100% EtOH for 3 min). The probe was denatured for 5 min at 76°C. 10μl of the denatured probe (diluted 1:2 in hybridization buffer) was spotted on a baked slide. The coverslips were air-dried and placed on the spotted probe. The coverslips were sealed with rubber cement and then incubated overnight at 37°C in a humid chamber. The next day, coverslips were washed twice with 50% Formamide/2X SSC for 5 min at 37°C and then once for 5 min with 2X SSC at room temperature. Counterstain with DAPI (0.1 μl/ml) in 2X SSC was performed for 4 min at room temperature. Coverslips were washed again once in 2xSSC and once 1X PBS and then mounted on microscopy slides on a drop of antifade medium (Vectashield, Vector Laboratories). Slides were image on a DeltaVision Multiplexed system provided by the ScopeM facility of ETH as above.

For image analysis, deconvolved images were processed with Fiji (Schindelin et al., 2012). A Z-projection of the Max intensity has been performed for each image and used for further analysis. RNA Foci were counted by eye. All Z-projected images of one replicate were opened using the Fiji software and were set to the same intensity.

DNA FISH

DNA FISH was performed exactly as described for the RNA FISH. The only differences, are that no RVC was used in the blocking solution and the samples were denatured by incubating the slides at 76°C for 5 min once the coverslips were placed on the spotted probe and sealed with rubber cement.

ChIP and ChIP-qPCR analysis

Four million cells were plated the night before into a gelatin-coated 60.1 cm^2 (B10) dish (TPP). For each condition, two B10 dishes were prepared in parallel. Cells were crosslinked with 1% formaldehyde in DMEM for 10 min at room temperature. The reaction was quenched with glycine (125 mM, PanReac Applichem) for 5 min at room temperature. Cells were washed once with ice-cold 1X PBS and then swelling buffer (5 mM HEPES pH 8, 85 mM KCl, 0.5% IGEPAL-CA630, protease inhibitor cocktail (Roche)) was added to the cells. Cells were scraped and transferred to a 15 ml falcon (Greiner), where they were incubated for 15 min on ice. Cells were centrifuged 5 min at 2000 rpm at 4°C, to pellet the nuclei. Afterwards, nuclei were washed again with swelling buffer followed by another centrifugation (250 g at 4°C for 5 min). The nuclei pellet was lysed in 400 μl RIPA buffer 1% SDS (1X PBS, 1% IGEPAL-C630, 0.5% sodium deoxycholate, 1% sodium dodecyl sulfate, protease inhibitor cocktail (Roche)) and incubated on ice for 10 min. The lysates were sonicated on a Bioruptor (Diagenode) for 30 min, 30s on and 30s off cycles at 4°C. Lysates were centrifuged at max speed for 15 min at 4°C. The supernatant was retrieved into a new 2 ml Eppendorf tube and diluted 10 times with RIPA buffer 0% SDS (1X PBS, 1% IGEPAL-C630, 0.5% sodium deoxycholate, protease inhibitor cocktail (Roche)) to obtain a concentration of 0.1% sodium dodecyl sulfate. 10% of chromatin was taken away for Input calculation, the rest of the chromatin was snap-frozen and stored at -80°C. Input DNA was treated with 10 μg RNase A for 1 h at 37°C followed by a proteinase K treatment (40 μg) for 1-2 h. DNA was extracted with phenol/chloroform (Sigma-Aldrich) and concentration was measured and used to calculate the total amount of chromatin in each sample.

For the pulldown, 20 μg of chromatin was precleared with 10 μl of Dynabeads protein G (ThermoFisher Scientific), previously washed three times with RIPA 0.1% SDS (1X PBS, 1% IGEPAL-C630, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, protease inhibitor cocktail (Roche)), for 2 h on the wheel at 4°C. 1/10 of the pre-cleared chromatin was taken away and stored temporarily at -20°C, this was later used as the Input. The rest of the pre-cleared chromatin was transferred into a new 1.5 ml

Eppendorf tube and incubated with 2 μg of antibody for each condition overnight at 4°C (H3K9me3: ab8898, rabbit-IgG: NI01). 10 μl of Dynabeads protein G (ThermoFisher Scientific) was added to the chromatin-antibody complexes and incubated 4 h on the wheel at 4°C. Samples were placed on the magnetic rack and the supernatant was discarded. Samples were washed twice with wash buffer 1 (16.7 mM Tris-HCL pH 8, 0.167 M NaCl, 0.1% SDS, 1% Triton X-100) for 5 min rotating at room temperature. Then they were washed once with wash buffer 2 (16.7 mM Tris-HCl pH 8, 0.5M NaCl, 0.1% SDS, 1% Triton X-100) for 5 min rotating at room temperature and twice in LiCl wash buffer (0.25 M LiCl, 0.5% sodium deoxycholate, 1 mM EDTA pH 8, 10 mM Tris-HCl pH 8, 0.5% IGEPAL-CA630) for 5 min rotating at room temperature. Finally, the samples were washed twice in TE buffer (10 mM Tris-HCl pH 8, 5 mM EDTA pH 8) for 5 min rotating at room temperature. Samples were incubated in 300 μl elution buffer (1% SDS, 100 mM $NaHCO_3$) for 30 min at 37°C shaking at 900 rpm. Samples were placed on a magnetic rack and the supernatant transferred into a new Eppendorf tube containing 38.5 μl Proteinase K mix (15 μl 1M Tris-HCl pH 8, 15 μl 5M NaCl, 7.5 μl 0.5M EDTA pH 8, 1 μl Proteinase K (20 mg/ml)). Also 300 μl elution buffer and 38.5 μl Proteinase K mix was added to the Inputs. Pull-downs and Inputs were incubated at 50°C for 3 h shaking at 1100 rpm and then at 65°C overnight. DNA was treated with 10 μg of RNase A for 45 min at 37°C and extracted with a phenol/chloroform extraction, followed by an ethanol precipitation.

ChIPed and Input DNA was diluted 1:10 prior to the qPCR for the control Primers (Dazl, MusD and the intergenic region) and 1:50 prior to the qPCR for the major satellites. The qPCR was performed with the KAPA SYBR Fast qPCR Kit (Kapa Biosystems) and analyzed on a LightCycler 480 (Roche). The enrichment was calculated with the $2^{-\Delta\Delta CT}$ method over input. Control regions (Dazl, MusD (Karimi et al., 2011)) were represented as enrichment over the intergenic region. The enrichment for the major satellites was represented as the enrichment compared to WT. Primers are listed in Table S2.

Major satellite computational analysis and binding site identification

RepeatMasker annotations were obtained from UCSC for the mm10 mouse reference genome (**http://hgdownload.soe.ucsc.edu/goldenPath/mm10/database/rmskOutCurrent.txt.gz**) and filtered for major satellites (GSAT_MM) (Bao et al., 2015; Smit et al.). Most regions annotated as major satellites were rather short (< 1000 bps) and we only considered 4 regions, where the genome

sequence contained > 20 Kbps long major satellite regions. One of them was mapped into the X-chromosome and another one to chromosome 9. The other two fell into genomic contigs that could not be assigned to any chromosome (JH584304.1 and GL456383.1). For these four annotated regions, 8mer binding sites were scanned and counted for each mESC-expressed miRNA (Table S3).

miRNA inhibitor transfection

400 000 WT mESCs were plated the night before into a gelatin-coated 60.1 cm^2 (B10) dish (TPP). The next day, WT mESCs were either transfected with RNAiMax (Invitrogen) and 30 nM of a negative control inhibitor (Ambion, #4464074) or with RNAiMax (Invitrogen) and a mix of miR-30a-3p, miR-30d-3p and miR-30e-3p inhibitors, 10 nM for each inhibitor (Ambion, #4464084). 36 h later, the cell pellet was collected. Briefly, cells were washed once with 1X PBS (Life Technologies), then trypsinized with 0.05% Trypsin-EDTA (Life Technologies) for 5 min at 37°C. Trypsinization was stopped, by adding medium and spinning the cells down for 5 min at 182xg. The cell pellet was washed once in 1X PBS (Life Technologies), spun down 5 min at 182xg and then stored at -80°C.

