Translation on Demand as a post-transcriptional regulation mechanism of ES cell differentiation

Master thesis for acquisition of the degree of Master of Science (M.Sc.) in Computational Biology of the faculty of mathematics and natural sciences at the University of Cologne

submitted by Elias Schwall

July 5, 2025

Reviewers:

Prof. Dr. Andreas Beyer

Prof. Dr. Achim Tresch

Supervised by:

Luke Brandwood

Declaration of originality

Hiermit versichere ich an Eides Statt, dass ich die vorliegende Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

I hereby declare that this thesis is my own work and that I have not used any sources other than those listed in the bibliography. Content from published or unpublished worksthat has been quoted directly or indirectly or paraphrased isindicated as such. The thesis has not been submitted in the same or similar form or in part for any other academic award. The electronic version I have submitted is completely identical to the hard copy version submitted.

Cologne, July 5, 2025 Elias Schwall

Acknowledgement

This thesis would not have been possible without the support and encouragement of numerous individuals. I would like to express my deepest gratitude to reviewers Prof. Dr. Andreas Beyer and Prof. Dr. Achim Tresch for granting me the opportunity to undertake my master's thesis at the Exzellenzcluster for Cellular Stress Responses in Aging-Associated Diseases (CECAD). Prof. Dr. Beyer consistently provided exceptional insights into grappling with scientific questions as well as questioning scientific assumptions and offered invaluable feedback. He always was avaible for addressing my problems.

I would also like to extend my heartfelt gratitude to Luke Brandwood, who supervised my project. He introduced me to the field of stem cell research and Translation on Demand by teaching me day after day during my time at the AG Beyer. Every day I could discuss new insights and strategies with him, and his mentoring and feedback were paramount for progressing in this project. I was extremely grateful for having had the chance to support him in his PhD project. My time at the AG Beyer was extremely fun not only because of learning new things everyday but also because of the people working there. Therefore, I would also like to thank all other members of AG Beyer.

I would like to offer a special note of thanks to my friends and mentors from Aimed Analytics, Dr. Kevin Baßler and Dr. Patrick Günther. They were very considerate during the time of my master's thesis and always offered me support and feedback.

The experience of my master's study program was significantly enriched by my friend and fellow student Amir Hyseni. Working alongside him towards the same goal made it all the more fun.

Lastly, I would be remiss if I didn't acknowledge the enormous support from my family, which was indispensable for the completion of this thesis. I extend my sincerest thanks to my father Robert and my mother Maria-Theresia, who deserve special recognition for supporting in every way and always being my biggest cheerleaders. My older brother Johannes, whose presence and readiness to help have been a constant in my life, and my partner Patricia, who constantly supported and encouraged me throughout this intensive period of work.

Table of contents

1 Abstract

2 Introduction

2.1 Mouse embryonic stem cells

Mouse embryonic stem cells (mESCs) are defined by their self-renewal capabilities and pluripotency, allowing them to differentiate into any cell type of the three germ layers. While this differentiation process is limited in vivo, pluripotent stem cells (PSCs) cultivated in vitro can self-renew indefinitely under specific culture conditions (Nichols and Smith 2009). One such condition is the 2i+LIF (two inhibitors with leukemia inhibitory factor) medium, which maintains the naive pluripotent state of ESCs, closely mimicking the inner cell mass of pre-implantation mouse embryos.

The continuum of pluripotency is described by a progression from the naive state to a primed state, characterized by the priming for differentiation (Nichols and Smith 2009). Naive mESCs serve as a model for the inner cell mass of the pre-implantation embryo, exhibiting minimal lineage commitment. In contrast, epiblast stem cells (EpiSCs), which are isolated from later developmental stages, correspond to the post-implantation epiblast and are primed for differentiation (Boroviak et al. 2014).

During the transition from the naive to the primed states, sequential changes occur at multiple molecular levels, encompassing alterations in signaling pathways, epigenetic markers, and gene expression (Dejosez and Zwaka 2012). Specifically, early modifications in the phosphoproteome serve as indicators of these state transitions, initiating a cascade of changes that ultimately impact the transcriptome and proteome (Needham et al. 2019). This progression underscores the dynamic nature of pluripotency and highlights the intricate molecular mechanisms governing cell fate decisions in early development.

It is worth noting that early mESCs development differs from human embryonic stem cells (hESCs). While the molecular mechanisms controlling developmental processes are evolutionarily conserved, the speed at which these processes occur can vary significantly between species. For instance, the differentiation of motor neurons from embryonic stem cells is more than twice as fast in mice compared to humans (Rayon et al. 2020).

Regardless of the species, however, early developing embryonic stem cells need to undergo rapid changes in their gene expression profiles to ensure proper development and

differentiation.

2.2 2i Condition for maintaining naive pluripotency

Since serum contains undefined factors that can induce differentiation, the 2i medium was developed to provide a defined, chemically controlled environment that preserves mESCs in the naive pluripotent state (Ying et al. 2008). The 2i medium relies on the synergy between two inhibitors, PD0325901 and CHIR99021, to sustain mESCs in an undifferentiated state. PD0325901, a MEK inhibitor, blocks the MEK/ERK signaling pathway, suppressing differentiation signals from FGF and ERK that would otherwise prompt mESCs to exit pluripotency (see Figure 1). Meanwhile, CHIR99021, a GSK3α/β inhibitor, stabilizes β-catenin which activates Wnt/β-catenin signaling, promoting the expression of self-renewal genes and reinforcing pluripotency (Ying et al. 2008). Together, these inhibitors create a controlled environment that preserves mESCs in a ground state resembling the undifferentiated pre-implantation epiblast (Ying et al. 2008). Leukemia inhibitory factor (LIF) activates the JAK–STAT3 pathway to suppress differentiation. When combined with the two inhibitors of the 2i medium, this signaling network maintains mESCs in a naïve pluripotent state, preserving their full, unbiased differentiation potential. (Ying et al. 2008).

The removal of the 2i medium triggers the transition from the naive pluripotent state by reactivating the MEK/ERK and GSK3 pathways, allowing ERK signaling to promote differentiation and reducing Wnt signaling to enable lineage specification. This shift initiates a cascade of biological changes, including the downregulation of naive markers (such as Nanog and Esrrb) and upregulation of formative markers (such as Otx2), along with epigenetic remodeling, altered chromatin accessibility, and a metabolic shift from glycolysis to oxidative phosphorylation (see Figure 1) (Wray et al. 2011).

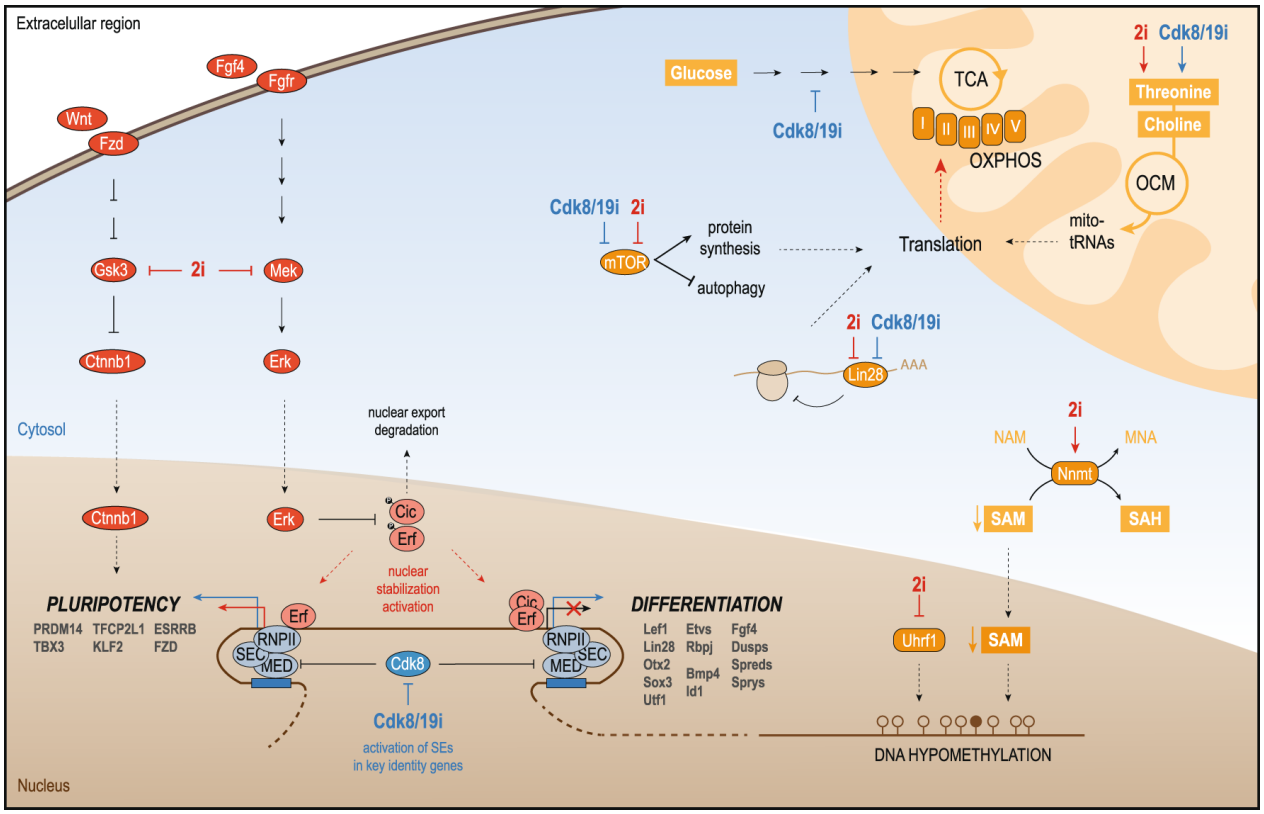

Figure 1: **2i mechanism of stabilizing naive pluripotency**
The inhibition of Mek/Erk through 2i-dependent mechanisms prevents the phosphorylation of Erf and Cic, resulting in the stabilization of these factors within the nucleus. These transcriptional effectors may play a role in activating genes associated with naïve pluripotency while repressing those involved in primed/formative states. Cdk8/19i influences the phosphorylation of proteins that are part of the elongation and transcription machinery, including the super elongation complex (SEC), RNA polymerase II (RNPII), and Mediator (MED). This could lead to the rapid activation of genes regulated by super-enhancers, which include factors that stabilize the naïve state as well as certain factors from the primed/formative states that are also upregulated in response to Cdk8/19i. The activation of the naïve circuitry through 2i and Cdk8/19i may support the enhanced mitochondrial oxidative phosphorylation (OXPHOS) capacity of these cells. The degradation of Lin28a might alleviate its repressive effects on target mRNAs, leading to the upregulation of mitochondrial proteins among others. The balance between autophagy and protein synthesis may sustain the translation of these mitochondrial components. Additionally, the re-wiring of mitochondrial sources for one-carbon units may further support the translation of mitochondrial-encoded proteins. Furthermore, the increased availability of S-adenosylmethionine (SAM) levels, due to NNMT upregulation, combined with the impairment of DNA methylation caused by Uhrf1 degradation, contributes to the characteristic hypomethylated DNA state typical of 2i conditions (Martinez-Val et al. 2021).

2.3 Translation on Demand

During this transitions, embryonic stem cells (ESCs) need to undergo drastic changes in protein expression to adapt and develop efficiently. There must be mechanisms, these cells undergoing fundamental changes in their transcriptome and proteome can employ, that allow for the rapid development. These mechanisms might also be one of the reasons why certain species exhibit a faster rate of development. One of these mechanisms might be Translation on Demand (ToD (Brockmann et al. 2007)). Mechanisms such as ToD, can allow for swift adjustments in protein levels to aid these rapid changes. This ensures that cells can quickly respond to developmental cues and environmental changes, maintaining the necessary pace of development. Such mechanisms are crucial for adapting to the specific temporal requirements, ensuring that ESCs can effectively progress through the stages of development.

We define ToD as an increase in protein abundance due to enhanced translation, occurring without any changes in mRNA levels of the corresponding gene. In other words, ToD is a regulatory mechanism where translation efficiency is selectively increased for transcripts of specific genes. This upregulation allows for rapid adjustments in protein levels, supporting cellular responses that need immediate protein synthesis without requiring new mRNA transcription. Such translation driven increases could be particularly useful in situations where quick adaptations are necessary, as it enables cells to respond to changes in the environment or developmental cues by rapidly elevating the protein output from existing mRNA pools.

In vivo, there are already described mechanisms that facilitate forms of selective, rapid translation of mRNAs. For example, neurons need local translation at active synapses for quick cellular responses (Holt and Schuman 2013). In early embryos, maternal mRNA is stored and activated after fertilization, enabling the embryo to rapidly synthesize proteins necessary for early development without new mRNA synthesis (De Leon, Johnson, and Bachvarova 1983). In stress response, rapid and selective synthesis of stress proteins, such as heat shock proteins (HSPs), is facilitated by ToD, allowing cells to quickly adapt and protect themselves under stress conditions without increasing mRNA levels (Holcik and Sonenberg 2005). ToD has also been shown in the Yeast Saccharomyces cerevisiae

as a post-transcriptional expression regulator (Beyer et al. 2004).

We know that cells use ToD as a tool to selectively and rapidly change their proteomic landscape. However, the mechanism by which cells facilitate ToD remains elusive. In the following sections, we explore potential mechanisms, whether specific or in combination, that might enable cells to facilitate ToD.

2.4 ERK and mTOR signalling pathway and its effects on translation

ERK signaling influences translation primarily by modulating the activity of eukaryotic translation initiation factor 4E (eIF4E) (Lavoie, Gagnon, and Therrien 2020). ERK activates MAPK-Interacting Kinases (MNK)1 and MNK2 kinases, which phosphorylate eIF4E, enhancing its binding to the mRNA cap structure and promoting the formation of the eIF4F complex (see Figure 2). This complex, consisting of eIF4E, eIF4G, and eIF4A, is essential for recruiting the ribosome to mRNA, scanning for the start codon, and initiating protein synthesis. ERK signaling can selectively enhance the translation of specific mRNAs, particularly those with complex 5′ UTRs, leading to the production of proteins involved in cell growth and survival. Additionally, ERK interacts with other pathways, such as mTOR, to coordinate translation regulation in response to various signals, thus optimizing cellular responses to environmental changes.

The mTOR pathway complements the role of ERK in translation regulation. It comprises two distinct complexes: mTORC1 and mTORC2. mTORC1 is primarily responsible for regulating protein synthesis, autophagy, and metabolism, while mTORC2 is involved in cell survival and cytoskeletal organization.

Activation of mTORC1 occurs through nutrient availability, with amino acids serving as key activators, and through growth factors that activate the PI3K-AKT pathway (see Figure 2) (Ma and Blenis 2009). mTORC1 influences translation by phosphorylating 4E-BP1, which releases eIF4E to initiate cap-dependent translation (see Figure 2) (Ma and Blenis 2009). Furthermore, mTORC1 activates S6K1, which phosphorylates ribosomal protein S6, enhancing the translation of ribosomal proteins and elongation factors. These actions collectively increase the global protein synthesis capacity of the cell (see Figure 2) (Ma

and Blenis 2009).

