# Gene regulation in embryonic development

Berg, P.R. van den

# 4 RNA-SEQUENCING VALIDATES MICRORNA-GENE PREDICTIONS FROM TRANSLATION MODEL


## Abstract

Micro-RNAs (miRs) are post-transcriptional regulators of gene expression, which play important roles in development and cancer. Current methods to discover the effects of miRs on particular genes are either purely computational, and prone to produce many false positive, or cumbersome and low-throughput biochemical assays. In Chapter 3 we modeled the kinetics of protein. In this way we reduced the large number of sequence-based predictions to a few highly likely candidates that can be validated experimentally. Here, we validate the effect of six miRs on their predicted three target genes. We optimize assays for introducing miRs mimics or miR inhibitors in mouse embryonic stem cells using two fluorescent reporter cell lines that indicate the activity of specific miR . We performed RNA-sequencing on mouse emrbyonic stem cells (ESCs) transfected with mimics of six miRs and found in four cases that the predicted target gene is differentially expressed to a comparable extent as known targets of the respective miRs. These results corroborate the use of the kinetic model from Chapter 3 as a tool to identify novel miR-gene interactions.

## 4.1 Introduction

Mature miRs are single-stranded RNAs of circa 22 nucleotides that mediate RNA interference in eukaryotes. miRs typically pair to the 3' untranslated region (3'-UTR) of an mRNA and silence gene expression either through RNA-induced silencing complex (RISC)-mediated cleavage, destabilization of the poly-A tail, or blocking the translational machinery. In animals, the binding region in miRs is short and mostly not 100% complementary to the mRNA binding site. This, combined with the fact that there are around 24,000 predicted miRs and 140,000 predicted mRNA transcripts makes predicting miR-gene a difficult task. One popular database of predicted miR-gene interactions is TargetScan [1]. TargetScan reports a score which reflects the likelihood of miR binding, however neither its accuracy nor its precision is well characterized. One obvious reason an interaction might not exist is the lack of co-occurrence in cells. In Chapter 3 of this thesis we pre-selected dozens of mir-gene interactions using a very lenient TargetScan score threshold and whittled down this selection using our protein turnover rate model. By definition of our model, these miRs and genes are expressed simultaneously. In this chapter we will investigate three genes for which we predicted interactions: *Acad8*, *Cdk7* and *Eif4h*.

*Acad8* is part of the acyl-CoA dehydrogenase family of enzymes and is not known to play a significant role in embryonic development or differentiation. However, mutations in *Acad8* are known to cause the rare genetic disease Isobutyryl-CoA dehydrogenase deficiency. *Acad8*'s predicted regulator miR-433-3p, on the other hand, is known to target genes involved in development: the transcription factor *CREB* [2], the *WNT* regulator *DKK1* [3] and the *Egfr* binding adaptor protein *GRB2* [4]. The second predicted regulator of *Acad8* is miR-23b-5p, which is known for its role in cancer via targets like *EBF3* [5], *FOXC1* [6] and *Hmgb2* [7]. We decided not to investigate the third miR, miR-5615-3p, due to its low expression levels.

The second gene is the cell cycle gene *Cdk7* which is part of the cyclin-dependent protein kinase family. *Cdk7* is essential during the very early stages of development [8, 9]. The two miRs that are predicted to regulate *Cdk7* are mir-99a-3p and mir-100-3p, which are very similar in sequence. Both are known to regulate *MTOR*/*Mtor* [10, 11], a serine-threonine kinase that is essential for growth and proliferation [12]. Other targets of mir-99a-3p and mir-100-3p include *NOX4* [13], *CDC25A* [14] and *Hoxa1* [15], and *RASGRP3* [16] and *IGF1R* [17] respectively.

The third gene with candidate interactions is *Eif4h*, a translation initiation factor. This gene is part of the machinery that recruits ribosomes to mRNA. Williams-Beuren syndrome is a rare genetic defect that results from a deletion of *Eif4h* [18]. We predicted regulation of the miRs miR-152-3p and miR-467e-3p, where, to the best of our knowledge, the latter has no experimentally validated targets. Some of miR-152-3p's known targets are the two cell cycle genes *CDKN1B* [19] and *CDK8* [20], the pluripotency inducing *KLF4* [21], and the DNA methyltransferase *Dnmt1* [22].

In this chapter we will demonstrate an miR mimic and inhibitor assay, which we optimized using two fluorescent reporter cell lines for miR activity [23]. By RNA-sequencing of

mouse ESCs transfected with miR mimics we validate four out of six predicted miR-gene interactions.

## 4.2 Results

In Chapter 3 we identified *Acad8*, *Cdk7* and *Eif4h* to be translationally regulated by miRs Fig 1a. TargetScan predicts one or two binding sites per miR of which only one is conserved between species Fig 1b. Based on the cumulative weighted context++ score (CWC++S) for these genes alone, these interactions do not stand out among the hundreds of predicted interaction each of these genes have with the exception for *Acad8* with miR-433-3p [1].

In order to investigate the effect of these miRs on mouse ESCs we set up a transfection assay with miR mimics and inhibitors Fig 2A. miR mimics (Pre-miR miRNA Precursors, Thermo Fisher Scientific) are double stranded RNAs designed to be processed by the cell into mature miRs that are identical to a mature miR. Their small size facilitates transfection, making them a potent tool to simulate miR overexpression in the cell. In contrast, miR inhibitors (miRCURY LNA miRNA Power Inhibitor, Qiagen) block miRs by complementary binding to the mature miR. The modified LNA bases have a higher binding affinity compared to RNA and therefore inhibit and degrade endogenous miR effectively. In order to evaluate the effectiveness of the transfections we created two fluorescent reporter cell lines Fig 2B. These cell lines have bi-directional CAG promoters with highly correlated transcription of two fluorescent proteins: mCherry and citrine. The citrine transcript has additionally been cloned with miR binding sites at its 3'-end, resulting in reduced expression of citrine relative to mCherry if the respective miR is present. We created reporter cell lines for a miR that is undetected in our system (mir-590-3p) and one that is highly expressed (miR292a-5p) in order to evaluate the mimic and inhibitors respectively. Flow cytometry measurements of the mimic transfection revealed a high percentage of positively transfected and regulated cells after 24h Fig 2C. Although the effect increased slightly over time, we picked 24h as the ideal time point in order to limit the amount of secondary effects of the mimic Fig 2D. Transfection of the inhibitor was slightly less effective even at higher doses Fig 2E. For the miR inhibitor transfection we selected 48h transfection with 2X the suggested dose to be ideal Fig 2F.

