Science in School
The European journal for science teachers

ISSUE 65 – November 2023 **Topics** Biology | Chemistry | Health

THAVIS 3D/unsplash

Not junk after all: the importance of non-coding RNAs

Zuzana Koskova, Miguel Hernandez

Originally assumed to be simply 'junk DNA', sections of the genome that don't encode proteins have been revealed as a source of many important non-coding RNA structures.

The central dogma of molecular biology is that DNA is used as a template to create messenger RNA (mRNA), which in turn is translated into proteins that build the tissues in our bodies and carry out the main functions of our cells and organs. In other words, DNA → mRNA → proteins. Interestingly, though, only 2% of the DNA in our whole genome codes for proteins! So, what does the other 98% of the human genome do? In the mid-1900s, it was widely believed that this was useless 'junk DNA'. However, in the late 20th century and the early 21st century, some of this this non-coding DNA has been shown to not only contain important regulatory elements for transcription, but also sequences that encode various non-coding RNAs that have functions in many cellular mechanisms.[1]

You might already know some non-coding RNAs

The first known classes of non-coding RNAs (ncRNAs) were ribosomal (rRNAs) and transfer RNAs (tRNAs), discovered in the 1950s. tRNAs are essential for protein synthesis, as they decode the mRNA sequence and translate it into an amino acid sequence. This translation of mRNAs into proteins occurs on the ribosome, which itself consists of large ribosomal RNAs and multiple ribosomal proteins (figure 1). In fact, ribosomes are so important for the cells that rRNAs are the most abundant RNAs in the cell!

Both tRNAs and rRNAs have complex 3D folded structures that are essential to their function, much like proteins.