The mTORC1 signaling pathway has a negative feedback regulation on itself and in consequence on the ERK signaling pathway through the inhibition of IRS-1 (see Figure 2). It is this negative feedback that balances translation rates to save cellular resources and prevent the accumulation of translational stress. The integration of these complex signaling pathways shows the tight regulation cells need to enforce on fundamental processes like translation to maintain homeostasis. This regulation becomes even more critical during development where ESCs need to differentiate and proliferate rapidly. During this phase, the precise control of translation rates is essential to ensure that energy and resources are optimally utilized.

In the naive state, mTORC signaling is low, which keeps global protein synthesis rates low. Upon the removal of 2i conditions, mTORC signaling rises, leading to an increase in global translation. This increase does not occur immediately, which suggests the presence of mTOR independent mechanisms, such as cap-independent translation. Furthermore, recent research indicates that non-canonical alternative mechanisms of cellular mRNA translation initiation, whether cap-dependent or independent, enable selective translation of mRNAs during physiological and pathological stress. These stress conditions often lead to a global downregulation of the canonical eIF4E1/cap-mediated mRNA translation and a targeted reprogramming of the cellular proteome, as seen in tumor development and malignant progression (Mahé et al. 2024). Cap-independent translation and other translational regulation mechanisms are tightly controlled by the make up of untranslated regions (UTRs) of mRNAs (Meyer et al. 2015).

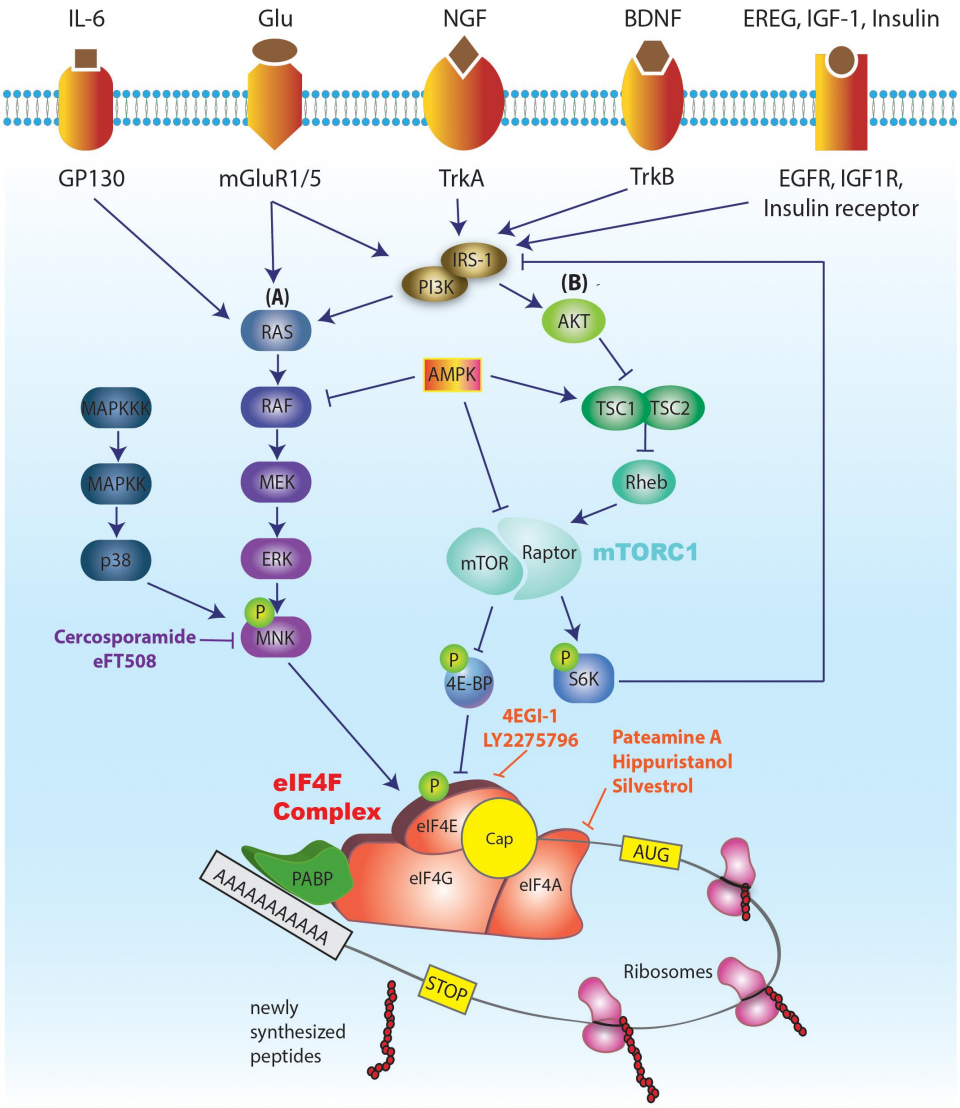

Figure 2: **ERK and mTORs effects on eIF4E regulation**
The ability of eIF4E to bind to the mRNA cap makes it a key regulator of translation. A crucial step in translation initiation involves eIF4E binding to the mRNA cap, facilitating the formation of the eIF4F complex on the cap structure (comprising a 7mGp linked to the first nucleotide). The eIF4F complex includes eIF4E, eIF4G (a scaffolding protein), and eIF4A (a helicase). The successful assembly of the eIF4F complex on the mRNA cap promotes the recruitment of the pre-initiation complex (PIC), followed by scanning of the 5′ UTR to locate the start codon AUG and the joining of the 60S ribosomal subunit, marking the completion of translation initiation. eIF4E acts as a downstream effector of both the mTORC1 pathway (via 4E-BP-dependent repression) and the ERK pathway (via eIF4E phosphorylation by MNK 1/2). The activities of mTORC1 and ERK signaling pathways are influenced by various external signals and internal cues (Uttam et al. 2018).

2.5 Untranslated Regions in mRNAs

mRNA is composed of coding sequences, containing exons and introns, interspersed by non coding segments known as UTRs. Despite not encoding any information for protein production itself, UTRs are pivotal for regulating gene expression post-transcriptionally. There are two types of UTRs. The 5' UTR located immediately upstream of the start codon, while the 3' UTR is positioned immediately downstream of the coding region. Together they integrate diverse cis-regulatory elements that dictate mRNA stability, localization, translation initiation, and subsequently protein functionality. UTRs can vary considerably in length and composition, with longer 3' UTRs often correlating with higher organismal complexity and providing more sites for regulatory factor binding.

Through secondary structures, upstream open reading frames (uORFs), internal ribosome entry sites (IRESs), microRNA (miRNA) binding sites, and binding sites for RNA-binding proteins (RBPs), UTRs provide cells with a modular toolbox to regulate gene expression in response to developmental cues, stress conditions, and disease processes (Mayr 2019; Leppek, Das, and Barna 2018).

The 3' UTR is rich in regulatory motifs that can influence mRNA stability and turn over. AU-rich elements (AREs) for example and other decay signals within the 3' UTR can either enhance or reduce mRNA half-life through interactions with specific RBPs, thereby indirectly controlling protein abundance (Shaw and Kamen 1986). 3' UTRs also function as critical determinants for mRNA localization within the cell by containing sequences that direct mRNAs to specific subcellular compartments (D. A. Melton 1987).

The 5' UTR frequently recruits RBPs and miRNAs that modulate translation efficiency by either stabilizing or repressing ribosome engagement. The formation of RNA secondary structures within the 5' UTRs, such as hairpins and G-quadruplexes, can impede ribosome scanning and modulate translation initiation (Leppek, Das, and Barna 2018). These structures can again either enhance or inhibit translation depending on their stability, location relative to the cap, and interactions with RNA helicases that resolve inhibitory folding.

2.6 RNA binding proteins and miRNAs

As stated earlier, RBPs and miRNAs are essential regulators for mRNA turnover as well as translational efficiency. RBPs constitute a large and diverse class of proteins that govern nearly every step of RNA metabolism, from splicing and transport to stability and translation (Corley, Burns, and Yeo 2020). RBPs bind mRNA through modular RNA-binding domains (RBDs), which primary sequence of amino acids dictates the specificity of the RBP (Lunde, Moore, and Varani 2007), as well as intrinsically disordered regions (IDRs) to allow dynamic interactions with numerous RNA substrates. Besides influencing the turnover rate of specific mRNAs, RBPs also regulate mRNA stability. They form complexes with other proteins to recruit for example ribonucleoproteins (RNPs) from the translation machinery, ensuring proper RNA folding, translocation and translational regulation (Glisovic et al. 2008). Another crucial way that RBPs regulate mRNA stability and translation efficiency is through the formation of closed-loop structures. This process relies on the interaction between factors attached to the 5' cap and the 3' poly(A) tail of mRNAs, which enhances translation efficiency by physically circularizing the mRNA. RBPs contribute to this mechanism by binding simultaneously to both the cap and tail regions or by recruiting bridging factors like PABP (Zhao and Fan 2021). This increases ribosome recycling and protects mRNAs from degradation enzymes.

Similar to RBPs another prominent example of post-transcriptional regulation through direct interaction with mRNAs are miRNAs. miRNAs are a ubiquitous class of short, non-coding RNA molecules. These about 21–24 nucleotide long transcripts are generated through a tightly controlled multistep biogenesis process and exert profound influence on the translation of mRNAs. miRNAs influence translation by primarily binding to partially complementary sequences in the 3' UTRs of target mRNAs, although binding sites can also reside in the coding region or 5' UTR in some cases (Filipowicz, Bhattacharyya, and Sonenberg 2008). The binding is largely governed by the "seed" region, comprising nucleotides 2–8 at the miRNA's 5' end, which is critical for target specificity (Jackson and Standart 2007). Once bound, the miRNA-RISC complex exerts its regulatory effects via several overlapping mechanisms. One major mechanism involves the inhibition of translation initiation. The miRNA-RISC can interfere with cap recognition by eIF4E or

disrupt the assembly of the initiation complex, thereby preventing the recruitment of ribosomal subunits to the target mRNA (Pillai, Bhattacharyya, and Filipowicz 2007). This blockade of initiation results in a halt of protein synthesis while often leaving the mRNA intact in a translationally silent state (Wilczynska and Bushell 2015). In mESCs specifically, families of miRNAs have been shown to maintain pluripotency and suppressing genes that drive differentiation (C. Melton, Judson, and Blelloch 2010). The fact that, ERK signalling globally downregulates miRNAs, further highlights the role of miRNAs to maintain pluripotency and ERKs to exit the naive state (Sun et al. 2016).

2.7 Epigenetic and Epitranscriptomic regulation of gene expression

Lastly epigenetic regulation, although indirectly, can influence translation. Classical epigenetic marks such as DNA methylation and histone modifications determine transcription start site selection and alternative exon usage, thereby dictating the 5' UTR features and coding potential of mRNAs, which in turn influence their translational competence and the diversity of protein isoforms produced (Leenen, Muller, and Turner 2016). In mESCs specifically, the incorporation of specific histone variants, such as H3.3, has been shown to modulate the chromatin environment. Regulatory complexes can modify chromatin to control gene expression during development by influencing the accessibility and stability of specific genomic regions (Banaszynski et al. 2013).

Additionally, post-transcriptional chemical modifications of RNA (the epitranscriptome) represent a dynamic and reversible system that regulates mRNA stability, translation initiation, elongation speed, and termination processes. Similar to epigenetic modifications, key modifications such as N6-methyladenosine (m6A), N1-methyladenosine (m1A), pseudouridine (Ψ), and 5-methylcytosine (m5C) are installed by dedicated writer enzymes, recognized by specific reader proteins and removed by erasers, collectively mediating changes in RNA structure and ribosome engagement (Peer et al. 2019).

Collectively, these mechanisms allow cells to continue efficiently managing protein synthesis in response to various internal and external stimuli. By finely tuning the translation process, cells can adapt to changing conditions, conserve resources, and ensure that

proteins are synthesized correctly and at the right time. This regulation is crucial for maintaining cellular homeostasis and supporting functions such as growth, differentiation, and response to stress. In this project, we propose another mechanism by which cells selectively and rapidly increase the translational efficiency of certain transcripts. ToD might be enabled by one or multiple of the stated mechanisms by which cells regulate translation.

2.8 Aim of this study

In this study, we set out to characterize ToD as a post-transcriptional regulator mechanism of mESC differentiation by addressing three core questions:

1. Which genes are subject to ToD in early development?

We integrate RNA-seq and quantitative proteomics time-course data to identify transcripts whose protein outputs change more rapidly than their mRNA levels (Yang et al. 2019).

2. Why are these genes regulated by ToD?

Alongside finding genes that are subject to ToD regulation, we want to illustrate why these specific genes are prioritized and to uncover the underlying mechanisms of this regulation. We analyze the functional roles and pathway enrichments of the candidate set to understand why rapid translational tuning is advantageous during early differentiation. Additionally, we are interested in understanding the gene regulatory network and proteome changes that occur as mESCs exit pluripotency, as these regulatory networks are most likely to be regulated by ToD.

3. How and when does ToD occur?

Furthermore, we want to understand which posttranscriptional regulation mechanism are essential for enabling mESCs to facilitate ToD. We interrogate each candidate's mRNA structural features 5' and 3' UTR sequences, upstream open reading frames,

and known RBP/miRNA binding sites to infer cis-regulatory mechanisms. By examining high-resolution differentiation time-courses, we pinpoint the developmental window during which translational shifts take place.

By analyzing the temporal dynamics within these data layers, we can distinguish cases where protein levels increase independently of mRNA changes, a hallmark of ToD regulation. This approach provides a comprehensive view of the post-transcriptional mechanisms that contribute to protein abundance, offering insights into gene-specific regulatory processes. Understanding the more nuanced relationship between RNA and protein expression, as well as the timing of regulation, will provide valuable insights into the fundamental principles of molecular biology and enhance our comprehension of cellular regulatory networks as well as early development.

Furthermore, this study should serve as a framework to detect ToD candidates in by leveraging the use of transcriptome and proteome data together. We also aim to validate our findings and approaches with two other datasets to check the validity of our methods and see if we obtain similar results. This validation will help ensure that the mechanisms identified are robust and applicable across different biological contexts. Together, these analyses will provide a comprehensive “who, why, how, and when” picture of Translation on Demand in mESC fate decisions.

3 Methods and tools

3.1 Experimental Approach and Workflow

To research ToD, we exploit the fact that protein translation is largely inhibited under 2i conditions. By sampling mouse embryonic stem cells under 2i conditions and then periodically after 2i removal, we can take snapshots of the early changes in the transcriptome and proteome (see Figure 3). Upon 2i removal, we trigger the transition from the naive pluripotent state to the primed formative state. This allows us to detect changes in the translational behavior of genes aiding this transition.