Having set up effective dose and timings for our transfection assays, we next set out to validate the predicted miR targets. Although our kinetic model was set up to predict regulation at the level of translation, we used mRNA abundance as a readout, reasoning that mRNA levels are likely affected as well. After 24h of exposure to the miR mimics, cell samples were collected and purified mRNA was subjected to RNA-seq Fig 3A. Differential expression analysis revealed significant downregulation of all three predicted target genes by at least one of the proposed miRs Fig 3B. *Acad8* is downregulated by miR-433-3p (P=1.7e–3), *Cdk7* is downregulated by miR-99a-5p (P=0.014) and *Eif4h* is downregulated by both miR-152-3p (P=0.015) and miR-467e-3p (P=1.21e–19). The observed downregulation supported the hypothesis that miR-gene binding takes place in these four cases.

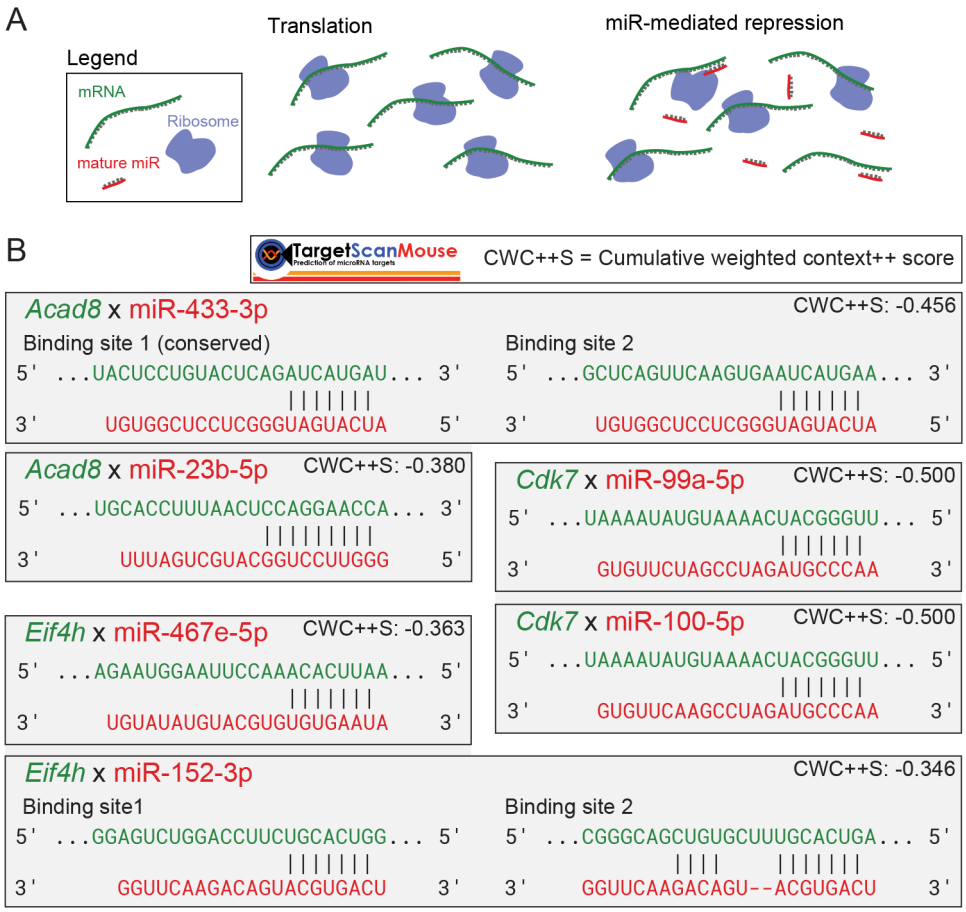



**Figure 1. Proposed miR-gene interactions from chapter 3 are not scored highly by TargetScan.** (A) Cartoon of miR-mediated translational regulation, miR-mediated mRNA degradation is not shown. (B) miR-gene interaction predictions from TargetScan 7.2 with their CWC++S [1]. Only one of eight predicted binding sites is preserved across species. Vertical lines indicate binding bases.

In order to assess the level of downregulation of our target genes, we compared these genes to some known targets in the context of the full transcriptome Fig 3C. The genes *Dkk1*, *Ebf3*, *Foxc1*, *Nox4*, *Hoxa1* and *Rasgrp3* were undetected in our system. Of the rest of the genes with known interactions most appear to be downregulated. Fixing the false discovery rate (FDR) at 0.1% only miR-100-5p–*Cdk7*, miR-152-3p–*Dnm1* and our own proposed miR–3p–*Eif4h*, are identified as interactions. Importantly, the lack of measured differential mRNA expression does not exclude binding of the mIRs and regulation of translation. These data show that our proposed significant miR-gene interactions, miR-433-3p–*Acad8*, miR-99a-5p–*Cdk7*, miR-152-3p–*Eif4h* and miR-467e-3p–*Eif4h*, lead to a similar extent of downregulation

as found in known targets.