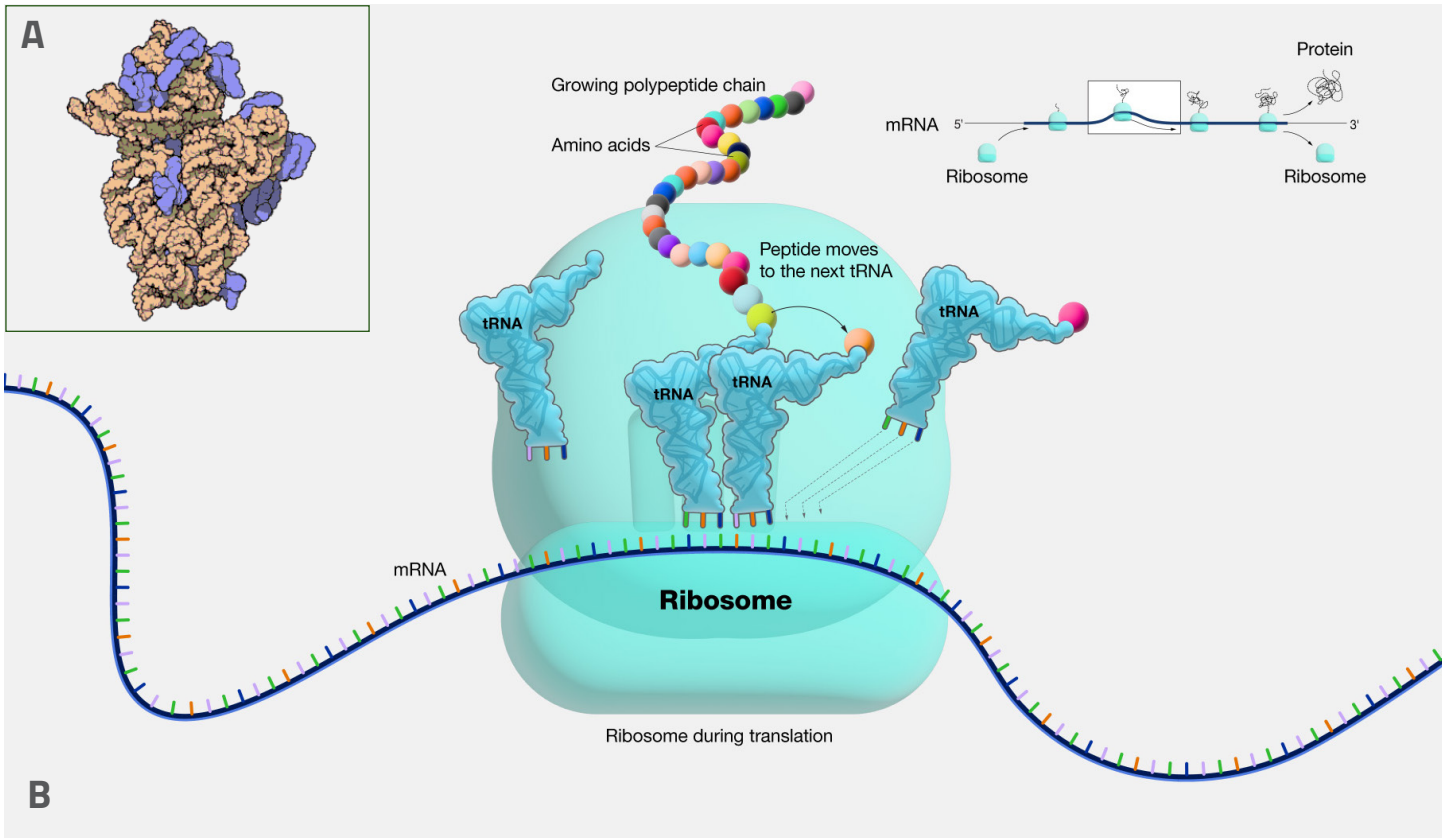

Figure 1: A) A ribosome consisting of large rRNAs (brown) and ribosomal proteins (purple).
B) On the ribosome, tRNAs decode the sequence of mRNAs and translate it into the amino acid sequence of a protein.
Images: Animation by David S. Goodsell, RCSB Protein Data Bank/Wikimedia, Public Domain (left) and National Human Genome Research Institute, Public Domain (right)

rRNAs and tRNAs form a group of so-called housekeeping ncRNAs because they are necessary to maintain fundamental cellular processes. A second group of non-coding RNAs was only discovered later in the 1990s. These RNAs, called regulatory ncRNAs, have diverse regulatory roles in the cell. Non-coding RNAs of less than 200 base pairs are called small non-coding RNAs (sncRNAs), while all non-coding RNAs longer than 200 base pairs are long non-coding RNAs (lncRNAs) (figure 3), and both can be further divided into groups based on their function.[2]

RNAs as regulators of gene expression

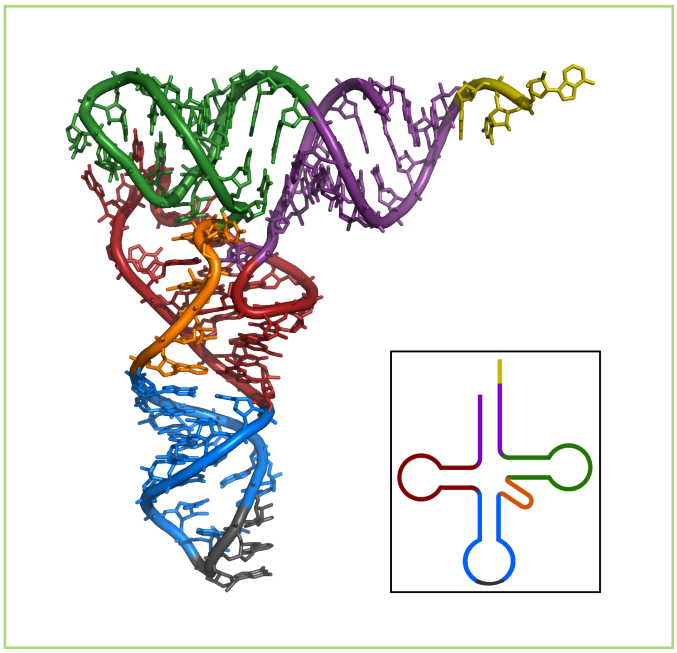

Figure 2: The 3D structure of a tRNA molecule from yeast, showing how the different loops in the primary structure (inset) fold together.
Image: Yikrazuul/Wikimedia, CC BY-SA 3.0