We connect transcriptomic and proteomic dynamics by comparing their Log Fold Change

(LFC) between time points. This way, we can compare qualitative trends for transcript and protein abundance rather than comparing them by count values. After identifying genes that are likely ToD-regulated, we assess their biological function and place them in the context of early development (see Figure 3).

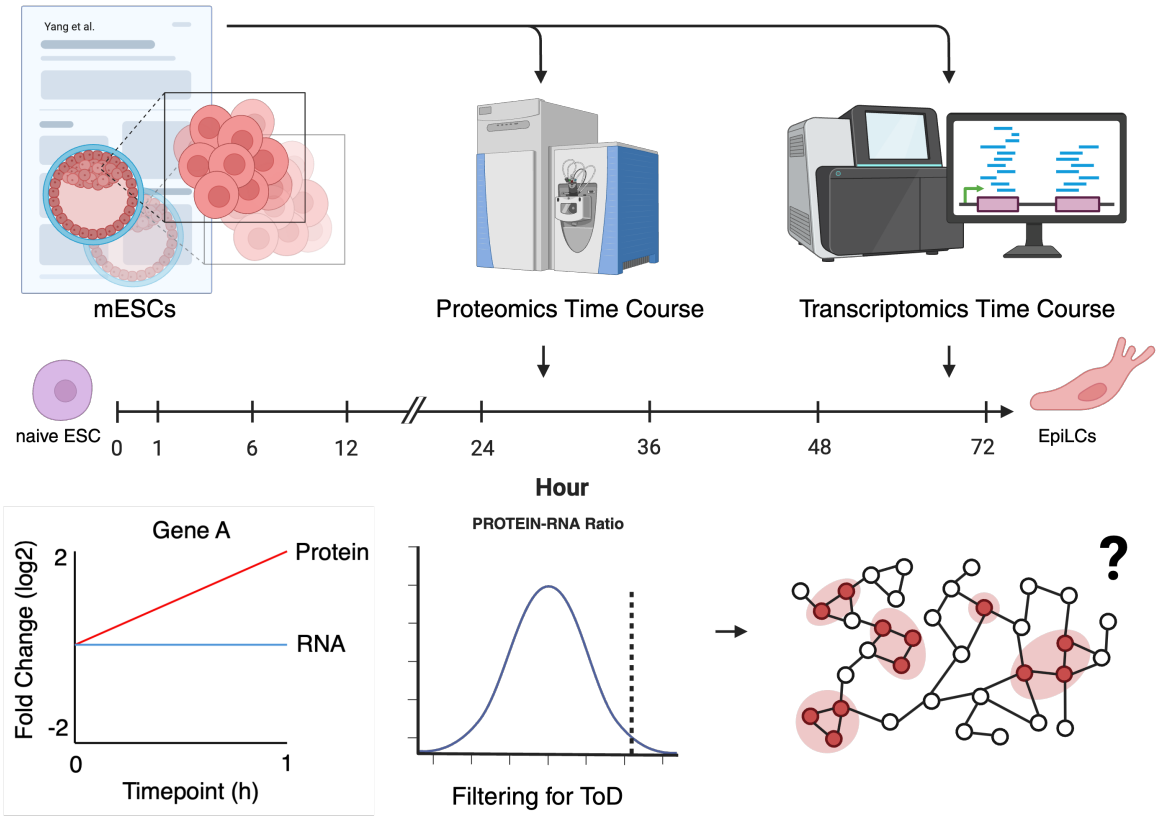

Figure 3: **Experimental Approach and Workflow** Exemplary workflow for the Yang dataset. Proteomic and transcriptomic time courses were used to examine LFC changes at different time intervals. Filtering criteria were applied to determine if a specific gene is likely to be ToD regulated. ToD candidates were tested for functional enrichment and other annotations methods to determine their role in early development.

3.2 Data overview and preprocessing

The data for this project is comprised of three datasets with our main focus being the Yang Data, due to its higher temporal resolution in early time points and more consistent time course in the proteomic data (see Table 1). This dataset was sourced from Supplementary Table S2 (temporal proteomic data) and Supplementary Table S3 (temporal transcriptomic

data) in the publication Multi-omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency (Yang et al. 2019).

The temporal proteomic data is in the form of Log2 Label-Free Quantification (LFQ) intensities and included the following time points: 0min, 30min, 1h, 6h, 12h, 24h, 36h, 48h, and 72h. Each time point had four biological replicates, which were also pooled. The temporal transcriptomic data consisted of non-normalized transcript counts and included the following time points: 0min, 1h, 6h, 12h, 24h, 36h, 48h, and 72h. Each time point had two biological replicates, which were averaged. Since the transcriptomic dataset did not include a 30min time point, this time point was removed from the proteomic data for further analysis (see Table 1).

To increase our accuracy in calling meaningful ToD candidate genes, we realigned the transciptomic data with the GRCm39 reference genome using STAR (Mudge et al. 2025; Dobin et al. 2013).

In the Leeb data, the temporal resolution spans from 0 to 32 hours in 4-hour intervals. For each time point, there are four biological replicates in the proteomic time course, while the transcriptomic time course does not include any replicates (see Table 1). Gaussian Process Regression (GPR) was employed to smooth out the transcriptomic trajectories as this approach leverages the more densely packed information from measurements taken at 2-hour intervals across the time course. This was a purposeful tradeoff, prioritizing a finer-grained time course over the inclusion of replicates, allowing GPR to effectively reduce noise and enhance the clarity of the temporal transcriptomic data (see Section 3.3.1).

The Leeb Medium Change (MC) data includes time course data for 0h, 4h and 8h for the purpose of calling ToD candidates. Additionally, it includes a time course with the same temporal resolution, where 2i conditions were kept while a medium change was still applied to the cells. Each time point has four biological replicates in the proteomic time course and three replicates in the transcriptomic time course (see Table 1).

Table 1: Datasets

Dataset	Temporal Resolution (h)	Replicates Transciptomics	Replicates Proteomics
Yang Data	0, 1, 6, 12, 24, 36, 48, 72	2	4
Leeb Data	0, 4, 8, 12, 24, 32	1	4
Leeb MC Data (Medium Change)	0, 4, 8	3	4

3.3 Transcriptomic Data Normalization

For the Yang dataset as well as the Leeb MC dataset we used DESeq2 to normalize the transciptomic data (Love, Huber, and Anders 2014). DESeq2's normalization procedure operates on the assumption that the majority of genes exhibit no differential expression. To normalize gene counts, DESeq2 uses the median of ratios method. This involves dividing the count of each gene in a sample by its geometric mean across all samples, resulting in a ratio (see Equation 1; Equation 2). The geometric mean serves as a measure of central tendency, representing the average gene count across samples, and provides a standard for comparing individual sample counts. The size factor for a given sample is then determined by calculating the median of these ratios (see Equation 3). This method is robust, accounting for sequencing depth and RNA composition, and effectively compensates for sample specific biases.

$$\text{Geometric Mean}_g = \left(\prod_{s=1}^{n} \text{Count}_{gs} \right)^{1/n} \tag{1}$$

- Geometric Mean_g : The geometric mean of counts for gene g
- Count_{gs} : The count of gene g in sample s
- n : The total number of samples

$$\text{Ratio}_{gs} = \frac{\text{Count}_{gs}}{\text{Geometric Mean}_{g}} \tag{2}$$

$$\text{Size Factor}_{s} = \text{median}\big(\text{Ratio}_{gs} \mid \text{for all } g\big) \tag{3}$$

3.3.1 Gaussian process regression

The negative binomial distribution is an effective model for count data derived from both bulk and single-cell RNA sequencing. Gaussian process regression (GPR) offers a non-parametric method for modeling temporal variations in gene expression (BinTayyash et al. 2021). GPR can be particularly useful when there are few or no replicates. By specifying a covariance function (kernel), GPR incorporates prior knowledge about the gene expression profiles. GPR naturally accounts for correlations between neighboring time points, which is beneficial in time-series data where observations are not independent. The model uses these correlations to infer missing data points and smooth out noise.

3.4 Calculation of Fold Changes and Ratios

LFC values between consecutive time points for proteomic and transcriptomic time courses were calculated with Equation 4.

$$\text{logFC}_{\text{between time points}} = \text{logFC}_{2}\big(\frac{\text{Expression}_{\text{time point}_{n+1}}}{\text{Expression}_{\text{time point}_{n}}}\big) \tag{4}$$

LFC values against the first time point (0h) for proteomic and transcriptomic time courses were calculated with Equation 5.

$$\text{logFC}_{\text{against time point 1}} = \text{logFC}_{2}\big(\frac{\text{Expression}_{\text{time point}_{n}}}{\text{Expression}_{\text{time point}_{1}}}\big) \tag{5}$$

We calculated a Protein-to-RNA-LFC-Ratio (LFC-Ratio) to highlight genes with greater

changes in protein than RNA, identifying possible ToD candidates (see Equation 6). A positive Protein-to-RNA FC Ratio indicates that protein expression has increased more than RNA expression, suggesting potential translational upregulation through ToD.

$$\text{LFC-Ratio} = \text{logFC}_{\text{between time points protein}} - \text{logFC}_{\text{between time points RNA}} \tag{6}$$

3.5 ToD criterias

To considere a gene potentially ToD regulated it has to pass three filtering criterias.

1. Stable transcript expression: The LFC value of a potentially ToD regulated genes has to be within a stable range R_s (e.g. -1 to 1) between two consequtive time points (see Equation 7).

$$-R_s \leq \text{LFC}_{\text{between time points}} \leq R_s \tag{7}$$

2. Protein abundance changes more than transcript abundance: The LFC-Ratio between protein and transcript has to be over a certain threshold (e.g. 1 would mean that the LFC value for the protein was bigger than the transcript LFC value by 1, see Equation 8).

$$\text{LFC-Ratio} \geq \text{ToD-threshold} \tag{8}$$

3. Protein trajectory has to stay up: After passing the ToD-threshold the protein LFC value is only allowed to decrease by a certain amount R_d (e.g 30%, see Equation 9)

$$\text{logFC}_{\text{protein against 0h after ToD call}} \geq \text{logFC}_{\text{protein against 0h during ToD call}} \times (1 - R_d) \tag{9}$$

3.6 Gene Overrepresentation Analysis with clusterProfiler

We performed Gene Overrepresentation Analysis (ORA) on potential ToD candidates at each time point using the R package clusterProfiler (Yu et al. 2012). This analysis aimed to identify overrepresented biological functions, molecular activities, or cellular localizations among ToD candidates compared to a background reference. Using a hypergeometric model, clusterProfiler evaluates the enrichment of gene set terms within each set of ToD candidates, generating p-values to quantify significance. To control for false discovery rates (FDR), we applied Benjamini-Hochberg correction to the resulting p-values (Benjamini and Hochberg 1995). This approach highlighted time-specific biological processes potentially regulated at the translational level among ToD candidates. The corresponding complete data set (genes both present in the proteomic and transcriptomic data) was used as a background (or universe) for the analysis.

3.7 Transciption Factor Acticity inference with decoupleR

We used the R package decoupleR to infere acticity of Transciption Factors (TFs) that might be subject to ToD (Badia-i-Mompel et al. 2022). decoupleR utilizes CollecTRI, a comprehensive resource containing a curated collection of TFs and their transcriptional targets compiled from 12 different resources (Müller-Dott et al. 2023). Interactions are weighted according to their mode of regulation, either as activators or inhibitors. To infer TF activity, decoupleR uses the Univariate Linear Model (ULM) method to calculate TF enrichment scores. For each sample in our datasets and each TF in the CollecTRI network, it fits a linear model that predicts the observed gene expression based solely on the TF's TF-Gene interaction weights. Once fitted, the obtained t-value of the slope is used as the score. A positive t-value indicates that the TF is active, while a negative t-value suggests that the TF is inactive.

3.8 Motif enrichment with MEME

For motif enrichment in UTRs of ToD candidates, we utilized AME from the MEME suite (McLeay and Bailey 2010; Bailey et al. 2015). To use AME for our analysis purpose we used the Memes R package and the MEME CLI tool (Nystrom and McKay 2021). AME works by comparing a set of target sequences, in this case, the UTRs of ToD candidates, against a set of background sequences to identify motifs that are significantly enriched in the target set. When control sequences are provided, AME disregards both the FASTA scores and the sequence order. Instead, the focus is on comparing the target sequences directly against the control sequences to identify motifs that are significantly enriched in the target set using fishers exact test. Using control sequences allows the allows us to focus solely on the differences between the target and control sequences, enhancing the detection of motifs that are distinctively enriched in the target sequences.

3.9 Software

Table 2: Software used for the analysis and visualizations

Software	Version	Provider	Reference
R	4.4.3 (2025-02-28)	R Foundation for Statistical Computing	R Core Team (2025)
STAR	2.7.11b	Alex Dobin	Dobin et al. (2013)
Docker	27.5.1	Docker Inc.	Merkel (2014)
MEME	5.5.7	University of Nevada, Washigton, Queensland and UCSD	Bailey et al. (2015)
BioRender	2025	BioRender	BioRender (2025)

3.9.1 R packages

Table 3: R packages used for the analysis and visualizations

Package Name	Version	Reference
DESeq2	1.46.0	Love, Huber, and Anders (2014)
biomaRt	2.62.1	Durinck et al. (2005)
clusterProfiler	4.14.6	Yu et al. (2012)
enrichplot	1.26.6	Yu (2025)
msigdbr	24.1.0	Dolgalev (2025)
memes	1.14.0	Nystrom and McKay (2021)
universalmotif	1.24.2	Tremblay (2024)
decoupleR	2.12.0	Badia-i-Mompel et al. (2022)
OmnipathR	3.14.0	Valdeolivas, Turei, and Gabor (2019)
stringr	1.5.1	Wickham (2023)
Biostrings	2.74.1	Pagès et al. (2024)
dplyr	1.1.4	Wickham et al. (2023)
tidyr	1.3.1	Wickham, Vaughan, and Girlich (2024)
tibble	3.2.1	Müller and Wickham (2023)
ggplot2	3.5.1	Wickham (2016)
ggsci	3.2.0	Xiao (2024)
ggrepel	0.9.6	Slowikowski (2024)
pheatmap	1.0.12	Kolde (2019)
ComplexHeatmap	2.22.0	Gu (2022)
UpSetR	1.4.0	Gehlenborg (2019)

4 Results

4.1 Detecting Translation on Demand Candidates

One hallmark of ToD regulation is a rapid increase in translation, while the rate of transcription stays stable. As a filtering criteria of ToD regulation, we selected only genes who showed RNA LFCs which stayed inside a stable range for each time interval (see Section 3.5). The stringency of this range depends on the biological question we want to answer with a specific list of ToD candidates. For example, for gene set enrichment it might make sense to broaden the range we define as stable to detect functional trends in a certain time interval which would not be detectable otherwise. If not stated otherwise we used a stable LFC RNA range from minus one to one. The next filtering step to determine if a gene might be subject to ToD regulation at a specific time interval was to filter for a LFC-Ratio of at least one to ensure a substantial increase in protein levels while maintaining a relatively stable RNA LFC, which filters for robust protein level changes independent of RNA fluctuations (see Section 3.5, Figure 4 A). Again, this LFC-Ratio threshold depends on the biological context and needs to be adapted for the dataset in question. If not stated otherwise, we used a LFC-Ratio threshold of one. This filtering step results in an initial ToD candidate list generated with the Yang Data of 113 ToD candidates for the first time interval (0h to 1h), 103 for the second (1h to 6h) and 78 for the third (6h to 12h) etc. (see Figure 4 B). Corrected for time, the number of ToD candidates are 103, 17 and 13 for the first three time intervals respectivly (see Figure 4 B). This trend continues in later time intervals.
To make the ToD candidate calling less subject to noise, we implemented a third filtering step. After a gene passes the LFC-Ratio threshold at a given time interval, we look at its protein LFC trajectory normalized for the first time point (0h) and see if the subsequent LFC value after the threshold passing is within a certain range defined by the time interval where it passed the LFC-Ratio threshold (Section 3.5). For example, if a gene passed the LFC-Ratio threshold at the 0h to 1h time interval, we look in the protein LFC trajectory of that gene for 1h vs 0h and the next LFC value at 6h vs 0h. The LFC value at the 6h vs 0h time interval is only allowed to go down by 30% compared to the previous value at 1h vs 0h to eliminate genes that showed higher levels of noise in their protein LFC trajectory.

