# introduction

Degradation of mRNA plays an important role in the cells. The decay of the mRNA has a regulatory role in gene expression during homeostasis and a critical role in the maternal-to-zygotic transition. The maternal-to-zygotic transition is a crucial process that takes place in embryos of multicellular organisms such as Drosophila, fish, humans, etc.[[1]](#endnote-1) during the very early stages of development. Embryonic genes are initially silenced before fertilization and become activated, initiating expression within the cell following fertilization. During the transition the maternal mRNAs are cleared by post-transcriptional mechanisms.

The post-transcriptional fate of the mRNA molecules is affected by both cis-acting elements and trans regulators such as RNA binding proteins[[2]](#endnote-2) and microRNAs[[3]](#endnote-3). For example, in zebra fish early development, 3' UTR sequences serve as degradation-inducing sites for the microRNA miR-430. The influence of the 3' UTR sequences on the degradation mechanism of mRNA molecules in embryos is well-studied and found to play a significant role in regulating the degradation rate.

It is known that also the 5' UTR contains many regulatory elements such as upstream ORFs (uORFs), internal ribosome entry sites (IRESs), microRNA binding sites, and structural components involved in the regulation of mRNA translation initiation and pre-mRNA splicing[[4]](#endnote-4). In our study, we wanted to examine whether sequence elements in the 5' UTR also affects the stability of the molecule, which, in turn, influence the mRNA degradation rate.

These kinds of discoveries could improve the efficiency of mRNA vaccines, as it has been shown that increasing the stability of mRNA can enhance the production of the desired protein, thereby improving the human body’s ability to overcome the targeted pathogen [[5]](#endnote-5).

Previous studies measured the mRNA degradation dynamics, of tens of thousands of 3'-UTR sequences, using a massively parallel reporter assay during the maternal-to-zygotic transition in zebrafish embryos. This data was then used to train artificial intelligent and deep learning models to predict and learn how the 3' UTR's sequence affect the degradation rate[[6]](#endnote-6) [[7]](#endnote-7). In this project, as mentioned above, we are interested in applying the same methodology to the 5' UTR, to identify regulatory sequence elements in the 5' UTR of mRNAs that affect the degradation rate of the molecule in the cell.

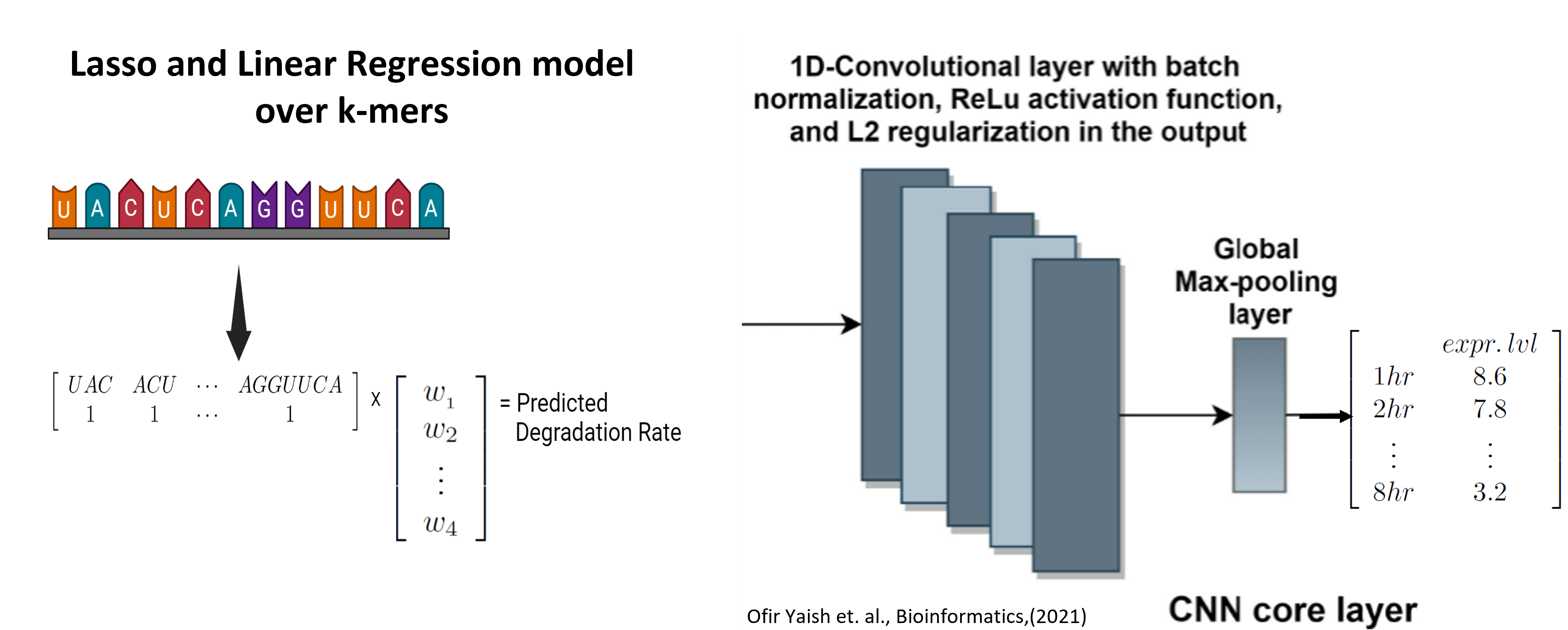
Our data was collected using UTR-seq which is a massively parallel reporter assay to measure RNA degradation. The reporter library contains 30,000 designed 110-nt long sequences. Two mRNA reporter libraries were generated with different poly(A) tail lengths: pre-adenylated reporters (A+), which were transcribed with a 40-nt long poly(A), and nonadenylated reporters (A–) without a poly(A) at the end of their 3′-UTR. The two mRNA reporter libraries were injected into one-cell zebrafish embryos, and the mRNA levels were measured in nine time points: 1-h to 8-h post-fertilization (hpf). Therefore, the dataset contains 30,000 sequences, and for each sequence, eight measured mRNA levels, both for A– and A+. Each sequence in the UTR-seq dataset has a read count (Unique Molecular Identifier count), by RNA-Seq of the library. Counts are converted into an mRNA level in logarithmic scale by normalization with spike-ins.