QUANTIFICATION AND STATISTICAL ANALYSIS

See Methods Details for details on quantification and statistical analysis. In general, statistical analysis was performed using PRIMS 8 as indicated in the figure legends.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal Anti-H3K9me3	Abcam	Ab8898
Rabbit Anti-HP1α	Cell Signaling	#2616
Rabbit monoclonal Anti-Dnmt3a	Abcam	Ab188470
Rabbit monoclonal Anti-Argonaute 1 (D84G10)	Cell Signaling	#5053
Rabbit monoclonal Anti-Argonaute 2 (C34C6)	Cell Signaling	#2897
Mouse monoclonal Anti-α-TUBULIN	Sigma-Aldrich	T6199
Rabbit polyclonal Anti-Lamin B1	Abcam	Ab16048
Rat Anti-HA	Roche	3F10
Normal Rabbit IgG	Sigma-Aldrich	NI01
Anti-rabbit IgG, HRP-linked	Cell Signaling	#7074
Anti-mouse IgG, HRP-linked	Cell Signaling	#7076

Anti-rat IgG, HRP-linked	Cell Signaling	#7077
Donkey Anti-rabbit 546	Invitrogen	A11040
Donkey Anti-mouse 488	Invitrogen	A21202
Goat Anti-rat 488	Invitrogen	A11006
Chemicals, Peptides, and Recombinant Proteins		
Trizol	Life Technologies	15596018
Ethanol	Merck	100983100
DMEM Media	Sigma-Aldrich	D6429-500ML
ESGRO recombinant mouse LIF protein	Millipore	ESG1107
Penicillin/Streptomycin	Sigma-Aldrich	P0781-100ML
0.05% Trypsin-EDTA	Life Technologies	25300054
1X PBS	Life Technologies	10010015
2-ß-mercaptoethanol	Life Technologies	31350010
FBS	Life Technologies	10270-106
Dimethyl Sulfoxid	Sigma-Aldrich	D8418
Phenol/Choloroform/Isoamyl Alcohol	Sigma-Aldrich	P2069
Tris(hydroxymethly)aminomethane	Biosolve	0020092391BS
Sodium chloride	Merck	1064045000
IGEPAL CA-630	Sigma-Aldrich	I3021
Sodium deoxycholate	Sigma-Aldrich	30970
Sodium dodecyl sulfate	Sigma-Aldrich	75746
Agarose	Sigma-Aldrich	A9539
Sodium hydrogen carbonate	Merck	1063290500
EDTA-free Protease inhibitor cocktail	Roche	5056489001
PhosSTOP	Roche	4906837001
Tween20	Sigma-Aldrich	P1379
Coomassie	VWR	443283M
Triton-X	Roth	9002-93-1
Lithium chloride	Merck	1056790100
EDTA	PanReac AppliChem	A2937
Dynabeads Protein G	ThermoFisher Scientific	1004D
M2 Anti-FLAG magnetic beads	Sigma-Aldrich	M8823
Proteinase K	PanReac AppliChem	A3830
RNase A	Roche	10109169001
Glycogen	Invitrogen	AM9510
HEPES	Sigma-Aldrich	H4034
PIPES	Sigma-Aldrich	P6757
Magnesium chloride	Fluka	P6757
Sucrose	Fluka	63072
Bovine Serum Albumin	PanReac Applichem	A1391
DAPI	Sigma-Aldrich	D9542
Formaldehyde	Sigma-Aldrich	47608-1L-F
Deionized Formamide	Eurobio	GHYFOR01-01
Sodium citrate tribasic dihydrate	Sigma-Aldrich	S4641
Dextran sulfate	Fluka	66786
Salmon Sperm	Invitrogen	15632011
Ribonucleoside Vanadyl Complex (RVC)	NEB	S1402S
Glycine	PanReac Applichem	A1067
Fibronectin	Merck	FC010
Methanol	Merck	1060072500
Vectashield	Vector Laboratories	H-1000
Potassium chloride	Merck	104936100

Lipofectamine 2000	Invitrogen	52887
Lipofectamine 3000	Invitrogen	100022052
Ammonium persulfate (APS)	Sigma-Aldrich	A3678
TEMED	Sigma-Aldrich	T9281
Gycerol	PanReac AppliChem	A1123
Bromophenol blue	Roth	T116.1
Urea	Eurobio	GEPURE00-67
Puromycin dihydrochloride from Streptomyces alboniger	Sigma-Aldrich	P8833
Hygromycin B	Invitrogen	10687010
Gelatin from porcine skin	Sigma-Aldrich	G1890
Acrylamide 4K Solution	PanReacAppliChem	A1672
Lipofectamine RNAiMax	Invitrogen	56532
mir Vana miRNA inhibitor negative control	Ambion	#4464074
mir Vana miRNA inhibitor hsa-miR-30a, d, e-3p	Ambion	#4464084
Critical Commercial Assays		
KAPA SYBR FAST for Roche LightCycler 480	Sigma-Aldrich	KK4611
GoScript Reverse Transcriptase	Promega	A5004
Nick translation kit	Abott	32-801300
RQ1 RNase-Free DNase Kit	Promega	M6101
DC Protein Assay Reagent	Bio-rad	5000113-5
Clarity Western ECL Substrate	Bio-rad	1705061
SuperSignal West Femto Maximum Sensitivity Substrate	ThermoFisher Scientific	34095
RNase-free DNase Set	Qiagen	79254
Direct-zol RNA miniprep Kit	Zymo Research	R2050
Zymoclean Gel DNA Recovery Kit	Zymo Research	D4001
In Fusion HD Cloning Kit	Takara	638920
Deposited Data		
Proteomics data of AGO1 and AGO2 interactome	This manuscript	ProteomeXchange PXD026786
Small RNA-seq WT mESCs	(Ngondo et al., 2018)	GSE80415
Small RNA-seq Ago1 IP	(Ngondo et al., 2018)	GSE80454
Experimental Models: Cell Lines		
WT Mouse ESCs (E14Tg2a)	ATCC	CRL-1821
*Dnmt*_TKO mESCs	(Tsumura et al., 2006)	N/A
Oligonucleotides		
Primers used for cloning, PCR and RT-qPCR	This manuscript	Table S2
Recombinant DNA		
pX458	(Ran et al., 2013)	Addgene #48138
pMSCV-PIG	(Mayr and Bartel, 2009)	Addgene #21654
pUC19	(Norrander et al., 1983)	Addgene #50005
pMSCV_hyg	Takara	634401
pMSCV_PIG_3XHA-Ago1	This manuscript	Addgene #170916
pDEST5-BirA-FLAG-pcDNA5-FRT-TO	(Couzens et al., 2013)	N/A
pMSCV_Hyg_FLAG-BirA-eGFP-NLS2	This manuscript	Addgene #170917
pUC19_HR-BirA-FLAG_Ago2	This manuscript	Addgene #170918
pUC19_HR-BirA-FLAG_Ago1	This manuscript	Addgene #170919
pX458_sgAgo2_Nterm	This manuscript	Addgene #170920
pX458_sgAgo1_Nterm	This manuscript	Addgene #170921
pX458_sgAgo1_5	This manuscript	Addgene #172470
pX458_sgAgo1_6	This manuscript	Addgene #172471

pCR4 Maj9-2	(Lehnertz et al., 2003)	N/A
Software and Algorithms		
CellProfiler (3.1.9)	(Mcquin et al., 2018)	https://cellprofiler.org/
Fiji ImageJ (2.0.0)	(Schindelin et al., 2012)	https://imagej.net/Fiji
Scaffold (v4.11.0)	Proteome Software Inc.	https://www.proteomesoftware.com
Perseus	(Tyanova et al., 2016)	https://maxquant.net/perseus/
Adobe Illustrator 2021	Adobe Inc.	https://www.adobe.com
Prism v8	GraphPad	https://www.graphpad.com
ImageLab v6.1.0	Bio-Rad Laboratories	https://www.bio-rad.com
Cytoscape v3.8.0	(Shannon et al., 2003)	https://cytoscape.org
Mascot	(Perkins et al., 1999)	http://www.matrixscience.com
Other		
PVDF membrane	Sigma-Aldrich	GE10600023
1.5 ml Safe-Lock Tubes	Eppendorf	0030120086
2 ml Safe-Lock Tubes	Eppendorf	0030120094
T25: 25 cm^2 Filter Flask	TPP	90026
T75: 75 cm^2 Filter Flask	TPP	90076
T150: 150 cm^2 Filter Flask	TPP	90151
B10: 60.1 cm^2 B10	TPP	93100
96-well plate	TPP	92096
6-well plate	TPP	92006
15 ml CELLSTAR Tubes	Greiner	188271

Acknowledgments

We would like to thank the members of the Ciaudo, Dr. Tobias Beyer and Prof. Madhav Jagannathan (ETH) for fruitful discussions and the critical reading of this manuscript. This work was supported by the Swiss National Science Foundation (grants 31003A_173120 and 310030_196861) to C.C. R.A and M.M were supported by the NCCR RNA and disease. We are also thankful to the Scientific Centre for Optical and Electron Microscopy (ScopeM, ETH Zurich) for their support for imaging. We also want to thank the Functional Genomics Center Zurich (FGCZ) for their support with the Mass Spectrometry.