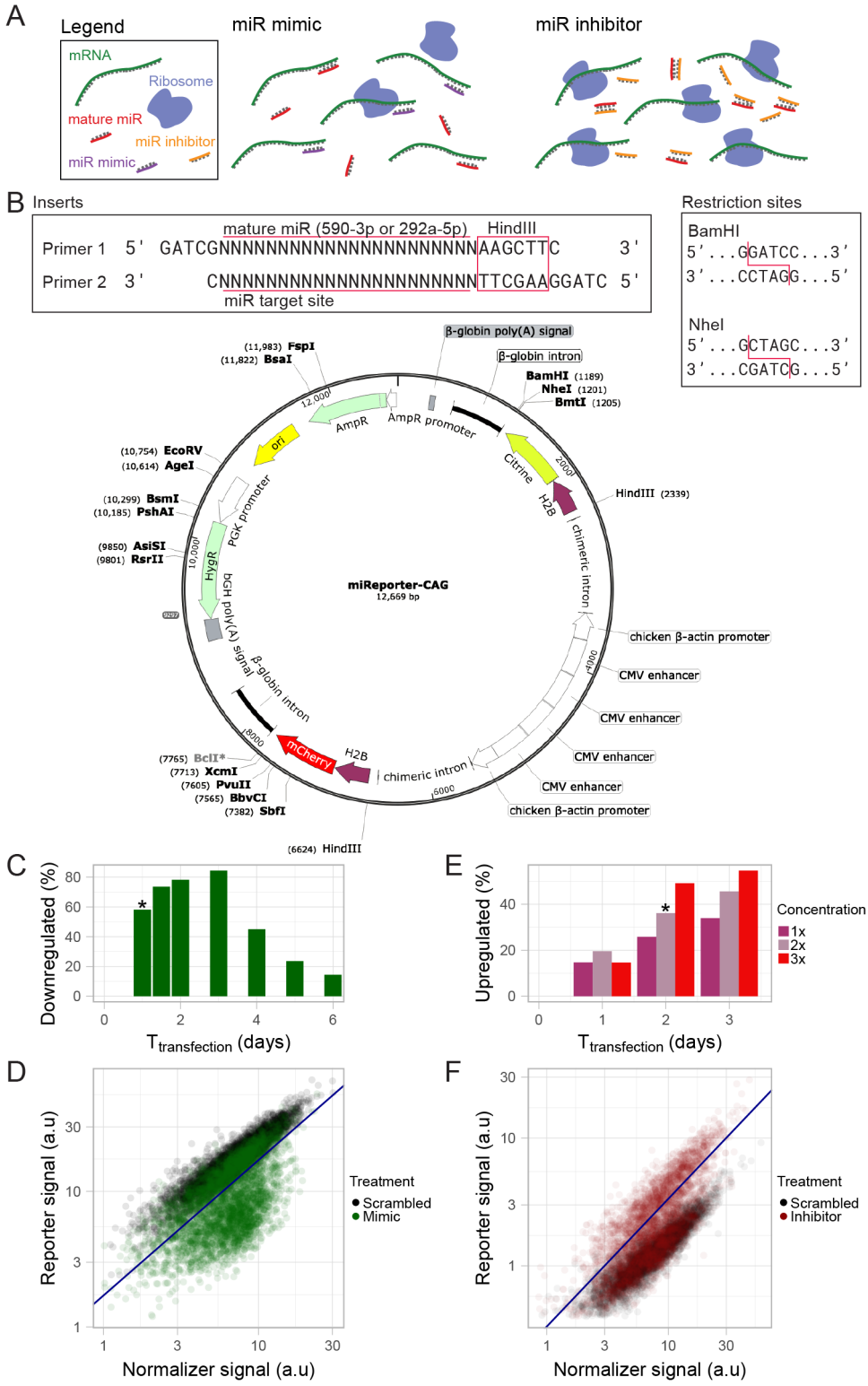
## 4.3 Discussion

Due to an enormous amount of possible combinations of miRs and genes it is a formidable challenge to identify true interactions. The kinetic model discussed in Chapter 3 whittles down potential interactions to a much smaller set of interactions that may have impact on translational regulation. We created reporter cell lines for activity of a lowly expressed and a highly expressed miR. Using these cell lines, we set up transfection assays for miR mimics and miR inhibitors that can enhance or annul a miRs effect on protein expression. We performed transcriptomic measurements of the mimic assay for six proposed miRs. These data showed that at least four out of six miRs bind and downregulate their predicted target mRNA, a surprisingly high fraction considering the search space. Moreover, *Eif4h* is, to the best of our knowledge, the first confirmed target for miR-467e-3p.

In this study we decided not to investigate the miRs using the dual luciferase assay. Although this is considered the gold standard of validating miR-gene interactions it only reports on the interaction between a specific miR-gene pair and does therefore not scale well to many possible pairs. By RNA-seq we can simultaneously observe the effects on the whole transcriptome. This can potentially reveal biological relevance of miRs in addition to providing controls in the form of known targets. On the other hand, some of the observed differential expression might be a secondary effect of direct miR targets. However, we start out with a clear hypothesis about the interaction from a very different source. In a follow up study we will complement the mimic experiments with the corresponding inhibitor assays. We hypothesize that quenching the miRs of miR-433-3p–*Acad8*, miR-99a-5p–*Cdk7*, miR-152-3p–*Eif4h* and miR-467e-3p–*Eif4h*, will result in higher expression of the respective mRNAs.

miR-433-3p is downregulated upon retinoic acid (RA) differentiation meaning any tar-