After applying this filtering method, the number of ToD candidates for the first three time intervals were 61, 52, and 34 respectively (see Figure 4 C). The trend of having more ToD candidates when correcting for time in the early time intervals remained evident after applying the filtering step.

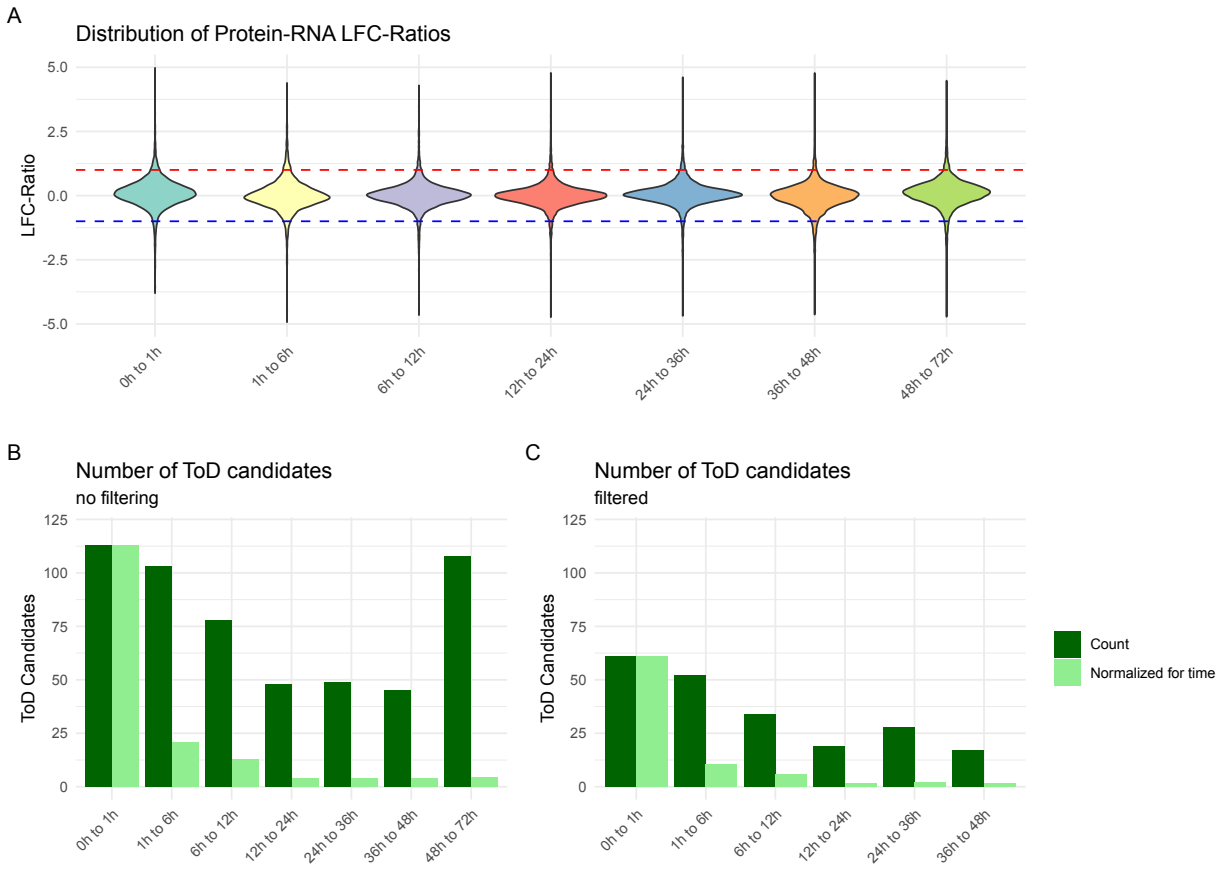

Figure 4: **Filtering for ToD candidates in the Yang Data**
(A) Distribution of LFC-Ratios at different time intervals. The dashed red line (LFC-Ratio of 1) shows the LFC-Ratio threshold a gene needs to pass. The dashed blue line (LFC-Ratio of -1) shows the opposite behaviore where we see a decreased protein expression.
(B) The barplot shows the number of ToD candidates at each time interval which passed the LFC-Ratio of one. The dark green bars show the absolute number of ToD candidates at a given time interval. The light green bars show the number of ToD candidates corrected for time.
(C) The barplot shows the number of ToD candidates at each time interval which passed the LFC-Ratio and which passed the third filtering step. The dark green bars show the absolute number of ToD candidates at a given time interval. The light green bars show the number of ToD candidates corrected for time.

4.2 Early development is shaped by Translation on Demand candidates

Upon examining the filtered ToD candidates across various temporal intervals, it is evident that these candidates include numerous genes crucial for early developmental processes. Specifically, during the initial interval from 0 to 1 hour, a significant proportion of the ToD candidates are associated with developmental functions. Notably, genes such as Agap1, Dnajc14, Emc7, Rab27a, Vti1a and Napb play roles in protein trafficking and vesicular transport.

Not suprisingly then, genes associated with cytoskeletal regulation like Cdc42bpa, Cep95, Plekhh1 and Stil are represented among ToD candidates in the first time interval.

Another group of genes, such as Celf2, Rex2, and Rsrp1, are involved in post-transcriptional regulation. Specifically, Celf2 encodes an RNA-binding protein that plays a critical role in regulating several post-transcriptional events. It is involved in pre-mRNA alternative splicing, mRNA translation, and stability. Celf2 mediates exon inclusion and/or exclusion in pre-mRNA, which is subject to tissue specific and developmentally regulated alternative splicing (Ladd, Charlet-B., and Cooper 2001). Lastly, another group of genes like Axin1, Csnk1g2, Jag1, Rap2c and Rgs19 contributes to regulatory networks that govern developmental processes.

ToD candidates contributing to the mentioned processes can also be found in subsequent time intervals, like the splicing factor Srsf11 or developmental regulators like Epha2 in time interval 1h to 6h among others (see Table 4).

4.3 Gene over-representation analysis reveals time dependent pathway regulation of developmental terms by Translation on Demand candidates

Gene over-representation analysis (see Section 3.6) revealed that ToD candidates are associated with with key developmental terms (see see Figure 5 B). In the 0h to 1h time interval, we identified a nuanced overrepresentation of Notch signaling as well as

Wnt/β-catenin signaling (see Figure 5 A).

Additionally, in the 1h to 6h time interval ToD candidates associated with TNFA signaling and TGF beta signaling where significantly over-represented (see Figure 5 B).

Next to these signalling Hallmark terms, we saw a significant over-representation of GO biological processes associated with development in that time interval (see Figure 5 C). These included tissue specific developmental terms such as tube morphogenesis, vasculature development, epithelial cell proliferation, skeletal system development, ossification and response to growth factor.

ToD candidates of the 6h to 12h time interval also displayed a significant over-representation in GO molecular functions, specifically in calcium ion binding and calcium dependent protein binding as well as a nuanced over-representation in calcium dependent phospholipid binding (see Figure 5 D).

Other significant gene over-representations were detected in the GO Cellular Components (CC) and Reactome during the 1h to 6h and 6h to 12h time intervals. CC terms included cellsubstrate junctions, golgi lumen as well as collagen containing extracellular matrix (see Figure 8). ToD candidate genes over-represented in reactome terms, where associated with extracellular matrix organization, glycoaminoglycan sythesis and apoptotic terms (see Figure 8).

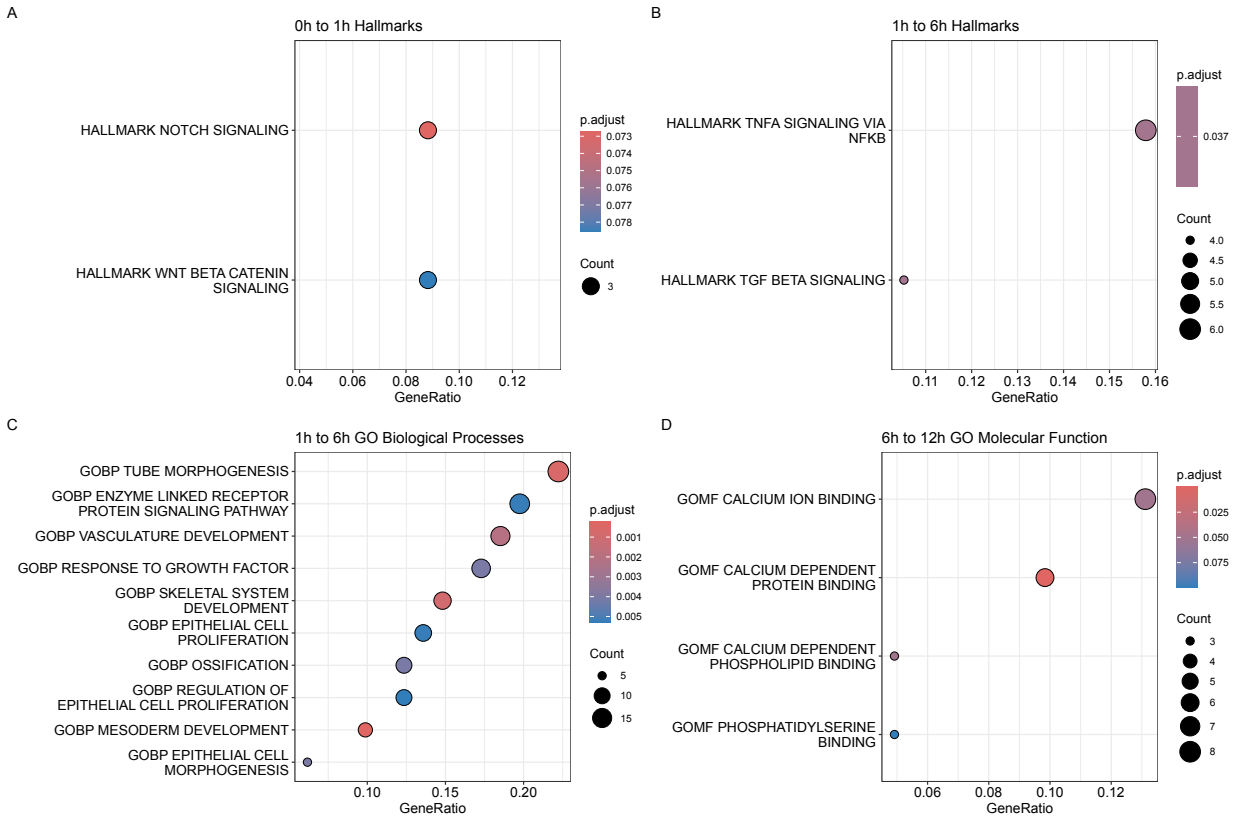

Figure 5: **Gene overrepresentation analysis of early ToD candidates in Yang Data**
(A) Hallmark term over-representation analysis at the 0h to 1h time interval. Adjusted p-value threshold of 0.1. Stable transcript range for ToD call of -2 to 2.
(B) Hallmark term over-representation analysis at the 1h to 6h time interval. Adjusted p-value threshold of 0.05. Stable transcript range for ToD call of -2 to 2.
(C) GO term over-representation analysis for biological processes (BP) at the 1h to 6h time interval. Adjusted p-value threshold of 0.05. Stable transcript range for ToD call of -2 to 2.
(D) GO term over-representation analysis for molecular function (MF) at the 6h to 12h time interval. Adjusted p-value threshold of 0.05. Stable transcript range for ToD call of -2 to 2.
The color gradient represents the adjusted p-value, with red indicating a lower adjusted p-value and blue indicating a higher adjusted p-value. The size of each dot corresponds to the number of genes present in the respective gene set from the dataset.

4.4 Translation on Demand candidates regulate gene expression as transcription factors and their activity increases upon ToD regulation

As we identified a significant part (Fisher's Exact Test with pvalue of 0.021) of ToD candidate genes as transcription factors (TFs) using AnimalTFDB (Shen et al. 2023), we

wanted to investigate if these TFs display their activity upon potential ToD regulation or perhaps with a temporal offset after upregulation through ToD. Out of 438 filtered ToD candidates, 46 (10.5%) were identified as TFs within a stable transcript range of -2 to 2 (see Table 4). From those 46 TFs, 22 had target annotation in CollecTRI (see Figure 6). Strikingly, most of the TF's activity derived from the transcriptomic data correlates with ToD regulation, meaning that they are most active upon ToD regulation and onward (see Figure 6). Furthermore, these TFs are involved in the transition from the naive pluripotent to the primed formative state.

Foxd3 for example, which expresses a moderate to high activity in early time intervals and a potential ToD regulation at the 1h to 6h time interval (see Figure 6). Klf10, which is potentially ToD regulated at the second time interval as well, exhibits a similar activity pattern (see Figure 6). Additionally, Pbx1, ToD regulated in the second time interval, shows its highest activity during the early time points from 0h to 6h (see Figure 6). Similarly, the expression pattern of Hmga1 aligns with that of Pbx1 (see Figure 6). Other key factors for guiding the naive to formative transition like Sall2, Tbx3 and Tead1 are also present in ToD regulated TFs (Xiong et al. 2024; Khan et al. 2020; Nishioka et al. 2008). Notably Hsf2, Nr2c1 and Nfe2l2 showed a possible ToD regulation during the 1h to 6h time interval. During that time interval their activity as TFs is still moderate. After ToD regulation however they show elevated activity during intermediate time points from 24h to 36h.