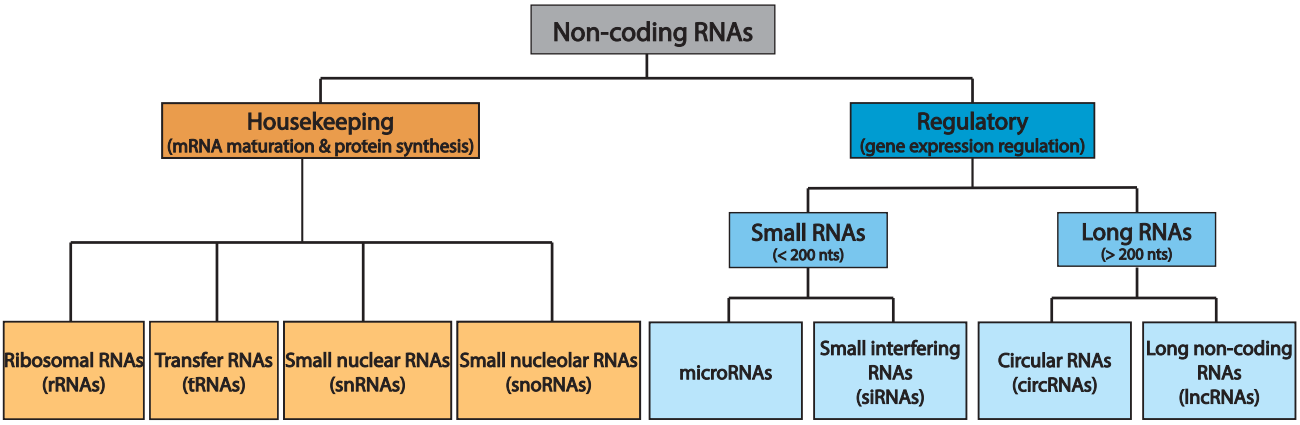

Figure 3: Classification of non-coding RNAs as housekeeping or regulatory, with some examples for each category. This is not a complete list of known ncRNAs, and new types of ncRNA are still being discovered.
Image courtesy of Zuzana Koskova

Instead of the expected uselessness of junk DNA, we understand now that non-coding RNAs are important for various cellular processes. But how do these RNAs fulfil their functions? In most cases, ncRNAs must interact with other molecules. For example, rRNAs interact with proteins during protein synthesis; however, it seems they can also work with almost any kind of molecule.

RNAs can even interact with other RNAs. One of the most exciting examples of RNAs interacting with each other are microRNAs, which are found in almost all organisms apart from bacteria and some yeast species. MicroRNAs consist of only 21–23 nucleotides and are described as having 'hairpin' structures because, before processing, they fold back on themselves like a hairpin. Despite their tiny size, they act as a guide for powerful scissors inside the cell by binding to a protein called RNA-induced silencing complex (RISC), which cuts up coding mRNAs. When the microRNA-bound RISC meets an mRNA with a complementary nucleotide sequence, it cuts the mRNA and prevents its translation into a protein (figure 4). Nowadays, scientists routinely use this RNA-silencing mechanism in their experiments to study how cells behave if a particular protein is removed.[3]

Another fascinating role of RNA has been discovered in the chromosomes of mammals. Human cells contain 23 pairs of chromosomes. For each chromosome pair, we receive one chromosome from our mother and one from our father. This means that we have two copies of each gene and normal cell function is based on the amount of protein that will be translated from the two gene copies. If a cell ends up with more or fewer than two of a particular chromosome, it can lead to serious health problems, such as Down Syndrome, which is caused by three copies of chromosome 21.

However, there is one special pair of chromosomes that doesn't always match: the sex chromosomes, which may be present as