**Figure 2 *(following page)*. Dose and timing for miR mimic and inhibitors transfection experiments can be obtained using fluorescent reporters of miR activity.** (A) Cartoon of miR mimic and miR inhibitor translational regulation. (B) miReporter plasmid, inserts and digestion sites (BamHI and NheI). The insert overhangs are compatible with BamHI and NheI, but block redigestion. See Methods for full cloning strategy. (C) Inhibition of the miR-590-3p reporter transcript by the miR-590-3p mimic for seven time points as measured by flow cytometry. The asterisk indicates the optimal transfection timing shown in D (1d). (D) Fluorescence signal of miR-590-3p reporter for miR-590-3p mimic or scrambled control at optimal transfection conditions. Blue line indicates 1$^{st}$ percentile of reporter/normalizer ratio of the scrambled control. (E) Relief of inhibition on the miR-292a-5p reporter transcript by the miR-292a-5p inhibitor for three time points at three transfection concentrations as measured by flow cytometry. The asterisk indicates the optimal transfection timing shown in F (2days, 2X). (F) Fluorescence signal of miR-292a-5p reporter for miR-292a-5p inhibitor or scrambled control at optimal transfection conditions. Blue line indicates 99$^{th}$ percentile of reporter/normalizer ratio of the scrambled control.


A
Legend
mRNA
Ribosome
mature miR
miR inhibitor
miR mimic
miR mimic
miR inhibitor
B
Inserts
mature miR (590-3p or 292a-5p)
HindIII
Primer 1 5' GATCGNNNNNNNNNNNNNNNNNNNNNNNAAGCTTC 3'
Primer 2 3' CNNNNNNNNNNNNNNNNNNNNNNNTTCGAAGGATC 5'
miR target site
Restriction sites
BamHI
5'...GGATCC...3'
3'...CCTAGG...5'
NheI
5'...GCTAGC...3'
3'...CGATCG...5'
miReporter-CAG
12,669 bp
β-globin poly(A) signal
β-globin intron
BamHI (1189)
NheI (1201)
BmtI (1205)
HindIII (2339)
chicken β-actin promoter
CMV enhancer
CMV enhancer
CMV enhancer
CMV enhancer
chicken β-actin promoter
(6624) HindIII
(7765) BclI*
(7713) XcmI
(7605) PvuII
(7565) BbvCI
(7382) SbfI
mCherry
H2B
chimeric intron
β-globin intron
bGH poly(A) signal
HygR
PGK promoter
ori
AmpR
AmpR promoter
Citrine
H2B
chimeric intron
(11,983) FspI
(11,822) BsaI
(10,754) EcoRV
(10,614) AgeI
(10,299) BsmI
(10,185) PshAI
(9850) AsiSI
(9801) RsrII
C
Downregulated (%)
Ttransfection (days)
D
Reporter signal (a.u)
Normalizer signal (a.u)
Treatment
Scrambled
Mimic
E
Upregulated (%)
Ttransfection (days)
Concentration
1x
2x
3x
F
Reporter signal (a.u)
Normalizer signal (a.u)
Treatment
Scrambled
Inhibitor

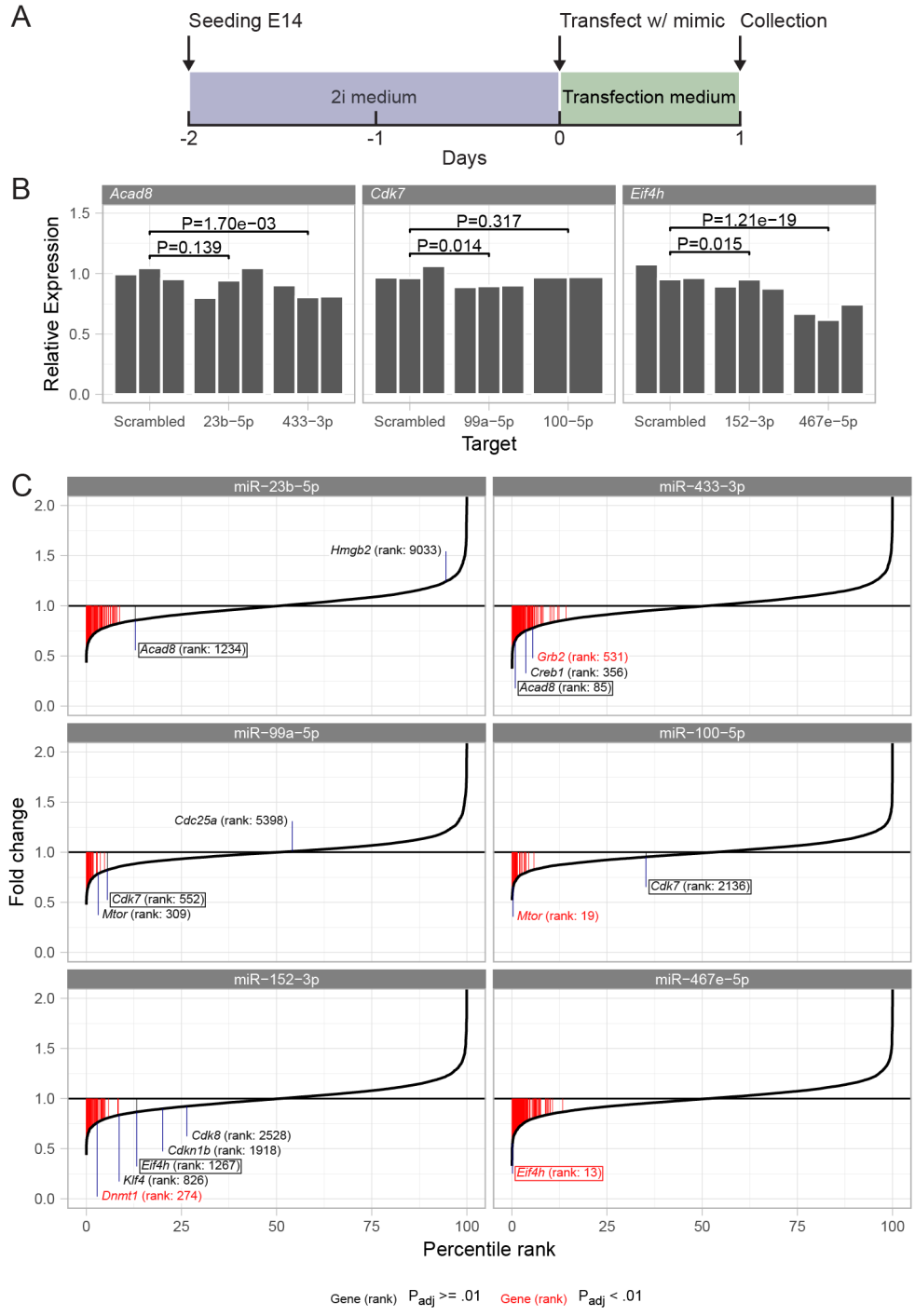
get's translational efficiency increases. Two of its known targets were present in our mouse ESCs, *Grb2* (downregulated FDR = 0.10%) and *Creb1* (downregulated FDR = 5.8%). *Grb2* is a known suppressor of *Nanog* so it may aid the exit of pluripotency [24]. Furthermore, *Creb* may be involved with the differentiation towards ectoderm or primitive endoderm, as it is involved in the differentiation towards many cell types like endothelial cells, brown adipocytes and osteoclasts [25, 26, 27]. Co-regulation of our discovered gene *Acad8* could suggest that miR-433-3p increases mitochondrial activity and simultaneously, reinforces the exit from pluripotency and differentiation.