Author Contributions

Conceptualization, MM, TF and CC.; laboratory experiments, MM, TF, JL, PS, RA, PN; computational analysis, MS.; writing—original draft preparation, MM, TF and CC.; writing—review and editing, CC.; expertise and editing, PN.; visualization, MM, TF, MS and CC.;

supervision, CC; funding acquisition, CC. All authors have read and agreed to the published version of the manuscript.

Declaration of Interests

The authors declare no financial and non-financial competing interests.

References

Agirre, E., Bellora, N., Alló, M., Pagès, A., Bertucci, P., Kornblihtt, A.R., and Eyras, E. (2015). A chromatin code for alternative splicing involving a putative association between CTCF and HP1aα proteins. BMC Biol. *13*.

Alisch, R.S., Jin, P., Epstein, M., Caspary, T., and Warren, S.T. (2007). Argonaute2 is essential for mammalian gastrulation and proper mesoderm formation. PLoS Genet. *3*, e227.

Alló, M., Agirreb, E., Bessonovc, S., Bertucci, P., Acuña, L.G., Buggiano, V., Bellorab, N., Singhb, B., Petrillo, E., Blaustein, M., et al. (2014). Argonaute-1 binds transcriptional enhancers and controls constitutive and alternative splicing in human cells. Proc. Natl. Acad. Sci. U. S. A. *111*, 15622–15629.

Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.C., et al. (2012). Argonaute proteins couple chromatin silencing to alternative splicing. Nat. Struct. Mol. Biol. *19*, 998–1005.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120–124.

Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA *6*, 4–9.

Bartel, D.P. (2018). Metazoan MicroRNAs. Cell *173*, 20–51.

Bernard, P., Maure, J.F., Partridge, J.F., Genier, S., Javerzat, J.P., and Allshire, R.C. (2001). Requirement of heterochromatin for cohesion at centromeres. Science. *294*, 2539–2542.

Bodak, M., and Ciaudo, C. (2016). Monitoring Long Interspersed Nuclear Element 1 Expression During Mouse Embryonic Stem Cell Differentiation. Methods Mol. Biol. *1400*, 237–259.

Bodak, M., Cirera-Salinas, D., Luitz, J., and Ciaudo, C. (2017). The Role of RNA Interference in Stem Cell Biology: Beyond the Mutant Phenotypes. J. Mol. Biol. *429*, 1532–1543.

Boroviak, T., Stirparo, G.G., Dietmann, S., Hernando-Herraez, I., Mohammed, H., Reik, W., Smith, A., Sasaki, E., Nichols, J., and Bertone, P. (2018). Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development. Development *145*, dev167833.
Bühler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a Nascent Transcript Initiates RNAi- and Heterochromatin-Dependent Gene Silencing. Cell *125*, 873–886.
Cheloufi, S., Dos Santos, C.O., Chong, M.M.W., and Hannon, G.J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature *465*, 584–589.
Chen, C.Y.A., Zheng, D., Xia, Z., and Shyu, A. Bin (2009). Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. Nat. Struct. Mol. Biol. *16*, 1160–1166.
Cho, S., Park, J.S., and Kang, Y.K. (2014). AGO2 and SETDB1 cooperate in promoter-targeted transcriptional silencing of the androgen receptor gene. Nucleic Acids Res. *42*, 13545–13556.
Couzens, A.L., Knight, J.D.R., Kean, M.J., Teo, G., Weiss, A., Dunham, W.H., Lin, Z.Y., Bagshaw, R.D., Sicheri, F., Pawson, T., et al. (2013). Protein interaction network of the mammalian hippo pathway reveals mechanisms of kinase-phosphatase interactions. Sci. Signal. *6*.
DeVeale, B., Swindlehurst-Chan, J., and Blelloch, R. (2021). The roles of microRNAs in mouse development. Nat. Rev. Genet. *22*, 307–323.
Fabian, M.R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J.T., Svitkin, Y. V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J.A., et al. (2009). Mammalian miRNA RISC Recruits CAF1 and PABP to Affect PABP-Dependent Deadenylation. Mol. Cell *35*, 868–880.
Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat. Cell Biol. *8*, 188–194.
Frohn, A., Eberl, H.C., Stöhr, J., Glasmacher, E., Rüdel, S., Heissmeyer, V., Mann, M., and Meister, G. (2012). Dicer-dependent and -independent argonaute2 protein interaction networks in mammalian cells. Mol. Cell. Proteomics *11*, 1442–1456.
Gagnon, K.T., Li, L., Chu, Y., Janowski, B.A., and Corey, D.R. (2014a). RNAi factors are present and active in human cell nuclei. Cell Rep. *6*, 211–221.
Gagnon, K.T., Li, L., Janowski, B.A., and Corey, D.R. (2014b). Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. Nat. Protoc. *9*, 2045–2060.
Gao, M., Wei, W., Li, M.M., Wu, Y.S., Ba, Z., Jin, K.X., Li, M.M., Liao, Y.Q., Adhikari, S., Chong, Z., et al. (2014). Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. Cell Res. *24*, 532–541.

Goto, D.B., and Nakayama, J. (2012). RNA and epigenetic silencing: Insight from fission yeast. Dev. Growth Differ. *54*, 129–141.
Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. (2004). Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. J. Cell Biol. *166*, 493–505.
Hall, L.E., Mitchell, S.E., and O'Neill, R.J. (2012). Pericentric and centromeric transcription: A perfect balance required. Chromosom. Res. *20*, 535–546.
Houlard, M., Berlivet, S., Probst, A. V., Quivy, J.P., Héry, P., Almouzni, G., and Gérard, M. (2006). CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. PLoS Genet. *2*, 1686–1696.
Hsieh, C.L., Lin, C.L., Liu, H., Chang, Y.J., Shih, C.J., Zhong, C.Z., Lee, S.C., and Tan, B.C.M. (2011). WDHD1 modulates the post-transcriptional step of the centromeric silencing pathway. Nucleic Acids Res. *39*, 4048–4062.
Hu, J., Chen, Z., Xia, D., Wu, J., Xu, H., and Ye, Z.Q. (2012). Promoter-associated small double-stranded RNA interacts with heterogeneous nuclear ribonucleoprotein A2/B1 to induce transcriptional activation. Biochem. J. *447*, 407–416.
Huang, V., Zheng, J., Qi, Z., Wang, J., Place, R.F., Yu, J., Li, H., and Li, L.C. (2013). Ago1 Interacts with RNA Polymerase II and Binds to the Promoters of Actively Transcribed Genes in Human Cancer Cells. PLoS Genet. *9*.
Hyun, K., Jeon, J., Park, K., and Kim, J. (2017). Writing, erasing and reading histone lysine methylations. Exp. Mol. Med. *49*.
Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Nordsell, R., Shames, D.S., Minna, J.D., and Corey, D.R. (2006). Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat. Struct. Mol. Biol. *13*, 787–792.
Kalantari, R., Hicks, J.A., Li, L., Gagnon, K.T., Sridhara, V., Lemoff, A., Mirzaei, H., and Corey, D.R. (2016). Stable association of RNAi machinery is conserved between the cytoplasm and nucleus of human cells. RNA *22*, 1085–1098.
Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev. *19*, 489–501.
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.
Kim, D.H., Villeneuve, L.M., Morris, K. V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat. Struct. Mol. Biol. *13*, 793–797.
Komissarov, A.S., Gavrilova, E. V., Demin, S.J., Ishov, A.M., and Podgornaya, O.I. (2011).

Tandemly repeated DNA families in the mouse genome. BMC Genomics *12*.
Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 120–124.
Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. Curr. Biol. *13*, 1192–1200.
Li, L.C., Okino, S.T., Zhao, H., Pookot, D., Place, R.F., Urakami, S., Enokida, H., and Dahiya, R. (2006). Small dsRNAs induce transcriptional activation in human cells. Proc. Natl. Acad. Sci. U. S. A. *103*, 17337–17342.
Li, X., Wang, X., Cheng, Z., and Zhu, Q. (2020). AGO2 and its partners: a silencing complex, a chromatin modulator, and new features. Crit. Rev. Biochem. Mol. Biol. *55*, 33–53.
Liu, J., Carmell, M.A., Rivas, F. V., Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 Is the Catalytic Engine of Mammalian RNAi. Science. *305*, 1437–1441.
Lykke-Andersen, K., Gilchrist, M.J., Grabarek, J.B., Das, P., Miska, E., and Zernicka-Goetz, M. (2008). Maternal Argonaute 2 Is Essential for Early Mouse Development at the Maternal-Zygotic Transition. Mol. Biol. Cell *19*, 4383–4392.
Martinez, N.J., Chang, H.M., De Riba Borrajo, J., and Gregory, R.I. (2013). The co-chaperones Fkbp4/5 control Argonaute2 expression and facilitate RISC assembly. Rna *19*, 1583–1593.
Mayr, C., and Bartel, D.P. (2009). Widespread Shortening of 3′UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. Cell *138*, 673–684.
Mcquin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, A., Karhohs, K.W., Doan, M., Ding, L., Rafelski, S.M., Thirstrup, D., et al. (2018). Cellprofiler 3.0: Next-generation image processing for biology. PLoS Biol. *16*, 1–17.
Meister, G. (2013). Argonaute proteins: Functional insights and emerging roles. Nat. Rev. Genet. *14*, 447–459.
Modzelewski, A.J., Holmes, R.J., Hilz, S., Grimson, A., and Cohen, P.E. (2012). AGO4 Regulates Entry into Meiosis and Influences Silencing of Sex Chromosomes in the Male Mouse Germline. Dev. Cell *23*, 251–264.
Morita, S., Horii, T., Kimura, M., Goto, Y., Ochiya, T., and Hatada, I. (2007). One Argonaute family member, Eif2c2 (Ago2), is essential for development and appears not to be involved in DNA methylation. Genomics *89*, 687–696.
Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789–802.