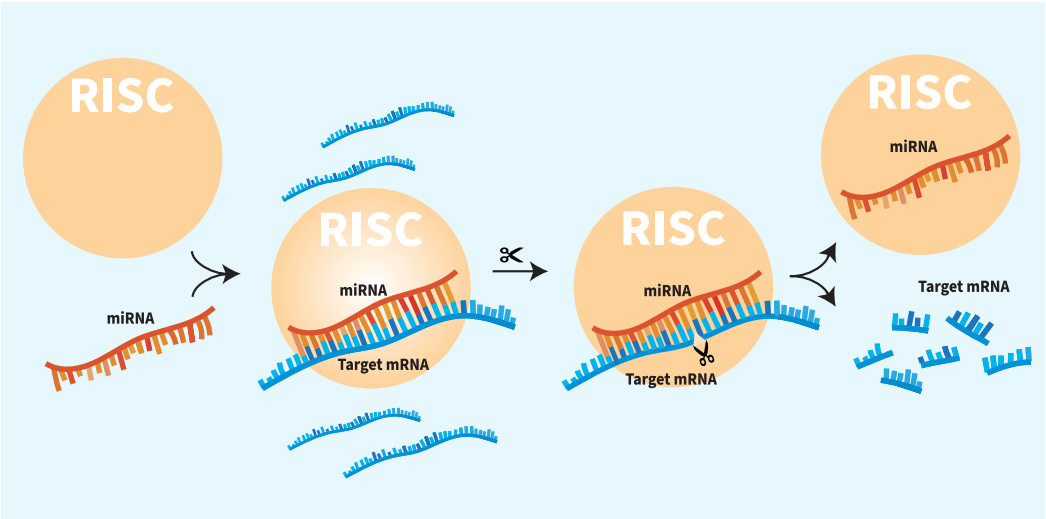

Figure 4: The RNA-induced silencing complex (RISC) protein binds to processed microRNAs, which guide RISC to mRNAs with the same sequence. RISC cuts these mRNAs and prevents their translation into proteins.
Image courtesy of Zuzana Koskova

two X chromosomes in female cells or one X and one Y in male cells. But how does the cell handle the different numbers of X chromosomes in different individuals? This problem is solved by setting the correct number of expressed X chromosomes to one in both sexes. In early development, the non-coding RNA Xist coats one of the X chromosomes in XX individuals and inactivates it (figure 5). In this way, female XX cells have as many active X chromosome genes as male XY cells.[4]

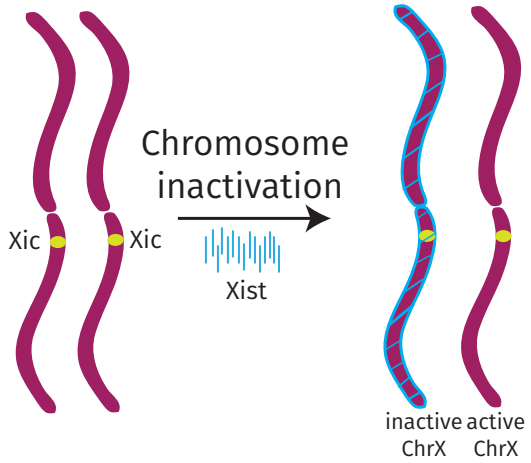

Figure 5: X chromosome inactivation is orchestrated by the non-coding RNA Xist. First, the X-inactivating centres (Xic), which are only found on X chromosomes, pair up. Next, the Xic on one of the X chromosomes is activated (although it's still not understood how the choice is made). The activated Xic enhances the production of Xist, which coats and inactivates that chromosome. This ensures that the number of proteins produced from genes located on the X chromosomes is equal between male and female cells.
Image courtesy of Zuzana Koskova

RNAs as data storage and catalysts

These examples of RNA's functions show the incredible complexity of our cells. How did this complexity evolve? In 1986, Walter Gilbert proposed that the first biomolecular processes were completely based on RNA, long before DNA and proteins evolved. This concept is called the 'RNA World' hypothesis.[5]

There are a number of arguments supporting this hypothesis. We already know that RNA molecules can carry genetic information just like DNA, although RNA degrades more easily so DNA provides more stable storage. However, some organisms, such as certain viruses, use the instability of RNA to their benefit by storing their genomes as RNA, which is more liable to mutate and thereby enables fast adaptation. Moreover, there are special RNA molecules called ribozymes, which have a folded structure like microRNAs but additionally possess a catalytic function like proteins: they can cut proteins and RNA molecules or accelerate enzymatic reactions. This means that although the fundamental principle of molecular biology is DNA → RNA → protein, RNA can fulfil the functions of all three units! The most famous ribozymes found in nature are the hammerhead ribozyme and hairpin ribozyme (figure 6). They were first identified in plant viruses,[6] where they are used to cut RNAs and even produce tiny infectious particles.[7] However, researchers have also found them in human cells, where they're probably involved in mRNA processing.[8] In the future, we might even be able to use them to treat cancer and other diseases.[9]

Another function that biomolecules based on RNA can perform in cells is mediating signalling within or between cells.

