SAPIENZA
Università di Roma
Facoltà di Scienze Matematiche Fisiche e Naturali

DOTTORATO DI RICERCA
IN GENETICA E BIOLOGIA MOLECOLARE

XXXV Ciclo
(A.A. 2022/2023)

Decoding the role of the lncRNA HOTAIRM1 in human motor neurons

Dottorando:

Paolo Tollis

Docente guida: Elisa Caffarelli

Dott.ssa Elisa Caffarelli

Coordinatore:

Prof. Fulvio Cruciani

Tutore:

Dott. Pietro Laneve

"Ever tried.
Ever failed.
No matter.
Try again.
Fail again.
Fail better."

(Samuel Beckett)

Index

Glossary

ALS: amyotrophic lateral sclerosis. A progressive neurodegenerative disease that affects nerve cells in the brain and spinal cord. The disease is characterized by the death of motor neurons, which are responsible for controlling voluntary muscle movement. As the motor neurons die, the muscles they control weaken and waste away, leading to a loss of movement and eventually to paralysis.

AMT: 4'-aminomethyl-4,5',8-trimethylpsoralen. Chemical compound used in RNA pull-down experiments to covalently link direct RNA-RNA interactions occurring in living cells.

ceRNA: competing endogenous RNA. RNA molecules that regulate other RNA transcripts by competing for shared microRNAs.

ChIP: chromatin immunoprecipitation. An antibody-based technology used to selectively enrich specific DNA-binding proteins along with their DNA targets.

circRNA: circular RNA. A type of single-stranded RNA which, unlike linear RNAs, forms a covalently closed continuous loop.

CLIP: crosslinking immunoprecipitation. A method used in molecular biology that combines UV cross-linking with immunoprecipitation to analyze protein interactions with RNA.

CNS: central nervous system. The portion of the nervous system consisting primarily of the brain and spinal cord.

CRISPR/Cas9: a gene-editing technology that can target and edit parts of the genome with high accuracy.

DIV: days in vitro. Number of days a neuronal cell culture has been maintained.

DOX: Doxycycline. Inducible Tet-On system widely used to control gene expression in mammalian cells.

EMT: epithelial–mesenchymal transition. Process by which epithelial cells lose their cell polarity as well as cell–cell adhesion and gain migratory and invasive properties to become mesenchymal stem cells; these are multipotent stromal cells that can differentiate into a variety of cell types.

eRNA: enhancer RNA. Class of relatively long non-coding RNA molecules (50-2000 nucleotides) transcribed from the DNA sequence of enhancer regions.

ESC: embryonic stem cell. Derived from the inner cell mass of a blastocyst are pluripotent stem cells with unique properties of pluripotency and self-renewal. They can divide indefinitely *in vitro*, while maintaining the capacity to generate all the cell types of an adult organism.

exRNA: extracellular RNA. A special form of RNA in the body. RNA carries information about genes and metabolic regulation in the body, which can reflect the real-time status of cells.

FDR: false discovery rate. Method for conceptualizing the rate of type I errors in null hypothesis testing when performing multiple comparisons.

GTEx: Genotype-Tissue Expression portal. A comprehensive public resource to study tissue-specific gene expression and regulation.

hnRNP: heterogeneous nuclear ribonucleoprotein. Molecular complexes of RNA and protein localized in the cell nucleus during gene transcription and subsequent post-transcriptional modification of the newly synthesized RNA.

HSC: hematopoietic stem cell. Cell that has the capacity to self-renew and the potential to differentiate into all types of the mature blood cell types.

IF: immunofluorescence. A histochemical staining technique used for demonstrating the presence of antibodies bound to antigens in tissues or serum. IF relies on the use of antibodies chemically labeled with fluorescent dyes to visualize molecules under a light microscope.

iPSC: induced pluripotent stem cell. Human adult cells that have been reprogrammed back into an embryonic-like pluripotent state through the exogenous expression of particular transcription factors.

KO: knockout. A genetic technique in which one of an organism's genes is made inoperative ("knocked out" of the organism).

lncRNA: long non-coding RNA. RNA molecule with length exceeding 200 nucleotides lacking protein-coding potential.

miRISC: miRNA-induced silencing effector complex. A multi-protein complex that uses microRNAs to identify mRNAs targeted for repression.

miRNA: microRNA. Small, single-stranded, non-coding RNA molecules containing 21 to 23 nucleotides. Found in plants,

animals and some viruses, miRNAs are involved in RNA silencing and post-transcriptional regulation of gene expression.

MN: motor neuron. Specialized neuronal cells located in the central nervous system (CNS) controlling a variety of downstream targets involved in voluntary muscle movement.

MND: motor neuron disease. A rare condition that progressively damages motor neurons. Believed to be caused because of a combination of environmental, lifestyle and genetic factors.

MNP: MN progenitor. A type of stem cell that can give rise to motor neurons.

mRNA: messenger RNA. A type of RNA molecule that carries genetic information from DNA to the ribosome, where it is used to synthesize proteins.

NAT: natural antisense transcript. RNAs encoded within a cell that have transcript complementarity to other RNA transcripts.

ncRNA: non-coding RNA. RNA molecule lacking protein-coding potential.

ORA: over representation analysis. A method used to identify significantly enriched Gene Ontology terms in a set of genes.

ORF: open reading frame. A continuous stretch of DNA or RNA that can be translated into a protein.

piRNA: piwi-interacting RNA. A type of small non-coding RNA molecule involved in the regulation of gene expression, named for their association with a family of proteins called piwi proteins, which are thought to play a role in their biogenesis.

PROMPT: promoter upstream transcript. Class of RNAs which are heterologous in length and produced only upstream of the promoters of active protein-coding genes.

RAP-MS: RNA antisense purification coupled with mass spectrometry. It allows the selective precipitation of endogenous RNA complexes from cell extracts through hybrid capture with biotinylated antisense oligonucleotides used to identify the protein interactome of a specific RNA of interest.

RBP: RNA-binding protein. Proteins that bind to the double or single stranded RNA in cells and participate in forming ribonucleoprotein complexes.

RIP: RNA immunoprecipitation. A powerful method to study the physical association between individual proteins and RNA molecules *in vivo*.

RNA-seq: RNA Sequencing. Approach for transcriptome profiling that uses next-generation sequencing (NGS) to reveal the presence and quantity of RNA in a biological sample at a given moment.

rRNA: ribosomal RNA. Class of non-coding RNAs which is the primary component of ribosomes, essential to all cells.

siRNA: short interfering RNA. Class of double-stranded non-coding RNA molecules, operating within the RNA interference pathway.

SMA: spinal muscular atrophy. A rare neuromuscular disorder that results in the loss of motor neurons and progressive muscle degeneration.

snoRNA: small nucleolar RNA. A class of small RNA molecules that primarily guide chemical modifications of other RNAs, mainly ribosomal RNAs, transfer RNAs and small nuclear RNAs.

snRNA: small nuclear RNA. A class of small RNA molecules that are found within the splicing speckles and Cajal bodies of the nucleus of eukaryotic cells.

spMN: spinal motor neuron. Cells found in the spinal cord, where they receive input from other nerve cells and send output to the muscles via axons that extend out from the spinal cord. The axons of spinal motor neurons form synapses with muscle fibers, and the activation of these synapses leads to muscle contraction. Spinal motor neurons are important for controlling movements of the limbs, trunk, and other skeletal muscles, and they are essential for the maintenance of posture and balance. Dysfunction of spinal motor neurons can lead to muscle weakness, atrophy, and other problems.

TF: transcription factor. A protein that controls the rate of transcription of genetic information from DNA to messenger RNA, by binding to a specific DNA sequence in the nucleus.

tRNA: transfer RNA. An adaptor molecule composed of RNA that serves as the physical link between the mRNA and the amino acid sequence of proteins.

UCSC Genome Browser: University of California Santa Cruz Genome Browser (genome.ucsc.edu); it is a popular web-based viewer for genome sequence data and annotations.

UPR: unfolded protein response. A cellular stress response related to the endoplasmic reticulum stress.

UTR: untranslated region. It refers to either of two ends, one on each side of a coding sequence on a strand of mRNA. If it is found on the 5' side, it is called the 5' UTR (or leader sequence), or if it is found on the 3' side, it is called the 3' UTR (or trailer sequence).

Summary

The mammalian genome produces thousands of long non-coding RNAs (lncRNAs), which have been demonstrated to be fundamental in the control of many biological processes. These molecules play a crucial role in the multilayered regulation of physiological and disease-related gene expression programs, having significant implications in shaping central nervous system (CNS) complexity.

Neuronal differentiation is a timely and spatially regulated process, relying on precisely orchestrated gene expression control. The coordinated activity of transcription factors and non-coding RNAs (ncRNAs), organized in intricate regulatory networks, drives cell fate specification ensuring correct and specific neuronal functions.

We previously described,[1] at both the molecular and functional level, the lncRNA *nHOTAIRM1* as a neuronal-enriched transcript, which is upregulated during *in vitro* neuronal differentiation and highly expressed in post-mitotic motor neurons (MNs). We demonstrated that the nuclear *nHOTAIRM1*, even if much less abundant than its cytoplasmic counterpart, it is involved in the achievement of correct neuronal differentiation timing as an epigenetic regulator of *NEUROG2* expression.[1] Remarkably, among all human brain tissues, *nHOTAIRM1* is specifically expressed in the spinal cord. Consistently, we found that *nHOTAIRM1* accumulates in MN-enriched ventral spinal cord lineages differentiated from human induced pluripotent stem cells (iPSCs).[1] All this evidence prompted us to further investigate the role of the highly expressed *nHOTAIRM1* specifically on MN generation and/or function, to ultimately determine whether its deregulation affects MN differentiation and activity. To experimentally address these questions, here we applied a genome editing-based loss-of-function approach to a model system that efficiently recapitulates spinal MN differentiation, and we identified key *nHOTAIRM1* target genes implicated in MN maturation, morphology and activity.

Our findings allowed us to conclude that *nHOTAIRM1* directs multiple crucial aspects of MN physiology, from their development to the acquisition of appropriate morphological features and motor function.

1. Introduction

1.1 Non-coding RNA: rewriting the rules

In the late 1950s, Elliot Volkin and Lawrence Astrachan thoroughly defined RNA as a DNA-like molecule generated from DNA itself, establishing for the first time the connection between DNA and RNA.

In the same years, when James Watson and Francis Crick discovered the double-helix structure of DNA, following the X-ray crystallographic studies of Rosalind Franklin, RNA was considered a mere unstable intermediating molecule, a transitory step in the information flow from DNA to proteins. The Central Dogma of Molecular Biology comprised the transcription of a DNA gene into RNA, containing the information program for the synthesis of a particular protein, that took place in the cytoplasm.[2]

Discovered in the early 1960s, messenger RNA (mRNA) was for a long time considered the most prevalent cellular RNA species, along with ribosomal RNA (rRNA)[3] and transfer RNA (tRNA), this latter being the first non-coding RNA (ncRNA) to ever be characterized,[4,5] able to form RNA-RNA base pairing interactions with mRNAs. At that time, the scientific research, deeply rooted in a protein-coding point of view, was devoted to the investigation of the molecular functions of proteins.

In the following decades, different small ncRNA species were discovered, such as small nuclear RNAs (snRNAs)[6] and small nucleolar RNAs (snoRNAs)[7], that were essential for the maturation of mRNA and rRNA, respectively. Later, in the early 1980s, Thomas Cech and Sidney Altman discovered ribozymes, demonstrating for the first time ever that RNA itself could act as the only one catalyst of an entire chemical reaction.[8]

Only in the early 1990s, the discovery of small RNAs, namely microRNAs (miRNAs) and short interfering RNAs (siRNAs), completely rewrote the rules of gene expression regulation, adding a further role to RNA as regulatory molecule, and providing a new

dimension to our understanding of complex gene regulatory networks.[9–12]

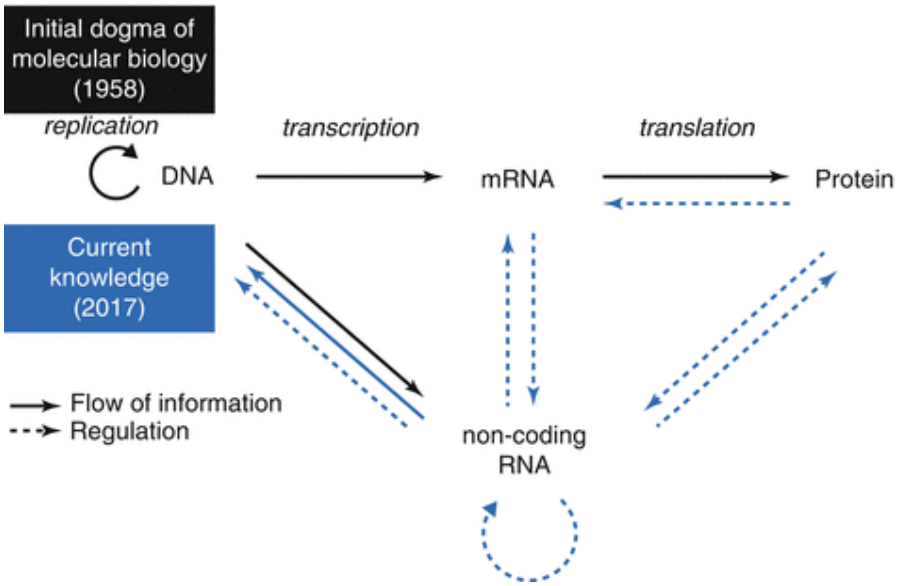

Figure 1. Initial and current dogma of molecular biology.[2]

These discoveries pointed the attention of the scientific community to the role of ncRNAs, which - in that scientific landscape – were commonly regarded as by-products of massive transcription with less biological meaning. Indeed, originally, most of the genome was referred to as "selfish" or "junk DNA" and have been considered as useless evolutionary fossils for almost 20 years.
In fact, only at the beginning of the twenty-first century, the pervasive transcription of genomes, the surprising number of long non-coding RNAs (lncRNAs) and the significance of non-coding sequences became evident.
Since then, ncRNAs started to mark their own progress into Cellular and Molecular Biology, so much that today the perspective on gene expression regulation has dramatically changed from a coding to a non-coding point of view. In this scenario, RNA has been the object of an unprecedented

reevaluation, becoming so much more than a simple sequence template for protein synthesis.[13]

The Human Genome Project represented a milestone in this direction, since it revealed that most of the genome is actively and pervasively transcribed, whether it encodes for proteins or not.[14]

The developments of more sensitive sequencing technologies, the improvements of machine learning methods for RNA secondary structure prediction[15] and the advances in data open sharing, which launched the "omics" revolution in the new millennium, made a huge contribution towards our modern conception of the RNA world.[16]

Today, thanks to the efforts of the FANTOM and the ENCODE Projects,[17–21] we know that only approximately 2% of the human genome encodes for proteins, while the remaining 98% is composed by non-coding sequences,[22] whose increasing number correlates with the complexity of organisms.[23] In this context, novel kinds of ncRNAs that exhibit remarkable developmental-, cell-type- and also disease-specific expression, such as miRNAs, piwi-interacting RNAs (piRNAs), lncRNAs – which include also circular RNAs (circRNAs) – and extracellular RNAs (exRNAs) have been identified as multifaceted players in gene expression regulation.[24–28]

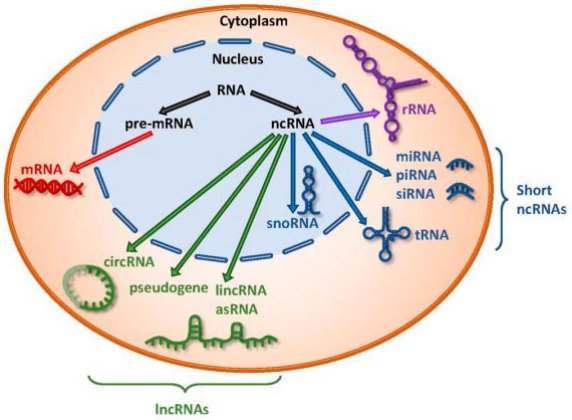

Figure 2. Different classes of ncRNAs.[29]

1.2. LncRNAs

Genomes are extensively transcribed and give rise to thousands of lncRNAs which are defined as RNA transcripts longer than 200 nucleotides that are not translated into functional proteins.[26,30] This broad definition includes a vast and extremely heterogeneous family of transcripts - each with a unique biogenesis, genetic origin, function and localization - that are involved in the regulation of a wide range of fundamental physiological processes, through different molecular mechanisms. This explains why their dysfunction has such a huge impact on several pathologies[31]. Today, in fact, they are known as directors of complex gene expression programs at the transcriptional and post-transcriptional level, so that they became the new research hotspot in the RNA world.[32–34]

1.2.1 General features

According to the GENCODE v7 catalog, to date we know that the human genome is estimated to produce ~60,000 lncRNAs.[19,27,35]
These molecules have little to no coding potential,[19] most of them are transcribed by RNA polymerase II[36] and, similarly to mRNAs, they undergo splicing, 5'-capping and 3'-polyadenylation,[37,38] and produce numerous splicing isoforms.[39,40] Differently from mRNAs, instead, lncRNAs tend to be shorter in length, have fewer but longer exons and are generally expressed at relatively low levels, enough to exert their specific regulative functions.
From the evolutionary point of view, it has been demonstrated that lncRNA main sequences evolve faster than those of protein-coding genes, but slower than those of intronic and intergenic transcripts.[19,41]
In fact, lncRNAs show a poor primary sequence conservation while still preserving a structural homology that ensures functional integrity throughout evolution. However, the discovery that the conservation of lncRNA promoter regions is higher than that of protein-coding genes tells us that lncRNA gene expression control has persisted throughout evolution.

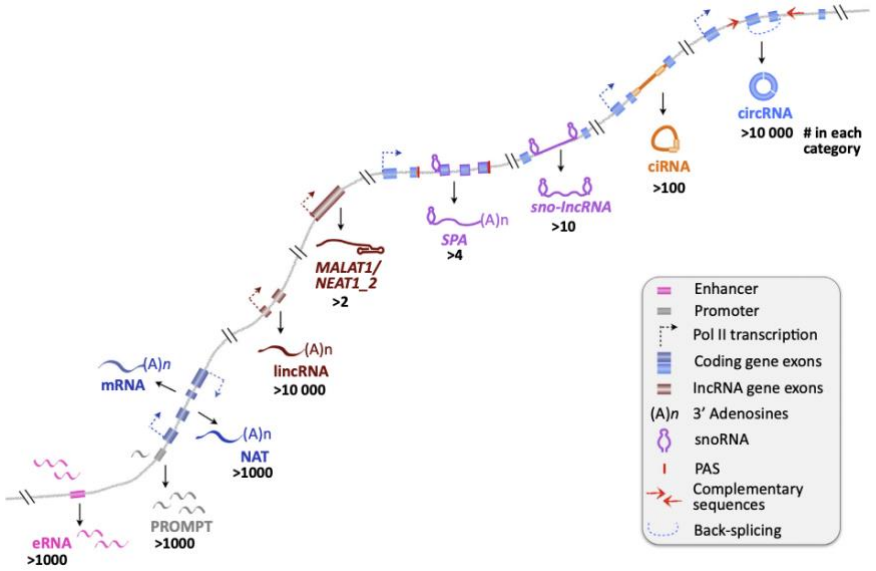

Figure 3. The vast diversity in the generation of lncRNAs. [42]

A general classification of lncRNAs can be made based on their relative position in the genome with respect to protein-coding genes and their different mechanisms of processing. Examples include promoter upstream transcripts (PROMPTs), enhancer RNAs (eRNAs), long intervening/intergenic ncRNAs (lincRNAs), and natural antisense transcripts (NATs), that are transcribed from promoter upstream regions, enhancers, intergenic regions, and the opposite strand of protein-coding genes, respectively[42] (Fig. 3).
These molecules follow unique biogenesis and processing, so that a variety of lncRNAs with non-canonical structures can also be generated from long primary transcripts that undergo unusual RNA processing pathways such as sno- and SPA-lncRNA (Fig. 3).[43,44]
Today it is well established that cells also express endogenous circRNAs along with their linear counterparts. CircRNAs are covalently closed single-stranded RNA molecules that arise from special back-splicing events.[45,46] It is interesting to note that while linear RNAs have a short half-life that enables quick modulation of their activity, circRNAs, which lack the 5' and 3' ends, have

greater stability and are able to sponge RNA-binding proteins (RBPs) and miRNAs, and some of them have been shown to be incredibly powerful protein translation platforms,[47] indicating that these species may be involved in a wide range of biological processes.