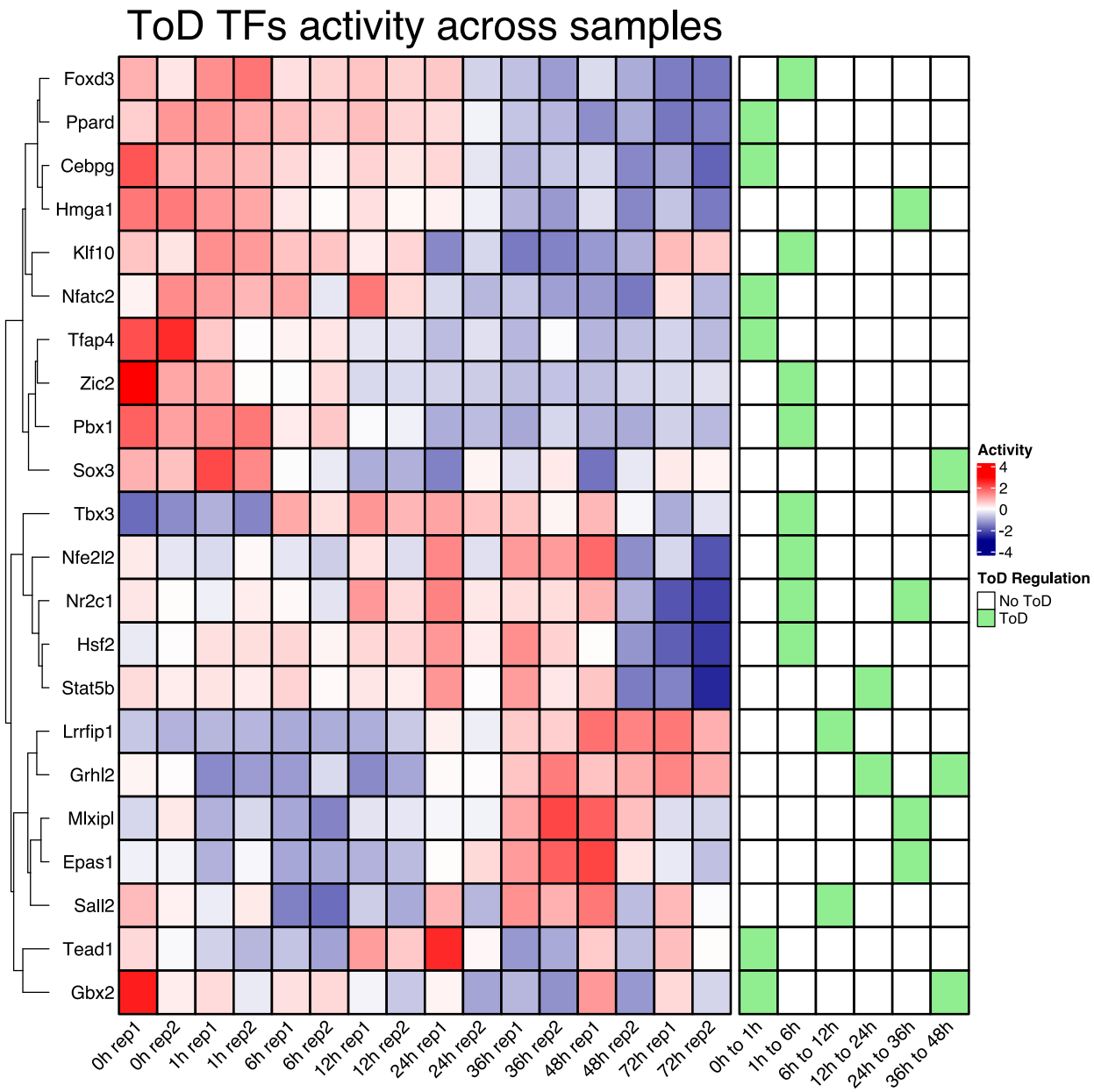

Figure 6: **ToD TFs activity across samples in Yang Data** Activity scores of ToD candidate TFs present in CollecTRI network. Stable transcript range for ToD call of -2 to 2. Positive activity scores are colored in red. Negative activity scores are colored blue (see Section 3.7). Annotation heatmap (right) shows at which time interval(s) a specific TF is potentially ToD regulated.

4.5 Translation on Demand candidates share RNA binding proteins target motifs

Now that we have established possible biological reasons why certain transcripts are subject to rapid selective translational upregulation, we wanted to understand mechanism underlying ToD. RBPs were a first possible contender since it is known that they influence mRNA translation by binding to transcripts inside or outside the nucleus.

To research if RBPs bind to our ToD candidates, we used AME from the MEME suite to do an enrichment analysis of RBP target motifs in 3' and 5' UTRs of ToD candidate transcripts across time intervals (see Section 3.8). We used the CISBP-RNA database to retrieve target motifs for RBPs (Ray et al. 2013). For the 3' UTR, 16 RBP target motifs were enriched in at least one time interval (see Figure 7 A), while for the 5' UTR, 32 RBP target motifs were enriched in at least one time interval (see Figure 7 B). Notably, most of the RBP target motif in the 3' UTR showed an enrichment in the early two time intervals (see Figure 7 A), while target motifs in the 5' UTR showed the vast majority of enrichment in the 24h to 26h time interval (see Figure 7 B). To verify the presence of these RBPs in the Yang Data and their potential impact on the ToD candidates, we examined the LFC time course of these specific proteins. In the proteomic data, 14 out of 16 RBPs for the 3' UTR and 23 out of 32 for the 5' UTR were identified (see Figure 7 C, D). Notably, RBPs who have their target motif enriched in 3' UTRs of ToD candidate genes, displayed a more stable protein expression throughout the time course compared to RBPs who have their target motif enriched in 5' UTRs of ToD candidate genes (see Figure 7 C, D). To check which ToD candidate genes' UTRs carry certain RBP target motifs and to see how the ratio of an enriched RBP target motif changes across time intervals, we searched for matching target motifs in the UTRs of ToD candidate genes. The results matched the enrichment behavior of the RBPs target motifs (see Figure 9).

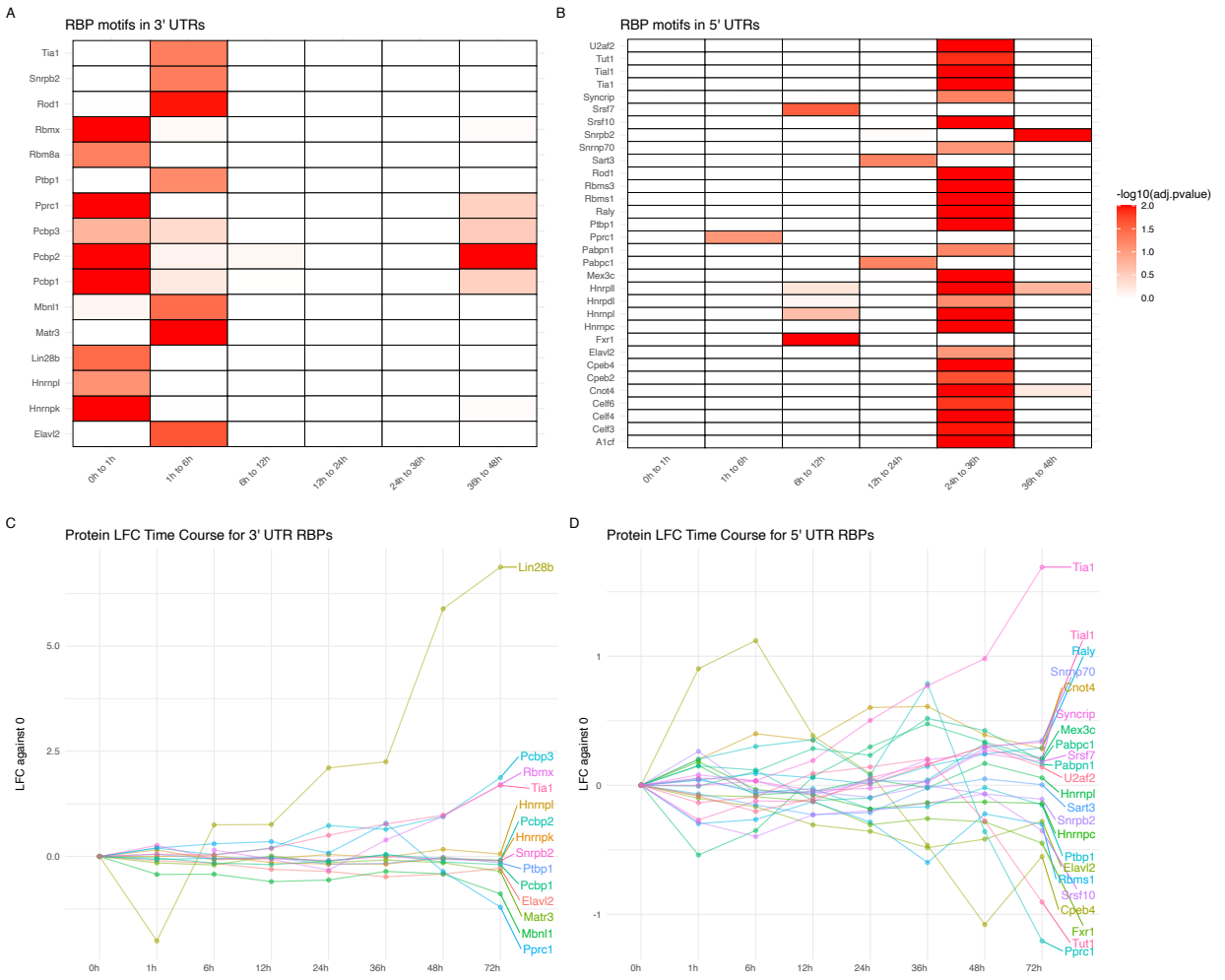

Figure 7: **RBPs motifs in ToD candidate UTRs in Yang Data**
(A) AME results for RBPs target motifs enriched in 3'UTRs of ToD candidates across time intervals. All 3'UTRs sequences from matched Yang Data (present in proteome and transcriptome time course) were used as a control. Stable transcript range for ToD call of -2 to 2.
(B) AME results for RBPs target motifs enriched in 5'UTRs of ToD candidates across time intervals. Stable transcript range for ToD call of -2 to 2. All 5'UTRs sequences from matched Yang Data (present in proteome and transcriptome time course) were used as a control.
(C) LFCs of 3'UTR RBPs against 0h
(D) LFCs of 5'UTR RBPs against 0h

4.6 Translation on Demand candidate calling in Leeb Data

To evaluate our method of calling ToD candidates, we wanted to apply our approach to other datasets to see if results are comparable and which challenges arise when dealing

with different datasets.

5 Discussion

With this study we aimed to identify which genes might be subject to ToD in early development. With our proposed filtering approach, we were able to identify genes, whos protein abundance rapidly increased during early development while transcript abundance remained stable. The fact, that we detect more ToD candidates per unit of time in early time intervals might be due to the relevance of ToD regulation during these early developmental time intervals, because rapid protein systhesis of specific transcipts is needed at that time to initiate priming for differentiation.

Our next objective was to illustrate why these genes are regulated by ToD. Looking at their gene function, it became clear that ToD candidates are deeply involved in key developmental processes. Among early ToD candidate genes we saw genes associated with protein trafficking, cytoskeletal regulation as well as posttranscriptional regulation like alternative splicing. This observation suggests that protein trafficking is likely subjected to substantial demand during this initial period, coinciding with the onset of protein translation following the removal of the 2i condition. The increased load on protein trafficking mechanisms may reflect the cellular necessity to rapidly establish the proteomic landscape required for subsequent developmental stages. These findings illustrate that the ToD candidates identified in the Yang dataset play a crucial role in facilitating the transition from the naive pluripotent state to the primed formative state. By contributing to essential processes such as protein trafficking, cytoskeletal regulation, post-transcriptional modification, and developmental signaling, these genes help orchestrate the complex molecular changes necessary for early development.

To gain more insight in which biological processes ToD candidate genes are involved in, we did gene ORA. In the first time interval from 0h to 1h, we detected a significant overrepresentation of ToD candidate genes in Notch signaling as well as Wnt/β-catenin signaling (see Figure 5 A). Both pathways are well known for their involvement in embryonic development (Zhou et al. 2022; J. Liu et al. 2022). The prominence of these pathways

during the initial time interval suggests that they are among the first to respond to the removal of the 2i conditions. This response likely initiates a cascade of molecular events that prepare the cells for subsequent differentiation, highlighting their importance in the early phases of transitioning from a naive to a primed state.

During the next time interval from 1h to 6h, ToD candidates associated with TNFA signaling and TGF beta signaling where significantly over-represented (see Figure 5 B). During embryogenesis and organogenesis, TNFA signaling plays a crucial role in mediating inflammatory responses and regulating apoptosis, which can influence tissue remodeling and development (You et al. 2021). TNFA's involvement in these processes underscores its importance in maintaining cellular homeostasis and promoting the appropriate differentiation pathways necessary for proper tissue formation. Similarly, TGF-beta signaling is pivotal in embryonic development, as it regulates cell proliferation, differentiation, and migration. TGF-beta signaling is essential for the establishment of the body plan and the development of various tissues and organs (Deng et al. 2024). The overrepresentation of TGF-beta signaling in this interval suggests that ToD candidates may be actively contributing to the cellular transitions required for mESCs to exit the naive state and progress towards lineage specification. Furthermore, an over-representation during this time interval of tissue specific developmental terms suggest the progression toward priming for lineage commitment. This priming requires sophisticated and rapid reprogramming of the epigenetic, gene expression, and protein landscapes of ESCs. ToD might be a central mechanism enabling cells to make these necessary rapid changes. Again, CC terms such as Golgi lumen and substrate junctions in the second and third time interval (1h to 6h and 6h to 12h) may indicate an increase in protein trafficking and the initial progression towards the EpiSCs phenotype, respectively. Meanwhile, Reactome terms related to glycosaminoglycan (GAG) synthesis support this perspective by showing modulation of the extracellular environment and cellular interactions. Additionally, GAGs have been shown to have essential complementary roles in embryonic development (Kramer 2010). The fact that we see an over-representation in calcium ion binding and calcium dependent protein binding, further paint the picture that ToD candidate genes aid mESC through the progression towards the primed formative state. Calcium-dependent signaling is a central and

conserved mechanism essential for development by orchestrating key processes such as cell proliferation, differentiation, and migration. It plays a crucial role in various stages of embryogenesis, including the establishment of body plans, organogenesis, and tissue morphogenesis. The ability of calcium to act as a versatile second messenger allows it to integrate and decode diverse extracellular signals, thereby regulating the dynamic cellular environments necessary for proper development (Stewart and Davis 2019).

These findings underscore the critical role that ToD candidate genes play in early development. The proteins encoded by these genes are essential for facilitating the transition from a naive to a formative primed state. This aligns with our initial hypothesis that ToD may serve as a cellular mechanism that enables swift alterations in the proteomic state of stem cells during development. By integrating cellular signaling cues, ToD effectively regulates gene networks to ensure a proper developmental transition.