A different types of symbols

Description automatically generated with medium confidence

After gathering the data, we applied computational methods such as linear regression models and deep learning networks to study sequence patterns in the 5' UTR that regulate the stability of the molecules during the Maternal-To-Zygotic transition in zebrafish. The massive degradation of maternal mRNAs is a key regulatory event in early embryos[[8]](#endnote-8) and a powerful system to study mRNA dynamics in the absence of de-novo transcription.

The linear model was used to identify 3 to 7 nucleotide-long subsequences in the 5' UTR that regulate mRNA stability. The subsequence length was limited to 3-7 nucleotides based on the assumption that most RNA-binding protein and micro RNAs binding sites are associated with sequences of these lengths. Using this type of module also assume that the presence of these binding sites has a linear nature, as the appearance of 2 identical sequences has a doubled effect on its decay mechanism. In previous studies, the use of linear regression models on 3' UTR sequences was successful, and we aimed to adapt a similar approach for the 5' UTR sequences [[9]](#endnote-9). Neural networks were also successfully used to investigate the influence of the 3' UTR on mRNA stability, offering the advantage of being non-linear and position dependent. We trained neural network models that were developed and published by Ofir Yaish and Yaron Orenstein. In their paper, two models were introduced: a convolution neural network (CNN), and a recurrent neural network (RNN). Although the RNN can achieve better predictions in some cases, the CNN is more interpretable. Therefore, it will be the main tool used in our study.



1. **References**

   Bazzini, Ariel A., et al. "Codon identity regulates mRNA stability and translation efficiency during the maternal‐to‐zygotic transition." *The EMBO journal* 35.19 (2016): 2087-2103.

   ‏ [↑](#endnote-ref-1)
2. Pique´ , M., Lo´ pez, J.M., Foissac, S., Guigo´ , R., and Me´ ndez, R. (2008). A combinatorial code for CPE-mediated translational control. Cell 132, 434–448. [↑](#endnote-ref-2)
3. Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K.,

   Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation

   and clearance of maternal mRNAs. Science 312, 75–79. [↑](#endnote-ref-3)
4. Jia, L., Mao, Y., Ji, Q. *et al.* Decoding mRNA translatability and stability from the 5′ UTR. *Nat Struct Mol Biol* **27**, 814–821 (2020). https://doi.org/10.1038/s41594-020-0465-x [↑](#endnote-ref-4)
5. Zhang, H., Zhang, L., Lin, A. et al. Algorithm for optimized mRNA design improves stability and immunogenicity. Nature 621, 396–403 (2023). [↑](#endnote-ref-5)
6. Rabani, Michal, et al. "A massively parallel reporter assay of 3′ UTR sequences identifies in vivo rules for mRNA degradation." *Molecular cell* 68.6 (2017): 1083-1094 [↑](#endnote-ref-6)
7. Ofir Yaish, Yaron Orenstein, Computational modeling of mRNA degradation dynamics using deep neural networks, Bioinformatics, Volume 38, Issue 4, February 2022, Pages 1087–1101 [↑](#endnote-ref-7)
8. Jukam, D., Shariati, S.A.M., and Skotheim, J.M. (2017). Zygotic genome activation

   in vertebrates. Dev. Cell 42, 316–332. [↑](#endnote-ref-8)
9. Rabani, Michal, et al. "A massively parallel reporter assay of 3′ UTR sequences identifies in vivo rules for mRNA degradation." *Molecular cell* 68.6 (2017): 1083-1094 [↑](#endnote-ref-9)