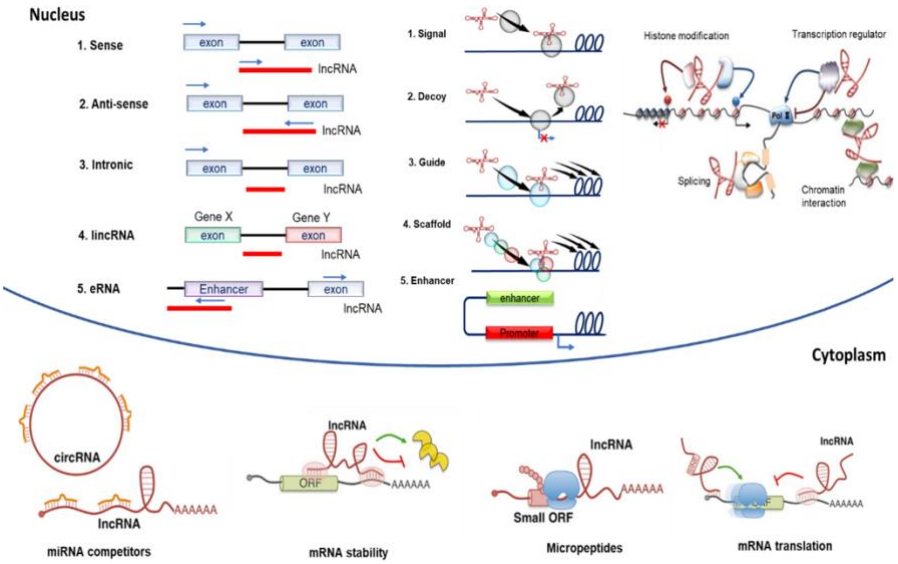

Figure 4. Classification, localization and functions of lncRNAs. [48]

Based on lncRNA genomic origin, these molecules can be divided into the following subgroups: i) sense lncRNAs, transcribed from the sense strand of protein-coding genes; ii) antisense lncRNAs, produced from the opposite strand of the coding region; iii) bidirectional lncRNAs, transcribed from the same promoter as a protein-coding gene but in the opposite direction; iv) sense-overlapping lncRNAs, that overlap regions that are potential hotspots for the splicing of their host protein-coding genes; v) intronic lncRNAs, produced in introns of a coding gene that does not overlap any exons; vi) intergenic lncRNAs, originated from

sequences which do not overlap protein-coding genes; and vii) enhancer lncRNAs, which are transcribed from enhancer genomic loci[48] (Fig. 4).

1.2.2 LncRNA function and mechanisms of action

LncRNAs are versatile regulators of a variety of fundamental biological processes, such as development, cell differentiation and growth, by controlling gene expression at the epigenetic, transcriptional, or post-transcriptional levels through a multitude of mechanisms, depending on their subcellular localization.[32–34]
In fact, these molecules can be exclusively located in the nucleus, or in the cytoplasm, or expressed in both compartments in which they exert different functions. This makes it crucial to consider their molecular context when attempting to investigate their mechanisms of action.

1.2.3 Nuclear lncRNAs

Nuclear lncRNAs are generally present in their free form in the nucleoplasm or can be linked to chromatin.[49,50]
In the latter case, the amount of RNA associated with chromatin is twice as high compared to DNA.[51] The negative charges of RNA molecules can neutralize the positively charged histone tails, leading to lncRNA-mediated attenuation of electrostatic compaction of chromatin structure, which is directly responsible for a rapid switch in gene expression.[52]
In this regard, lncRNAs are conventionally classified as *cis*- or *trans*-acting molecules, when they operate in proximity or far from their site of transcription, respectively.
The perfect example of a *cis*-acting lncRNA that operates at the epigenetic level is represented by the lncRNA *XIST* (X-inactive specific transcript),[53] that is involved in dosage compensation in

mammals. In this process, the lncRNA, that is expressed from one of the two X chromosomes in female cells, acting as a scaffold (Fig. 4) for the recruitment of chromatin remodelers and repressor complexes, is able to alter the entire chromatin structure of the chromosome, resulting in its total transcriptional silencing.[54]

Another leading character in the regulation of nuclear architecture is the lncRNA *MALAT1* (Metastasis Associated Lung Adenocarcinoma Transcript 1).[55,56] This lncRNA is particularly abundant in nuclear speckles, in which it acts as a molecular scaffold (Fig. 4) interacting both with Cbx4, a component of an epigenetic repressor, and with the SR splicing factors, thus regulating both transcription and splicing of pre-messenger RNAs.[56]

Along with nuclear speckles, another noteworthy subnuclear compartment formed around lncRNAs are paraspeckles. Today these ribonucleoprotein bodies that originate from liquid-liquid phase separation are a paradigm for subnuclear bodies involved in the regulation of gene expression. Within these structures, paraspeckle proteins are organized by the lncRNA *NEAT1* (Nuclear Enriched Abundant Transcript 1) to finely regulate the expression of certain genes by nuclear retention of their mRNAs.[57–59]

Nuclear lncRNAs can also act as molecular decoy (Fig. 4). That is the case of the lncRNA *PANDAR* (Promotor of CDKN1A Antisense DNA damage Activated RNA)[60], that reduces the apoptotic response modulating PTBP1 protein, a member of the heterogeneous nuclear ribonucleoproteins (hnRNP) family, known to be involved in alternative splicing regulation.[61]

An exquisite example of *trans*-acting lncRNA is given by the lncRNA *HOTAIR* (*HOX* transcript antisense RNA). This transcript, that is dysregulated in the majority of human cancers, is transcribed from the *HoxC* gene cluster, and it has been reported to directly recruit PRC2 (Polycomb repressive complex 2) and CoREST/LSD1 epigenetic complexes in order to mediate transcriptional silencing of *HoxD* gene cluster. [62–65]

An active regulatory role in hematopoietic stem cell (HSC) fate regulation has been lately ascribed to *LncHSC-2* (HSC-enriched lncRNA 2). This nuclear *trans*-acting lncRNA works as an enhancer lncRNA (Fig. 4), tuning the balance between self-renewal and differentiation in HSC[66]. Its mechanism of action implicates to bundle together both the transcription factor TCF3/E2A and promoter-proximal regions of distant genes related to maturation of myelo-lymphoid progenitors, through which *lncHSC-2* controls HSC fate[66].

1.2.4 Cytoplasmic lncRNAs

While nuclear lncRNAs and their mechanisms of action have been extensively studied, cytoplasmic lncRNAs are substantially less well understood. Following export to the cytoplasm, lncRNAs can associate with RBPs originating cytoplasmic lncRNA-associated ribonucleoprotein complexes (lncRNPs) that govern cytoplasmic events with dramatical effects on gene expression, cellular structure maintenance and cellular functions.

These cytoplasmic lncRNPs: i) regulate the stability and/or translation of specific mRNAs; ii) act as RBP decoy and/or miRNA decoy, in order to reduce the cytosolic availability of specific cytoplasmic factors and/or miRNAs; iii) regulate protein turnover by serving as platforms that facilitate the presentation of specific RBPs to the protein degradation machinery; iv) can regulate localization and dynamics of specific membrane(less) organelles and/or extracellular vesicles; v) lead to conformational changes that activate signaling molecules/molecular pathways (Fig. 5).[67]

A particularly well-known example is the lncRNA *NORAD* (Non-coding RNA activated by DNA damage), an abundant and highly conserved cytoplasmic lncRNA that regulates the stability of several mRNAs. Thanks to its capacity to sequester a significant fraction of the total cellular pool of PUMILIO proteins PUM1 and

PUM2 - which are RBPs that act as negative regulators of gene expression[68] - *NORAD* acts as an RBP decoy to finely regulate a large set of target transcripts that play a critical role in maintaining the fidelity of chromosome transmission and genome stability.[69]
A typical role for cytoplasmic lncRNAs is the competing endogenous RNA (ceRNA) involved in miRNA sponge activity. This is employed by the intriguing cytoplasmic lncRNA *PNUTS* (Phosphatase 1 Nuclear Targeting Subunit), a non-coding isoform of the protein-coding gene *PPP1R10*, produced by HNRNPE1-mediated alternative splicing.[70] *PNUTS* contains seven miR-205 binding sites, which reduce the availability of miR-205 to bind and suppress the zinc finger E-box-binding homeobox 1 (ZEB1) and ZEB2 mRNAs during epithelial–mesenchymal transition (EMT).[71]
Another emerging role for cytoplasmic lncRNAs is that of regulators of mRNA stability. This is the case of the lncRNA *TINCR* (tissue differentiation-inducing non-protein coding RNA) that positively regulates the expression of its mRNAs targets through a STAU1-mediated stabilization mechanism[72], and that of the lncRNAs *1/2-sbsRNAs* (half-STAU1-binding site RNAs) that instead are responsible for STAU1-mediated mRNA decay.[73]
Regarding lncRNA-mediated regulation of mRNA translation, a well-known example is given by the human lincRNA *p21*, also known as tumor protein p53 pathway corepressor, that can negatively regulate the translation of *CTNNB1* (β-catenin) and *JUNB* (AP-1 Transcription Factor Subunit) transcripts by imperfect base pairing (both at 5′ and 3′-untranslated regions (UTRs)) when the levels of RNA-binding protein HuR (human antigen R) are reduced.[74]
Cytoplasmic lncRNAs are also capable of modulating signal transduction pathways by binding specific signaling molecules. An amazing example is given by the lncRNA *NKILA* (NF-κB-interacting lncRNA), that regulates NF-κB signaling and represses cancer-associated inflammation. This lncRNA works forming two distinct hairpins, hairpin A (nucleotides 322–359) and hairpin B (nucleotides 395–418), which both bind to p65, in order to

modulate T cell activation-induced cell death by inhibiting NF-κB activity.[75]

Finally, some lncRNAs may contain short open reading frames (sORFs) (Fig. 4) which allow them to encode small, bioactive peptides, like the micropeptide myoregulin (MLN), encoded by a muscle-specific lncRNA, that directly interacts with the SERCA membrane pump to inhibit the transport of Ca^{2+} into the sarcoplasmic reticulum (SR) of skeletal muscle cells.[76]

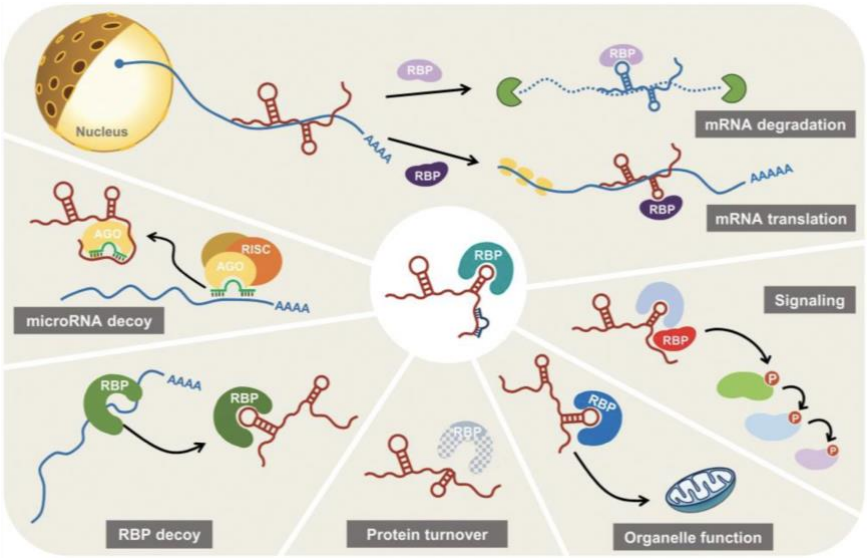

Figure 5. Regulation of gene expression by cytoplasmic lncRNAs.[67]

1.3 LncRNAs in motor neuron development and disease

Mammalian nervous system is one of the most sophisticated and complex systems existing in living organisms. It is composed by an incredible array of neuronal and glial cell types, all of them subject to an extremely fine and well-orchestrated gene expression regulation occurring at multiple levels, from transcription to RNA processing, translation, and decay. [77]

In this scenario, lncRNAs are the main actors that direct the regulatory complexity that becomes highly specific among different brain tissues and cell types.

According to the most recent discoveries in transcriptomics, there is a direct relationship between the evolutionarily increased complexity of the human brain and the expanding number of lncRNAs, which are responsible for the regulation of a wide range of biological processes including neuronal development, differentiation, and function.[77,78]

Remarkably, 40% of the lncRNAs identified in the human genome (corresponding to 4,000–20,000 lncRNA genes) are specifically expressed in the brain.[19,79,80]

During development, different classes of lncRNAs exhibit distinct spatiotemporal patterns to promote neuronal differentiation and maturation, which leads to the generation of functional subsets of neurons.

Motor neurons (MNs) are neuronal cells located in the central nervous system (CNS) controlling a variety of downstream targets. They can be divided in two categories: upper MNs, which originate from the cerebral cortex and establish glutamatergic connections, and lower MNs, which are located in the brainstem and the spinal cord and have a cholinergic activity; both of them control essential functions such as movement, breathing and swallowing.[81] Among lower MNs, spinal MNs (spMNs) have undergone extensive research in the last decade since their degeneration is a pathological hallmark of severe neurodegenerative diseases such as

amyotrophic lateral sclerosis (ALS) and spinal muscular atrophy (SMA).[82]

These highly polarized cells are located in the ventral horn of the spinal cord. Thanks to their axonal projections that extend through several meters - making them the longest known cell type – they are able to control effector muscles in the periphery, representing the ultimate and irreplaceable component of the locomotor neuronal circuitry.[83]

In the last decade, considerable focus has been placed on MN formation and function, since these cells are particularly sensitive to severe degenerative diseases, commonly known as motor neuron diseases (MNDs), where their degeneration produces a variety of debilitating behavioral outcomes.[84]

LncRNAs influence several aspects of MN pathophysiology, such as development, neurite outgrowth, chemical synaptic transmission, and even memory consolidation and ageing.[19,85–88]

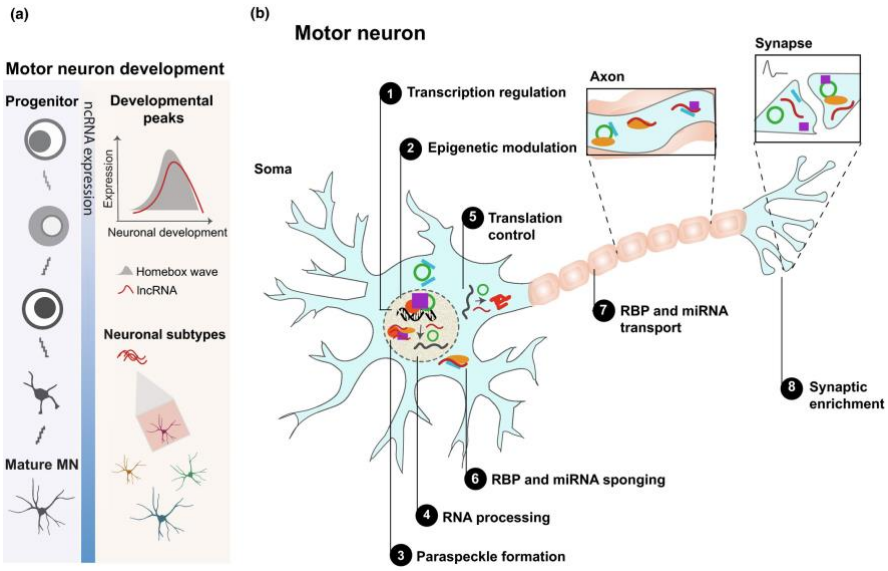

Figure 6. The role of lncRNAs in motor neuron development. (Adapted from [82])

An exquisite example of epigenetic regulation by lncRNAs in MNs (Fig. 6) is the lncRNA *Meg3* (maternally expressed gene 3), produced by the imprinted mammalian *Dlk1-Dio3 locus*.[89] This lncRNA is induced in differentiating embryonic stem cell (ESC)-derived MNs and also particularly enriched in the mouse spinal cord. Through the interaction with the PRC2–Jarid2 complex, it regulates the expression of important Homeobox (*Hox*) genes. These are transcription factors that regulate MN fate along the rostro-caudal axis of the CNS and *Meg3* has been shown to perpetuate rostral MN cell fate by maintaining the silenced epigenetic state of MN progenitor genes, thus promoting differentiation in post-mitotic MNs.
Similarly to *Meg3*, the lncRNA *Cat7* (chromatin-associated transcript 7) finely controls the levels of *MNX1* (Hb9), a transcription factor essential for MN formation.[90] This lncRNA guides the PRC1 repressive machinery to ensure that *MNX1* is expressed in a specific time-window during MN differentiation of human embryonic stem cells (hESCs), thus playing a pivotal role in MN fate determination.
Another functional role played in MNs *in vivo* is that of the lncRNA *lncrps25* (intergenic long non-coding RNA close to ribosomal protein S25).[91] Upon *lncrps25* knockdown, zebrafish primary MNs showed reduced axon length and branching, while its complete depletion led to a reduced expression of the oligodendrocyte transcription factor 2 (*Olig2*) in brain and spinal cord. *Olig2* is part of the *Olig* gene family that encode for basic helix-loop-helix transcription factors required for MN differentiation and oligodendrocyte development.[92,93]
Strikingly, the lncRNA *lncrps25* alone plays an essential role in the development of MNs in zebrafish.[91]
LncRNAs can also regulate axonal transport and RBP localization (Fig. 6). This is the case of the lncRNA *RMST* (rhabdomyosarcoma 2-associated transcript) that has been shown to localize in distal axonal compartments in mouse MNs,[94] and given its direct interaction with hnRNPA2/B1,[95] it has been speculated to carry

RNA-protein complexes to mediate subcellular trafficking in specific axonal districts.[94]

An excellent example of miRNA sponge activity in MNs (Fig. 6) is represented by the lncRNA *lncMN2*.[96] This lncRNA controls the transition from dividing progenitors to post-mitotic stage and was shown to be critical for the maturation of mouse embryonic stem cell (mESC)-derived MNs. It functions by sponging miR-466i-5p that allows the upregulation of important genes involved in MN differentiation and function.[96]

Recently, it has been discovered that lncRNAs can also regulate mouse MN homeostasis, which is fundamental for the ionic balance of these highly polarized cells. This is the case of the lncRNA *Lhx1os*, which directly interacts with the endoplasmic reticulum (ER)-associated protein PDIA3, a disulfide isomerase important to mediate the unfolded protein response (UPR) that is part of the (ER)-stress response pathway.[97] *Lhx1os* KO mice showed locomotor deficits as well as postnatal reduction of mature MNs, demonstrating the crucial role played by the lncRNA for MN normal physiological activity.[97]

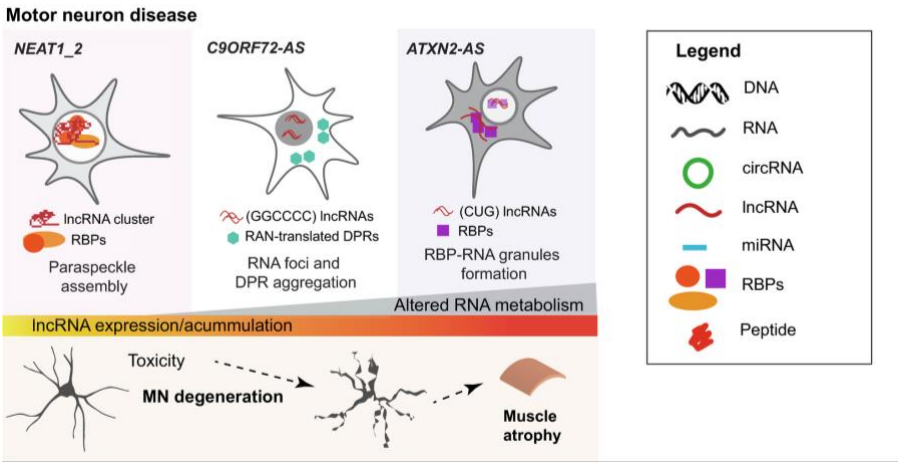

Figure 7. LncRNAs in MNDs. (Adapted from [82])

Given the emerging role of lncRNAs in MN development and function, it is not surprising that lncRNA dysregulation has been shown to have profound implications in MNDs such as ALS and SMA, that are indeed associated with alterations of RNA metabolism.