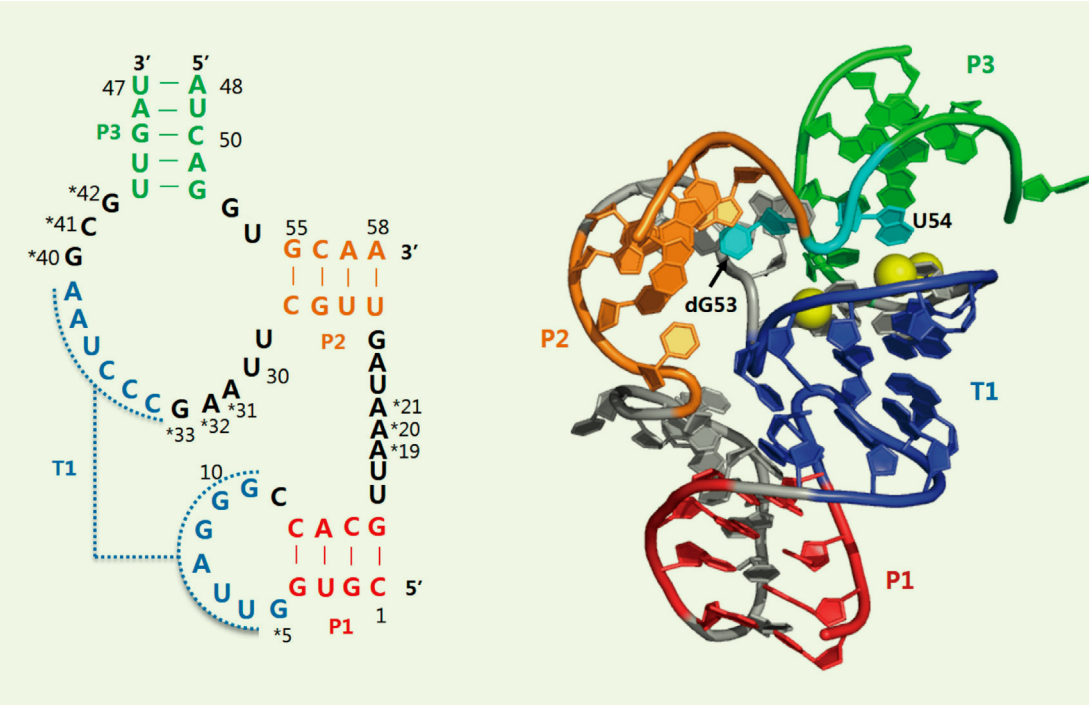

Figure 6: Secondary (left) and tertiary (right) structures of the env25 pistol ribozyme. Stem (P1–P3) and pseudoknot (T1) structure elements are indicated.
Image adapted from Ref. [10]

Proteins can perform this role but there are also a number of small signalling molecules, and some very important ones, such as cyclic AMP and cyclic di-GMP, are made from RNA building blocks.[11] Together with ribozymes, these molecules might represent the last remnants of the RNA World!

Sadly, while these remnants in our cells are fascinating to think about, it's incredibly difficult to collect direct evidence for or against the RNA World, which would have existed around 4 billion years ago.

This is just the beginning

We are only starting to reveal the functions and importance of RNA molecules in cellular processes. My own research project explores the function of RNAs on the surface of chromosomes during cell division. Other research groups are exploring the role of ncRNAs in diseases and how to use them in medical treatments.[12] Scientists have only recently finished sequencing the 3.2 billion nucleotides of the human genome[13] and there is still a great deal of work to do to understand it.

The story of non-coding RNAs shows how we thought we understood the genome, but then realized that nature is still much more complex than we ever expected. It is crucial to stay curious and have an open mind to notice all the unexpected and beautiful things hidden behind nature's complexity. ‹‹

Editor's note: Some parts of the introduction and conclusion were rephrased to avoid any misunderstanding concerning the nature of 'junk DNA', which is not the focus of this article.

Glossary

DNA: Deoxyribonucleic acid sequence carries the genetic information of the organism inside each cell

RNA: Ribonucleic acid is a less-stable nucleic acid than DNA, but it fulfils various important functions in the cell

mRNAs: Messenger RNAs serve as a transcript of the DNA and a template for its translation into proteins

tRNAs: Transfer RNAs decode the mRNA template into an amino acid sequence to make proteins into ribosomes

rRNAs: Ribosomal RNAs are the most abundant RNA species in the cell; together with ribosomal proteins they make up the structural components of the ribosomes

ncRNAs: Non-coding RNAs are RNA molecules that are not translated into proteins. These include tRNAs and rRNAs, as well as regulatory ncRNAs

Ribozymes: These RNAs have a catalytic function like enzymes; not to be confused with ribosomes.