This is further highlighted by the finding that TFs especially are subject to ToD. TFs like Foxd3, Klf10 and Pbx1, which are early ToD candidate genes play a critical role in progressing cell development. Foxd3 has been shown to promote exit from naive pluripotency by decommissioning enhancers and inhibiting germline specifications (Respuela et al. 2016). Klf10 regulates target genes that impact tissue formation and may function in conjunction with other KLFs to sustain the stem cell state or propel differentiation (Bialkowska, Yang, and Mallipattu 2017). The temporal activity behavior of Pbx1 agrees with the established relationship to NANOG, a naive marker. Pbx1 upregulates Nanog expression to sustain self-renewal and maintaining pluripotency (Chan et al. 2009). Later ToD candidate genes like Hmga1 are also involved in various aspects of development and differentiation, including cell proliferation and the regulation and maintenance of pluripotency (Vignali and Marracci 2020). This intricate regulation is crucial as mESCs transition from naive pluripotency towards the formative state. During this process, they must integrate developmental cues, which involves selectively dismantling the naive transcription factor circuitry while simultaneously acquiring competence for lineage commitment. It is therefore not surprising that ToD regulated TFs are also involved in this process, modulating the pluripotency network as the cells acquire lineage competence. TFs such as Hsf2, Nr2c1 and Nfe2l2 also play crucial roles next to developemntal regulation. During early development, ESCs are

under increased cellular stress, and the selective upregulation of Hsf2 protein production could enhance the cell's ability to cope with these stressors by increasing the expression of heat shock proteins (Åkerfelt, Morimoto, and Sistonen 2010). Similar to Hsf2, Nfe2l2 is part of the cellular stress response, specifically in sensing environmental and endogenous oxidative stress. Deficiency of Nrf2 (the TF protein encoded by Nfe2l2) can lead to early embryonic lethality (Leung et al. 2003). As part of the highly conserved nuclear hormone receptors (NHRs) superfamily, Nr2c1 is involved in many fundamental biological processes. Next to developmental and cell fate regulation, NHRs influence reproduction, metabolism, the circadian cycle, and immunological functions ("The Nuclear Hormone Receptor Gene Nr2c1 (Tr2) Is a Critical Regulator of Early Retina Cell Patterning" 2017).

In summary, these results again show that ToD candidates fill critical roles in early ESC development. Most importantly by directly promoting the exit from naive pluripotency or supporting pluripotency to aid a proper transition to the primed formative state. Additionally genes like Hsf2, Nr2c1, and Nfe2l2 regulate cellular stress responses and developmental processes. Hsf2 and Nfe2l2 are particularly crucial for managing stress through heat shock and oxidative stress responses, respectively. These findings underscore the importance of ToD fine tuning gene expression to support critical transition mechanisms during early development.

Lastly, we wanted to explore which mechanisms might enable mESCs to facilitate ToD. Therefore we looked for RBP target motifs in UTRs of ToD candidate genes. Notably, these RBPs carry out essential roles in post-transcriptional regulation like mRNA stability, alternative splicing, translational efficiency and mRNA localization. In the first time interval from 0h to 1h we detected an enrichment for target motifs of the RBPs Rbmx, Rbm8a, Pprc1, Pcbp1, Pcbp2, Pcbp3, Lin28b and Hnrnpk (see Figure 7 A). For example, among many other developmental functions, Lin28 interacts within large ribonucleoprotein complexes containing multiple RBPs and helicases to modulate translation of key developmental mRNAs, such as Dnmt3a and Hmga2 (a member of the same family we discovered as a ToD regulated TF Figure 6) (Parisi et al. 2021). Pcbp1-3 are part of the PCBP family, which are known to play a critical role in modulating mRNA stability, alternative splicing as well as translation regulation (Bakhmet et al. 2024). Furthermore, Pcbp1 was identified to timely

downregulate Oct4 upon differentiation signals during the ESC transition from the naive to formative state (Bakhmet et al. 2024). In the second time interval from 1h to 6h we detected an enrichment for target motifs of the RBPs Tia1, Snrpb2, Rod1, Ptbp1, Mbnl1, Matr3 and Elavl2 (see Figure 7 A). Again, RBPs like Ptbp1 orchestrate early embryonic development by integrating developmental cues in translational regulation and shape post-transcriptional modification by regulating alternative splicing (H.-L. Liu et al. 2023).
Prominent groups of RBP target motifs enriched in 5' UTRs for example were PABPs like Pabpc1 and Pabpn1 which protect poly(A) tails from premature degradation, simultaneously stimulating translation initiation by facilitating the formation of a closed loop mRNA structure through interactions with cap binding proteins (Zhao and Fan 2021).
These results show that ToD candidates share RNA-binding protein (RBP) target motifs, which are crucial for rapid selective translational upregulation. The RBPs identified, such as Lin28b, Pcbp1-3, and Ptbp1, play pivotal roles in post-transcriptional regulation, affecting mRNA stability, alternative splicing, and translational efficiency. These proteins are integral to the regulatory networks that guide early embryonic development, responding to developmental cues to ensure precise gene expression and there might be a driving force behind ToD.

6 Conclusion

7 References

Åkerfelt, Malin, Richard I. Morimoto, and Lea Sistonen. 2010. "Heat Shock Factors: Integrators of Cell Stress, Development and Lifespan." *Nature Reviews. Molecular Cell Biology* 11 (8): 545–55. https://doi.org/10.1038/nrm2938.

Badia-i-Mompel, Pau, Jesús Vélez Santiago, Jana Braunger, Celina Geiss, Daniel Dimitrov, Sophia Müller-Dott, Petr Taus, et al. 2022. "decoupleR: Ensemble of Computational Methods to Infer Biological Activities from Omics Data." *Bioinformatics Advances* 2 (1): vbac016. https://doi.org/10.1093/bioadv/vbac016.

Bailey, Timothy L., James Johnson, Charles E. Grant, and William S. Noble. 2015. "The MEME Suite." *Nucleic Acids Research* 43 (W1): W39–49. https://doi.org/10.1093/nar/gkv416.

Bakhmet, E. I., A. S. Zinovyeva, A. A. Kuzmin, D. V. Smirnova, M. N. Gordeev, E. E. Petrenko, N. D. Aksenov, and A. N. Tomilin. 2024. "Pcbp1 Constrains Oct4 Expression in the Context of Pluripotency." bioRxiv. https://doi.org/10.1101/2024.09.07.611681.

Banaszynski, Laura A., Duancheng Wen, Scott Dewell, Sarah J. Whitcomb, Mingyan Lin, Nichole Diaz, Simon J. Elsässer, et al. 2013. "Hira-Dependent Histone H3.3 Deposition Facilitates PRC2 Recruitment at Developmental Loci in ES Cells." *Cell* 155 (1): 107–20. https://doi.org/10.1016/j.cell.2013.08.061.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B (Methodological)* 57 (1): 289–300. https://www.jstor.org/stable/2346101.

Beyer, Andreas, Jens Hollunder, Heinz-Peter Nasheuer, and Thomas Wilhelm. 2004. "Post-Transcriptional Expression Regulation in the Yeast Saccharomyces Cerevisiae on a Genomic Scale." *Molecular & Cellular Proteomics: MCP* 3 (11): 1083–92. https://doi.org/10.1074/mcp.M400099-MCP200.

Bialkowska, Agnieszka B., Vincent W. Yang, and Sandeep K. Mallipattu. 2017. "Krüppel-Like Factors in Mammalian Stem Cells and Development." *Development* 144 (5): 737–54. https://doi.org/10.1242/dev.145441.

BinTayyash, Nuha, Sokratia Georgaka, S. T. John, Sumon Ahmed, Alexis Boukouvalas, James Hensman, and Magnus Rattray. 2021. “Non-Parametric Modelling of Temporal and Spatial Counts Data from RNA-seq Experiments.” *Bioinformatics (Oxford, England)* 37 (21): 3788–95. https://doi.org/10.1093/bioinformatics/btab486.

BioRender. 2025. “Scientific Image and Illustration Software | BioRender.”

Boroviak, Thorsten, Remco Loos, Paul Bertone, Austin Smith, and Jennifer Nichols. 2014. “The Ability of Inner-Cell-Mass Cells to Self-Renew as Embryonic Stem Cells Is Acquired Following Epiblast Specification.” *Nature Cell Biology* 16 (6): 513–25. https://doi.org/10.1038/ncb2965.

Brockmann, Regina, Andreas Beyer, Jürgen J. Heinisch, and Thomas Wilhelm. 2007. “Posttranscriptional Expression Regulation: What Determines Translation Rates?” *PLOS Computational Biology* 3 (3): e57. https://doi.org/10.1371/journal.pcbi.0030057.

Chan, Ken Kwok-Keung, Jingyao Zhang, Na-Yu Chia, Yun-Shen Chan, Hui Shan Sim, Ker Sin Tan, Steve Kah-Weng Oh, Huck-Hui Ng, and Andre Boon-Hwa Choo. 2009. “KLF4 and PBX1 Directly Regulate NANOG Expression in Human Embryonic Stem Cells.” *Stem Cells (Dayton, Ohio)* 27 (9): 2114–25. https://doi.org/10.1002/stem.143.

Corley, Meredith, Margaret C. Burns, and Gene W. Yeo. 2020. “How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms.” *Molecular Cell* 78 (1): 9–29. https://doi.org/10.1016/j.molcel.2020.03.011.

De Leon, Victor, Andrew Johnson, and Rosemary Bachvarova. 1983. “Half-Lives and Relative Amounts of Stored and Polysomal Ribosomes and Poly(A)+ RNA in Mouse Oocytes.” *Developmental Biology* 98 (2): 400–408. https://doi.org/10.1016/0012-1606(83)90369-X.

Dejosez, Marion, and Thomas P. Zwaka. 2012. “Pluripotency and Nuclear Reprogramming.” *Annual Review of Biochemistry* 81 (Volume 81, 2012): 737–65. https://doi.org/10.1146/annurev-biochem-052709-104948.

Deng, Ziqin, Tao Fan, Chu Xiao, He Tian, Yujia Zheng, Chunxiang Li, and Jie He. 2024. “TGF-β Signaling in Health, Disease and Therapeutics.” *Signal Transduction and Targeted Therapy* 9 (1): 61. https://doi.org/10.1038/s41392-024-01764-w.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali

Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. “STAR: Ultrafast Universal RNA-seq Aligner.” *Bioinformatics* 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Dolgalev, Igor. 2025. *Msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format*. Manual.

Durinck, Steffen, Yves Moreau, Arek Kasprzyk, Sean Davis, Bart De Moor, Alvis Brazma, and Wolfgang Huber. 2005. “BioMart and Bioconductor: A Powerful Link Between Biological Databases and Microarray Data Analysis.” *Bioinformatics (Oxford, England)* 21: 3439–40.

Filipowicz, Witold, Suvendra N. Bhattacharyya, and Nahum Sonenberg. 2008. “Mechanisms of Post-Transcriptional Regulation by microRNAs: Are the Answers in Sight?” *Nature Reviews Genetics* 9 (2): 102–14. https://doi.org/10.1038/nrg2290.

Gehlenborg, Nils. 2019. *UpSetR: A More Scalable Alternative to Venn and Euler Diagrams for Visualizing Intersecting Sets*. Manual.

Glisovic, Tina, Jennifer L. Bachorik, Jeongsik Yong, and Gideon Dreyfuss. 2008. “RNA-binding Proteins and Post-Transcriptional Gene Regulation.” *FEBS Letters* 582 (14): 1977–86. https://doi.org/10.1016/j.febslet.2008.03.004.

Gu, Zuguang. 2022. “Complex Heatmap Visualization.” *iMeta*. https://doi.org/10.1002/imt2.43.

Holcik, Martin, and Nahum Sonenberg. 2005. “Translational Control in Stress and Apoptosis.” *Nature Reviews Molecular Cell Biology* 6 (4): 318–27. https://doi.org/10.1038/nrm1618.

Holt, Christine E., and Erin M. Schuman. 2013. “The Central Dogma Decentralized: New Perspectives on RNA Function and Local Translation in Neurons.” *Neuron* 80 (3): 648–57. https://doi.org/10.1016/j.neuron.2013.10.036.

Jackson, Richard J., and Nancy Standart. 2007. “How Do MicroRNAs Regulate Gene Expression?” *Science’s STKE* 2007 (367): re1–1. https://doi.org/10.1126/stke.3672007re1.

Khan, Saif F., Victoria Damerell, Rehana Omar, Michelle Du Toit, Mohsin Khan, Hapiloe Mabaruti Maranyane, Mihlali Mlaza, et al. 2020. “The Roles and Regulation of TBX3

in Development and Disease." *Gene* 726 (February): 144223. https://doi.org/10.1016/j.gene.2019.144223.

Kolde, Raivo. 2019. *Pheatmap: Pretty Heatmaps*. Manual.

Kramer, Kenneth L. 2010. "Specific Sides to Multifaceted Glycosaminoglycans Are Observed in Embryonic Development." *Seminars in Cell & Developmental Biology* 21 (6): 631–37. https://doi.org/10.1016/j.semcdb.2010.06.002.

Lackner, Andreas, Robert Sehlke, Marius Garmhausen, Giuliano Giuseppe Stirparo, Michelle Huth, Fabian Titz-Teixeira, Petra van der Lelij, et al. 2021. "Cooperative Genetic Networks Drive Embryonic Stem Cell Transition from Naïve to Formative Pluripotency." *The EMBO Journal* 40 (8): e105776. https://doi.org/10.15252/embj.2020105776.

Ladd, Andrea N., Nicolas Charlet-B., and Thomas A. Cooper. 2001. "The CELF Family of RNA Binding Proteins Is Implicated in Cell-Specific and Developmentally Regulated Alternative Splicing." *Molecular and Cellular Biology* 21 (4): 1285–96. https://doi.org/10.1128/MCB.21.4.1285-1296.2001.

Lavoie, Hugo, Jessica Gagnon, and Marc Therrien. 2020. "ERK Signalling: A Master Regulator of Cell Behaviour, Life and Fate." *Nature Reviews Molecular Cell Biology* 21 (10): 607–32. https://doi.org/10.1038/s41580-020-0255-7.

Leenen, Fleur A. D., Claude P. Muller, and Jonathan D. Turner. 2016. "DNA Methylation: Conducting the Orchestra from Exposure to Phenotype?" *Clinical Epigenetics* 8 (1): 92. https://doi.org/10.1186/s13148-016-0256-8.

Leppek, Kathrin, Rhiju Das, and Maria Barna. 2018. "Functional 5′ UTR mRNA Structures in Eukaryotic Translation Regulation and How to Find Them." *Nature Reviews Molecular Cell Biology* 19 (3): 158–74. https://doi.org/10.1038/nrm.2017.103.