NEAT1 is a well-characterized lncRNA that has a role in ALS MN neurotoxicity. So far, two isoforms of *NEAT1*, *NEAT1_1* and *NEAT1_2*, have been described that are spliced by alternative processing of the 3'-UTR,[57,98,99] with *NEAT1_2* playing a predominant role in paraspeckle formation.[100]

Paraspeckle formation is a tightly regulated process that occurs in parallel to RNA Pol II transcription of *NEAT1_2* and binding of several paraspeckle proteins to *NEAT1,*[57,101] most of which are RBPs.[102]

So, the size and number of these RNP assemblies are directly influenced by the expression levels of the lncRNA *NEAT1*, which interacts with several RBPs involved in ALS (i.e. FUS and TDP-43).[103] It has been demonstrated that, in ALS MNs, paraspeckle formation is augmented as a cellular compensatory mechanism to increase MN survival, and the expression levels of the lncRNA *NEAT1_2* are augmented as well. Unfortunately, in ALS MNs this leads to defects in paraspeckle formation and has been shown to have a neurotoxic effect, causing cell death and neurodegeneration.[104]

One of the causative mechanisms that leads to MN degeneration and ALS is the expansion of a repeat region of six-nucleotide motifs $(GGGGCC)_n$ (G_4C_2) in the 5' region of the *C9ORF72* gene.[105–107]

This causes *C9ORF72* loss-of-function and toxic gain-of-function mediated by the repeat expansions that are implicated in *C9ORF72*-associated ALS cases. In addition to different *C9ORF72* sense isoforms, more antisense transcripts that arise from the same promoter have been detected. Among them there is *C9ORF72-AS*, (*C9ORF72* antisense lncRNA). The functional relevance of *C9ORF72-AS* is not well understood, but it has been discovered

that this lncRNA originates a higher number of *C9ORF72-AS* RNA *foci* that are found in peri-nucleolar regions and are suggested to be responsible for nucleolar defects and stress during C9-ALS pathogenesis.
The ALS risk gene *ataxin-2* also gives rise to antisense transcripts. It encodes for the cytoplasmic ATXN2 protein, that is ubiquitously expressed in neuronal and non-neuronal tissues[108] and that associates with polyribosomes in the ER, regulating mRNA translation.[109]
An increase in the length of a polyQ repeat in ATXN2 (from 22 normal to 27–33 glutamines in ALS) is significantly associated with enhanced risk for developing ALS.[110]
ATXN2-AS (antisense ATXN2-lncRNA) is a lncRNA that also undergoes repeat expansion, and its expression is increased in ALS patients. Expanded *ATXN2-AS* transcripts have been shown to have a neurotoxic effect in SK-N-MC neuroblastoma cells[111] since they aberrantly interact with RBPs involved in rRNA processing and splicing.[112]
Overall, growing experimental evidence suggests that lncRNAs play significant roles in the development and pathology of MNs. Although there is still much to learn about how lncRNAs control MN pathophysiology and why lncRNA dysregulation leads to MNDs, it is clear that this knowledge could be useful in the development of diagnostic and therapeutic strategies for treating MNDs like ALS and SMA.

1.4 The lncRNA *HOTAIRM1*

HOTAIRM1 (*HOXA* transcript antisense RNA, myeloid-specific 1) is a lincRNA transcribed from the 3' region of the *HOXA* gene cluster, located at chromosome 7 in humans (Fig. 8). It is adjacent to the *HOXA1* gene and shares with it a promoter segment that contains a CpG island, typically associated with bidirectional transcription, and it is transcribed antisense by the RNA Pol II.[113]

HOTAIRM1 gene sequence is not conserved across species as for most lncRNAs; however, in other species, such as *Mus musculus*, it shows syntenic conservation.[114] The transcription of this lncRNA is specifically induced by the exposure to retinoic acid (RA), a derivative of vitamin A involved in pivotal biological processes such as differentiation, development and homeostasis.[115]

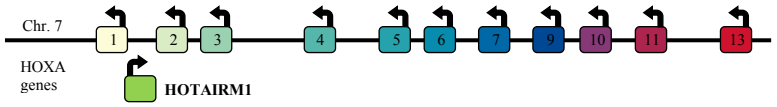

Figure 8. Localization of *HOTAIRM1*. The lncRNA is located at chromosome 7 in humans, between the *HOXA1* and *HOXA2* genes of the *HOXA* cluster and transcribed antisense to the *HOXA* genes. Two alternative RefSeq splicing isoforms are produced: HOTAIRM1_1 (NR_038366.1) comprising three exons, and HOTAIRM1_2 (NR_038367.1) having two exons.

RA-mediated signaling pathway is activated by RA binding of nuclear RA receptors, which undergo heterodimerization with retinoid X receptors and bind to specific RA response elements within the promoter of target genes. These sequences are also found in the *HOXA1 locus*, responsible for RA-dependent expression of *HOTAIRM1*.[116]
This lncRNA has been shown to have a role in different physiological processes, such as osteogenic,[117] myeloid[118] and neuronal[119] differentiation. Moreover, it has been implicated in diseases such as different types of cancer, namely colorectal cancer, in which it is downregulated and acts as a tumor suppressor[120], and recurrent gliomas, in which it is upregulated compared to primary gliomas.[121]
A previous study showed[122] that *HOTAIRM1* acts as a miRNA sponge for miR-20a, mir-106b and miR-125b during myeloid differentiation. More recently, the lncRNA has also been identified as one of the 3 main trans-acting-lncRNAs that modulate the expression of protein-coding genes during SARS-CoV-2 infection in peripheral blood mononuclear cells.[123]

NCBI's Reference Sequence Database (RefSeq) reports two alternative splicing isoforms of *HOTAIRM1* in the latest human genome assembly (GRCh38/hg38) according to the UCSC Genome Browser: the transcript variant 1 (*HOTAIRM1*_1; NR_038366.1), which comprisew three exons, and the transcript variant 2 (*HOTAIRM1*_2; NR_038367.1), which consists of exon 1 and 3 and lacks exon 2.
Recently,[1] the research group I joined for my PhD project performed the molecular characterization of *HOTAIRM1*. Besides the already known myeloid-specific isoform, a neuronal-enriched isoform - named *nHOTAIRM1*, for neuronal *HOTAIRM1* - was characterized.
Notably, *nHOTAIRM1* is induced during neuronal differentiation of RA-treated SH-SY5Y neuroblastoma cells as well as in human iPSCs induced to differentiate into ventral spinal cord lineages. Using these *in vitro* neuronal differentiation systems, we observed that *nHOTAIRM1* is expressed in both the nuclear and cytoplasmic compartments, and it was demonstrated to epigenetically regulate the expression of *NEUROG2*, a master gene of neuronal differentiation, which is why this lncRNA plays an important role in the regulation of neurogenesis,

1.4.1 *HOTAIRM1* in myeloid differentiation

HOTAIRM1 was first identified in NB4 cells as a ncRNA composed of 2 exons which is upregulated during granulocytic differentiation after RA treatment. In this context, the expression of this lincRNA is controlled by PU.1, a transcription factor activated as a downstream target of RA signaling pathway.[124]
It has been observed that knockdown of *HOTAIRM1* quantitatively reduced RA-induced expression of *HOXA1* and *HOXA4* genes during myeloid differentiation of NB4 cells, as well as attenuating the induction of CD11b and CD18, two important genes with a role in myeloid differentiation. Noticeably, this did not impact the more distal *HOXA* genes, suggesting that *HOTAIRM1* plays a role in

myelopoiesis through the modulation of gene expression in the proximal *HOXA* genes.[114] Another work in the same myeloid cell system[125] showed that RA-induced granulocytic differentiation was delayed upon *HOTAIRM1* knockdown. This resulted in the promotion of cell proliferation, preserving the expression of genes involved in DNA replication, which are normally downregulated during differentiation. This was followed by an alteration of the expression of *CD49d* and *CD11c*, two integrins α chains transcripts involved in the repression of cell cycle progression, resulting in a significantly larger population of immature and proliferating cells.
This evidence showed how *HOTAIRM1* regulates granulocytic maturation in NB4 cells and may affect cell fate during myeloid differentiation by regulating cell cycle progression.
Moreover, *HOTAIRM1* has been shown to regulate the expression of *HOXA* genes in human peripheral blood monocytes by competitively binding to miR-396, a negative regulator of monocyte-related *HOXA1* gene expression. This revealed a novel regulatory mechanism for this lincRNA, which maintains monocytes identity and prevents their differentiation into dendritic cells.[126]

1.4.2 *HOTAIRM1* in neuronal differentiation

The first evidence of *HOTAIRM1* involvement in neuronal differentiation emerged in 2011, when an RNA-seq analysis revealed that its gene expression levels dramatically changed in human neurons derived from iPSCs.[119] Among the 50047 genes identified in the RNA-seq analysis, 3055 lncRNAs have been identified as upregulated from day 0 to day 10 of differentiation. Interestingly, several lncRNAs mapping to *HOX* gene clusters showed increased levels in day 10 neurons, among which *HOTAIRM1*, that was upregulated by about 54-fold.

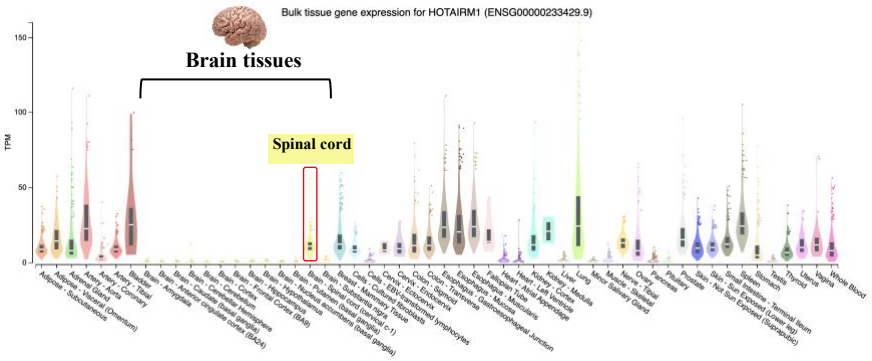

Figure 9. *HOTAIRM1* tissue gene expression reported in the GTEx Portal. Among thirteen different brain tissues, the lncRNA is particularly and specifically expressed in the spinal cord.

Moreover, the Genotype-Tissue Expression (GTEx) portal Release V8 (dbGaP Accession dbGaP Accession phs000424.v8.p2)[127] reported that, among thirteen different brain tissues, *HOTAIRM1* is specifically expressed in the spinal cord (Figure 9), suggesting it may exert specific functions in the regulation of spinal neuron differentiation.

Neuronal differentiation is a timely and spatially regulated process that relies on precisely orchestrated gene expression control. The sequential activation/repression of genes driving cell fate specification is achieved by complex regulatory networks, where transcription factors and ncRNAs work in a coordinated manner. In this context, my research group identified *nHOTAIRM1* as a new player in neuronal differentiation, demonstrating that it epigenetically controls the expression of the proneural TF NEUROGENIN 2 that is key to neuronal fate commitment and critical for brain development.

This was achieved through chromatin immunoprecipitation (ChIP) assays, in which it emerged that H3K27me3 repressive mark increased on *NEUROG2* promoter upon *nHOTAIRM1* induction in differentiating SH-SY5Y cells, while *HOTAIRM1* depletion caused a significant reduction of H3K27me3 deposition. These results indicated the implication of nuclear *HOTAIRM1* in the control of the epigenetic status of *NEUROG2*.

To further corroborate this evidence, they monitored the ability of the lncRNA to associate with PRC2 by RNA immunoprecipitation (RIP) assays and by RNA FISH combined with immunofluorescence (IF) analysis. Immunoprecipitation with an antibody against SUZ12, a component of the PRC2, was performed in nuclear extracts from differentiating SH-SY5Y cells. RIP assay revealed the interaction between SUZ12 protein and *nHOTAIRM1*. In line with this result, a combined RNA FISH/IF approach highlighted the partial overlap between *nHOTAIRM1* and PRC2 signals.
Also, it was observed that *nHOTAIRM1* activity impacts on *NEUROD* and *ASCL1*, two downstream targets that are part of NEUROGENIN 2 regulatory cascade, thus contributing to the achievement of proper neuronal differentiation timing.
This study[1] allowed us to add a new tile to the mosaic of *NEUROG2* regulation and revealed a novel role for the lncRNA *nHOTAIRM1* in shaping the outcome of neurogenesis (Fig. 10).

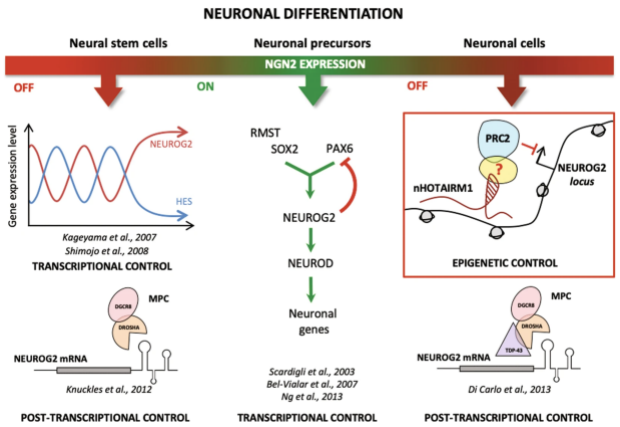

Figure 10. Epigenetic, transcriptional and post-transcriptional regulatory mechanisms converging on *NEUROG2* during neuronal differentiation. The role of *nHOTAIRM1* as a scaffold that coordinates the recruitment of the repressive epigenetic machinery PRC2 on *NEUROG2* gene promoter is highlighted by the red square.

2. Aim of the project

The mammalian brain is a transcriptionally highly complex organ, expressing approximately 40% of annotated long non-coding RNAs (lncRNAs), the most abundant class of regulatory non-coding RNAs (ncRNAs) that crucially participate in every stage of neuronal differentiation and function. Notably, lncRNA repertoire is the greatest element that differentiate humans from other primates and vertebrates.
Recent studies have shed light on the role of lncRNAs in several aspects of motor neuron (MN) homeostasis and activity. MNs are highly specialized neuronal cells located in the ventral horn of the spinal cord. They propagate nerve impulses from the central nervous system (CNS) into peripheral tissues, where they are translated into muscle contraction, which makes them an indispensable part of the locomotor neural circuitry. Remarkably, disruption of lncRNA-mediated regulative networks is considered relevant to the pathogenesis underlying motor neuron diseases (MNDs).
In this context, my PhD project aims at decoding the role of the previously characterized neuronal-enriched lncRNA *nHOTAIRM1* in differentiation and function of MNs. Notably, this lncRNA is exclusively expressed in the spinal cord, among several brain tissues, and we demonstrated that it accumulates in human induced pluripotent stem cells (iPSC)-derived MNs.[1]
With the purpose to contribute to the study of lncRNAs in MN pathophysiology, I planned to: i) perturbate *HOTAIRM1* expression, through a genome editing-based loss-of-function approach, in a model system that efficiently recapitulates spMN differentiation; ii) identify *nHOTAIRM1* target genes through transcriptome analyses in electro-physiologically active iPSC-derived spMNs and iii) uncover its mechanism of action by combining RNA-pulldown techniques and advanced imaging assays.

3. Results

3.1 Choosing the appropriate differentiation model system: iPSC-derived spinal motor neurons

The choice of the most appropriate *in vitro* differentiation model system is crucial to investigate the function and mechanism of action of a long non-coding RNA (lncRNA).
Interestingly, the GTEx Database reports that among 13 brain tissues, the lncRNA *HOTAIRM1* is specifically expressed in the spinal cord. For this reason, human induced pluripotent stem cells (iPSCs) were efficiently differentiated into electro-physiologically active spinal motor neurons (spMNs) following an established protocol that makes use of lineage-specific inducible transcriptional programming modules. In particular, it employs the construct (Fig. 11) expressing the transcription factors (TFs) Ngn2, Isl1 and Lhx3 (NIL module), that induces the spMN identity.[128] The ectopic expression of the NIL module induced by Doxycycline (DOX) is able to induce spMN formation within five days.[84]

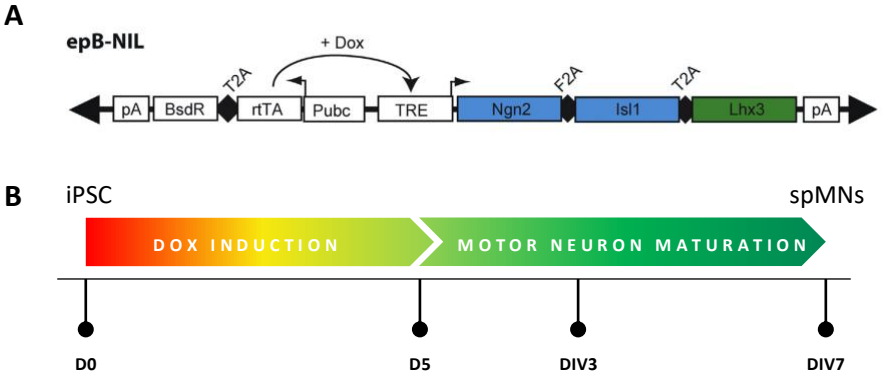

Figure 11. A. (adapted from[84]). Schematic representation of the epB-NIL construct. pA: polyadenylation signal; BsdR: blasticidin resistance gene; T2A: self-cleavage peptide; rtTA: TET transactivator protein gene; Pubc: human Ubiquitin C constitutive promoter; TRE: TET responsive element; Dox: doxycycline. Black triangles represent terminal repeats of the transposon. B. Timeline of the iPSC-to-spMN differentiation protocol. After five days (D0-D5) of DOX induction, spMN maturation occurs within seven days in vitro (DIV).

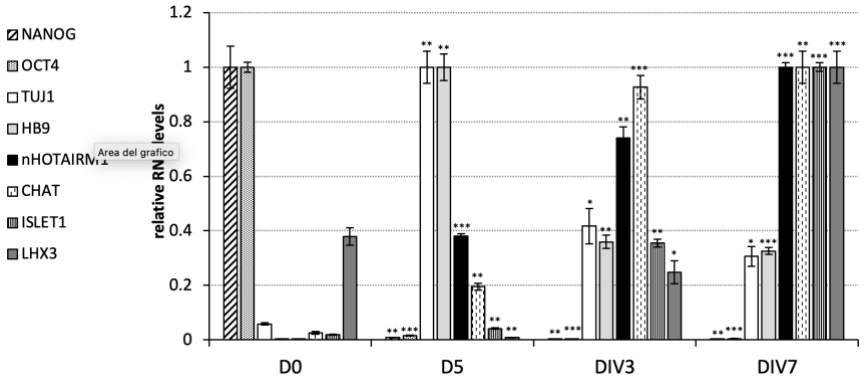

Figure 12. qRT-PCR gene expression analysis of specific markers along differentiating iPSC-NIL. Day 0 has been used as the calibrator sample and set as 1. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test, referred to D0 as reference sample for statistical tests.

The succession of differentiation stages induced by DOX treatment was monitored by following the expression of specific markers via qRT-PCR (Fig. 12). The undifferentiated iPSCs (day 0) were labeled by the pluripotency markers *NANOG* and *OCT4*, whose expression dropped down as the differentiation proceeds. The conversion of iPSCs into MN progenitors (MNPs) (day 5), is signed by the consistent expression of the pan-neuronal marker *TUJ1* and the early marker of cholinergic motor neurons *HB9* (or *MNX1*).[129] MNPs were dissociated at day 5 and re-plated for further maturation. After 3 days *in vitro* (DIV3), cells displayed consistent expression of *ChAT* (choline acetyl transferase), an enzyme required for the synthesis of the neurotransmitter acetylcholine (ACh) that is found upregulated as soon as iPSCs are committed to a neuronal lineage.[130] After 7 days *in vitro* (DIV7) a marked increase of *LHX3* (LIM/homeobox 3) and *ISLET1* (Insulin gene enhancer) was observed. Notably, the coordinated expression of these two TFs directs spMN identity and specification.[131] Accordingly, at this stage (DIV7), cell populations are composed almost exclusively of post-mitotic spMNs that display functional properties typical of electro-physiologically active MNs.[132]

In this set of differentiation, the expression of *nHOTAIRM1* was profiled. As shown in Fig. 12, it is not expressed (or expressed at very low levels) in iPSCs (day 0), it is gradually induced in MNPs (day 5) reaching its maximum level at the stage of post-mitotic spMNs (DIV7). The significant increase in gene expression levels suggests that the lncRNA might be involved in spMN differentiation, whereas its high expression level in post-mitotic cells also indicates a potential role in MN activity.