In Chapter 3 we have observed that miR-99-5p expression increases in RA differentiation. Therefore we expect an increase in downregulation of *Ckd7* and *Mtor* as the cells differentiate. By targeting both of these genes miR-99-5p may simultaneously regulate the cell cycle, proliferation and cell growth. *Mtor*$^{-/-}$ mice embryos have shown to be embryonic lethal with a disorganized visceral endoderm at E5.5 [28]. Visceral endoderm is a derivative of primitive endoderm, which is one of the two cell types emerging in RA differentiation. Perhaps *Mtor*, and by extension *Ckd7*, are involved with the specification of the cell types induced by RA. *Ckd7*, despite being a cell cycle gene, is known to be involved in many differentiation processes.

The observation that the translational regulator *Eif4h* is itself translationally regulated demonstrates the complexity of translation regulation. One notable, known target of miR-152-3p is *Klf4*, which is part of the pluripotency gene regulatory network [29] and one of the four factors that were originally used to induce pluripotency in adult fibroblast cells [30]. In chapter 3 we observed that miR-152-3p, as well miR-467e-5p, increase over time around the exit from pluripotency during RA differentiation (Chapter 3 Fig S3). Perhaps miR-152-3p co-regulates the exit from pluripotency (by down-regulating *Klf4*, while simultaneously tweaking translation by downregulating *Eif4h*. Interestingly, according to our data, miR-467e-5p is more likely to downregulate *Klf4* expression than miR-152-3p (FDR = 1.4% vs 3.2%), but the validity of this interaction would require further investigation.

Using miRs the cell has many possible ways to subtly change its state. Unfortunately, many of the interactions between miRs and their target genes remain unconfirmed. The methods presented here, will serve as tools to discover bona fide interactions in a more scalable way than before. Future studies will have to reveal the functional relevance of these miRs for embryonic stem cell differentiation.

**Figure 3 *(following page)*. Four out of six miR mimic transfections downregulate their proposed mRNA targets.** (A) Schematic of the transfection protocol. (B) Expression levels (regularized counts scaled to scrambled control) of the proposed targets *Acad8*, *Cdk7* and *Eif4h* after miR mimic transfection and scrambled control. P-value shown is for an uncorrected one-sided test (see Methods). (C) Expression fold changes relative to scrambled control of six miR mimic transfections. Text annotations are known or proposed targets. The boxed genes are our proposed targets. Red lines indicate significantly downregulated genes. $P_{adj}$ = Benjamini-Hochberg corrected p-value.


A
Seeding E14
Transfect w/ mimic
Collection
2i medium
Transfection medium
-2
-1
0
1
Days
B
Acad8
Cdk7
Eif4h
Relative Expression
P=1.70e−03
P=0.139
P=0.317
P=0.014
P=1.21e−19
P=0.015
Scrambled
23b−5p
433−3p
99a−5p
100−5p
152−3p
467e−5p
Target
C
miR−23b−5p
miR−433−3p
miR−99a−5p
miR−100−5p
miR−152−3p
miR−467e−5p
Fold change
Hmgb2 (rank: 9033)
Acad8 (rank: 1234)
Grb2 (rank: 531)
Creb1 (rank: 356)
Acad8 (rank: 85)
Cdc25a (rank: 5398)
Cdk7 (rank: 552)
Mtor (rank: 309)
Cdk7 (rank: 2136)
Mtor (rank: 19)
Cdk8 (rank: 2528)
Cdkn1b (rank: 1918)
Eif4h (rank: 1267)
Klf4 (rank: 826)
Dnmt1 (rank: 274)
Eif4h (rank: 13)
Percentile rank
Gene (rank) Padj >= .01
Gene (rank) Padj < .01

## 4.4 Materials and Methods

### 4.4.1 Cell culture

E14 mouse embryonic stem cells were maintained as described in Chapter 3 of this thesis. During transfections, cells were temporarily cultured in serum+LIF medium (10% ES certified FBS, 1X non-essential amino acids, 0.1mM β-mercaptoethanol, 1X pen/strep, 2mM L-glutamine, 10,000U/ml mLIF, mLIF from Merck, rest from Thermo Fisher Scientific). Furthermore, miR reporter cell line clone selection took place on homegrown mouse embryonic fibroblast feeders.