Leung, Laura, Mandy Kwong, Stephen Hou, Candy Lee, and Jefferson Y. Chan. 2003. "Deficiency of the Nrf1 and Nrf2 Transcription Factors Results in Early Embryonic Lethality and Severe Oxidative Stress." *The Journal of Biological Chemistry* 278 (48): 48021–29. https://doi.org/10.1074/jbc.M308439200.

Liu, Hui-Lin, Xiu-Min Lu, Hai-Yan Wang, Kai-Bin Hu, Qing-Yun Wu, Ping Liao, Sen Li, Zai-Yun Long, and Yong-Tang Wang. 2023. "The Role of RNA Splicing Factor PTBP1

in Neuronal Development." *Biochimica Et Biophysica Acta. Molecular Cell Research* 1870 (7): 119506. https://doi.org/10.1016/j.bbamcr.2023.119506.

Liu, Jiaqi, Qing Xiao, Jiani Xiao, Chenxi Niu, Yuanyuan Li, Xiaojun Zhang, Zhengwei Zhou, Guang Shu, and Gang Yin. 2022. "Wnt/β-Catenin Signalling: Function, Biological Mechanisms, and Therapeutic Opportunities." *Signal Transduction and Targeted Therapy* 7 (1): 3. https://doi.org/10.1038/s41392-021-00762-6.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-seq Data with DESeq2." *Genome Biology* 15 (12): 550. https://doi.org/10.1186/s13059-014-0550-8.

Lunde, Bradley M., Claire Moore, and Gabriele Varani. 2007. "RNA-binding Proteins: Modular Design for Efficient Function." *Nature Reviews Molecular Cell Biology* 8 (6): 479–90. https://doi.org/10.1038/nrm2178.

Ma, Xiaoju Max, and John Blenis. 2009. "Molecular Mechanisms of mTOR-mediated Translational Control." *Nature Reviews Molecular Cell Biology* 10 (5): 307–18. https://doi.org/10.1038/nrm2672.

Mahé, Mélanie, Tiffany Rios-Fuller, Olga Katsara, and Robert J. Schneider. 2024. "Non-Canonical mRNA Translation Initiation in Cell Stress and Cancer." *NAR Cancer* 6 (2): zcae026. https://doi.org/10.1093/narcan/zcae026.

Martinez-Val, Ana, Cian J. Lynch, Isabel Calvo, Pilar Ximénez-Embún, Fernando Garcia, Eduardo Zarzuela, Manuel Serrano, and Javier Munoz. 2021. "Dissection of Two Routes to Naïve Pluripotency Using Different Kinase Inhibitors." *Nature Communications* 12 (1): 1863. https://doi.org/10.1038/s41467-021-22181-5.

Mayr, Christine. 2019. "What Are 3′ UTRs Doing?" *Cold Spring Harbor Perspectives in Biology* 11 (10): a034728. https://doi.org/10.1101/cshperspect.a034728.

McLeay, Robert C., and Timothy L. Bailey. 2010. "Motif Enrichment Analysis: A Unified Framework and an Evaluation on ChIP Data." *BMC Bioinformatics* 11 (1): 165. https://doi.org/10.1186/1471-2105-11-165.

Melton, Collin, Robert L. Judson, and Robert Blelloch. 2010. "Opposing microRNA Families Regulate Self-Renewal in Mouse Embryonic Stem Cells." *Nature* 463 (7281): 621–26. https://doi.org/10.1038/nature08725.

Melton, D. A. 1987. “Translocation of a Localized Maternal mRNA to the Vegetal Pole of Xenopus Oocytes.” *Nature* 328 (6125): 80–82. https://doi.org/10.1038/328080a0.

Merkel, Dirk. 2014. “Docker: Lightweight Linux Containers for Consistent Development and Deployment.” *Linux Journal* 2014 (239): 2.

Meyer, Kate D., Deepak P. Patil, Jun Zhou, Alexandra Zinoviev, Maxim A. Skabkin, Olivier Elemento, Tatyana V. Pestova, Shu-Bing Qian, and Samie R. Jaffrey. 2015. “5′ UTR m6A Promotes Cap-Independent Translation.” *Cell* 163 (4): 999–1010. https://doi.org/10.1016/j.cell.2015.10.012.

Mudge, Jonathan M, Sílvia Carbonell-Sala, Mark Diekhans, Jose Gonzalez Martinez, Toby Hunt, Irwin Jungreis, Jane E Loveland, et al. 2025. “GENCODE 2025: Reference Gene Annotation for Human and Mouse.” *Nucleic Acids Research* 53 (D1): D966–75. https://doi.org/10.1093/nar/gkae1078.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. Manual.

Müller-Dott, Sophia, Eirini Tsirvouli, Miguel Vazquez, Ricardo O Ramirez Flores, Pau Badia-i-Mompel, Robin Fallegger, Dénes Türei, Astrid Lægreid, and Julio Saez-Rodriguez. 2023. “Expanding the Coverage of Regulons from High-Confidence Prior Knowledge for Accurate Estimation of Transcription Factor Activities.” *Nucleic Acids Research* 51 (20): 10934–49. https://doi.org/10.1093/nar/gkad841.

Needham, Elise J., Benjamin L. Parker, Timur Burykin, David E. James, and Sean J. Humphrey. 2019. “Illuminating the Dark Phosphoproteome.” *Science Signaling* 12 (565): eaau8645. https://doi.org/10.1126/scisignal.aau8645.

Nichols, Jennifer, and Austin Smith. 2009. “Naive and Primed Pluripotent States.” *Cell Stem Cell* 4 (6): 487–92. https://doi.org/10.1016/j.stem.2009.05.015.

Nishioka, Noriyuki, Shinji Yamamoto, Hiroshi Kiyonari, Hiroko Sato, Atsushi Sawada, Mitsunori Ota, Kazuki Nakao, and Hiroshi Sasaki. 2008. “*Tead4* Is Required for Specification of Trophectoderm in Pre-Implantation Mouse Embryos.” *Mechanisms of Development* 125 (3): 270–83. https://doi.org/10.1016/j.mod.2007.11.002.

Nystrom, Spencer L., and Daniel J. McKay. 2021. “Memes: A Motif Analysis Environment in R Using Tools from the MEME Suite.” *PLOS Computational Biology* 17 (9): e1008991. https://doi.org/10.1371/journal.pcbi.1008991.

Pagès, Hervé, Patrick Aboyoun, Robert Gentleman, and Saikat DebRoy. 2024. *Biostrings: Efficient Manipulation of Biological Strings*. Manual. https://doi.org/10.18129/B9.bioc.Biostrings.

Parisi, Silvia, Daniela Castaldo, Silvia Piscitelli, Chiara D'Ambrosio, Giuseppina Divisato, Fabiana Passaro, Rosario Avolio, et al. 2021. "Identification of RNA-binding Proteins That Partner with Lin28a to Regulate Dnmt3a Expression." *Scientific Reports* 11 (1): 2345. https://doi.org/10.1038/s41598-021-81429-8.

Peer, Eyal, Sharon Moshitch-Moshkovitz, Gideon Rechavi, and Dan Dominissini. 2019. "The Epitranscriptome in Translation Regulation." *Cold Spring Harbor Perspectives in Biology* 11 (8): a032623. https://doi.org/10.1101/cshperspect.a032623.

Pillai, Ramesh S., Suvendra N. Bhattacharyya, and Witold Filipowicz. 2007. "Repression of Protein Synthesis by miRNAs: How Many Mechanisms?" *Trends in Cell Biology* 17 (3): 118–26. https://doi.org/10.1016/j.tcb.2006.12.007.

R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. Manual. Vienna, Austria: R Foundation for Statistical Computing.

Ray, Debashish, Hilal Kazan, Kate B. Cook, Matthew T. Weirauch, Hamed S. Najafabadi, Xiao Li, Serge Gueroussov, et al. 2013. "A Compendium of RNA-binding Motifs for Decoding Gene Regulation." *Nature* 499 (7457): 172–77. https://doi.org/10.1038/nature12311.

Rayon, Teresa, Despina Stamataki, Ruben Perez-Carrasco, Lorena Garcia-Perez, Christopher Barrington, Manuela Melchionda, Katherine Exelby, et al. 2020. "Species-Specific Developmental Timing Is Associated with Differences in Protein Stability in Mouse and Human." *Science (New York, N.Y.)* 369 (6510): eaba7667. https://doi.org/10.1126/science.aba7667.

Respuela, Patricia, Miloš Nikolić, Minjia Tan, Peter Frommolt, Yingming Zhao, Joanna Wysocka, and Alvaro Rada-Iglesias. 2016. "Foxd3 Promotes Exit from Naive Pluripotency Through Enhancer Decommissioning and Inhibits Germline Specification." *Cell Stem Cell* 18 (1): 118–33. https://doi.org/10.1016/j.stem.2015.09.010.

Shaw, Gray, and Robert Kamen. 1986. "A Conserved AU Sequence from the 3′ Untranslated Region of GM-CSF mRNA Mediates Selective mRNA Degradation." *Cell* 46 (5):

659–67. https://doi.org/10.1016/0092-8674(86)90341-7.

Shen, Wen-Kang, Si-Yi Chen, Zi-Quan Gan, Yu-Zhu Zhang, Tao Yue, Miao-Miao Chen, Yu Xue, Hui Hu, and An-Yuan Guo. 2023. "AnimalTFDB 4.0: A Comprehensive Animal Transcription Factor Database Updated with Variation and Expression Annotations." *Nucleic Acids Research* 51 (D1): D39–45. https://doi.org/10.1093/nar/gkac907.

Slowikowski, Kamil. 2024. *Ggrepel: Automatically Position Non-Overlapping Text Labels with 'Ggplot2'*. Manual.

Stewart, Teneale A., and Felicity M. Davis. 2019. "An Element for Development: Calcium Signaling in Mammalian Reproduction and Development." *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research*, 15th European Symposium on Calcium, 1866 (7): 1230–38. https://doi.org/10.1016/j.bbamcr.2019.02.016.

Sun, Hui-Lung, Ri Cui, JianKang Zhou, Kun-yu Teng, Yung-Hsuan Hsiao, Kotaro Nakanishi, Matteo Fassan, et al. 2016. "ERK Activation Globally Downregulates miRNAs Through Phosphorylating Exportin-5." *Cancer Cell* 30 (5): 723–36. https://doi.org/10.1016/j.ccell.2016.10.001.

"The Nuclear Hormone Receptor Gene Nr2c1 (Tr2) Is a Critical Regulator of Early Retina Cell Patterning." 2017. *Developmental Biology* 429 (1): 343–55. https://doi.org/10.1016/j.ydbio.2017.05.021.

Tremblay, Benjamin JM. 2024. "Universalmotif: An R Package for Biological Motif Analysis." *Journal of Open Source Software* 9: 7012. https://doi.org/10.21105/joss.07012.

Uttam, Sonali, Calvin Wong, Theodore J. Price, and Arkady Khoutorsky. 2018. "eIF4E-Dependent Translational Control: A Central Mechanism for Regulation of Pain Plasticity." *Frontiers in Genetics* 9 (October). https://doi.org/10.3389/fgene.2018.00470.

Valdeolivas, Alberto, Denes Turei, and Attila Gabor. 2019. "OmnipathR: Client for the OmniPath Web Service."

Vignali, Robert, and Silvia Marracci. 2020. "HMGA Genes and Proteins in Development and Evolution." *International Journal of Molecular Sciences* 21 (2): 654. https://doi.org/10.3390/ijms21020654.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

———. 2023. *Stringr: Simple, Consistent Wrappers for Common String Operations*. Manual.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. Manual.

Wickham, Hadley, Davis Vaughan, and Maximilian Girlich. 2024. *Tidyr: Tidy Messy Data*. Manual.

Wilczynska, A., and M. Bushell. 2015. "The Complexity of miRNA-mediated Repression." *Cell Death & Differentiation* 22 (1): 22–33. https://doi.org/10.1038/cdd.2014.112.

Wray, Jason, Tüzer Kalkan, Sandra Gomez-Lopez, Dominik Eckardt, Andrew Cook, Rolf Kemler, and Austin Smith. 2011. "Inhibition of Glycogen Synthase Kinase-3 Alleviates Tcf3 Repression of the Pluripotency Network and Increases Embryonic Stem Cell Resistance to Differentiation." *Nature Cell Biology* 13 (7): 838–45. https://doi.org/10.1038/ncb2267.

Xiao, Nan. 2024. *Ggsci: Scientific Journal and Sci-Fi Themed Color Palettes for 'Ggplot2'*. Manual.

Xiong, Hui, Bowen Lin, Junyang Liu, Renhong Lu, Zheyi Lin, Chengwen Hang, Wenjun Liu, et al. 2024. "SALL2 Regulates Neural Differentiation of Mouse Embryonic Stem Cells Through Tuba1a." *Cell Death & Disease* 15 (9): 710. https://doi.org/10.1038/s41419-024-07088-5.

Yang, Pengyi, Sean J. Humphrey, Senthilkumar Cinghu, Rajneesh Pathania, Andrew J. Oldfield, Dhirendra Kumar, Dinuka Perera, et al. 2019. "Multi-Omic Profiling Reveals Dynamics of the Phased Progression of Pluripotency." *Cell Systems* 8 (5): 427–445.e10. https://doi.org/10.1016/j.cels.2019.03.012.

Ying, Qi-Long, Jason Wray, Jennifer Nichols, Laura Batlle-Morera, Bradley Doble, James Woodgett, Philip Cohen, and Austin Smith. 2008. "The Ground State of Embryonic Stem Cell Self-Renewal." *Nature* 453 (7194): 519–23. https://doi.org/10.1038/nature06968.

You, Kai, Hui Gu, Zhengwei Yuan, and Xuewen Xu. 2021. "Tumor Necrosis Factor Alpha Signaling and Organogenesis." *Frontiers in Cell and Developmental Biology* 9 (July): 727075. https://doi.org/10.3389/fcell.2021.727075.

Yu, Guangchuang. 2025. *Enrichplot: Visualization of Functional Enrichment Result*. Manual. https://doi.org/10.18129/B9.bioc.enrichplot.

Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He. 2012. "clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters." *Omics: A Journal of Integrative Biology* 16 (5): 284–87. https://doi.org/10.1089/omi.2011.0118.

Zhao, Long-Wen, and Heng-Yu Fan. 2021. "Revisiting Poly(A)-Binding Proteins: Multifaceted Regulators During Gametogenesis and Early Embryogenesis." *BioEssays* 43 (6): 2000335. https://doi.org/10.1002/bies.202000335.

Zhou, Binghan, Wanling Lin, Yaling Long, Yunkai Yang, Huan Zhang, Kongming Wu, and Qian Chu. 2022. "Notch Signaling Pathway: Architecture, Disease, and Therapeutics." *Signal Transduction and Targeted Therapy* 7 (1): 95. https://doi.org/10.1038/s41392-022-00934-y.