3.2 The lncRNA *nHOTAIRM1* is mainly cytoplasmic and localized both in the soma and the neurites of spMNs

To gain insight into the role of *nHOTAIRM1* in spMNs, we first analyzed its subcellular localization. Cellular fractionation carried out in iPSC-derived spMNs (DIV7) revealed that *nHOTAIRM1* is predominantly cytoplasmic (about 83%, Fig. 13B), as we already established in MN-enriched ventral spinal cord lineages.[1] This finding suggests a predominant role of the lncRNA in post-transcriptional control of gene expression.
We also investigated the distribution of *nHOTAIRM1* in sima and neurites, the two compartments of spMNs that constitute these highly polarized cells. This was achieved by spMN soma/neurite separation[133] followed by qRT-PCR analysis, which allows the quantification of the relative enrichment of the lncRNA in the two compartments. The enrichment of *GNG3* mRNA in the soma (Fig. 13C) and the enrichment of the well-characterized neuronal projection marker *COL3A1* in the neurite compartment (Fig. 13D) were consistent with the known localization of these transcripts in neurons.[133] The results revealed the presence of *nHOTAIRM1* in both districts at comparable amounts, with the cell body being contributed also by the nuclear counterpart. (Fig. 13B). Next, we sought to confirm this distribution by mapping *nHOTAIRM1* localization through imaging assays. RNA fluorescence *in situ* hybridization (FISH) targeting *nHOTAIRM1*, combined with

immunofluorescence (IF), was performed in wild type (WT) spMNs (DIV7) (Fig. 14). As indicated by the red signals, *nHOTAIRM1* spots are enriched in the soma but also accumulated in the peripheral regions of the neurites, thus confirming its subcellular localization already observed through biochemical assays (Fig 13B). Yellow arrowheads point to *nHOTAIRM1* signals detected in distal axonal segments. These findings imply potential roles for the neuronal lncRNA in a variety of functions such as neuritogenesis (axon branching and elongation, neurite projection), axonal transport, local translation of axon-resident mRNAs and regulation of synaptic activity.

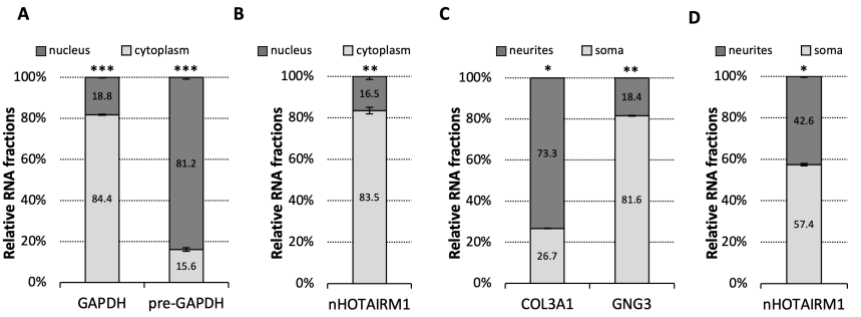

Figure 13. A. qRT-PCR analysis of *GAPDH* and *pre-GAPDH* RNA levels in nuclear (dark grey bars) and cytoplasmic (light grey bars) fractions obtained from iPSC-derived WT spMNs. Normalization was performed on total RNA. Data (means ± SEM) are expressed as percentage of total *GAPDH* or *pre-GAPDH* expression levels. N = 3. **B.** qRT-PCR analysis of *nHOTAIRM1* RNA levels in nuclear (dark grey bars) and cytoplasmic (light grey bars) fractions obtained from iPSC-derived WT spMNs. Normalization was performed on total RNA. Data (means ± SEM) are expressed as percentage of total *nHOTAIRM1* expression levels. N=3. **C.** qRT-PCR analysis of *GNG3* and *COL3A1* RNA levels in neuritic (dark grey bars) and somatic (light grey bars) fractions derived from compartmentalization of iPSC-derived WT spMNs. Normalization was performed on total RNA. Data (means ± SEM) are expressed as percentage of total *GNG3* or *COL3A1* expression levels. N = 3. **D.** qRT-PCR analysis of *nHOTAIRM1* RNA levels in neuritic (dark grey bars) and somatic (light grey bars) fractions derived from iPSC-derived WT spMNs. Normalization was performed on total RNA. Data (means ± SEM) are expressed as percentage of total *nHOTAIRM1* expression levels. N=3 *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test.

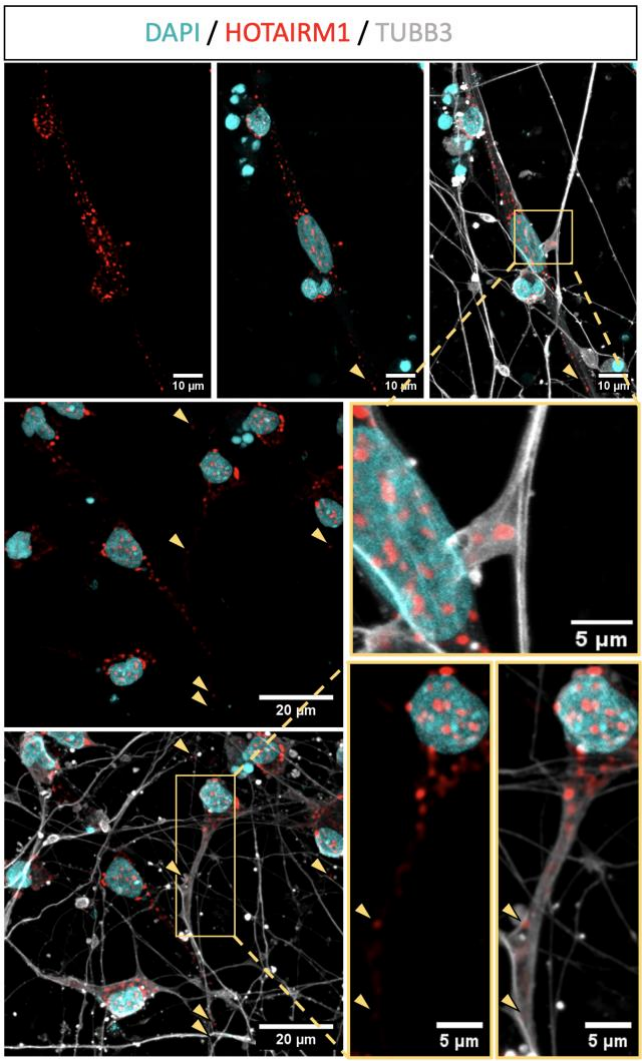

Figure 14. Different magnifications of *nHOTAIRM1* (red spots) localization in iPSC-derived WT spMNs through RNA fluorescence *in situ* hybridization (FISH), combined with immunofluorescence (IF) for TUBB3 staining (grey). As indicated by the yellow arrowheads, the lncRNA is localized both in the soma and in the peripheral regions of the neurites of spMNs (DIV7).

3.3 Identification of *nHOTAIRM1* target genes through transcriptome analysis of WT and *HOTAIRM1* KO spMNs

An essential step for decoding the function(s) of a lncRNA is the identification of its target genes. A comprehensive knowledge of *nHOTAIRM1* downstream targets was obtained by comparing the transcriptomes of WT and *nHOTAIRM1* knockout (KO) iPSC-derived spMNs (DIV7).
The KO iPSC lines were obtained through CRISPR/Cas9 genome editing approach by inserting a poly-A signal in the 5'end of the first exon of *HOTAIRM1* gene to stop its transcription (Fig. 15).

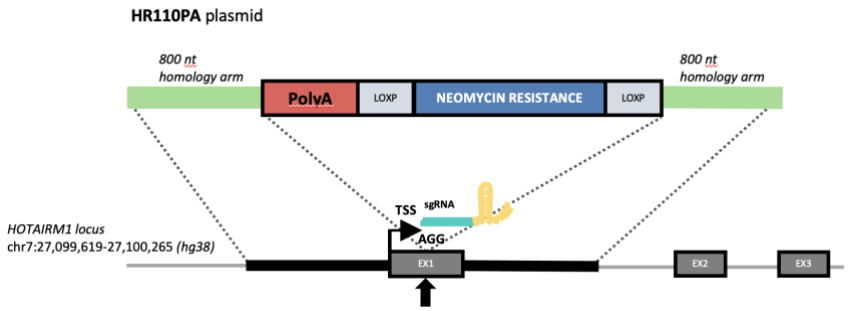

Figure 15. Schematical representation of the CRISPR/Cas9 genome editing approach to completely abolish *HOTAIRM1* transcription.

Two independent *HOTAIRM1* functional KO homozygous iPSC clones (KO#1 and KO#2) were generated. The differentiated KO#1 and KO#2 spMNs resulted to be drastically depleted of *nHOTAIRM1* expression (Fig. 16), showing comparable levels of reduction (about 97% and 95%, respectively) with respect to WT spMNs. Moreover, spMN differentiation of both KO iPSC clones showed that depletion of the lncRNA was maintained throughout all the differentiation stages (Fig. 17).

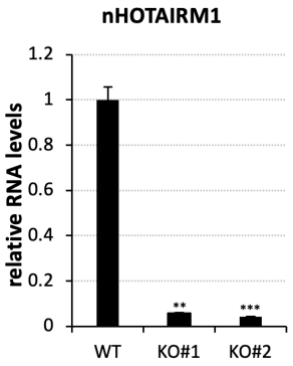

Figure 16. qRT-PCR analysis of *nHOTAIRM1* RNA levels in WT compared to KO#1 and KO#2 spMNs (DIV7). Normalization was performed using the WT sample as the calibrator sample and set as 1. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test referred to WT as reference sample for statistical tests.

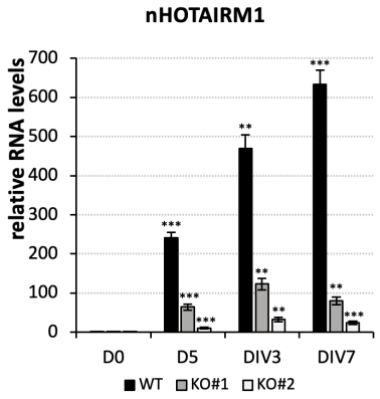

Figure 17. qRT-PCR analysis of *nHOTAIRM1* RNA levels in WT compared to KO#1 and KO#2 spMNs throughout the differentiation. Normalization was performed using the D0 sample as the calibrator sample and set as 1 for each cell line. Data (means ± SEM) are expressed in arbitrary units relative to ATP5O mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test referred to D0 as reference sample for statistical tests.

High-throughput transcriptome analyses were performed on three independent biological replicates of spMNs derived from WT and the *HOTAIRM1* KO#2 iPSC clone. Data mining of PolyA+ RNA-Seq analyses showed that both WT and KO samples clustered

homogeneously, as indicated by the Euclidean distance heatmap (Fig. 18A), that provides a graphical representation of the similarity of the samples, indicating reproducibility of expression profiles.

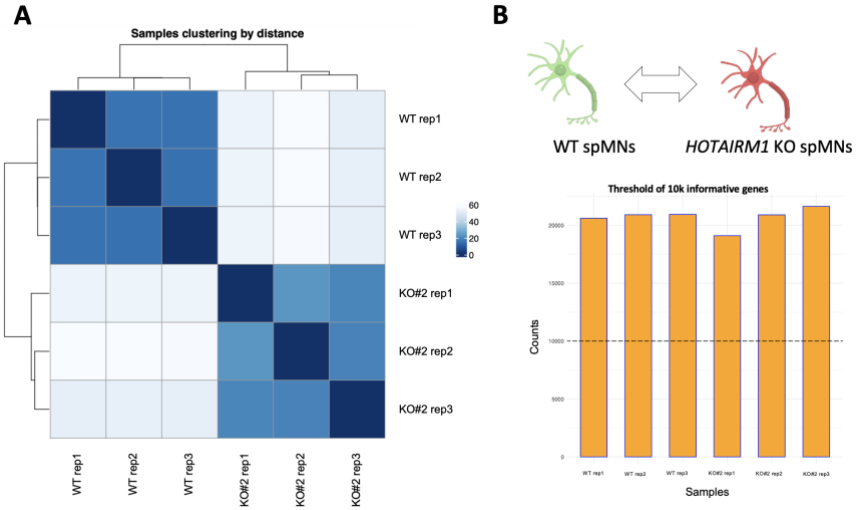

Figure 18. A. Euclidean distance heatmap showing the strong correlation among WT or KO#2 samples within the dataset. The darker the color, the smaller the Euclidean distance, indicating a strong correlation between the two samples considered. **B.** Upper area: graphical representation of WT and KO spMNs subject to RNA-seq analysis. Lower area: plot indicating read counts obtained from the sequencing for each sample.

Moreover, we obtained read counts for a large number of genes (Fig. 18B), and the control samples versus the KO samples correctly clustered between the two groups (Fig. 18A). We detected 16504 genes expressed in at least one sample (Fig. 19A).

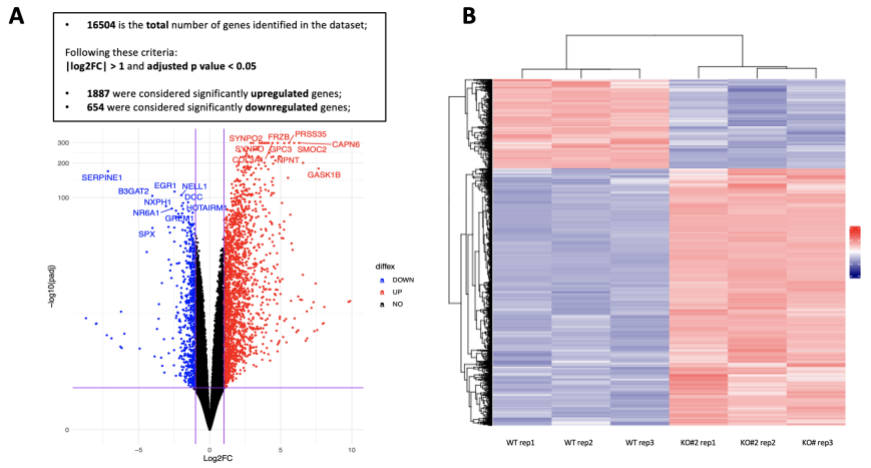

Figure 19. A. Volcano plot visualizing the criteria of selection of the significantly upregulated and downregulated genes upon *HOTAIRM1* KO in the differential expression analysis. The x-axis represents the log2 fold change (log2FC) between the groups, which is a measure of the relative difference in expression levels. The y-axis represents the adjusted p-value (-log10(padj)), which is a measure of the statistical significance of the difference in expression. **B.** Heatmap of the RNA-seq dataset, with the rows representing the genes and the columns representing the samples of the dataset. The expression levels are encoded as colors on the heatmap, with blue representing low expression and red representing high expression levels.

Among them we considered as differentially expressed those having |log2FC| > 1 and padj < 0.05 (Fig. 19A). Applying these filters, we found 1887 genes upregulated and 654 genes downregulated in KO versus WT spMNs (Fig. 19A).
To identify the biological processes in which *nHOTAIRM1* could play an important role, we performed a gene ontology over representation analysis (ORA) using the WebGestaltR tool on the lists of differentially expressed genes, assumed to be direct or indirect targets of *nHOTAIRM1*. From this test we considered significant only the gene sets having FDR < 0.05. Figure 20 shows the first 30 biological processes ranked by enrichment ratio for upregulated and downregulated genes. Of note, while the upregulated genes were not strictly related to MN biology, most of the downregulated ones were linked to MN differentiation, morphology and activity.

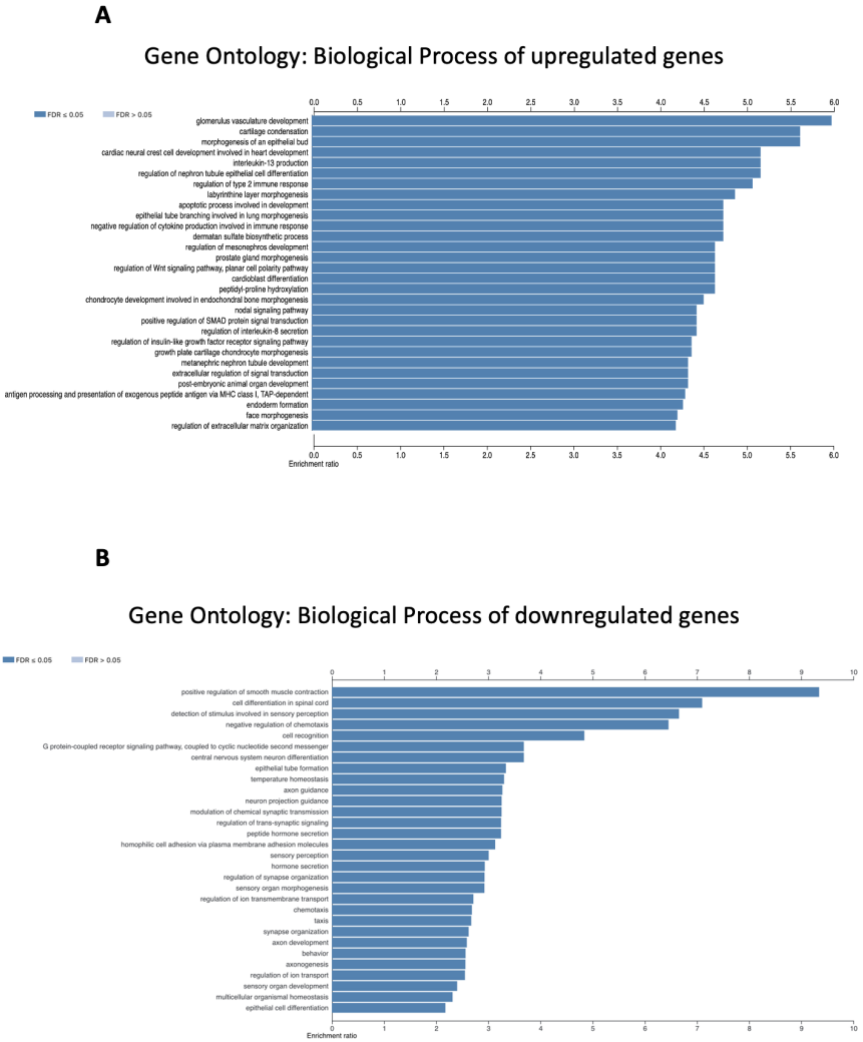

Figure 20. Gene ontology over representation analysis (ORA) plots produced by WebGestaltR tool. The categories of biological processes of both **A.** upregulated or **B.** downregulated genes are ranked by enrichment ratio, which indicates the overrepresentation of a particular GO term within a set of genes compared to its representation in the entire genome, used to the significantly enriched biological processes in each gene list. All of them have false discovery rate (FDR) ≤ 0.05.

To validate the RNA-seq analysis, we selected subsets of ten upregulated (Fig. 21A) and ten downregulated (Fig. 21B) genes and analyzed their expression by qRT-PCR in WT and KO spMNs. For a deeper analysis, we chose candidates with a

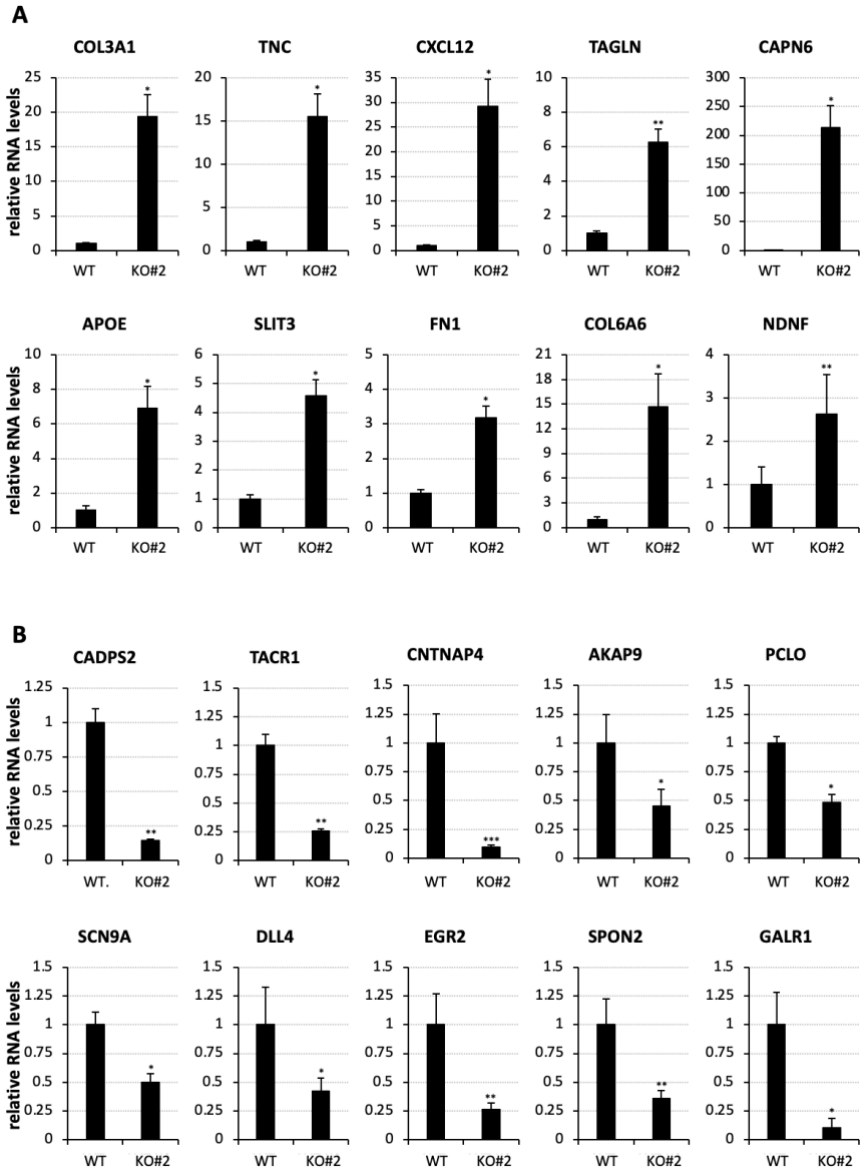

Figure 21. RNA-seq validation to provide additional confidence on RNA-seq analysis results through qRT-PCR in spMNs (DIV7) of **A.** ten significantly upregulated and **B.** ten significantly downregulated genes. Normalization was performed using the WT sample as the calibrator sample and set as 1. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test.

padj <0.05 and having different significance values ranging from log2FC -2.72 to log2FC 7.65. Notably, all of them followed the trend of expression derived from the RNA-seq analysis.
For the following functional studies, we focused on three main categories of downregulated genes that were implicated in fundamental aspects of MN biology. The first class includes over-represented genes implicated in "cell differentiation in spinal cord" and in "CNS neuron differentiation". The analysis of the expression profiles of these genes over the course of MN differentiation would provide us information about the implication of *nHOTAIRM1* in the entire process and/or in specific differentiation transition stages.
The second class comprises a conspicuous number of genes implicated in "neuron projection guidance", "axonogenesis" and "axon guidance". Comparing their expression in WT and *HOTAIRM1* KO spMNs can shed light on the regulatory role of the lncRNA on neurite outgrowth, a process that is fundamental for both differentiation and function of spMNs.
The third class encompasses genes crucially involved in "synapse organization" and "modulation of chemical synaptic transmission". It can be supposed that *nHOTAIRM1* dysregulation, leading to altered expression of these genes, may be associated with MN disease.
Remarkably, many *nHOTAIRM1* target genes may act at multiple levels and in an intertwined manner for ensuring proper differentiation and function of MNs and the achievement of homeostatic plasticity, which is important for excitability, synapses and the release of neurotransmitters.