Müller, M., Fazi, F., and Ciaudo, C. (2020). Argonaute Proteins: From Structure to Function in Development and Pathological Cell Fate Determination. Front. Cell Dev. Biol. *7*, 1–10.
Ngondo, R.P., Cirera-Salinas, D., Yu, J., Wischnewski, H., Bodak, M., Vandormael-Pournin, S., Geiselmann, A., Wettstein, R., Luitz, J., Cohen-Tannoudji, M., et al. (2018). Argonaute 2 Is Required for Extra-embryonic Endoderm Differentiation of Mouse Embryonic Stem Cells. Stem Cell Reports *10*, 461–476.
Nishi, K., Nishi, A., Nagasawa, T., and Ui-Tei, K. (2013). Human TNRC6A is an Argonaute-navigator protein for microRNA-mediated gene silencing in the nucleus. RNA *19*, 17–35.
Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S.I.S., and Watanabe, Y. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. Nat. Cell Biol. *4*, 89–93.
Norrander, J., Kempe, T., and Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene *26*, 101–106.
Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: Improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.
Perkins, D.N., Pappin, D.J.C., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis *20*, 3551–3567.
Portnoy, V., Lin, S.H.S., Li, K.H., Burlingame, A., Hu, Z.H., Li, H., and Li, L.C. (2016). SaRNA-guided Ago2 targets the RITA complex to promoters to stimulate transcription. Cell Res. *26*, 320–335.
Probst, A. V., and Almouzni, G. (2008). Pericentric heterochromatin: Dynamic organization during early development in mammals. Differentiation *76*, 15–23.
Probst, A. V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat. Rev. Mol. Cell Biol. *10*, 192–206.
Probst, A. V, Okamoto, I., Casanova, M., Marjou, F. El, and Le Baccon, P. (2010). A Strand-Specific Burst in Transcription of Pericentric Satellites Is Required for Chromocenter Formation and Early Mouse Development. Dev. Cell *19*, 625–638.
Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308.
Rudert, F., Bronner, S., Garnier, J.W., and Dollé, P. (1995). Transcripts from opposite strands of γ satellite DNA are differentially expressed during mouse development. Mamm. Genome *6*, 76–83.
Sarshad, A.A., Juan, A.H., Muler, A.I.C., Anastasakis, D.G., Wang, X., Genzor, P., Feng, X., Tsai, P.-F., Sun, H.-W., Haase, A.D., et al. (2018). Argonaute-miRNA Complexes Silence

Target mRNAs in the Nucleus of Mammalian Stem Cells. Mol. Cell *71*, 1040-1050.e8.
Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: An open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.
Schopp, I.M., Amaya Ramirez, C.C., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., and Béthune, J. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. Nat. Commun. *8*, 15690.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models. Genome Res. *13*, 2498–2504.
Shuaib, M., Parsi, K.M., Thimma, M., Adroub, S.A., Kawaji, H., Seridi, L., Ghosheh, Y., Fort, A., Fallatah, B., Ravasi, T., et al. (2019). Nuclear AGO1 Regulates Gene Expression by Affecting Chromatin Architecture in Human Cells. Cell Syst. *9*, 446-458.e6.
Smit, A., Hublex, R., and Green, P. RepeatMasker Open-4.0.
Storer, J., Hubley, R., Rosen, J., Wheeler, T.J., and Smit, A.F. (2021). The Dfam community resource of transposable element families, sequence models, and genome annotations. Mob. DNA *12*, 1–14.
Van Stry, M., Oguin, T.H., Cheloufi, S., Vogel, P., Watanabe, M., Pillai, M.R., Dash, P., Thomas, P.G., Hannon, G.J., and Bix, M. (2012). Enhanced Susceptibility of Ago1/3 Double-Null Mice to Influenza A Virus Infection. J. Virol. *86*, 4151–4157.
Su, H., Trombly, M.I., Chen, J., and Wang, X. (2009). Essential and overlapping functions for mammalian argonautes in microRNA silencing. Genes Dev. *23*, 304–317.
Takahashi, M., Obayashi, M., Ishiguro, T., Sato, N., Niimi, Y., Ozaki, K., Mogushi, K., Mahmut, Y., Tanaka, H., Tsuruta, F., et al. (2013). Cytoplasmic Location of α1A Voltage-Gated Calcium Channel C-Terminal Fragment (Cav2.1-CTF) Aggregate Is Sufficient to Cause Cell Death. PLoS One *8*, 1–11.
Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S.I., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., et al. (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes to Cells *11*, 805–814.
Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.
Vasudevan, S., and Steitz, J.A. (2007). AU-Rich-Element-Mediated Upregulation of Translation by FXR1 and Argonaute 2. Cell *128*, 1105–1118.
Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I.S., and Moazed, D. (2004). RNAi-Mediated Targeting of Heterochromatin by the RITS. Science. *303*, 672–676.

Vissel, B., and Choo, K.H. (1989). Mouse major (γ) satellite DNA is highly conserved and organized into extremely long tandem arrays: Implications for recombination between nonhomologous chromosomes. Genomics *5*, 407–414.
Wang, Q., and Goldstein, M. (2016). Small RNAs recruit chromatin-modifying enzymes MMSET and Tip60 to reconfigure damaged DNA upon double-strand break and facilitate repair. Cancer Res. *76*, 1904–1915.
Wettstein, R., Bodak, M., and Ciaudo, C. (2016). Generation of a Knockout Mouse Embryonic Stem Cell Line Using a Paired CRISPR/Cas9 Genome Engineering Tool. Methods Mol. Biol. *1341*, 321–343.
Zamudio, J.R., Kelly, T.J., and Sharp, P.A. (2014). Argonaute-bound small RNAs from promoter-proximal RNA polymerase II. Cell *156*, 920–934.