8 Appendices

8.1 Supplementary figures

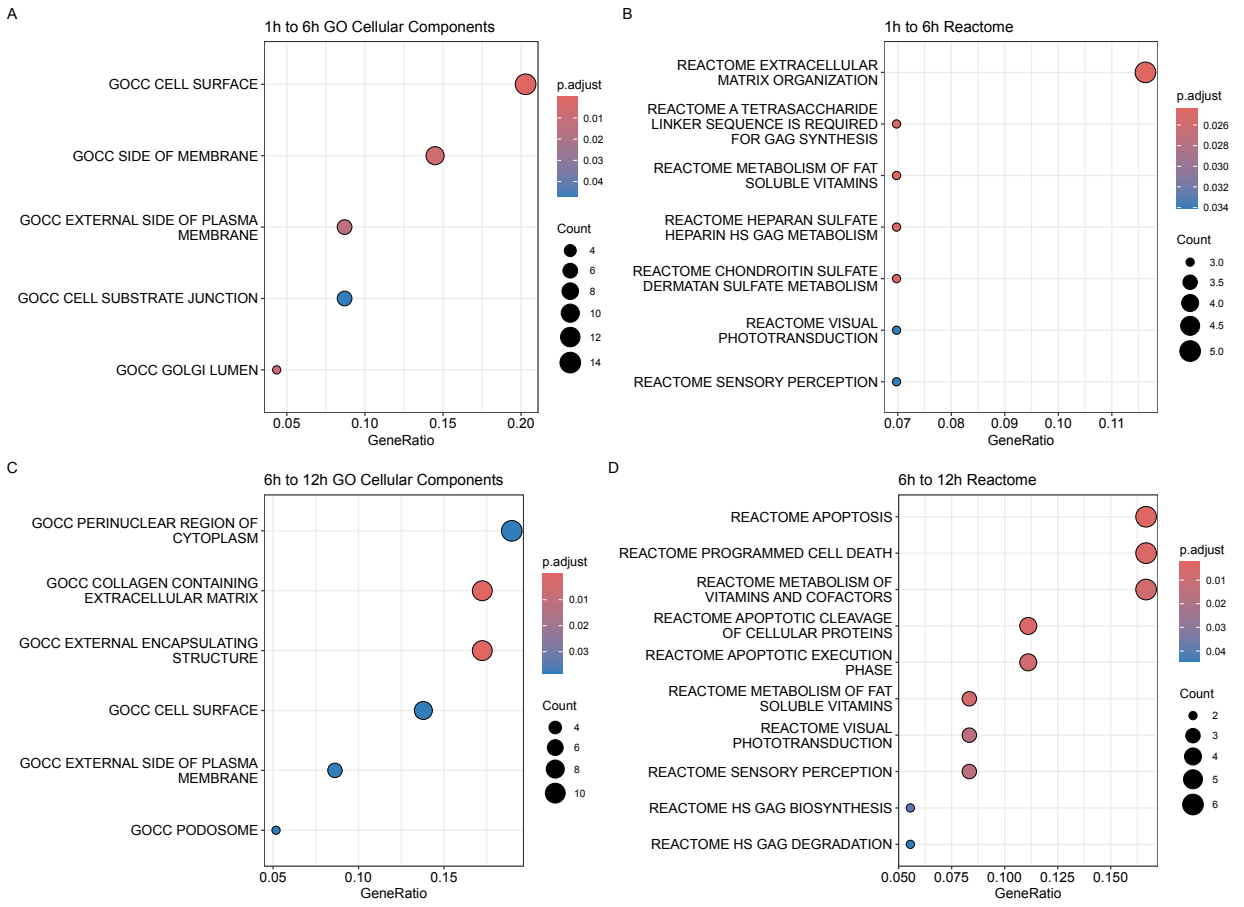

Figure 8: **Gene overrepresentation analysis of early ToD candidates in Yang Data**
(A) GO term over-representation analysis for cellular processes (CC) at the 1h to 6h time interval. Adjusted p-value threshold of 0.05. Stable transcript range for ToD call of -2 to 2.
(B) Reactome term over-representation analysis at the 1h to 6h time interval. Adjusted p-value threshold of 0.05. Stable transcript range for ToD call of -2 to 2.
(C) GO term over-representation analysis for cellular processes (CC) at the 6h to 12h time interval. Adjusted p-value threshold of 0.05. Stable transcript range for ToD call of -2 to 2.
(D) Reactome term over-representation analysis at the 6h to 12h time interval. Adjusted p-value threshold of 0.05. Stable transcript range for ToD call of -2 to 2.
The color gradient represents the adjusted p-value, with red indicating a lower adjusted p-value and blue indicating a higher adjusted p-value. The size of each dot corresponds to the number of genes present in the respective gene set from the dataset.

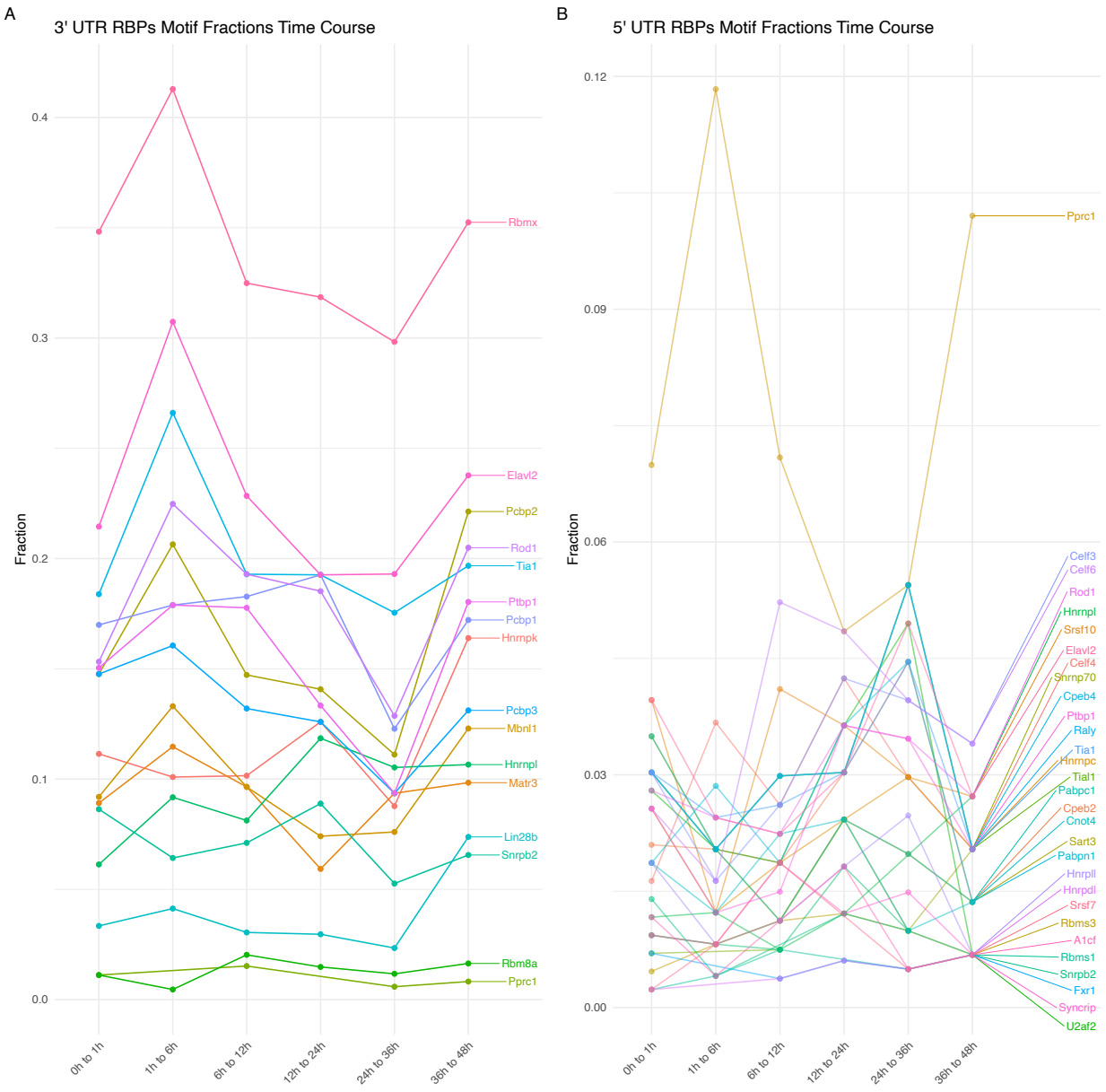

Figure 9: **RBPs motifs fractions in ToD candidate UTRs in Yang Data**
(A) RBPs target motifs fraction enriched in 3'UTRs of ToD candidates across time intervals. All 3'UTRs sequences from matched Yang Data (present in proteome and transcriptome time course) were used as a control. Stable transcript range for ToD call of -2 to 2. Fractions were calculated by dividing the number of 3' UTRs from ToD candidates carrying a specific motif at a given time interval by the total number of 3' UTRs from ToD candidates at that same time interval.
(B) RBPs target motifs enriched in 5'UTRs of ToD candidates across time intervals. Stable transcript range for ToD call of -2 to 2. All 5'UTRs sequences from matched Yang Data (present in proteome and transcriptome time course) were used as a control. Fractions were calculated by dividing the number of 5' UTRs from ToD candidates carrying a specific motif at a given time interval by the total number of 5' UTRs from ToD candidates at that same time interval.

8.2 Supplementary tables

Table 4: ToD candidate genes from the Yang Data

Time Interval	ToD candidate genes
0h to 1h	Agap1, AI987944, Arhgap27, Axin1, Bbs5, Bbs9, Bex4, Btbd9, Cacna1a, Cbx6, Ccdc91, Ccnf, Cdc42bpa, Celf2, Cep95, Cgn, Csnk1g2, Csrnp2, Dcbld1, Dnajc14, Ebpl, Emc7, Enox2, Epha1, Fbxo9, Gba2, Gm5141, Gstp2, Hps4, Jag1, Napb, Nudt22, Odc1, Otud5, Plch1, Plekhh1, Rab27a, Rap2c, Rassf1, Rex2, Rgs19, Rhbdd3, Rnf187, Rsrp1, Slc38a2, Smim20, Sowahb, Stil, Taf11, Tex15, Tm2d3, Tmc6, Tmub2, Tnks, Tpst2, Vangl1, Vti1a, Zcchc2, Zfp850, Zik1, Zmym6
1h to 6h	Agtrap, Alg3, Arhgap30, Avpi1, Bmpr1a, Capn3, Ccdc112, Cdh1, Coro2b, Ctdsp2, Dip2c, Epha2, Fam168b, Fat1, Fzr1, G2e3, Hsf2, Ifrd1, Irs2, Itgb3, Jak1, Kifc5b, Lgals8, Lin28b, Maml1, Mapkapk2, Micall1, Nfe2l2, Nptx2, Nr2c1, Pbx1, Pde8a, Pecam1, Rimklb, Rnf187, Rps6ka6, Slc19a2, Slc9a8, Sptb, Sqstm1, Srsf11, Ss18l2, Sulf1, Suv39h1, Syde2, Tex15, Tmem106b, Traip, Trim37, Trmt44, Zfand5, Zfp84
6h to 12h	Adck2, Akap11, Bbx, Bcl2l11, Bex4, Birc2, Bnc2, Cenpj, Cers5, Cox7a2, Dzip1l, Eif1b, Enox2, Gpc4, Htra1, Kctd9, Klhdc10, Lnx2, Mpped2, Myo1d, Ormdl3, Parp9, Pnp, Rab40c, Rapgef2, Rbm44, Rims2, S100a10, Scd1, Sdc1, Slc25a31, Tmem17, Ttll4, Ube2q2
12h to 24h	Alox15, Arhgef25, Bbc3, Camkv, Dapk1, Dgkd, Mapk6, Mief1, Mthfsd, Nuak2, Phf6, Pop5, Rab27a, Sh3kbp1, Slc35b1, Spop, Tktl2, Trim23, Zfp644

Time Interval	ToD candidate genes
24h to 36h	Actn3, Asnsd1, Capn3, Chchd7, Cnn2, Dgkd, Enpp3, Fam83b, Fbxw17, Foxp1, Gng3, Gstm4, Hmga1, Hmox1, Lnx2, Mrpl52, Nr2c1, Otub2, Peli1, Plce1, Pop5, Rsl24d1, Siva1, Spta1, Srgap1, Sumo3, Syne1, Zfp644
36h to 48h	B4galt1, Clip2, Cthrc1, Fam222b, Flnc, Flywch1, Glt1d1, Irs2, Mlxip, Mutyh, Phykpl, Ptpn13, Rbm47, Rragd, Tmem245, Trio, Wwp1

8.3 Code and Data availability

8.3.1 Datasets

The transcriptomic time course data from the Yang Data is available on the Gene Expression Omnibus database under: GSE117896. The corresponding proteome time course data is available on the ProteomeXchange Consortium under: PXD010621 (Yang et al. 2019). Transciptomic time course data from the Leeb Data is available on the Gene Expression Omnibus database under: GSE145653 (Lackner et al. 2021). The corresponding time course proteome data and the Leeb Medium Change dataset can be requested.

8.3.2 MEME analysis docker container

For the motif analysis we used the MEME suite alongside R. Therefore we set up a docker which had the MEME CLI tool installed as well as R and the required R packages to run the motif analysis. The docker image file is accessible at the Beyer Group GitHub repository.

8.3.3 R package for calling Translation on Demand candidates: ToDcallR

Because we had to do many iterations of how to filter for ToD candidates with different parameters and datasets, we developed ToDcallR alongside to help us with our analysis,

making our methods more accessible and to make it easier for future projects to call potential ToD regulated genes. The package is still under development, because the constant refinement of our calling methods. If you want to use the package or want to contribute, you can access the package on my GitHub repository.

8.4 Abbreviations

Table 5: List of abbreviations

Abbreviation	Definition
ToD	Translation on Demand
mESCs	Mouse embryonic stem cells
PSCs	Pluripotent stem cells
EpiSCs	Epiblast stem cells
hESCs	Human embryonic stem cells
LIF	Leukemia inhibitory factor
ESCs	Embryonic stem cells
HSPs	Heat shock proteins
eIF4E	Eukaryotic translation initiation factor 4E
MNKs	MAPK-Interacting Kinases
RNPs	Ribonucleoproteins
UTR	Untranslated region
uORFs	upstream Open Reading Frames
IRESs	Internal Ribosome Entry Sites
miRNAs	mircoRNAs
RBPs	RNA-binding proteins
AREs	AU-rich elements
RBDs	RNA-binding domains
IDRs	Intrinsically Disordered Regions
RBPs	Poly(A)-Binding proteins

Abbreviation	Definition
LFC	Log Fold Change
GPR	Gaussian Process Regression
LFC-Ratio	Protein-to-RNA-LFC-Ratio
GO	Gene ontology
BP	Biological processes
MF	Molecular function
CC	Cellular components
GAG	Glycosaminoglycan
TFs	Transcription factors
ULM	Univariate Linear Model
NHRs	Nuclear Hormone Receptors