3.4 *nHOTAIRM1* regulates binary fate decision between MNs and interneurons

Analysis of gene pathways significantly repressed in *HOTAIRM1* KO spMNs compared to WT (DIV7) highlighted functional categories related to "cell differentiation in spinal cord" and "CNS neuron differentiation" biological processes. This result intriguingly parallels the previous one (Fig. 12) showing a significant induction of *nHOTAIRM1* along spMN differentiation. Together, these findings suggest that the lncRNA may play a role in this process.

To verify this hypothesis, we extended RNA-seq data by analyzing through qRT-PCR the expression of genes belonging to those ontological clusters during motor neurogenesis. By comparing MN differentiation profiles between WT and *HOTAIRM1* KO cell lines, we confirmed altered expression patterns of several genes implicated in CNS formation, such as: the developing brain transcription factor and cofactor *SALL1*[134] and *LMO4*[135]; the nuclear receptor *RORα*[136,137], which identifies excitatory neurons in the spinal cord; and the receptor *EPHB1*, important for the regulation of contact-dependent cell-to-cell interactions (Fig. 22A).[138,139] Even more interestingly, we also proved the deregulation of genes linked to spinal cord or MN development (Fig. 22B), such as *DLL4*,[140] with a role in spinal cord neuronal subtype specification, *PROX1*, a regulator of ventral spinal cord patterning,[141] *LHX4*, required for ventral MN differentiation[142] and *DCC*, associated with MN migration.[143] The levels of the genes analyzed diminished along differentiation of *HOTAIRM1* KO iPSCs compared to WT, with a range of decrease spanning from -54% to -90%. This supports the idea that the lncRNA is required for (motor) neuron development.

During the conversion of iPSCs into MNs, the most notable downregulation was appreciated in the expression of *HB9* and *OLIG2* (Fig. 22C), two genes encoding key regulators for MN

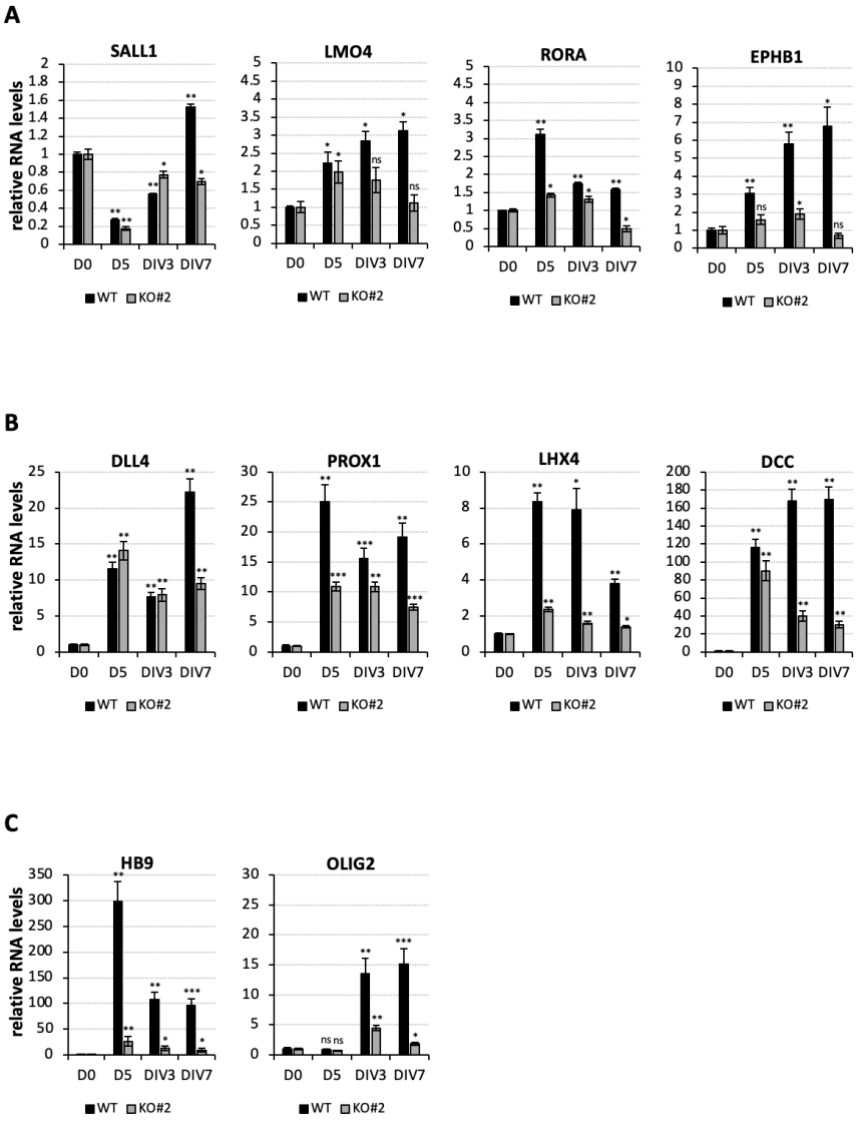

Figure 22. qRT-PCR analysis of genes related to **A**. CNS formation, **B.** MN development and **C.** MN fate specification in WT compared to KO#2 spMNs throughout the differentiation. Normalization was performed using the D0 sample as the calibrator sample and set as 1 for each cell line. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test referred to D0 as reference sample for statistical tests.

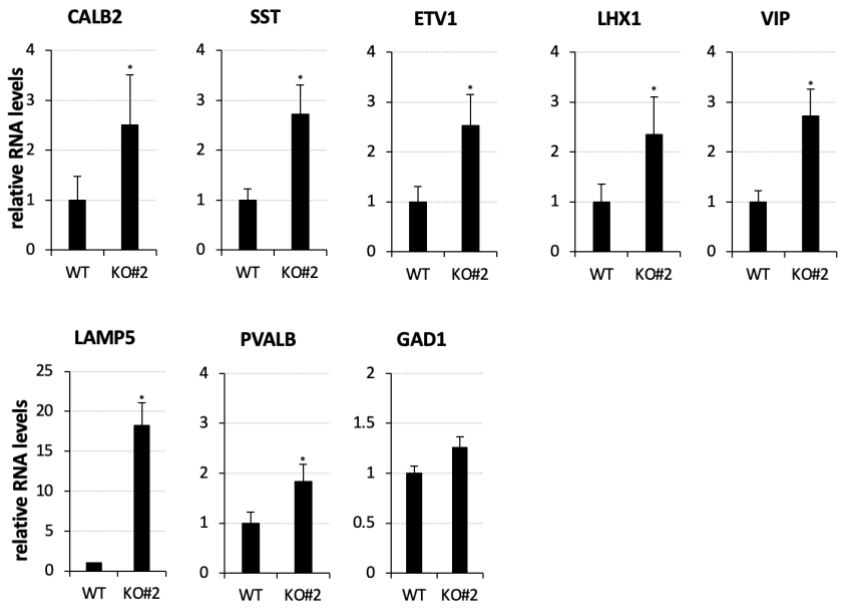

Figure 23. qRT-PCR analysis in WT compared to KO#2 spMNs (DIV7). Normalization was performed using the WT sample as the calibrator sample and set as 1. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test referred to D0 as reference sample for statistical tests.

specification. In particular, the homeobox gene *MNX1* (*HB9)*, a selective MN marker in the developing spinal cord,[144] is known to affect MN differentiation program, with a critical role in the consolidation of MN specification and in the repression of interneuron (IN) identity. Similarly, the TF OLIG2 induces MN specification and inhibits IN identity.[141] Being the activity of both genes relevant in the choice between MN and IN cell fate, we speculated that *nHOTAIRM1* might positively control the generation of MNs at expenses of INs. If true, we would expect to appreciate an increase of IN gene markers in *HOTAIRM1* KO spMNs. This possibility was checked analyzing the genes induced upon *HOTAIRM1* KO in the RNA-seq dataset (Fig. 23), among which we selected the IN markers *CALB2*,[145] *SST*,[145,146] *ETV1*,[146,147] *LHX1*,[148] *VIP*,[149] *LAMP5*,[150] *PVALB*[151] and *GAD1*[152,153] for qRT-PCR validation.

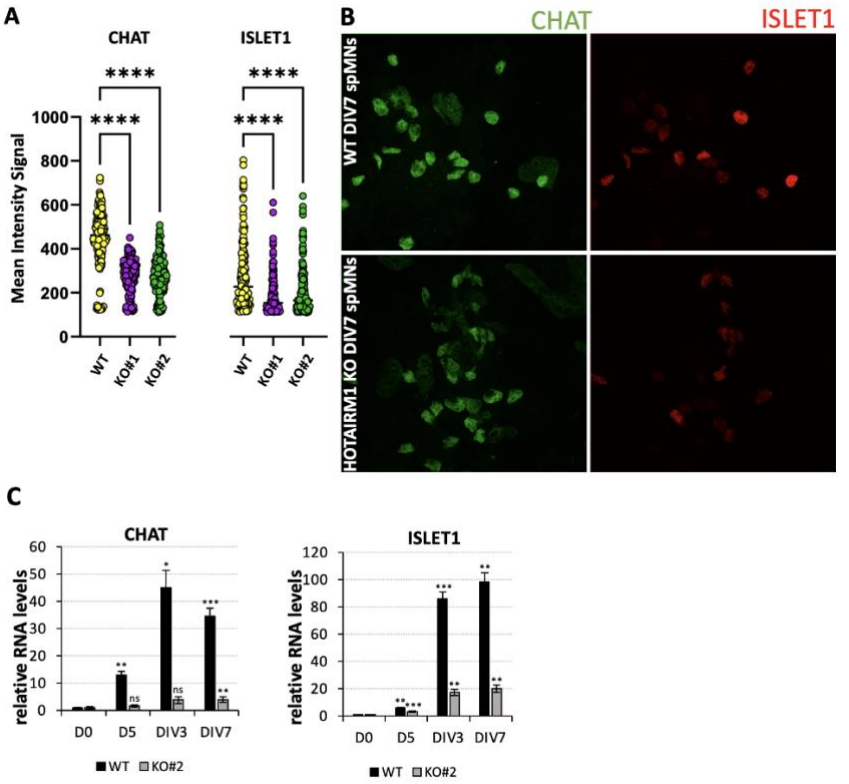

Figure 24. A Quantification of intensity signal of ChAT and ISLET1 immunostaining in WT compared to KO#1 and KO#2 spMNs (DIV7). N = 9 different acquisitions ****P ≤ 0.0001 **B.** Immunostaining for ChAT and ISLET1 proteins in in WT compared to KO#1 and KO#2 spMNs (DIV7). **C.** qRT-PCR analysis in WT compared to KO#2 spMNs throughout the differentiation. Normalization was performed using the D0 sample as the calibrator sample and set as 1 for each cell line. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test referred to D0 as reference sample for statistical tests.

We found an increased expression of such IN gene signature (Fig. 23), which paralleled the decrease of MN specific gene levels.
To definitively validate the impact of *HOTAIRM1* depletion on MN generation, we set up an alternative approach, based on immunofluorescence. We visualized WT and *HOTAIRM1* KO#1 and KO#2 spMNs (DIV7) by staining the major MN marker Islet1[144] and the cholinergic neuron marker *ChAT*.[154] A significant decrease (around -40%) of intensity signal was measured for both

factors (Fig. 24A), along with a decline in the number of *Islet1* cells (around -40%). Together, these results uphold the role of *nHOTAIRM1* in the control of MN differentiation/maturation (Fig. 24B). Of note, we also discovered that *HOTAIRM1* KO suppressed *Islet1* and *ChAT* RNA levels (Fig. 24C) throughout spMN differentiation, revealing that its activity is additionally required for the expression control of two crucial MN genes. Altogether, these results indicate that *nHOTAIRM1* participates in the regulation of the balance between neuronal cell types in the developing spinal cord, as demonstrated by its functional KO that impairs the production of MNs while favoring the formation of INs.

3.5 *nHOTAIRM1* is required for proper MN neurite outgrowth

Proper neurite outgrowth is controlled by several genes whose correct expression promotes the formation of functional neuronal connections. Interestingly, comparative transcriptome analysis between WT and *HOTAIRM1* KO spMNs revealed a list of genes downregulated upon *HOTAIRM1* depletion which are critically involved in neurite elongation and branching. We focused on five of them, namely *NrCAM, UNC5A, DCC, ROBO1, SEMA6D, SEMA3E* (Fig. 25A) that were enriched in three ontological gene categories i.e. “neuron projection guidance”, “axon guidance” and “axonogenesis”. In particular, *NrCAM* (neuronal cell adhesion molecule) has been crucially implicated in neurite outgrowth activity *in vitro*, whereas its localization to the developing synapses of the hippocampus in vivo strongly suggests that it plays also a role in synapse formation and/or function.[155–157] Notably, in spMNs depleted for *HOTAIRM1*, the expression of *NrCAM* declined by about 53% with respect to WT spMNs.
Other genes significantly downregulated following the depletion of *HOTAIRM1* in spMNs were *UNC5A*, involved in axonal pathfinding mechanisms and neuronal differentiation,[158,159]

ROBO1 (Roundabout Guidance Receptor 1), a receptor expressed by spMN cell bodies and their axons playing a role in axon guidance through interaction with Slit ligands.[160,161] Notably, its expression levels decrease by about 66% in *HOTAIRM1* KO spMNs. *Slit/ROBO* signaling is well known for playing a role in axon guidance by mediating axon repulsion in developing nervous system.[160,161] Remarkably, *Slit/ROBO* signaling intersects with the *Netrin-1/DCC* (Deleted in Colorectal Cancer) axis that, unlike *Slit/ROBO*, constitutes attractive signals for setting the position of motor exit points.[162] *DCC* gene is expressed in newly born MNs[163,164] and encodes for the receptor of Netrin-1, displaying outgrowth-promoting activity as a major ventral attractant.[165] We found that in *HOTAIRM1* KO spMNs, the expression of *DCC* receptor gene was reduced by about 80% (Fig 25A.)

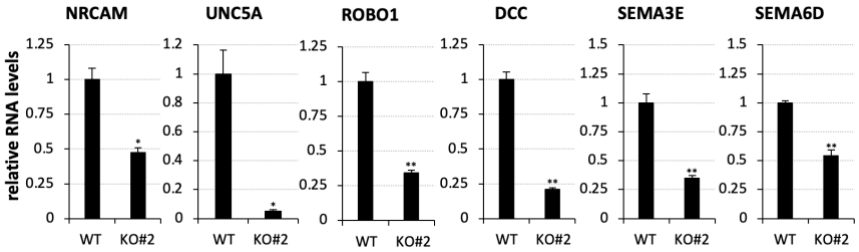

Figure 25. qRT-PCR analysis in WT compared to KO#2 spMNs (DIV7). Normalization was performed using the WT sample as the calibrator sample and set as 1. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test.

In line with the fine balance that must exist between the two guidance complexes, altered expression of the two receptors may cause misguiding in motor axon trajectories.
Other genes targeted by *nHOTAIRM1* with a relevant role in dendritic arborization and in synapse formation are *SEMA3E* (Semaphorin 3E) and *SEMA6D* (Semaphorin 6D). They belong to Semaphorins, a gene family encoding for proteins involved in axon guidance and in wiring decision for the organization of precise neuronal connections in the developing nervous system. Based on

the cellular context, they may act both as potent repellent or attractors.[166,167] We observed a reduction of *SEMA3E* and *SEMA6D* by about 66% and 46%, respectively in *HOTAIRM1* KO spMNs compared to WT spMNs. These data indicate that *nHOTAIRM1* regulates the activity of a number of genes encoding axon guidance molecules that are crucially implicated in the assembly of neuronal connections. Therefore, we sought to verify whether a defective branching phenotype mirrored what suggested by molecular analyses.

For monitoring the neuronal morphology, we took advantage of advanced imaging techniques.

Images of neurite networks composed of WT or KO spMNs (DIV7) were analyzed as described in Pani et al., 2014[168] and the Fiji-ImageJ MorphoNeuroNet script was exploited (Fig. 26A,B).

Starting from an equal number of cells among all the samples (Fig. 26D), we counted the total number of branches (defined as slab segments connecting end-points, end-points and junctions or junctions and junctions) (Fig. 26C) and of junctions (voxels with more than 2 neighbors) (Fig. 26E) and we measured both the average branch length and the total branch length per image.

A statistically significant reduction in the counts of both total branch number ($P(\text{branches}) < 0.0001$) and number of junctions

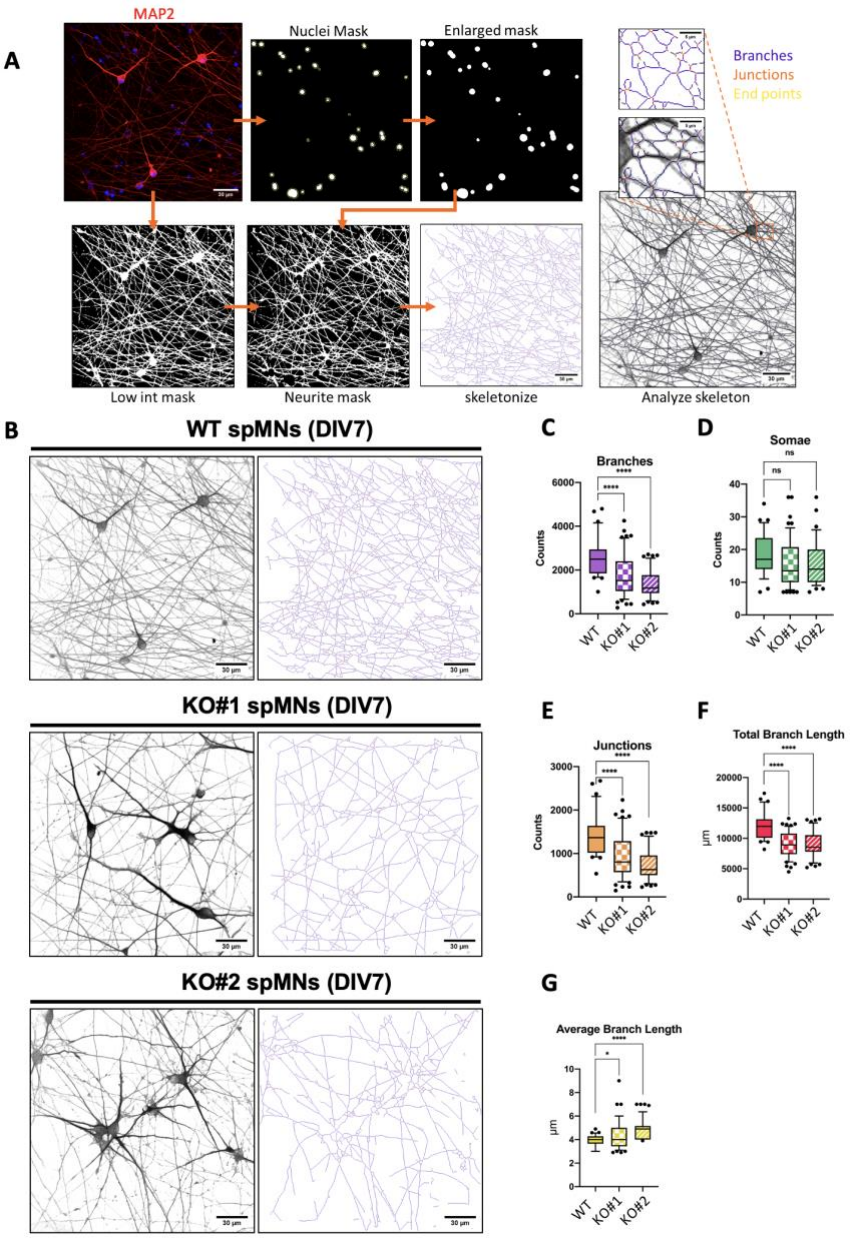

Figure 26. A. Workflow of the neurite branching analysis performed using ImageJ MorphoNeuroNet in WT, KO#1 and KO#2 spMNs. Starting from nuclei detection, a neurite network mask was generated and analyzed. **B.** Representative neurite networks of WT, KO#1 and KO#2 spMNs. **C.** Number of neurite branches of WT, KO#1 and KO#2 spMNs. **D.** Somae count of WT, KO#1 and KO#2 spMNs. **E.** Number of neurite junctions of WT, KO#1 and KO#2 spMNs. **F.** Total neurite branch length calculated in WT, KO#1 and KO#2 spMNs. **G.** Average neurite branch length

calculated in WT, KO#1 and KO#2 spMNs. Data are expressed as means ± SEM. *P ≤ 0.05, ****P ≤ 0.0001.