### 4.4.2 Cloning

The miReporter backbone (AddGene, Plasmid #82478) was transformed into DH5a competent cells (Cat. 18265017, Thermo Fisher Scientific) as per manufacturer's instructions. Then, transformed cells were expanded and harvested for miReporter backbone by miniprep (Qiaprep, Qiagen). A set of two oligos were synthesized for each of the two reporter cell lines: miR-590-3p-fwd: 5'-GATCG TAATTTTATGTATAAGCTAGT AAGCTTC-3', miR-590-3p-rev: 5'-CTAGGAAGCTT ACTAGCTTATACATAAAATTA C-3', miR-292a-5p-fwd: 5'-GATCG ACTCAAACTGGGGGCTCTTTTG AAGCTTC-3', miR-292a-5p-rev: 5'-CTAGGAAGCTT CAAAAGAGCCCCCAGTTTGAGT C-3' (Integrated DNA Technologies, see Fig 2b). Pairs of oligos were annealed and phosphorylated in a thermocycler: 30m at 37°C, 5m at 95°C, for 12 cycles (1μM fwd oligo, 1μM rev oligo, 1X T4 buffer, 1U/μl T4 Polynucleotide Kinase; buffer and enzyme from New England Biolabs). Next, backbone digestion and ligation was performed in one step in a thermocycler, which was facilitated by the ligated inserts destroying the restriction sites for the enzymes (See Fig 2b): 5m at 37°C, 5m at 23°C, for 12 cycles (1:2500 dilution of phosphorylated oligo duplex, 2.5ng/μl backbone, 5% v/v DTT, 0.15U/μl BamHI, 0.5U/μl NheI, 1U/μl T4 ligase, 1X restriction buffer; T4 from New England Biolabs, rest from Thermo Fisher Scientific). Plasmids were then amplified in the same manner as the backbone: Plasmids were transformed into DH5a cells, which were then expanded and used for midiprep extraction (Plasmid Midi, Qiagen).

### 4.4.3 miReporter cell lines creation

miReporter-590-3p and miReporter-292a-5p plasmids were transfected into ESCs with lipofectamine 3000 (Thermo Fisher Scientific) as per manufacturer's instruction. Briefly, 125μl DMEM (Sigma-Aldrich) was mixed with 5μl lipofectamine 3000 and briefly vortexed. Separately 125μl was mixed with 5μl p3000 reagent and 5μg of plasmid and also briefly vortexed. Both mixtures were combined and incubated at room temperature for 5 minutes to create DNA-lipid complexes. 2i medium was removed from pre-seeded ESCs at a confluency of about 70-90% and replaced with serum+LIF medium. DNA-lipid complexes were added to the medium for 24 h. Medium was then aspirated, cells washed twice with PBS, and cells were

left to grow for two days in 2i. Transfected cells were selected for by hygromycin (100 μg/ml in 2i) for three days. Single clones were selected differently for the two miReporter cell lines. Double-positive, single cells of the miReporter-590-3p cell line were sorted by Fluorescence-activated cell sorting in 96-well feeder-coated plates and expanded thereafter. miReporter-292a-5p cells however, were strongly diluted upon passage and single colonies with double-positive cells were picked by hand using a benchtop microscope and a 200μl pipette. Double positive colonies were left to expand on feeders in 48-well plates. Clones were grown for at least two passages to ascertain the stability of the transfection. The reporter activity was confirmed by flow cytometry.

### 4.4.4 Mimic and inhibitor transfection

ESCs and both miReporter cell lines were transfected with miR mimic and inhibitors in an identical fashion. Cells were seeded 48 h prior to transfection in 12-well plates. DNA-lipid complexes (Lipofectamine RNAiMax, Thermo Fisher Scientific) were prepared at the ratios recommended by the manufacturer but siRNA was replaced with either miR mimic or miR inhibitor (see Table 4.1). We considered 1X concentration of mimic/inhibitor to be 100nM in the DNA-lipid mixture. We used 37.5μl of DNA-lipid mixture per well. After 24h transfection cells were either harvested or washed with PBS and left to grow in 2i.

| miR-target | Our name | Company name | Company | Catalog number |
|---|---|---|---|---|
| None | mimic scrambled | Pre-miR™miRNA Precursor Negative Control #1 | Thermo Fisher Scientific | AM17110 |
| miR-590-3p | miR-590-3p mimic | Pre-miR miRNA Precursor | Thermo Fisher Scientific | AM17100-PM20392 |
| miR-23b-5p | miR-23b-5p mimic | Pre-miR miRNA Precursor | Thermo Fisher Scientific | AM17100-PM15655 |
| miR-433-3p | miR-433-3p mimic | Pre-miR miRNA Precursor | Thermo Fisher Scientific | AM17100-PM10774 |
| miR-99a-5p | miR-99a-5p mimic | Pre-miR miRNA Precursor | Thermo Fisher Scientific | AM17100-PM10719 |
| miR-100-5p | miR-100-5p mimic | Pre-miR miRNA Precursor | Thermo Fisher Scientific | AM17100-PM10188 |
| miR-152-3p | miR-152-3p mimic | Pre-miR miRNA Precursor | Thermo Fisher Scientific | AM17100-PM12269 |
| miR-467e-5p | miR-467e-5p mimic | Pre-miR miRNA Precursor | Thermo Fisher Scientific | AM17100-PM12611 |
| None | scrambled inhibitor | miRCURY LNA miRNA Inhibitor Control: Negative control A | Qiagen | YI00199006-DDA |

*Continued on next page*

*Continued from previous page*

| miR-target | Our name | Company name | Company | Catalog number |
|---|---|---|---|---|
| miR-292a-5p | miR-292a-5p inhibitor | miRCURY LNA miRNA Power Inhibitors | Qiagen | YI04101165-DDA |

**Table 4.1. Overview of miR mimic and inhibitors.**

### 4.4.5 Flow cytometry

Transfected cells were harvested for flow cytometry by washing with PBS and detachment using Accutase (Sigma-Aldrich). Detached cells were washed and resuspended in 2i. Cells were fixed in 4%Formaldehyde in medium (Cat. 43368, Alfa Aesar) for 15 min at room temperature. Cells were then centrifuged and the supernatant was removed. Cells were resuspended in 1% BSA (Cat. A2153, Sigma-Aldrich) in PBS and stored at 4°C until the measurement.