Figure & Figure Legends

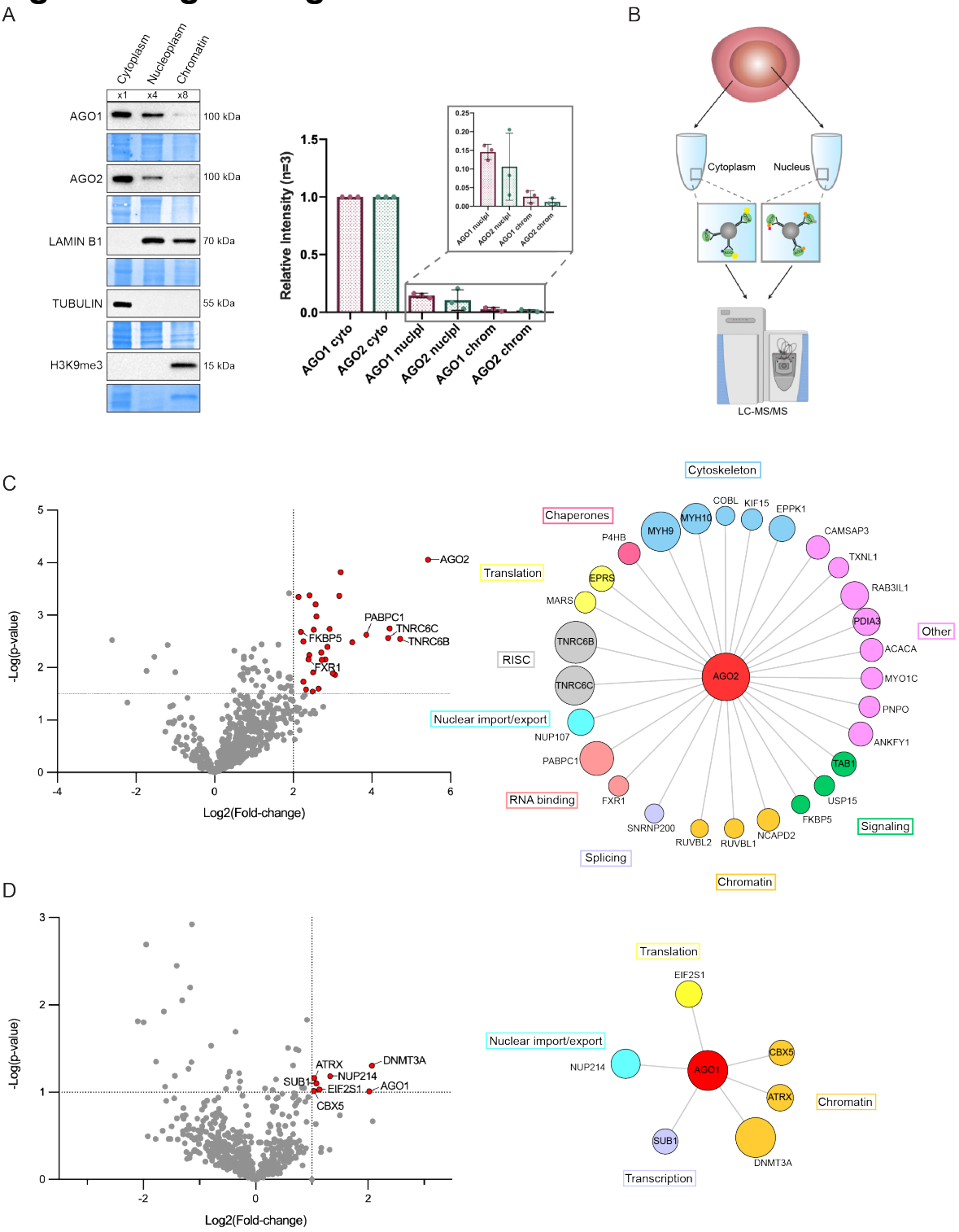

Figure 1. Interactome of AGO1 and AGO2 in different subcellular fractions.

(A) Representative Western blots for the fractionation of WT mESCs to visualize AGO1 and AGO2 subcellular localization and quantification of n=3 independent experiments. LAMIN B1 (nucleoplasm, chromatin), TUBULIN (cytoplasm), H3K9me3 (chromatin) were used as subcellular markers.

(B) Schematic representation of the experiment to identify AGO interactors. Partially created with bioRender.com

(C) Results of t-test used to analyze MS data from cytoplasmic AGO2 IP. Left: Volcano plot with significance thresholds, Log_2(Fold-change) >2, -Log(p-value)>1.5. Labelled are proteins that have been previously identified to interact with AGO2 (Table S1). Right: Cytoscape representation of significant proteins grouped by their functions. The size of the circles is proportional to the Log_2(Fold-change).
(D) Results of t-test used to analyze MS data from nuclear AGO1 IP. Left: Volcano plot with significance thresholds, Log_2(Fold-change) >1, -Log(p-value)>1. Right: Cytoscape representation of significant proteins grouped by their functions. The size of the circles is proportional to the Log_2(Fold-change).

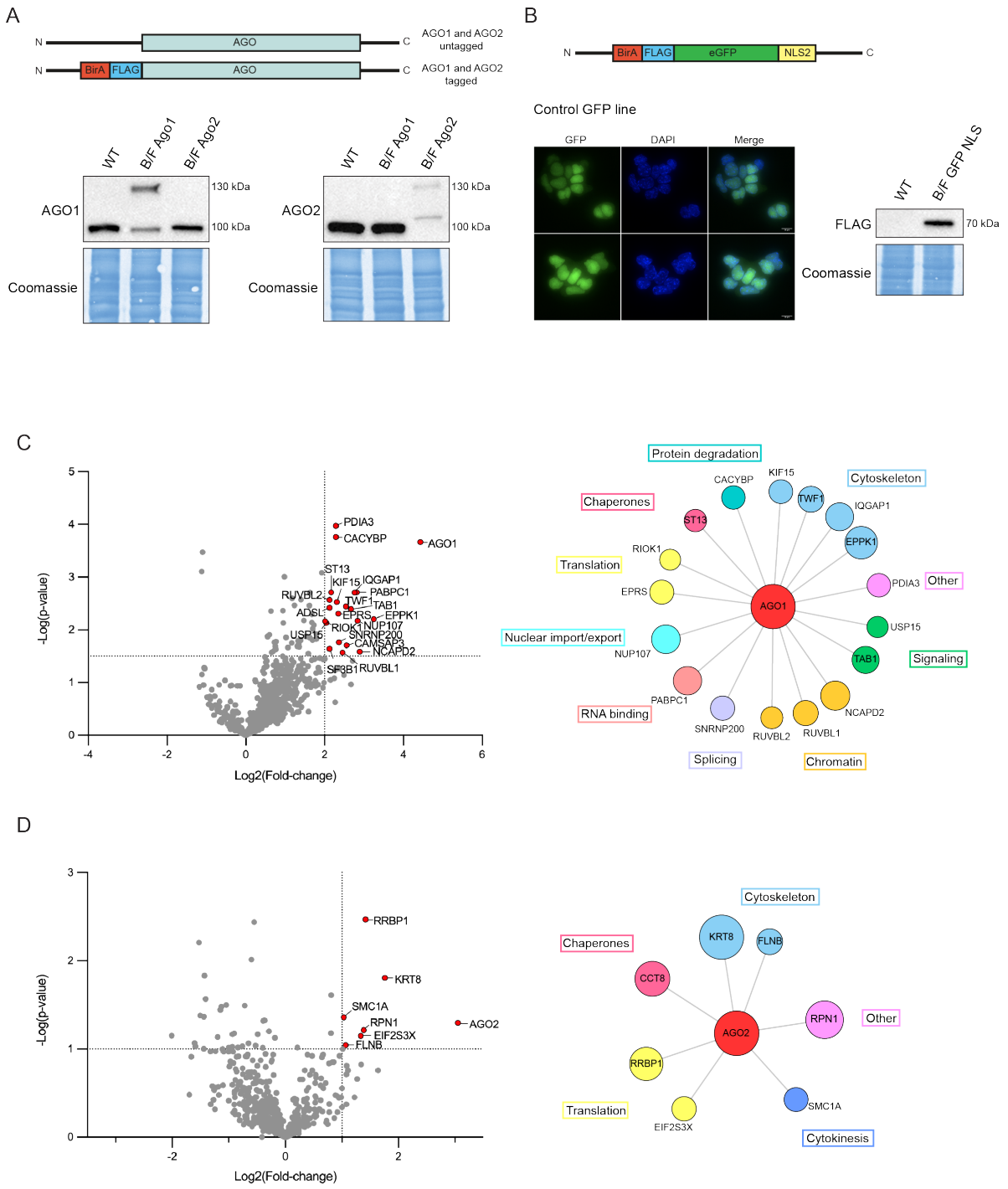

Figure S1. Interactome of AGO1 and AGO2 in different subcellular fractions. Related to Figure 1.

(A) Schematic representation of heterozygous endogenously tagged *Ago1* and *Ago2* loci and representative Western blots showing the tagged (130 kDa) and untagged (100 kDa) AGO1 and AGO2 proteins.

(B) Schematic representation of BirA-FLAG tagged GFP-NLS2 cell line and representative immunofluorescence images showing both nuclear and cytoplasmic localization of GFP, as well as a representative Western blot showing the tagged GFP (70 kDa).

(C) Results of t-test used to analyze MS data from cytoplasmic AGO1 IP. Left: Volcano plot with significance thresholds, Log_2(Fold-change) >2, -Log(p-value)>1.5. Right: Cytoscape representation of significant proteins grouped by their functions. The size of the circles is proportional to the Log_2(Fold-change).

(D) Results of t-test used to analyze MS data from nuclear AGO2 IP. Left: Volcano plot with significance thresholds, Log_2(Fold-change)>1, -Log(p-value)>1. Right: Cytoscape representation of significant proteins grouped by their functions. The size of the circles is proportional to the Log_2(Fold-change).

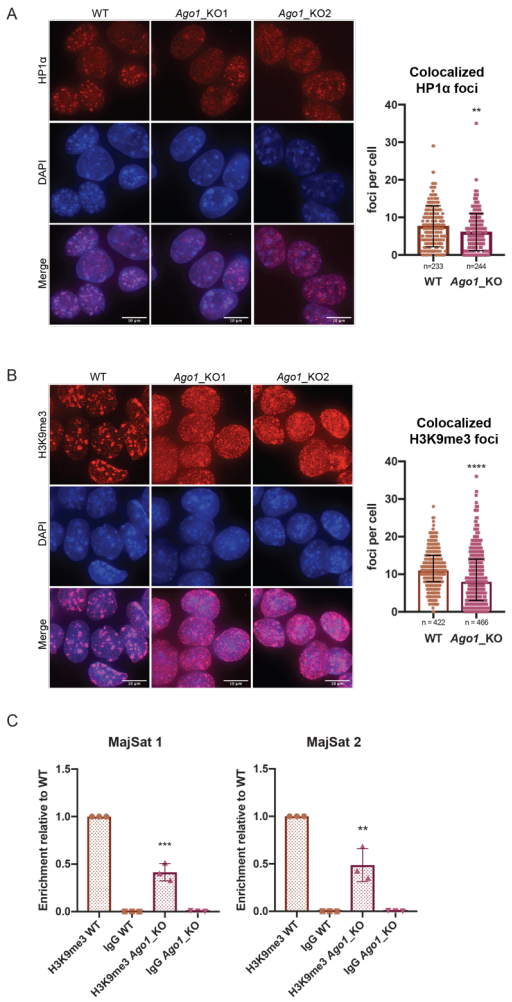

Figure 2. Distribution of HP1α and H3K9me3 at pericentromeric regions in WT versus *Ago1*_KO mESCs.