(P(junctions) < 0.0001) was found in the two KO neurite networks with respect to the WT one. Overall, these results indicate that *nHOTAIRM1* is important for correct outgrowth of neurites. In line with these observations, a lower total branch length (Fig. 26F) per field was calculated in the KO samples compared with the WT (P(length) <0.0001), whilst no significant difference was calculated between the two KO samples as control (P(length) = 0.9850). On the contrary, the average branch length (Fig. 26G) is higher in the KO samples compared to WT (P(avlength) <0.0121), as the lower density of the KO neurite network results in a lower number of intersections and, therefore, in longer slab segments. Overall, these results reveal that *nHOTAIRM1* is important for the correct neurite outgrowth, elongation and branching.

3.6 *nHOTAIRM1* controls genes involved in synaptic transmission

Other categories of genes affected by *HOTAIRM1* depletion in post-mitotic spMNs were those specifically involved in "synapse organization" and in "modulation of chemical synaptic transmission". They include *CACNA1D, CNR1, GRIA1, GRIK3* and *SHANK2* genes, all of which are significantly downregulated in KO spMNs (Fig. 27A).
CACNA1D gene encodes for Cav 1.3 belonging to the family of voltage-gated L-type Ca2+ channels.[169] Cav1.3 signaling is a requisite for normal neuronal development, synapse maturation as well as synaptic pruning.[170] Notably, we observed that the expression levels of *CACNA1D* declined by about 60% in *HOTAIRM1* KO post-mitotic spMNs.

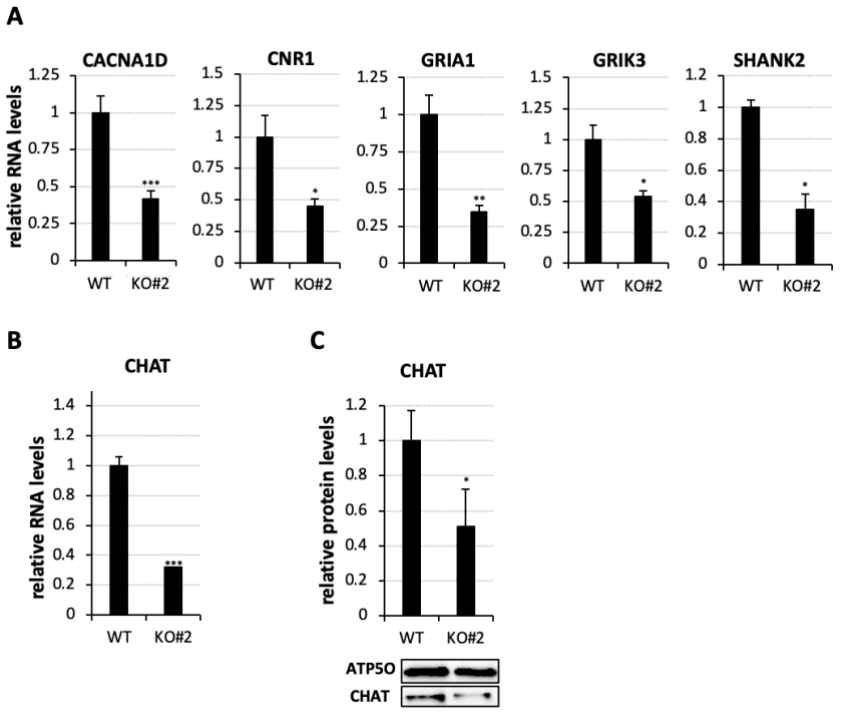

Figure 27. A,B. qRT-PCR analysis in WT compared to KO#2 spMNs (DIV7). Normalization was performed using the WT sample as the calibrator sample and set as 1. Data (means ± SEM) are expressed in arbitrary units relative to *ATP5O* mRNA and represent three independent experiments. N = 3, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001 two-tailed Student's t-test. **C.** Immunoblot analysis of ChAT protein in WT compared to KO#2 spMNs (DIV7). Normalization was performed relative to *ATP5O* protein levels. N = 3, *P ≤ 0.05 two-tailed Student's t-test.

CNR1 gene expression dropped down by about 56% in spMNs lacking *HOTAIRM1*. *CNR1* gene encodes for the type 1 cannabinoid receptor (CB1), that is expressed in both the CNS and peripheral nervous system, mainly in presynaptic terminals where it controls synaptic transmission by modulating the release of neurotransmitters.[171] *GRIA1* and *GRIK3* gene expression decreased by about 66% and 47%, respectively, in spMNs depleted for *HOTAIRM1*. *GRIA1*, also known as *GluR1*, encodes a glutamate receptor subunit and is expressed at high levels in neonatal MNs,[172] where it represents a determinant in defining MN dendritic

architecture.[173] *GRIK3* (glutamate ionotropic receptor kainate type subunit 3) encodes a principal subunit of the kainite-type ionotropic glutamate receptor, which is relevant in synaptic plasticity and potentiation.[174] Accordingly, it has been demonstrated that *GRIK3*-/- mice display reduced short- and long-term synaptic potential.[175] *SHANK2* (SH3- and multiple ankyrin repeats protein 2) encodes an important scaffolding protein that has been shown to affect synaptic connectivity.[176]
Altogether, these findings strongly support a role for *nHOTAIRM1* as a critical regulator of synaptic activity.
Besides the genes mentioned above, we also focused on *ChAT* (choline O-acetyltransferase) gene that encodes an enzymatic activity responsible for the biosynthesis of the neurotransmitter acetylcholine in cholinergic neurons. It is a well-known marker of mature MNs and it is essential for their functioning.[177] Evaluation of *ChAT* expression by qRT-PCR (Fig. 27B) revealed a dramatic reduction by about 70% in KO compared to WT spMNs (DIV7). Given the relevance of *ChAT* activity, we also measured the levels of ChAT protein (Fig. 27C) in protein lysates extracted from WT compared to KO#1 and KO#2 spMNs (DIV7) by Western Blot. Remarkably, found a significant downregulation of ChAT protein of about 50% in KO#2 spMNs compared to WT. Overall these results demonstrate that the lncRNA *nHOTAIRM1* is required for proper synaptic activity and neurotransmission of post-mitotic spMNs.

3.7 *nHOTAIRM1* controls a number of target genes through lncRNA-mRNA interactions

Given its predominantly cytoplasmic localization (Fig. 13B), we investigated how *nHOTAIRM1* exerts post-transcriptional control of genes involved in MN differentiation and function in this compartment. One of the main mechanisms played by lncRNAs in

the cytoplasm is that of competing endogenous RNA (ceRNA) that, sequestering microRNAs (miRNAs) from their mRNA targets, could cause translational de-repression. Alternatively, they can be engaged in the formation of specific RNA-protein complexes or RNA-RNA duplexes with their mRNA targets, modulating their stability and/or translation.

To test the first mechanism, we carried out a preliminary *in silico* analysis interrogating LncBase database (https://diana.e-ce.uth.gr/lncbasev3) that reports experimentally supported lncRNA-miRNA interactions. Notably, this revealed the occurrence of 56 miRNAs that are reported to bind *nHOTAIRM1* (Fig. 28A). Additional *in silico* analyses were performed exploiting the catRAPID algorithm[178,179] which predicts potential protein interactors of a given RNA based on their sequence in order to verify the interaction with Argonaute2 (AGO2) protein, a major component of the miRNA-induced silencing effector complex (miRISC). However, this analysis did not reveal any binding propensity between *nHOTAIRM1* and AGO2 protein. In line with this evidence, AGO2 was not even identified in *nHOTAIRM1* RNA antisense purification coupled with mass spectrometry (RAP-MS) experiment performed in spMNs.[1] To further explore the crosstalk between *nHOTAIRM1* and AGO2, we set up a crosslinking immunoprecipitation (CLIP) assay in cytoplasmic extracts from human-derived Neuroblastoma SH-SY5Y cells treated with retinoic acid (RA). These cells represent a suitable model system to mimic neuronal differentiation *in vitro*. The CLIP assay showed that *nHOTAIRM1* does not directly interact with AGO2 protein, thus corroborating the previous evidence that the lncRNA does not function as a sponge for miRNAs in the neuronal context (Fig. 28B).

We next explored whether *nHOTAIRM1* could control its target gene(s) through RNA-RNA interaction in the cytoplasm.

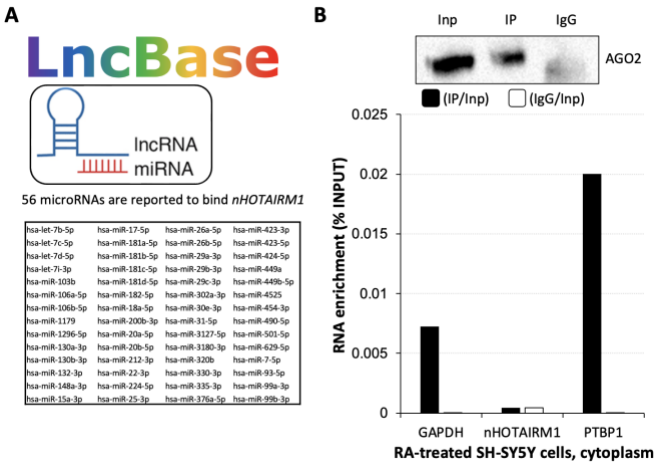

Figure 28. A. 56 miRNAs reported to bind *nHOTAIRM1* according to the LncBase database to date. **B.** CLIP assay for AGO2 in the cytoplasmic fraction of 10-day RA-treated SH-SY5Y cells. Upper panel: immunoblot analysis of AGO2 in Input (Inp) extract, immunoprecipitated (IP) and IgG (IgG) protein fractions. Lower panel: qRT-PCR analysis of RNA enrichment over Input, IP and IgG fractions. Data are expressed as Input percentage. GAPDH was used as a negative control. N = 1.

To this aim, *nHOTAIRM1* RNA pull-down assays were carried out in cytoplasmic extracts of WT DIV7 spMNs, under native conditions. Eighteen biotinylated DNA probes (Table 1) were designed and divided into two pools: even-numbered and odd-numbered. These probes were used separately to copurify endogenous *nHOTAIRM1* and its RNA interactors. As a negative control, probes targeting LacZ were used. Subsequent qRT-PCR analysis confirmed the specific enrichment of *nHOTAIRM1* in the pull-down fraction whereas GAPDH was used as a negative control (Fig. 29A). We analyzed by qRT-PCR the mRNAs found to be downregulated in *HOTAIRM1* KO compared to WT spMNs. Notably, among the tested candidates, eight transcripts were found significantly enriched in the *nHOTAIRM1* pull-down fraction. They are: *PROX1*, that participate in spinal cord development, *ROBO1*, involved in neurite outgrowth, *ChAT*, *GRIK3* and *CNR1*, playing a role in synaptic activity, *SEMA6D*, that contributes to the maintenance of the appropriate neuronal connections, *UNC5A*, involved in axonal pathfinding mechanisms and neuronal

differentiation[158,180] and *SHANK2* (SH3- and multiple ankyrin repeats protein 2), which affects synaptic connectivity.[176] To further detect direct and specific RNA-RNA interactions occurring *in vivo*, 4'-aminomethyl-4,5',8-trimethylpsoralen (AMT)-crosslinked RNA pull-down assays were performed in WT spMNs (DIV7). Notably, two out of the eight *nHOTAIRM1* mRNA targets identified in the native RNA pull-down assays, namely *ROBO1* and *SHANK2*, were identified as direct RNA interactors of the lncRNA (Fig. 29B). This suggests *nHOTAIRM1* might mechanistically and functionally control these genes, important for spMN physiology, regulating their mRNA stability and/or translation.[67]

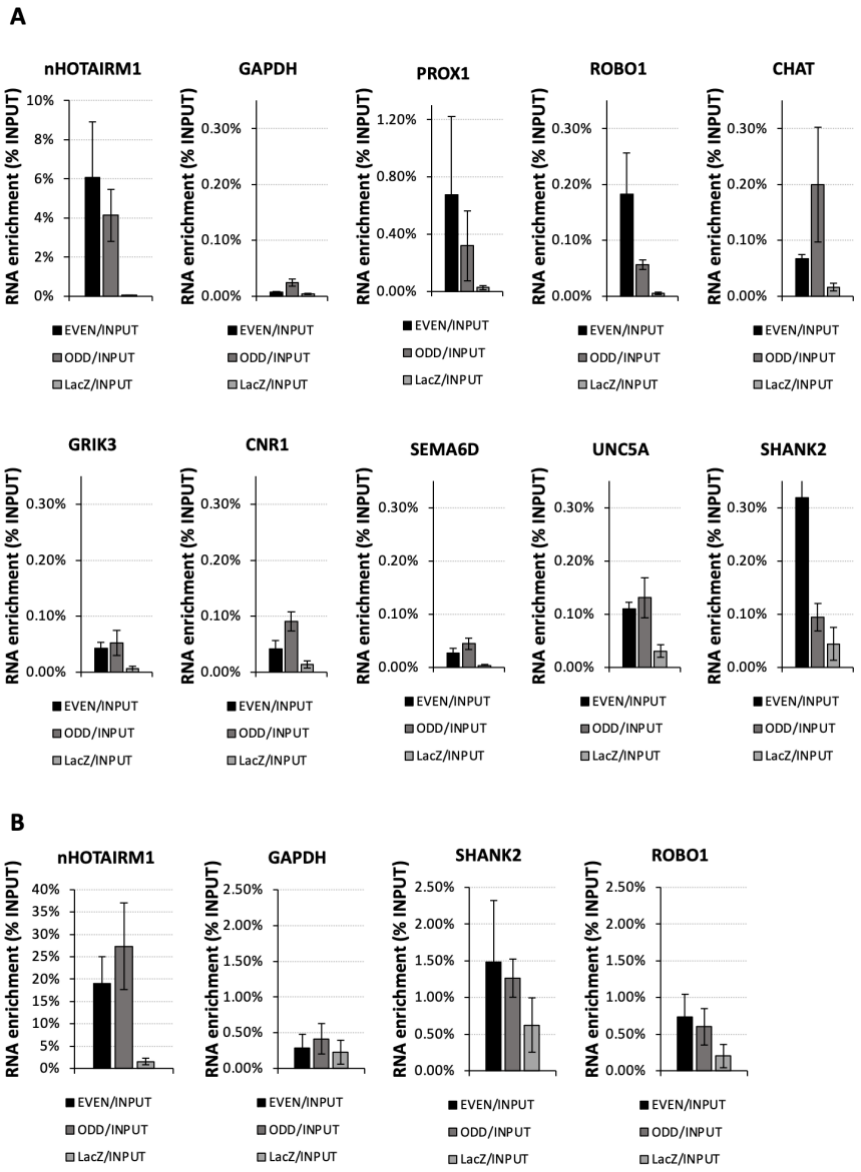

Figure 29. A. qRT-PCR analysis of *nHOTAIRM1* RNA pull-down experiments performed in cytoplasmic extracts of WT spMNs (DIV7) under native conditions. Data (means ± SEM) are expressed as relative RNA enrichments of Input percentage (%) and represent three independent experiments. N = 3. **B.** qRT-PCR analysis of *nHOTAIRM1* AMT-crosslinked RNA pull-down experiments performed in total extracts of WT spMNs (DIV7) to detect direct RNA-RNA hybridizations *in vivo*. Data (means ± SEM) are expressed as relative RNA enrichments of Input percentage (%) and represent three independent experiments. N = 3.

3.8 Mapping lncRNA-mRNA direct interaction: *ROBO1* and *SHANK2*

To map lncRNA-mRNA interactions between *nHOTAIRM1* and *ROBO1* or *SHANK2* mRNAs, we made use of the IntaRNA algorithm.[181] This bioinformatic tool was applied to the *ROBO1* and *SHANK2* protein-coding splicing isoforms that are highly expressed in our system, according to RNA-seq results in spMNs.
The energy maps of Fig. 30A, which take into account all the predicted alternative interactions between *nHOTAIRM1* and its mRNA targets *ROBO1* and *SHANK2* identified *in vivo*, shows the free energy for each intermolecular pair. For each lncRNA-mRNA pair, we focused on the region with minimal interaction energy (dark blue area) to visualize the most stable IntaRNA-predicted base pairings (Fig.13B)
Concerning *nHOTAIRM1-ROBO1* mRNA direct interaction, the analysis predicted: i) a putative region of interaction between the nucleotides 645-674 of *nHOTAIRM1* and the nucleotides 96-119 of *ROBO1-202* isoform (ENST00000464233.6), with an interaction energy of -22.07 kcal/mol and ii) a region of interaction between the nucleotides 644-689 of *nHOTAIRM1* and the nucleotides 18-66 of *ROBO1-204* isoform (ENST00000467549.5), with an interaction energy of -21.23 kcal/mol (Fig. 30B).
Regarding *nHOTAIRM1-SHANK2* mRNA direct interaction, IntaRNA predicted: i) a putative region of interaction between the nucleotides 662-671 of *nHOTAIRM1* and the nucleotides 4-13 of *SHANK2-203* isoform (ENST00000409161.5), with an interaction energy of -14.99 kcal/mol and ii) a region of interaction between the nucleotides 646-683 of *nHOTAIRM1* and the nucleotides 86-115 of *SHANK2-206* isoform (ENST00000424924.5), showing an interaction energy of -19.68 kcal/mol. (Fig. 30B).
Interestingly, the region of *nHOTAIRM1* establishing RNA-RNA interaction with mRNA isoforms encompasses the 5'end of exon 3 (nucleotides 644-689) that includes a G-rich tract predicted as putative G-quadruplex forming sequence according to the QGRS

mapper tool[182] (Fig. 30C). On the other hand, according to Ensemble assembly 104, the regions of *SHANK2* and *ROBO1* mRNA targets (isoforms *ROBO1-202*; *ROBO1-204*; *SHANK2-203*; *SHANK2-206*) that interact with the lncRNA are positioned in their 5'untranslated (5'-UTR) regions, well-known regulatory sequences implicated in the control of translational efficiency.[183]
These findings lead us to hypothesize that the RNA-RNA interaction between *nHOTAIRM1* and its RNA targets may play a role in the control of mRNA stability and/or translatability.