Fixed cells were measured on a BD LSRFortessa X-20. Forward and side scatter was measured as well as Citrine fluorescence (488nm laser, 530/30nm emission filter) and mCherry fluorescence (561nm laser, 610/20 emission filter). Live cell gating and Citrine/mCherry positive selection was achieved using custom R scripts (FlowCore v1.52.1 [31]). To determine relative down- or upregulation of citrine expression we calculated the ratio mCherry/citrine for each cell in the mimic/inhibitor assays and scrambled controls. Mimic cells with a ratio lower than the $1^{st}$ percentile of scrambled control ratios were deemed transfected with successful citrine inhibition. Inhibitor cells with a ratio higher than the $99^{th}$ percentile of scrambled control ratios were deemed transfected with successful miR inhibition.

### 4.4.6 RNA-sequencing

RNA-sequencing was performed as described in Chapter 3 of this thesis with a minimum of 10 million raw reads per sample.

### 4.4.7 RNA-sequencing analysis

RNA-sequencing data was preprocessed as described in Chapter 3 of this thesis. Sample 3 of the miR-100p-5p mimic was excluded due to low complexity. Only genes with expression counts larger than 20 were kept and were subsequently processed using DEseq2 (v1.26.0, [32]). Regularized counts are defined as the 2-base exponent of the *rlog*-values. DESeq2 was also used to identify differentially expressed genes and obtain $\log_2$ fold-changes. 14 genes were differentially expressed between the scrambled control and no treatment control: *Tjp2*, *Serp1*, *Rap1b*, *Cdk2ap1*, *Ssr2*, *B230219d22rik*, *Cyld*, *Fam98b*, *Tpm3*, *Uggt1*, *Snx6*, *Rfc4*, *Tcf7l1* and *Adam10*. This list of genes was excluded in other comparisons.

## Acknowledgements

Stefan Semrau and Patrick van den Berg conceived the project. S.S. acquired funding. P.v.d.B., Noémie Bérenger-Currias and Marleen Feliksik performed the experiments. P.v.d.B analyzed, interpreted and modeled the data. P.v.d.B. and S.S. wrote the manuscript.

# Acronyms

**CWC++S** cumulative weighted context++ score

**ESC** emrbyonic stem cell

**FDR** false discovery rate

**miR** micro-RNA

**RA** retinoic acid

**RISC** RNA-induced silencing complex

## 4.5 References

[1] Vikram Agarwal et al. “Predicting effective microRNA target sites in mammalian mRNAs”. In: *eLife* 4 (2015). DOI: 10.7554/eLife.05005.

[2] Shupeng Sun et al. “MiR-433-3p suppresses cell growth and enhances chemosensitivity by targeting CREB in human glioma”. In: *Oncotarget* 8.3 (2016), pp. 5057–5068. DOI: 10.18632/oncotarget.13789.

[3] Xiaolin Tang et al. “MicroRNA-433-3p promotes osteoblast differentiation through targeting DKK1 expression”. In: *PLOS ONE* 12.6 (2017), e0179860. DOI: 10.1371/journal.pone.0179860.

[4] Qizhong Shi et al. “MiR-433-3p Inhibits Proliferation and Invasion of Esophageal Squamous Cell Carcinoma by Targeting GRB2.” In: *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* 46.5 (2018), pp. 2187–2196. DOI: 10.1159/000489548.

[5] Jing Zhang et al. “MiR-23b-3p induces the proliferation and metastasis of esophageal squamous cell carcinomas cells through the inhibition of EBF3”. In: *Acta Biochimica et Biophysica Sinica* 50.6 (2018), pp. 605–614. DOI: 10.1093/abbs/gmy049.

[6] Guo dong Hu et al. “Long noncoding RNA CCAT2 functions as a competitive endogenous RNA to regulate FOXC1 expression by sponging miR-23b-5p in lung adenocarcinoma”. In: *Journal of Cellular Biochemistry* 120.5 (2018), pp. 7998–8007. DOI: 10.1002/jcb.28077.

[7] Diafara Boureima Oumarou et al. “Involvement of microRNA-23b-5p in the promotion of cardiac hypertrophy and dysfunction via the HMGB2 signaling pathway”. In: *Biomedicine & Pharmacotherapy* 116 (2019), p. 108977. DOI: 10.1016/j.biopha.2019.108977.

[8] Miguel Ganuza et al. “Genetic inactivation of Cdk7 leads to cell cycle arrest and induces premature aging due to adult stem cell exhaustion”. In: *The EMBO Journal* 31.11 (2012), pp. 2498–2510. DOI: 10.1038/emboj.2012.94.

[9] Shetal A Patel et al. “Functional analysis of the Cdk7.cyclin H.Mat1 complex in mouse embryonic stem cells and embryos.” In: *Journal of Biological Chemistry* 285.20 (2010), pp. 15587–15598. DOI: 10.1074/jbc.M109.081687.

[10] Junhong Cai et al. “MiR-100-5p, miR-199a-3p and miR-199b-5p induce autophagic death of endometrial carcinoma cell through targeting mTOR.” In: *International Journal of Clinical and Experimental Pathology* 10.9 (2017), pp. 9262–9272.