(A) Left: Representative IF images of HP1α in WT and *Ago1*_KO mESCs. scale bar = 10 μm. Right: Quantification of foci count for HP1α that colocalizes with DAPI regions in WT and *Ago1*_KO mESCs. The graph shows the mean distribution with the standard deviation. ** = pvalue < 0.01, unpaired t-test for n=3 independent experiments. For the quantification *Ago1*_KO1 and *Ago1*_KO2 were combined.

(B) Left: Representative IF images of H3K9me3 in WT and *Ago1*_KO mESCs. scale bar = 10 μm. Right: Quantification of foci count for H3K9me3 that colocalizes with DAPI regions in WT and *Ago1*_KO mESCs. Due to the bimodal distribution of H3K9me3 foci in *Ago1*_KO mESCs, the graph shows the median distribution with the interquartile range. **** = pvalue < 0.0001, Mann-Whitney test for n=3 independent experiments.

(C) ChIP-qPCR in WT and *Ago1*_KO mESCs. Pulldowns were performed with an antibody against H3K9me3 and a negative control IgG antibody. qPCR has been performed on major satellite primer set 1 and 2 (Table S2). The enrichment was calculated over input and represented relative to the WT H3K9me3 pulldown. *** = pvalue < 0.001 and ** = pvalue < 0.01, unpaired t-test for n=3 independent experiments. The *Ago1*_KO1 and *Ago1*_KO2 have been combined in this experiment.

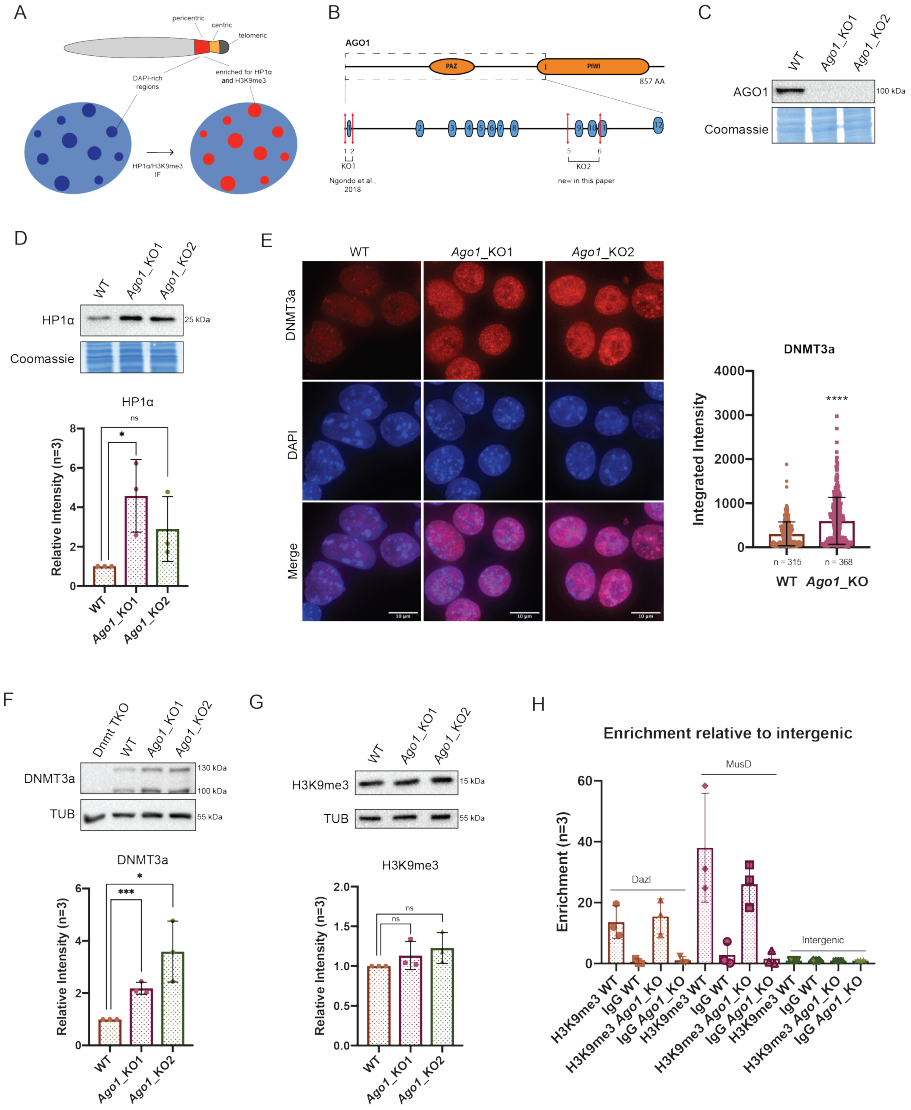

Figure S2. Assessment of nuclear AGO1 interactors and H3K9me3 in WT versus *Ago1*_KO mESCs. Related to Figure 2.

(A) Schematic representation of mouse pericentromeric regions. Shown are the DAPI-rich regions that are enriched for HP1α and H3K9me3 and visualized by IF.

(B) Schematic representation of the *Ago1*_KO paired CRISPR/Cas9 approach. Numbers 1, 2 and 5, 6 represent sgRNAs used to generate the KO lines. *Ago1*_KO1 has already been published by (Ngondo et al., 2018).

(C) Representative Western blots for AGO1 in WT, *Ago1*_KO1 and *Ago1*_KO2 mESCs out of n=3 independent experiments.

(D) Representative Western blots for HP1α in WT, *Ago1*_KO1 and *Ago1*_KO2 mESCs and quantification. * = pvalue < 0.05 and ns = non-significant, unpaired t-test for n=3 independent experiments.

(E) Left: Representative IF images of DNMT3a in WT and *Ago1*_KO mESCs. scale bar = 10 μm. Right: Quantification of DNMT3a integrated intensity measurement in WT, *Ago1*_KO mESCs. The graph is

showing the mean distribution with the standard deviations **** = pvalue < 0.0001, unpaired t-test for n=3 independent experiments. For the quantification *Ago1*_KO1 and *Ago1*_KO2 were combined.

(F) Representative Western blots for DNMT3a in WT, *Dnmt*_TKO (Tsumura et al., 2006), *Ago1*_KO1 and *Ago1*_KO2 mESCs and quantification. * = pvalue < 0.05 and *** = pvalue < 0.001, unpaired t-test for n=3 independent experiments.

(G) Representative Western blot for H3K9me3 in WT, *Ago1*_KO1 and *Ago1*_KO2 mESCs and quantification. ns = non-significant, unpaired t-test for n=3 independent experiments.

(H) ChIP-qPCR in WT and *Ago1*_KO mESCs. Pulldowns were performed with an antibody against H3K9me3 and a negative control IgG antibody. qPCR was performed on two positive control regions (Dazl and MusD (Karimi et al., 2011)) and a negative intergenic control region. The enrichment was calculated over input and represented relative to the intergenic region.