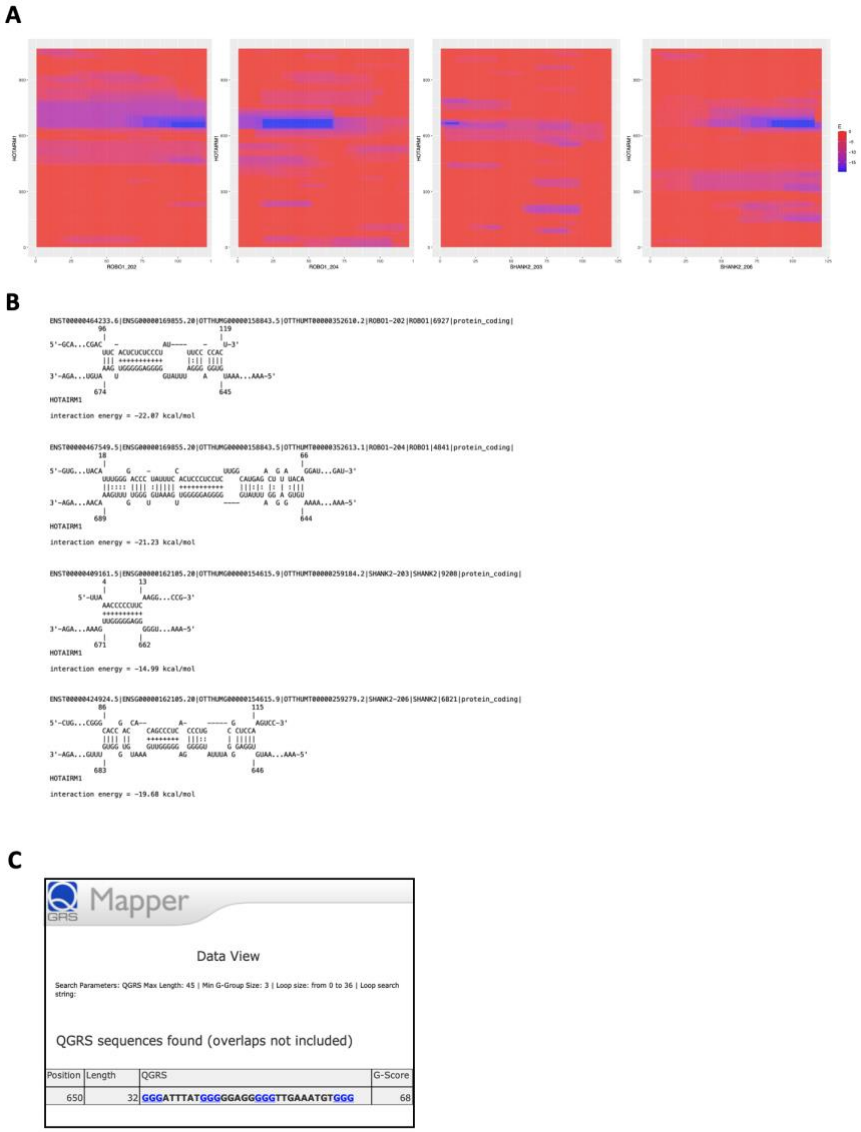

Figure 30. A. IntaRNA energy maps of the predicted stability of RNA-RNA interaction between *nHOTAIRM1* and *ROBO1* or *SHANK2* protein coding isoforms expressed in the RNA-seq in spMNs. On the x-axis is positioned the mRNA target sequence and on the y-axis is positioned the lncRNA sequence. The predicted free energy interactions at each position are shown in blue. The darker the color, the more stable the region of interaction based on free energy. The energy map can be used to identify the thermodynamically favored regions of the interaction and to understand how the stability of the interaction changes along the length of the RNA molecules. **B.** Detailed representation of the nucleotides involved in the thermodynamically favored RNA-RNA interactions predicted by

IntaRNA. **C.** Putative quadruplex forming G-rich sequence predicted by QGRS Mapper tool on *nHOTAIRM1* RNA sequence. This region is always included in the thermodynamically favored region of lncRNA-mRNA interaction predicted by IntaRNA.

4. Discussion

Neuronal differentiation is an intricate physiological process that relies on the combined action of several transcription factors (TFs) that control the expression of multiple target genes and drive precise neuronal programs for the specification of distinct cell types. During this process, the activity of TFs is integrated with several molecular signals to drive transcriptional, morphological and electrophysiological changes in a timely and spatially regulated manner.

During the last decades, it strongly emerged that these complex regulatory networks are finely coordinated by non-coding RNAs (ncRNAs) that have been shown to play a pivotal role at every level of gene expression control.

Long non-coding RNAs (lncRNAs), extraordinarily abundant in the brain, demonstrated to be critical regulators not only of differentiation, but also of the proper neuronal function, becoming prime actors in neuronal pathophysiology. The analyses carried out during my PhD project demonstrated that the neuronal-enriched lncRNA *nHOTAIRM1*, which is highly expressed in spinal motor neurons (spMNs), affects mRNA metabolism impacting motor neuron (MN) differentiation, morphology and function.

Originally described as a myeloid-specific transcript, *HOTAIRM1* was later found to be one of the most upregulated lncRNAs during the transition from iPSCs to glutamatergic neurons,[119] thereby pointing to its potential neuronal function. Recently, we characterized the neuronal isoform of *HOTAIRM1*, that we referred to as *nHOTAIRM1*, and demonstrated that the less abundant nuclear counterpart contributes to neuronal differentiation acting as an epigenetic regulator of *NEUROG2*, which is key to neuronal fate commitment.[1]

Due to its restricted expression to the spinal cord, among thirteen different brain tissues (GTEx portal Release V8 dbGaP Accession phs000424.v8.p2)[127] as well as its abundant expression in post-mitotic MNs,[1] here we wondered whether *nHOTAIRM1* can also

be a key player in differentiation and function of spMNs, a neuronal subtype whose vulnerability underlies several neurodegenerative diseases.
To place *nHOTAIRM1* in multilayered hierarchical gene regulatory networks, we exploited human iPSC-derived spMNs as a paradigm for decoding its function in the cytoplasm, where it is mainly localized. For unveiling its target genes and investigating its mechanism of action, we exploited a reverse genetics approach combined with FISH/IF assays and RNA pull-down experiments.
The discovery that *nHOTAIRM1* is functional in MN development and activity derives from loss-of-function studies carried out in post-mitotic spMNs. The identification of its downstream target genes at a genomic scale revealed its capacity to act as a multitask regulator, ensuring the production of functional spMNs. We obtained evidence that *nHOTAIRM1* participates in: i) spMN differentiation, affecting MNP formation; ii) neurite outgrowth and iii) modulation of synaptic transmission.
The role of *nHOTAIRM1* in early MN genesis was assessed during differentiation of WT vs *HOTAIRM1* KO iPSCs. It emerged that *nHOTAIRM1* promotes the expression of genes triggering MN differentiation program, as *HB9* (*MNX1*) and *OLIG2*, and determining MN identity, as *LHX4*, *ISLET1* and *DCC*. Notably, the induction of *HB9* expression marks the transition from iPSCs to MNPs whereas its inactivation induces a switch towards the interneuronal fate.[144,184] Similarly, the expression of *OLIG2* induces MN specification, while inhibiting interneuron identity.[141] In this scenario, *nHOTAIRM1* becomes a determinant of the cellular choice between alternative neuronal fates. This conclusion was further supported by the significant reduction of the levels of specific MN markers (Fig. 22C) paralleled by an increase of interneuronal markers (Fig. 23) observed upon *HOTAIRM1* depletion. Altogether, these findings led us to assign a novel role to the lncRNA in the neuronal cell fate decision as a pro-MN factor.
The involvement of *nHOTAIRM1* in the control of proper neuronal connections was assessed at both the molecular and phenotypic

levels. Loss of *HOTAIRM1* function led to significant downregulation of genes as *DCC* and *ROBO1* (Fig. 25) that are respectively involved in attractive (*Netrin-1*/*DCC*) and repulsive (*Slit*/*ROBO*) signaling pathways, whose correct balance is essential for guiding proper motor axon trajectories out of the spinal cord.[143] Notably, *HOTAIRM1* depletion phenotypically resulted in defects of neuronal morphology, with a significant reduction of the number of branches and junctions and of the total branch length (Fig. 26C,E,F) and an increased average branch length (Fig. 26G). Other relevant *nHOTAIRM1* targets are *NrCAM*, whose alteration has a profound effect on development as well as in wiring and targeting of neurons,[185,186] and the members of Semaphorin gene family, *SEMA3E* and *SEMA6D* (Fig. 25). Sema3E functions as a classical axon repellent in the spMN context,[167] whereas *SEMA6D* contributes to eliminate ipsilateral projections of corticospinal neurons (CSNs) in the spinal cord, therefore allowing neurons to maintain only the appropriate connections.[187]

Moreover, recent studies on zebrafish highlighted the essential role of *SEMA6D* in nervous system development. In particular, it was demonstrated that the deficiency of this protein caused dramatic developmental defects of primary MNs in embryonal spinal cord.[188]

Closely related to this aspect is the control exerted by *nHOTAIRM1* on genes encoding modulators of synaptic activity, whose dysfunction has been closely associated with several CNS disorders. For example, lower expression of *CNR1* has been associated with schizophrenia, and major depression disorder (MDD)[171] whereas diminished activity of *ChAT* was observed in the spinal cord of patients with amyotrophic lateral sclerosis (ALS) and was supposed to be implicated in loss-of-function in MNs[189] (Fig. 27A,B).

These considerations imply that *nHOTAIRM1*, besides contributing to correct MN homeostasis, may also take on a pathological significance.

Mechanistically, *nHOTAIRM1* appears to regulate a number of its target genes through lncRNA-mRNA interactions. For some of them - namely *PROX1*, *ChAT, GRIK3, CNR1, SEMA6D, UNC5A* - the interaction is indirect and mediated by other molecular intermediates (Fig. 29A), whereas the lncRNA engages direct lncRNA-mRNA interactions with *ROBO1* and *SHANK2* mRNAs *in vivo* (Fig. 29B).
Intriguingly, predictions of RNA-RNA interactions indicate that *nHOTAIRM1* binding may occur through a G-rich motif (Fig. 30) that may be folded into a G-quadruplex structure, known to be involved also in mRNA targeting. The finding that the binding counterparts on mRNA targets are within the 5'untranslated regions (UTRs), that play a major role in the control of translation efficiency[183], allows us to speculate that the lncRNA may exert post-transcriptional control of gene expression at the stability/translational level.
Taken together, the evidence collected within this study paints a picture in which the lncRNA *nHOTAIRM1* emerges as post-transcriptional director of the expression of genes fundamental for MN development, morphology and activity.

5. Materials and Methods

Cell culture and MN differentiation

Human induced pluripotent stem cells (iPSCs) NIL used in this study were derived, maintained and induced to differentiate into spinal motor neurons (spMNs) following the methods described in this work.[132]

Generation of HOTAIRM1 KO iPSC-NIL lines

sgRNAs (Table 3) were designed using Benchling design tool (https://www.benchling.com/) targeting the regions of *HOTAIRM1 locus*. The target regions for Cas9 double strand break were selected taking into consideration FANTOM5 TSS data in Zenbu (https://fantom.gsc.riken.jp/zenbu/). PX333 plasmid, encoding the WT Cas9 protein, was purchased from Addgene while sgRNAs were ordered as single-strand DNA probes (Biofab) and cloned as recommended by the Zhang Lab Protocol (https://media.addgene.org/data/plasmids/62/62987/62987-attachment_KiOWQSPn2egB.pdf) resulting in a vector identified as PX333-sgRNAs. HR110PA (System Biosciences) was used as a backbone to create the donor vector (DONOR). A Poly-adenylation sequence (PAS) was cloned into the DONOR vector followed by a Neomycin resistance cassette using In-Fusion® HD Cloning Plus Kit (Cat. #638910). Two homology arms (HA) HA1 and HA2 with a length of 800 nt were amplified by PCR on iPSC-NIL gDNA (Kapa HiFi, Takara). HA1 was cloned upstream of the PAS and HA2 was cloned downstream of the PAS. iPSCs were transfected on matrigel-coated dishes using the Neon Transfection System (Life Technologies), using 100 µl tips in R buffer and the following settings: 1200 V, 30 ms, 1 pulse. Selection was carried out in 800 µg/ml G418 for 5 days. Single *HOTAIRM1* KO clones were amplified and genotyped.

Cell fractionation
iPSC-derived spMNs were fractionated by the Ambion PARIS Kit (AM1921, Life Technologies). After RNA extraction, equal volumes of cytoplasmic or nuclear RNA were retro-transcribed and analyzed by qRT-PCR. Normalizations were based on the total amount of RNA.

Soma/neurite separation
Soma/neurite separation in iPSC-derived spMNs was performed as in this work.[190] Correct compartmentalization of soma and neurites was assessed by immunostaining of TUBB3 protein. RNA samples collected from spMN soma and neurites were then analyzed by qRT-PCR. Enrichment of the well-characterized neuronal projections marker *COL3A1* in the neurite compartment and of *GNG3* in the soma were consistent with the known localization of these transcripts in neurons.[133]

RNA-seq analysis
TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA, USA) was used to obtain sequencing libraries from polyA+ RNA extracted from iPSC-derived WT and *HOTAIRM1* KO spMNs (3 independent biological replicates).
The sequencing reaction produced 100 nucleotide long paired end reads and was performed on a Novaseq 6000 sequencing system (Illumina) with a depth of more than 20M. To remove adapter sequences and low-quality end bases Trim Galore[191] (version 0.6.4_dev) software was used; the minimum read length after trimming was set to 20. Alignment to human GRCh38 genome primary assembly was performed using the STAR version 2.7.9a[192] Of the total reads, 85% or more were successfully mapped to the human genome and most of them were aligned to unique locations. The quantMode TranscriptomeSAM option was used to generate alignments translated into transcript coordinates. The RSEM method was used to quantify the expression levels of the transcripts.[193]

Differential expression analysis was performed using the Bioconductor package DESeq2[194] that fits a negative binomial generalized linear model on each gene. We decided to retain only those genes that have an absolute expression of at least 10 in minimum three of the six samples as suggested by the developer of DESeq2 package. Shrinked log fold change values were obtained by the function lfcShrink[195] that looks at the largest fold changes that are not due to low counts and uses these to inform a prior distribution. The large fold changes from genes with lots of statistical information were not shrunk, while the imprecise fold changes were shrunk. Padj < 0.05 and absolute logFC greater than 1 were considered differentially expressed and used for further analysis.

Gene ontology analysis

Gene ontology (GO) over representation analysis (ORA) was performed using the http://geneontology.org/ database with WebGestalt R tool[196]. GO analysis was performed using significantly upregulated and downregulated genes with padj < 0.05 and absolute logFC greater than 1. Fig. 20 shows the first 30 biological processes ranked by enrichment ratio for upregulated and downregulated genes.

RNA-RNA interaction prediction

IntaRNA version 3.3.2[181] was used to map the binding regions between *nHOTAIRM1* and the mRNA interactors found in the AMT-crosslinked RNA pull-down experiments. To obtain *in silico* predictions of RNA-RNA interactions, from the RNA-seq results we filtered out the RNA transcripts that were not expressed in our experiment, by looking at the RSEM output isoforms.

Among the expressed splicing isoforms, we considered the TPM expression values in the WT and KO samples in order to select the protein-coding splicing isoforms that were highly expressed in our experiment. We subsequently retrieved the FASTA files of these

transcripts and we launched IntaRNA (Version 3.3.2) with default parameters to predict mRNA targets sites for *nHOTAIRM1*.

RNA extraction and analysis

Total RNA was extracted by Direct-zol RNA MiniPrep (R2052, Zymo Research). For quantitative real-time PCR (qRT-PCR) assay, cDNA was synthetized by Takara PrimeScript RT Reagent Kit (RR037A, Takara-bio). qPCR detection was performed using SensiFAST SYBR Lo-ROX Kit (BIO-94020, Bioline) on a 7500 Fast Real-Time PCR (Applied Biosystem). *ATP5O* mRNA was used as a reference target.

Immunoblotting

Protein samples for immunoblotting were extracted from iPSC-derives spMNs in RIPA Buffer (50 mM Tris–HCl [pH 8], 150 mM EGTA, 150 mM NaCl, 50 mM NaF, 10 % glycerol, 1.5 mM MgCl2, 1 % Triton). Lysates were separated on gradient poly-acrylamide gels and transferred to Amersham Protran 0.45 um nitrocellulose membrane (10600002, GE Healthcare Life Sciences), through the NuPAGE System (EI0002, Invitrogen). Immunoblots were incubated with the following antibodies: anti-ATP5O (A305-419A, Bethyl); anti-ChAT (ab223346, Abcam).

Protein staining was performed by WesternBright ECL (K-12045-D50, Advansta), detected by ChemiDoc XRS+ Molecular Imager (Bio-Rad) and quantified through the Image Lab Software (release 3.0.1).

Cross-linking immunoprecipitation (CLIP) assay

CLIP assay was performed in cytoplasmic extracts of RA-treated SH-SY5Y neuroblastoma cells as described in this work.[1]

RNA Fluorescence *in situ* hybridization (FISH) and immunofluorescence (IF)

iPSC-derived motor neurons were plated on 12mm diameter coverslip coated with Matrigel (hESC-qualified Matrix Corning 354277) and fixed in 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA) for 10 min at room temperature. Dehydration step with ice-cold Ethanol series (50%, 70%, 100%) was performed in order to store cells at -20°C in absolute ethanol until use.

nHOTAIRM1 was detected via Fluorescent *in situ* Hybridization (FISH) with a mix of 18 biotinylated probes (see Table 1) as described in these works.[197,198] Briefly, motor neurons were rehydrated by descendent ice-cold ethanol series (100%, 70%, 50%) and permeabilized in a solution of 0,05% Triton X-100 and 2 mM VRC (Sigma-Aldrich, R3380) in DPBS for 5 min. Cells were then washed three times in DPBS before replacing with 2X SSC buffer (3 M NaCl; 0,3 M sodium citrate in nuclease free water for a 20X stock solution). 5 min incubation in SSC was followed by incubation with pre-hybridization buffer (10% deionized formamide, Sigma-Aldrich, 47671; 2X SSC in nuclease free water) for 15 min at 37°C. Motor neurons were then incubated over night at 37°C in a slide hybridizer machine (ACD HybEZ™ II Hybridization System) with hybridization buffer (10% deionized formamide; 2X SSC; 10% w/v Dextran sulfate, (Sigma), 2 mM VRC in nuclease free water) completed with the biotinylated probes at a final concentration of 50 nM each. The next day cells were washed twice with 2X SSC for 5 min first at 37°C and then at RT. SSC buffer was then discarded and coverslips were incubated with TN buffer (Tris HCl pH 7,5 1 M; NaCl 5 M in nuclease free water) at RT for 10 min. Finally, biotinylated oligoes were stained incubating with 1:200 diluted 568-conjugated streptavidin (Invitrogen™ S11226) in 4% w/v BSA/TN buffer for 1-2 h at RT in a humid box.

When FISH staining was combined with Immunofluorescence (IF), or to perform IF alone, cells were washed twice with TN buffer

(only when coupling FISH with IF) and once with DPBS for 5 min at RT and then were incubated with primary antibodies (anti-βTUBIII, Sigma T2200; anti-GAP43, Santa Cruz Biotechnology sc-17790; anti-Map2, Proteintech 17490-1-AP) diluted in 1% w/v BSA/DPBS for 1 hour at RT. Subsequently, samples were washed three times with DPBS for 5 min at RT and incubated with 1:300 diluted secondary antibodies (Goat anti-Mouse 488, Invitrogen A-11029; Goat anti-rabbit 488, Invitrogen A-11008; Donkey anti-Rabbit 594, Immunological Sciences IS-20152-1) in 1% w/v BSA/DPBS for 45 min at RT. Lastly, cells were washed three times with DPBS for 5 min at RT, nuclei were counterstained with DAPI solution (Sigma, D9542; 1ug/ml/PBS) for 5 min at RT and coverslips were mounted with Prolong Diamond Mounting Media (ThermoFisher Scientific, P-36961).

Neurite branching analysis

Neurite branching analysis was performed in iPSC-derived WT, KO#1 and KO#2 spMNs as described in this work with minor modifications.[168]

A Fiji-imageJ macro for semi-automated analysis of motor neurons neurite networks was created adapting the workflow described in this work.[168]

Briefly, Z-projections of confocal images of WT and KO motor neurons stained with the neurite marker Map2[199–201] were generated.

To segment somas starting from the DAPI channel, a nuclei mask was generated with the Huang thresholding algorithm and outliers with a radius ≤ 20 px were removed. A selection was generated starting from the nuclei mask that was then expanded by 5 pixels (1 px = 0.207 micron) using the "Enlarge selection" function of ImageJ, to account for the small cytoplasmic portion typical of mature motor neurons (ref.).