References

[1] A BBC Future article on the role of the 'dark genome': https://www.bbc.com/future/article/20230412-the-mystery-of-the-human-genomes-dark-matter

[2] Palazzo AF, Lee EW (2015) Non-coding RNA: what is functional and what is junk? *Frontiers in Genetics* **6**: 2. doi: 10.3389/fgene.2015.00002

[3] Filipowicz W et al. (2005) Post-transcriptional gene silencing by siRNAs and miRNAs. *Current Opinion in Structural Biology* **15**: 331–341. doi: 10.1016/j.sbi.2005.05.006

[4] An article on the inactivation of the X chromosome on the UCLA News website: https://newsroom.ucla.edu/releases/rna-silences-x-chromosome

[5] Higgs P, Lehman N (2015) The RNA World: molecular cooperation at the origins of life. *Nat Rev Genet* **16**: 7–17. doi: org/10.1038/nrg3841

[6] Fedor MJ (2000) Structure and function of the hairpin ribozyme. *J Mol Biol* **297**: 269–291. doi: 10.1006/jmbi.2000.3560

[7] Hammann C et al. (2012) The ubiquitous hammerhead ribozyme. RNA 18: 871–885. doi: 10.1261/rna.031401.111

[8] de la Peña M, García-Robles I (2010). Intronic hammerhead ribozymes are ultraconserved in the human genome. EMBO Rep. 11: 711–716. doi: 10.1038/embor.2010.100

[9] An article on the trans-splicing ribozyme as genetic therapy on *Advanced Science News*: https://www.advancedsciencenews.com/therapeutic-applications-group-intron-based-trans-splicing-ribozymes/

[10] Lee K-Y, Lee B-J (2017) Structural and Biochemical Properties of Novel Self-Cleaving Ribozymes. *Molecules* **22**: 678. doi: 10.3390/molecules22040678

[11] Nelson JW, Breaker RR (2017) The lost language of the RNA World. *Sci Signal* **10**: eaam8812 doi:10.1126/scisignal.aam8812

[12] Bhatti GK et al. (2021) Emerging role of non-coding RNA in health and disease. *Metab Brain Dis* **36**: 1119–1134. doi: 10.1007/s11011-021-00739-y

[13] A BBC Future article on the first sequencing of a complete human genome: https://www.bbc.com/future/article/20230210-the-man-whose-genome-you-can-read-end-to-end

Resources

- Watch a nice 3D animation of transcription and translation.
- View a 3D animation of the production and function of miRNA.
- Read a Nature News article highlighting the fact that the RNA World is just one hypothesis in a field of research that is still very much active.
- Read an article on the key role of cyclic dinucleotide used as second messenger signalling molecules and its potential in healthcare: Römling U, Gekara N (2022) Ancient signal-sensing mechanisms based on cyclic dinucleotide molecules may lead to breakthroughs in human healthcare. *Science in School* **56**.
- Learn more about Early Earth research: Extance A (2020) Finding the recipe for life on Earth. *Science in School* **49**: 13–19.
- Read an article on different techniques to resolve and predict protein structures: Heber S (2021) From gaming to cutting-edge biology: AI and the protein folding problem. *Science in School* **52**.
- Read about click chemistry and its use on biomolecules: Godinho T (2022) Click does the trick: understanding the 2022 Nobel Prize in Chemistry. *Science in School* **60**.
- Play a game where jigsaw puzzles are used to explain genome reconstruction: Mastrorilli E (2021) Microbial genome puzzles. *Science in School* **51**.
- Try this role-playing activity to understand how research projects are funded and the importance of basic research: McHugh M (2022) What is it good for? Basic versus applied research. *Science in School* **55**.
- Teach students how to collect relevant data regarding a gene from biological databases: Grazioli C, Viale G (2022) A chromosome walk. *Science in School* **57**.

AUTHOR BIOGRAPHY

Zuzana Koskova is a PhD student at the European Molecular Biology Laboratory (EMBL) in Heidelberg. She holds a Bachelor's degree in Biosciences and a Master's degree in Molecular and Cell Biology from the University Heidelberg, Germany. In her scientific projects, she has worked on RNA-binding proteins and non-coding RNAs involved in cell division.

Miguel Hernandez is a PhD student in a computational lab focused on systems biomedicine at the University Hospital, Heidelberg. He holds a Bachelor's degree in Biology from the University of Guadalajara, Mexico, and a Master's degree in Molecular and Cellular Biology from University Heidelberg, Germany. Miguel is passionate about biology and his work focuses on immunology and gene regulation in the context of human diseases.

The authors of the article