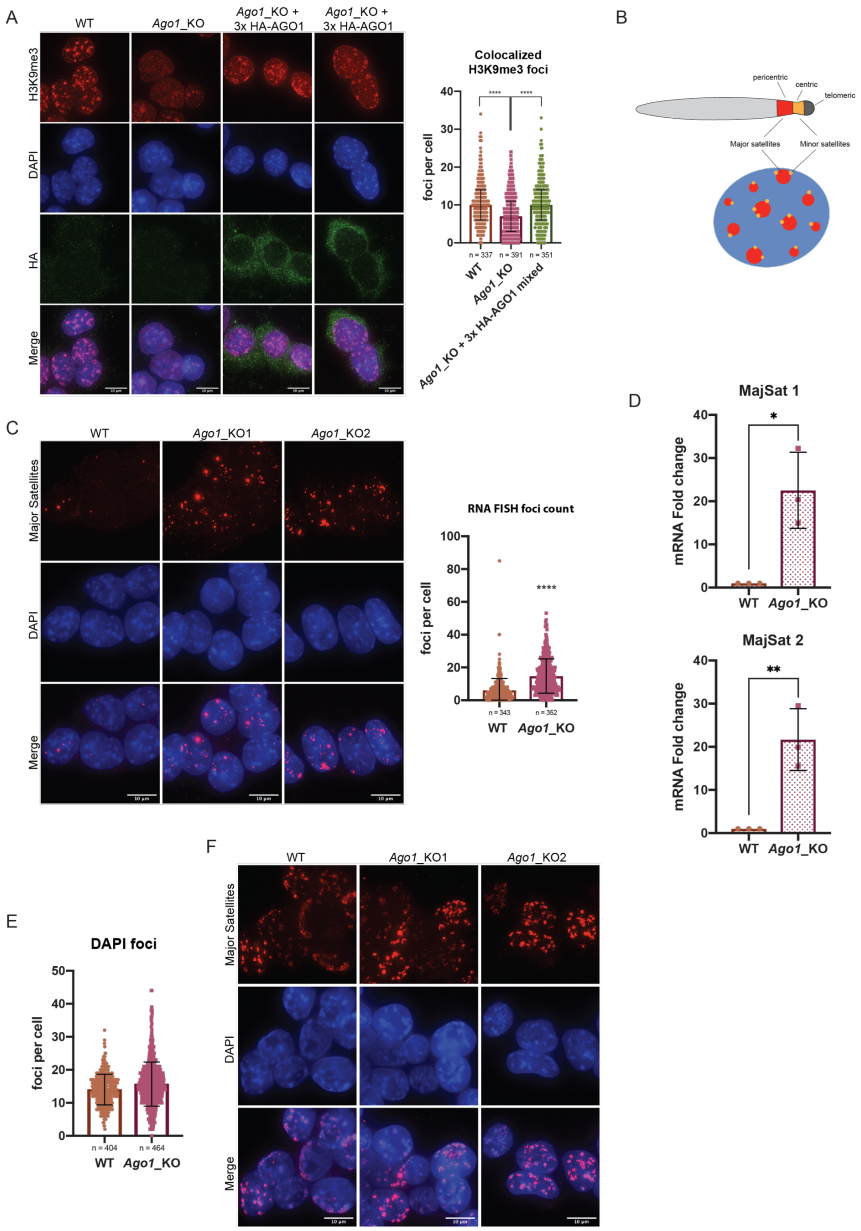

Figure 3. AGO1 rescue experiments and the assessment of major satellite transcripts in WT versus *Ago1*_KO mESCs.

(A) Left: Representative IF images of H3K9me3 in WT, *Ago1*_KO and two representative images of *Ago1*_KO + 3x HA-AGO1 mixed population mESCs. scale bar = 10 μm. Right: Quantification of foci count for H3K9me3 that colocalizes with DAPI regions in WT, *Ago1*_KO and *Ago1*_KO + 3x HA-AGO1 mixed population mESCs. The graph shows the median distribution with the interquartile range. **** = pvalue < 0.0001, Mann-Whitney test for n=3 independent experiments.

(B) Schematic representation of mouse pericentric and centric regions. Pericentric regions are associated with major satellite repeats. Centric regions are associated with minor satellite repeats.

(C) Left: Representative images of Major Satellite RNA FISH in WT and *Ago1*_KO mESCs. scale bar = 10 μm. Right: Quantification of Major Satellite RNA FISH foci count in WT and *Ago1*_KO mESCs. The

graph shows the mean distribution with standard deviations. **** = pvalue < 0.0001, unpaired t-test for n=3 independent experiments. For the quantification *Ago1*_KO1 and *Ago1*_KO2 were combined.

(D) RT-qPCR results for major satellite primer set 1 and 2 (Table S2) in WT and *Ago1*_KO mESCs. * = pvalue < 0.05 and ** = pvalue < 0.01, unpaired t-test for n=3 independent experiments.

(E) Quantification of the DAPI foci count in WT and *Ago1*_KO mESCs. The graph shows the mean distribution with standard deviations.

(F) Representative images of Major Satellite DNA FISH in WT and *Ago1*_KO mESCs. scale bar = 10 μm.

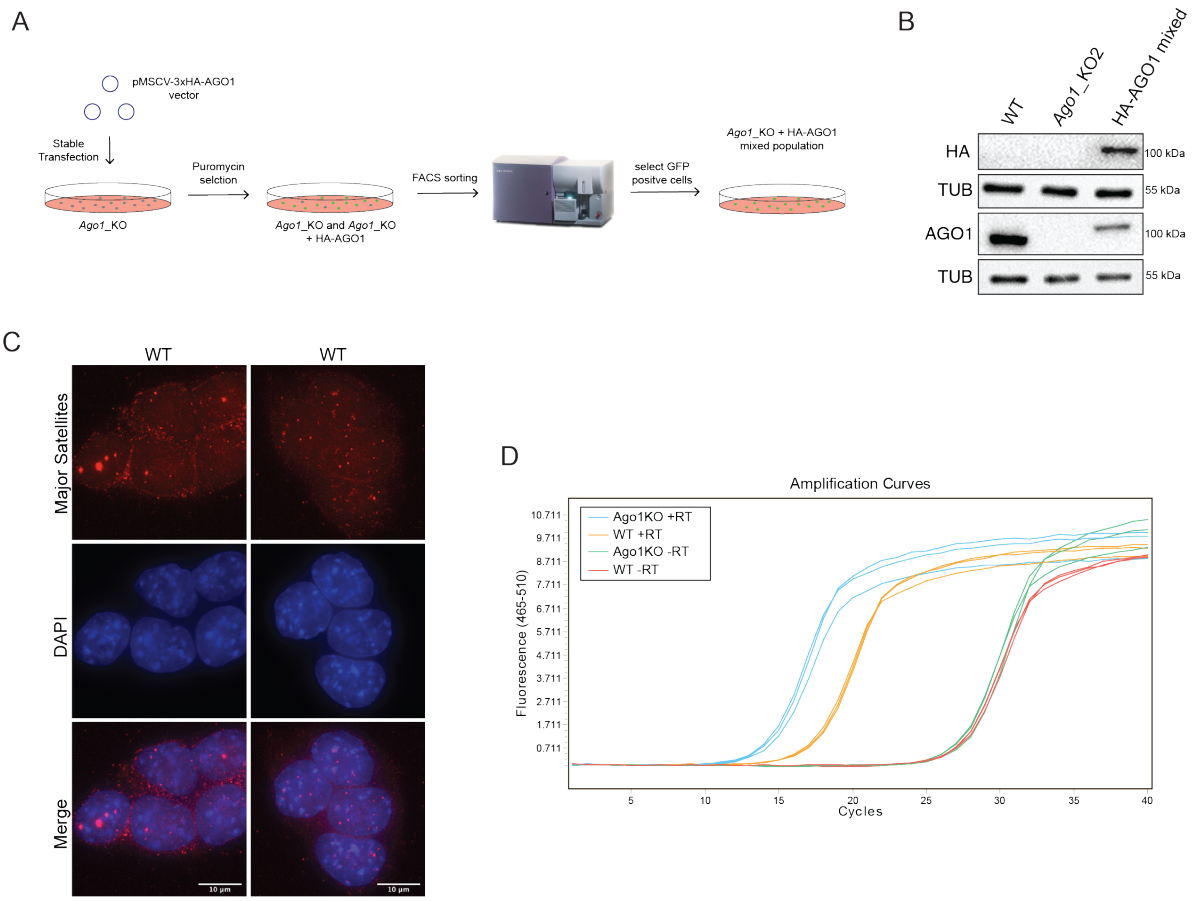

Figure S3. AGO1 rescue experiments and the assessment of major satellite transcripts in WT versus *Ago1*_KO mESCs. Related to Figure 3.

(A) Schematic representation of the experimental design to reintroduce AGO1 in *Ago1*_KO2 mESCs.

(B) Representative Western blots for HA and AGO1 in WT, *Ago1*_KO2 and *Ago1*_KO2 + 3x HA-AGO1 mixed population mESCs.

(C) Two representative images of Major Satellite RNA FISH in WT, showing that Major Satellite RNA is detected also in WT mESCs. scale bar = 10 μm.

(D) Amplification curves of WT and *Ago1*_KO qRT-PCR on Major Satellites in one replicate, showing that the – and + RT are well separated.