[11] Xiaoyang Ye et al. “MicroRNAs 99b-5p/100-5p Regulated by Endoplasmic Reticulum Stress are Involved in Abeta-Induced Pathologies”. In: *Frontiers in Aging Neuroscience* 7 (2015), p. 210. DOI: 10.3389/fnagi.2015.00210.

[12] Mirei Murakami et al. “mTOR Is Essential for Growth and Proliferation in Early Mouse Embryos and Embryonic Stem Cells”. In: *Molecular and Cellular Biology* 24.15 (2004), pp. 6710–6718. DOI: 10.1128/MCB.24.15.6710-6718.2004.

[13] Y Shi et al. “MiR-99a-5p regulates proliferation, migration and invasion abilities of human oral carcinoma cells by targeting NOX4”. In: *Neoplasma* 64.05 (2017), pp. 666–673. DOI: 10.4149/neo_2017_503.

[14] Hongzhen Qin et al. “MicroRNA-99a-5p suppresses breast cancer progression and cell-cycle pathway through downregulating CDC25A”. In: *Journal of Cellular Physiology* 234.4 (2019), pp. 3526–3537. DOI: 10.1002/jcp.26906.

[15] Faman Xiao et al. “Downregulation of HOXA1 gene affects small cell lung cancer cell survival and chemoresistance under the regulation of miR-100”. In: *European Journal of Cancer* 50.8 (2014), pp. 1541–1554. DOI: 10.1016/j.ejca.2014.01.024.

[16] Qian Peng et al. “FOXA1 Suppresses the Growth, Migration, and Invasion of Nasopharyngeal Carcinoma Cells through Repressing miR-100-5p and miR-125b-5p”. In: *Journal of Cancer* 11.9 (2020), pp. 2485–2495. DOI: 10.7150/jca.40709.

[17] Hongliang Zhang et al. “miR-100-5p Inhibits Malignant Behavior of Chordoma Cells by Targeting IGF1R”. In: *Cancer Management and Research* 12 (2020), pp. 4129–4137. DOI: 10.2147/CMAR.S252185.

[18] Rossella De Cegli et al. “A transcriptomic study of Williams-Beuren syndrome associated genes in mouse embryonic stem cells”. In: *Scientific Data* 6.1 (2019), pp. 1–5. DOI: 10.1038/s41597-019-0281-5.

[19] L Wang et al. “MiR-152-3p promotes the development of chronic myeloid leukemia by inhibiting p27.” In: *European review for medical and pharmacological sciences* 22.24 (2018), pp. 8789–8796. DOI: 10.26355/eurrev_201812_16646.

[20] Tao Yin et al. “miR-152-3p Modulates hepatic carcinogenesis by targeting cyclin-dependent kinase 8”. In: *Pathology - Research and Practice* 215.6 (2019), p. 152406. DOI: 10.1016/j.prp.2019.03.034.

[21] Feng Feng et al. “miR-148-3p and miR-152-3p synergistically regulate prostate cancer progression via repressing KLF4”. In: *Journal of Cellular Biochemistry* 120.10 (2019), pp. 17228–17239. DOI: 10.1002/jcb.28984.

[22] Jin Sun et al. “Regulation of human glioma cell apoptosis and invasion by miR-152-3p through targeting DNMT1 and regulating NF2”. In: *Journal of Experimental & Clinical Cancer Research* 36.1 (2017), p. 1061. DOI: 10.1186/s13046-017-0567-4.

[23] Hanna L Sladitschek et al. “Bidirectional Promoter Engineering for Single Cell MicroRNA Sensors in Embryonic Stem Cells”. In: *PLOS ONE* 11.5 (2016), e0155177. DOI: 10.1371/journal.pone.0155177.

[24] Takashi Hamazaki et al. “The Grb2/Mek Pathway Represses Nanog in Murine Embryonic Stem Cells”. In: *Molecular and Cellular Biology* 26.20 (2006), pp. 7539–7549. DOI: 10.1128/MCB.00508-06.

[25] Kohei Yamamizu et al. “PKA/CREB Signaling Triggers Initiation of Endothelial and Hematopoietic Cell Differentiation via Etv2 Induction”. In: *STEM CELLS* 30.4 (2012), pp. 687–696. DOI: 10.1002/stem.1041.

[26] Aaron M Cypess et al. “Insulin/IGF-I Regulation of Necdin and Brown Adipocyte Differentiation Via CREB- and FoxO1-Associated Pathways”. In: *Endocrinology* 152.10 (2011), pp. 3680–3689. DOI: 10.1210/en.2011-1229.

[27] Kojiro Sato et al. “Regulation of osteoclast differentiation and function by the CaMK-CREB pathway”. In: *Nature Medicine* 12.12 (2006), pp. 1410–1416. DOI: 10.1038/nm1515.

[28] Yann-Gaël Gangloff et al. “Disruption of the Mouse mTOR Gene Leads to Early Postimplantation Lethality and Prohibits Embryonic Stem Cell Development”. In: *Molecular and Cellular Biology* 24.21 (2004), pp. 9508–9516. DOI: 10.1128/MCB.24.21.9508-9516.2004.

[29] Mo Li et al. “Deconstructing the pluripotency gene regulatory network”. In: *Nature Cell Biology* 20.4 (2018), pp. 382–392. DOI: 10.1038/s41556-018-0067-6.

[30] Kazutoshi Takahashi et al. “Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors”. In: *CELL* 126.4 (2006), pp. 663–676. DOI: 10.1016/j.cell.2006.07.024.

[31] B Ellis et al. *flowCore: flowCore: Basic structures for flow cytometry data*. 2019.

[32] Michael I. Love et al. “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2”. In: *Genome Biology* 15.12 (2014), p. 31. DOI: 10.1186/s13059-014-0550-8.