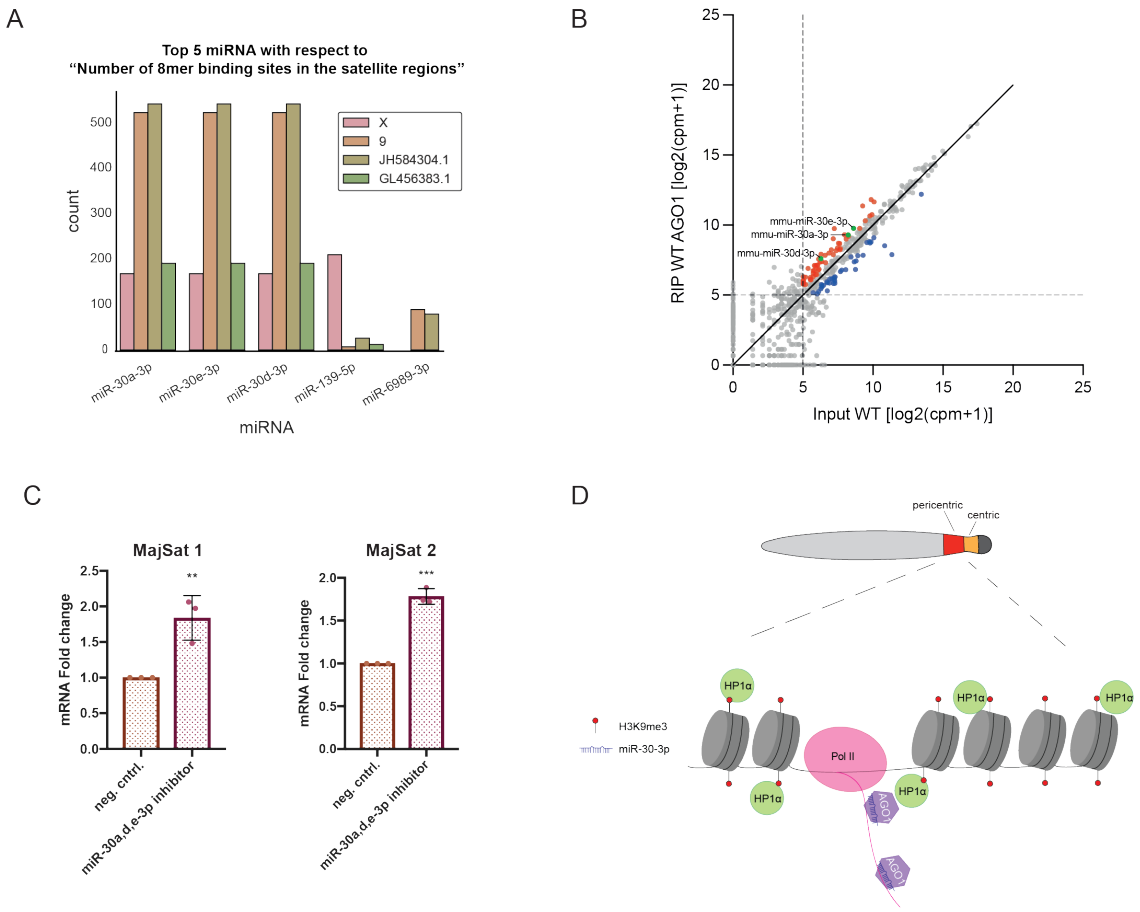

Figure 4. Identification of miRNAs targeting major satellite transcripts

(A) Representation of the top 5 microRNAs with 8mer binding sites targeting major satellite sequences located on chromosomes X, 9 and the genomic contigs JH584304.1 and GL456383.1.

(B) Scatterplot showing the log2(cpm+1) counts of miRNAs loaded in AGO1 versus miRNAs expressed in WT mESCs. The top 20% AGO1 loaded miRNAs versus Input with a minimal expression of log2(cpm+1) are marked in red and the lowest 20% of AGO1 loaded miRNAs versus Input with a minimal expression of log2(cpm+1) are marked in blue. Highlighted are the miR-30a, d, e-3p.

(C) qPCR results for major satellite primer sets 1 and 2 (Table S2) in WT mESCs transfected with a negative control inhibitor and a pool of inhibitors against miR-30a, d, e-3p. ** = pvalue < 0.01 and *** = pvalue < 0.001, unpaired t-test for n=3 independent experiments.

(D) Model showing the regulation of pericentromeric regions by AGO1 loaded with miR-30a, d, e-3p.

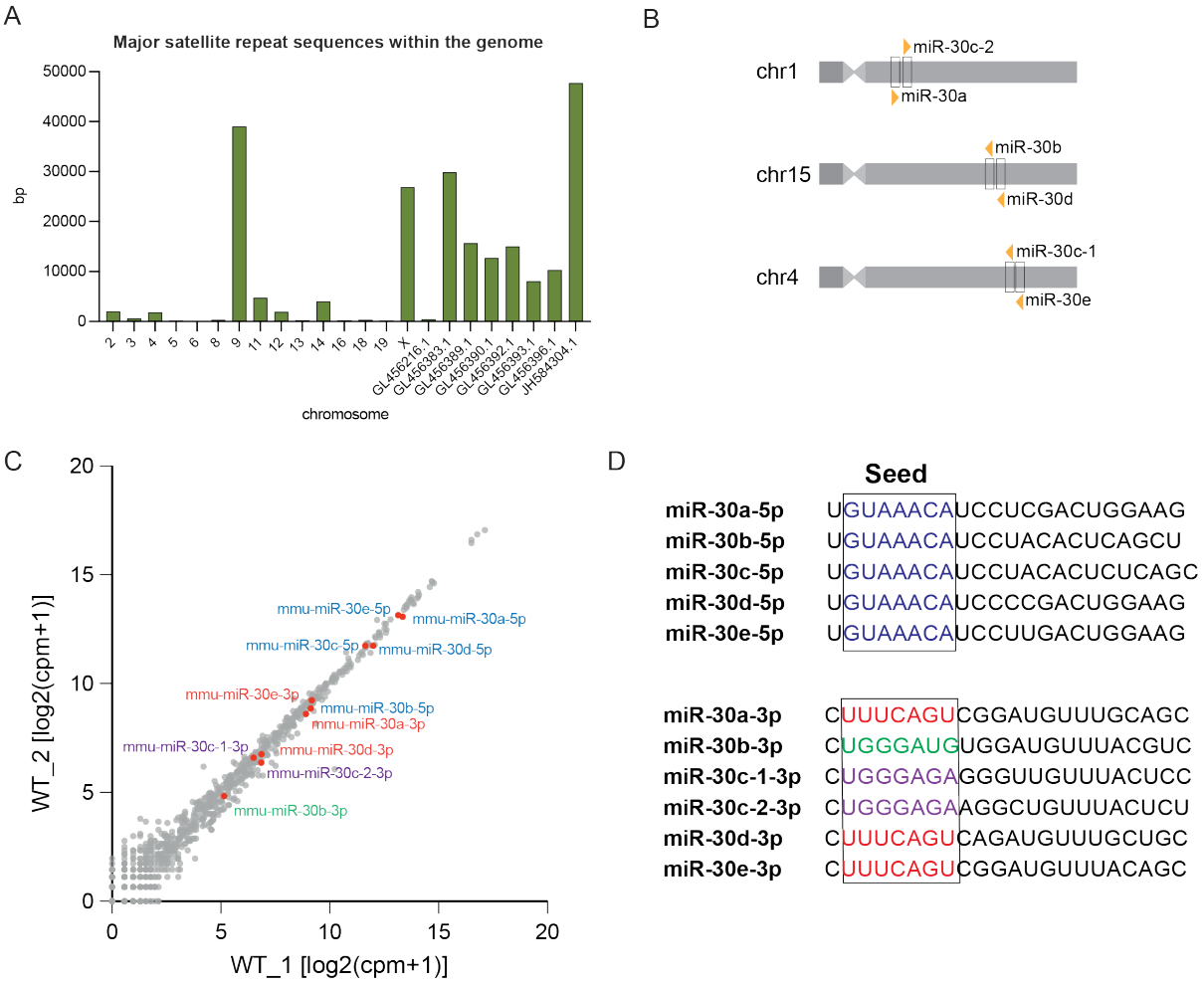

Figure S4. Identification of miRNAs targeting major satellite transcripts. Related to Figure 4.

(A) Summed length per chromosome or unmapped contig of annotated major satellite sequences within the mouse genome (mm10).

(B) Schematic representation of the miR-30 family in mouse.

(C) Scatterplot showing the log2(cpm+1) of expressed miRNAs in WT mESCs. Highlighted in red are the miR-30a, d, e-3p, in blue the miR-30a, b, c, d, e-5p, in violet miR-30c-1-3p and miR-30c-2-3p and in green miR-30b-3p

(D) Mature miRNA sequences of the miR-30-5p and miR-30-3p's. Highlighted in different colors are the different seed sequences.

ref of images:

FACS machine in Figure S3: **https://fluorofinder.com/cytometer-facsaria/**

Table S1 Mass spectrometry data. Related to Figure 1.

Table S2 List of all primers used in the study.

Table S3 Small RNA-seq CPM values related to Figure 4.