Subsequently, the Map2 channel was duplicated, and two different masks were generated. From the first copy, a high intensity mask

was created thresholding with Moments algorithm upon contrast enhancement (saturated pixels = 0.1%) and Gaussian Blur (sigma = 2 μm) filtering. This mask accounts for high intensity parts of the image, including somas, axon hillock and neurite edges. The second copy was used to generate a low intensity mask, the LoG mask, to account for thinner parts of neurites. To create the LoG mask, the contrast was enhanced (saturated pixels = 0.1 %), the LoG filter from the Process > Math ImageJ menu was applied and a threshold was applied to the image using Moments algorithm.
The "Image Calculator" function of ImageJ was then used to combine the high intensity mask with the LoG mask and to subtract the nuclei mask to obtain a final neurite mask.
Finally, the neurite mask was skeletonized and the "Analyze skeleton" function was exploited to determine number of branches, junctions and end-points per image, as well as total branch length and average branch length.
Number of cells per image was manually counted starting from the nuclei masks taking advantage of the multipoint tool.

Image acquisition and analysis
Samples were imaged on a Nikon Instrument A1 Confocal Laser Microscope equipped with a 1.49 NA 100x objective (Apo TIRF 100x Oil, Nikon, Tokyo, Japan) and with a 60x objective. Confocal images were collected with NIS-Elements AR software (Nikon): ND Acquisition module was used for multipoint acquisition of Z-stack images (150-175nm Z-spacing) of 4 um thickness.

Native RNA pull-down
Native RNA pull-down experiments were performed on cytoplasmic extracts from iPSC-derived spMNs (DIV7). Cells were harvested in PBS and centrifuged at 400 g for 5 min. Cell pellets were lysed in a buffer containing Tris-HCl pH 7.5 50 mM, NaCl 150 mM, MgCl2 3 mM, NP40 0.5%, EDTA 2 mM, DTT 1

mM; 1× PIC, and RNase inhibitors. After lysis and clearing by centrifugation, 1 mg of extract was diluted in a 1:2 ratio with hybridization buffer containing Tris-HCl pH 7.5 100 mM, NaCl 300 mM, MgCl2 1 mM, SDS 0.2%, Formamide 15%, NP40 0.5%, EDTA 10 mM, DTT 1 mM, 1× PIC, and RNase inhibitors.
10% of the total extract was collected for Input (INP). 100 pmol of previously heat-denatured DNA biotinylated probes were added. After a 4h incubation at 4°C, 0.1 ml of streptavidin Magnasphere paramagnetic beads (Promega) were added to pull down the lncRNA, and the mixture was incubated for 1 h at room temperature. After RNA pull-down, beads were washed 4 times with hybridization buffer and RNA was extracted and DNase treated. Subsequent RNA pull-down qRT-PCR results were represented normalized on the Input sample and expressed as a percentage of Input.

AMT-crosslinked RNA pull-down
AMT-crosslinked RNA pull-down experiments were performed as described in this work.[202]

Statistical analyses
Histograms show the mean ± SEM of 1 to 3 independent biological replicates. N is indicated in Figure Legends. Errors were calculated from relative quantities and then opportunely propagated; statistical significance was determined by two-tailed paired Student's t test. A p-value < 0.05 was considered statistically significant. *P<0.05, **P<0.01, ***<0.001, ****<0.0001.

Table 1: biotinylated DNA probe list

nHOTAIRM1 probe 01	TCGTCCTACGCTCATAAATC
nHOTAIRM1 probe 02	GGACTATGGCTGGTTTCTGG
nHOTAIRM1 probe 03	CTTCCTCCGCTAAATCTCAG
nHOTAIRM1 probe 04	GGAAGTTCCAATGACAACGC
nHOTAIRM1 probe 05	TGGCTCTTAACAGCAAAGGC
nHOTAIRM1 probe 06	GGGGCGGGTTGATTTAAGAA
nHOTAIRM1 probe 07	GCAGCATGTAAGCAACATGT
nHOTAIRM1 probe 08	GGCAAAACAGACCGTGAGAA
nHOTAIRM1 probe 09	GAGCGCCGGGGATTTAAATG
nHOTAIRM1 probe 10	CAAATCGGCCTTTGCAGTCG
nHOTAIRM1 probe 11	AGAACGCAGCTTTTGCTCTT
nHOTAIRM1 probe 12	GCCAGTTCATCTTTCATTGA
nHOTAIRM1 probe 13	CCATAAATCCCTCCACATTT
nHOTAIRM1 probe 14	GTTTCAAACACCCACATTTC
nHOTAIRM1 probe 15	CAGTCTCCAGGTCAATAACT
nHOTAIRM1 probe 16	GAGTAACACGGAGTTTCTTT
nHOTAIRM1 probe 17	CCAAGCCCAAGCTCTTGAAA
nHOTAIRM1 probe 18	AGAGGCAGAATTGGACAGTC
LacZ probe 01	AATGTGAGCGAGTAACAACC

LacZ probe 02	AATAATTCGCGTCTGGCCTT
LacZ probe 03	AATTCAGACGGCAAACGCT
LacZ probe 04	ATCTTCCAGATAACTGCCGT

Table 2: list of oligonucleotides used in this study

AKAP9_FW	GCTGAACGAGATGCCATAGAC
AKAP9_REV	GTTCTTGTAGGCGGGTAGTAGA
APOE_FW	CACTGGGTCGCTTTTGGG
APOE_REV	GGGTCAGTTGTTCCTCCAGT
ATP5O_FW	ACTCGGGTTTGACCTACAGC
ATP5O_REV	GGTACTGAAGCATCGCACCT
CACNA1D_FW	GCTGAAGCGAGAATAAGGGC
CACNA1D_REV	AGGAAGTCTGGTGCCTCTTG
CADPS2_FW	CATCCAGGGCACAGAGTTTG
CADPS2_REV	TGGCCAGAACAGGAATTTGC
CALB2_FW	CAAGAGCTGGAGAAGGCAAG
CALB2_REV	GCCATCTCGATTTTCCCATCT
CAPN6_FW	CCATCCAGGGCCTCATAACA
CAPN6_REV	GAACAGGAATGGGACCCTCA
CHAT_FW	CTCAGCTACAAGGCCCTGCT

CHAT_REV	ACCAGCGTGTCCTGGGTATG
CNR1_FW	ATTTGAGCCCACGTACAGGA
CNR1_REV	CTAACCCTGGTGAGCAGTGA
CNTNAP4_FW	ACAGAGAAGAAGGAGTGGTCT
CNTNAP4_REV	CCTGCCCAATACCAGAGATTTG
COL3A1_FW	AGGAAGCTGTTGAAGGAGGATG
COL3A1_REV	GGTTGGGGCAGTCTAATTCTTG
COL6A6_FW	TCGTGGAGACTTTTGGAGGT
COL6A6_REV	GTCTTGAACAACAGATGCCAGA
CXCL12_FW	AGTGGGTCTAGCGGAAAGTC
CXCL12_REV	CACTCCAAACTGTGCCCTTC
DCC_FW	TATGCAAACGGTCCAGTCCA
DCC_REV	ATCATCAGTAGAGACGCCCG
DLL4_FW	TACTTGTGATGAGGGCTGGG
DLL4_REV	CACAGTAGGTGCCCGTGAAT
EGR2_FW	GCCCCTTTGACCAGATGAAC
EGR2_REV	AAAGCTGCTGGGATATGGGA
EPHB1_FW	ATCTCTGGTGATTGCTCGGG
EPHB1_REV	CAGTGGGAGCAGCCTTCAG
ETV1_FW	TGGTAGCTCTTCTGGATGACC

ETV1_REV	AGTAATAGCGGAGTGAACGG
FN1_FW	ACTTCCTCCAGAGCAAAGGG
FN1_REV	GGCCAGTCCTACAACCAGTA
GAD1_FW	CCTCCAAGAACCTGCTTTCC
GAD1_REV	GTGGGTGATGAAAGTCCAGC
GALR1_FW	GCCAGCAACCAGACCTTCT
GALR1_REV	GGATGCTTCAGACTTCTTTGACA
GAPDH_FW	CCAAAATCAAGTGGGGCGAT
GAPDH_REV	GGCAGAGATGATGACCCTTT
GNG3_FW	CCTTCAGGTACCAGCCATCC
GNG3_REV	GGGTCAGTGGAGGGTACCAA
GRIA1_FW	GTCTGCTTCATTACGCCGAG
GRIA1_REV	TGTCACCTGCCAGTTCTTCT
GRIK3_FW	AGGATGGCGATGGTGATCTT
GRIK3_REV	TTCGAGAAGATGTGGGCCTT
HB9_FW	GAGACCCAGGTGAAGATTTG
HB9_REV	CCTTCTGTTTCTCCGCTTCC
ISL1_FW	AAGGTGGAGCTGCATTGGTTTG
ISL1_REV	TAAACCAGCTACAGGACAGGCC
LHX1_FW	CTACACCCAAGCCCACCC

LHX1_REV	TGCTTCATCCTCCGCTCC
LHX3_FW	CGGACCCAGTTCTGACCTATCC
LHX3_REV	GTGTGAGGTGCAGGGTGGAG
LHX4_FW	GCTCCGAGATGATGCAGAGT
LHX4_REV	TGTGCCAGTGTCTGTCCAG
LMO4_FW	TCGATTCCTGCGAGTGAACT
LMO4_REV	GCAGTAGTGGATTGCTCTGAAG
NANOG_FW	CCAAATTCTCCTGCCAGTGAC
NANOG_REV	CACGTGGTTTCCAAACAAGAAA
NDNF_FW	CGGATGGGGAACTAGACGAT
NDNF_REV	CTTTGGAGTGGAAGCTGAGC
nHOTAIRM1_FW	GTTGCTTACATGCTGCGTTTTC
nHOTAIRM1_REV	TTTCAAACACCCACATTTCAACC
NRCAM_FW	AGAGGCTGAGGTGAGAGGAT
NRCAM_REV	GAGGGGATAGCTTCGTCGAG
OCT4_FW	ATGCATTCAAACTGAGGTGCCTGC
OCT4_REV	AACTTCACCTTCCCTCCAACCAGT
OLIG2_FW	GCTCCTCAAATCGCATCCAG
OLIG2_REV	CTGCTGCCCTTACTCCGG
PCLO_FW	AGCGTTCTATGTCTGACCCC

PCLO_REV	AGGAGGTGGATTTGGCAGAG
pre-GAPDH_FW	CTGGGGGTAAGGAGATGCTG
pre-GAPDH_REV	TTACCAGAGTTAAAAGCAGCCC
PROX1_FW	CCACCACCCTTGTTCACCAG
PROX1_REV	AAGTAGGTCTTCAGCATATTGGA
ROBO1_FW	CTGGCGTCATGTGTCATCTG
ROBO1_REV	CCCAGAGATCCCAGTTCCTC
RORA_FW	AGCGGGAGGTGATGTGGC
RORA_REV	GTCGGGGCTGGCATACTTC
SALL1_FW	ATGTGGAATAGCACCCCTGC
SALL1_REV	CGATCTCCTTGCTGTCCTCC
SCN9A_FW	GCGAGCACATGAAAAGAGGT
SCN9A_REV	CATCGGCAAATTCAGTCTCAGA
SEMA3E_FW	TCTCCTCTTCCACTACCTCCA
SEMA3E_REV	CCCCACGATCTTTACAAGCG
SEMA6D_FW	CGGTGCTGAGATGTGTTACG
SEMA6D_REV	CCCTGCTCAGAAATGCCAC
SHANK2_FW	GCATTATTGAGGAGAAGACGGT
SHANK2_REV	CATCCACGGACTCCAGGTAC
SLIT3_FW	GTCCCAGTTGCCACACATTT

SLIT3_REV	CATCAACAACGAGCTGCAGG
SPON2_FW	GCTTCACCTTCTCCTCCCC
SPON2_REV	CAGCCGCACCAGTGTCAC
SST_FW	CAGAAGTCCCTGGCTGCT
SST_REV	CTCAAGCCTCATTTCATCCTGC
TACR1_FW	GGGCAGGAGGAAGAAGATGT
TACR1_REV	CCTGCTGGTGATTGGCTATG
TAGLN_FW	GGCTGGTGGAGTGGATCATA
TAGLN_REV	ACCTGCTCCATCTGCTTGAA
TNC_FW	TCCCTGGAATTTATGCCCGT
TNC_REV	GAACTGTCACCGTGTCAACC
TUJ1_FW	CCCGGAACCATGGACAGTGT
TUJ1_REV	TGACCCTTGGCCCAGTTGTT
UNC5A_FW	CAACGGAGGGGAGGAGTG
UNC5A_REV	ACGAGGATGAGGACAAGCAG

Table 3: sgRNAs sequences for CRISPR/Cas9 genome editing

sgRNA01	TCATCTTTCATTGAACGGTGGGG
sgRNA02	GTTCATCTTTCATTGAACGGTGG

6. References

(1) Rea, J.; Menci, V.; Tollis, P.; Santini, T.; Armaos, A.; Garone, M. G.; Iberite, F.; Cipriano, A.; Tartaglia, G. G.; Rosa, A.; Ballarino, M.; Laneve, P.; Caffarelli, E. HOTAIRM1 Regulates Neuronal Differentiation by Modulating NEUROGENIN 2 and the Downstream Neurogenic Cascade. *Cell Death Dis.* **2020**, *11* (7), 527. https://doi.org/10.1038/s41419-020-02738-w.

(2) Jarroux, J.; Morillon, A.; Pinskaya, M. History, Discovery, and Classification of LncRNAs. *Adv. Exp. Med. Biol.* **2017**, *1008*, 1–46. https://doi.org/10.1007/978-981-10-5203-3_1.

(3) Scherrer, K.; Darnell, J. E. Sedimentation Characteristics of Rapidly Labelled RNA from HeLa Cells. *Biochem. Biophys. Res. Commun.* **1962**, *7* (6), 486–490. https://doi.org/10.1016/0006-291X(62)90341-8.

(4) HOAGLAND, M. B.; STEPHENSON, M. L.; SCOTT, J. F.; HECHT, L. I.; ZAMECNIK, P. C. A Soluble Ribonucleic Acid Intermediate in Protein Synthesis. *J. Biol. Chem.* **1958**, *231* (1), 241–257. https://doi.org/10.1016/s0021-9258(19)77302-5.

(5) Argetsinger Steitz, J.; Jakes, K. How Ribosomes Select Initiator Regions in MRNA: Base Pair Formation between the 3' Terminus of 16S RRNA and the MRNA during Initiation of Protein Synthesis in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72* (12), 4734–4738. https://doi.org/10.1073/pnas.72.12.4734.

(6) Weinberg, R. A.; Penman, S. Small Molecular Weight Monodisperse Nuclear RNA. *J. Mol. Biol.* **1968**, *38* (3), 289–304. https://doi.org/10.1016/0022-2836(68)90387-2.

(7) Reddy, R.; Busch, H. Small Nuclear RNAs: RNA Sequences, Structure, and Modifications. *Struct. Funct. Major Minor Small Nucl. Ribonucleoprotein Part.* **1988**, 1–37. https://doi.org/10.1007/978-3-642-73020-7_1.

(8) Carlson, J. R. Sidney Altman and the RNA Revolution.

Proc. Natl. Acad. Sci. U. S. A. **2022**, *119* (32), e2211692119. https://doi.org/10.1073/pnas.2211692119.

(9) Lee, R. C.; Feinbaum, R. L.; Ambros, V. The C. Elegans Heterochronic Gene Lin-4 Encodes Small RNAs with Antisense Complementarity to Lin-14. *Cell* **1993**, *75* (5), 843–854. https://doi.org/10.1016/0092-8674(93)90529-Y.

(10) Lagos-Quintana, M.; Rauhut, R.; Lendeckel, W.; Tuschl, T. Identification of Novel Genes Coding for Small Expressed RNAs. *Science (80-.).* **2001**, *294* (5543), 853–858. https://doi.org/10.1126/science.1064921.

(11) Lau, N. C.; Lim, L. P.; Weinstein, E. G.; Bartel, D. P. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis Elegans. *Science (80-.).* **2001**, *294* (5543), 858–862. https://doi.org/10.1126/science.1065062.

(12) Martone, J.; Mariani, D.; Desideri, F.; Ballarino, M. Non-Coding RNAs Shaping Muscle. *Front. Cell Dev. Biol.* **2020**, *7*, 394. https://doi.org/10.3389/fcell.2019.00394.

(13) Chowdhary, A.; Satagopam, V.; Schneider, R. Long Non-Coding RNAs: Mechanisms, Experimental, and Computational Approaches in Identification, Characterization, and Their Biomarker Potential in Cancer. *Front. Genet.* **2021**, *12*, 649619. https://doi.org/10.3389/fgene.2021.649619.

(14) Jarroux, J.; Morillon, A.; Pinskaya, M. History, Discovery, and Classification of LncRNAs. *Adv. Exp. Med. Biol.* **2017**, *1008*, 1–46. https://doi.org/10.1007/978-981-10-5203-3_1.

(15) Zhao, Q.; Zhao, Z.; Fan, X.; Yuan, Z.; Mao, Q.; Yao, Y. Review of Machine Learning Methods for RNA Secondary Structure Prediction. *PLoS Comput. Biol.* **2021**, *17* (8), e1009291. https://doi.org/10.1371/journal.pcbi.1009291.

(16) Micheel, J.; Safrastyan, A.; Wollny, D. Advances in Non-Coding RNA Sequencing. *Non-coding RNA* **2021**, *7* (4), 70. https://doi.org/10.3390/ncrna7040070.

(17) Niu, D. K.; Jiang, L. Can ENCODE Tell Us How Much Junk DNA We Carry in Our Genome? *Biochemical and*

Biophysical Research Communications. January 25, 2013, pp 1340–1343. https://doi.org/10.1016/j.bbrc.2012.12.074.

(18) Dunham, I.; Kundaje, A.; Aldred, S. F.; Collins, P. J.; Davis, C. A.; Doyle, F.; Epstein, C. B.; Frietze, S.; Harrow, J.; Kaul, R.; Khatun, J.; Lajoie, B. R.; Landt, S. G.; Lee, B. K.; Pauli, F.; Rosenbloom, K. R.; Sabo, P.; Safi, A.; Sanyal, A.; Shoresh, N.; Simon, J. M.; Song, L.; Trinklein, N. D.; Altshuler, R. C.; Birney, E.; Brown, J. B.; Cheng, C.; Djebali, S.; Dong, X.; Ernst, J.; Furey, T. S.; Gerstein, M.; Giardine, B.; Greven, M.; Hardison, R. C.; Harris, R. S.; Herrero, J.; Hoffman, M. M.; Iyer, S.; Kellis, M.; Kheradpour, P.; Lassmann, T.; Li, Q.; Lin, X.; Marinov, G. K.; Merkel, A.; Mortazavi, A.; Parker, S. C. J.; Reddy, T. E.; Rozowsky, J.; Schlesinger, F.; Thurman, R. E.; Wang, J.; Ward, L. D.; Whitfield, T. W.; Wilder, S. P.; Wu, W.; Xi, H. S.; Yip, K. Y.; Zhuang, J.; Bernstein, B. E.; Green, E. D.; Gunter, C.; Snyder, M.; Pazin, M. J.; Lowdon, R. F.; Dillon, L. A. L.; Adams, L. B.; Kelly, C. J.; Zhang, J.; Wexler, J. R.; Good, P. J.; Feingold, E. A.; Crawford, G. E.; Dekker, J.; Elnitski, L.; Farnham, P. J.; Giddings, M. C.; Gingeras, T. R.; Guigó, R.; Hubbard, T. J.; Kent, W. J.; Lieb, J. D.; Margulies, E. H.; Myers, R. M.; Stamatoyannopoulos, J. A.; Tenenbaum, S. A.; Weng, Z.; White, K. P.; Wold, B.; Yu, Y.; Wrobel, J.; Risk, B. A.; Gunawardena, H. P.; Kuiper, H. C.; Maier, C. W.; Xie, L.; Chen, X.; Mikkelsen, T. S.; Gillespie, S.; Goren, A.; Ram, O.; Zhang, X.; Wang, L.; Issner, R.; Coyne, M. J.; Durham, T.; Ku, M.; Truong, T.; Eaton, M. L.; Dobin, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Xue, C.; Williams, B. A.; Zaleski, C.; Röder, M.; Kokocinski, F.; Abdelhamid, R. F.; Alioto, T.; Antoshechkin, I.; Baer, M. T.; Batut, P.; Bell, I.; Bell, K.; Chakrabortty, S.; Chrast, J.; Curado, J.; Derrien, T.; Drenkow, J.; Dumais, E.; Dumais, J.; Duttagupta, R.; Fastuca, M.; Fejes-Toth, K.; Ferreira, P.; Foissac, S.; Fullwood, M. J.; Gao, H.; Gonzalez, D.; Gordon, A.;

Howald, C.; Jha, S.; Johnson, R.; Kapranov, P.; King, B.; Kingswood, C.; Li, G.; Luo, O. J.; Park, E.; Preall, J. B.; Presaud, K.; Ribeca, P.; Robyr, D.; Ruan, X.; Sammeth, M.; Sandhu, K. S.; Schaeffer, L.; See, L. H.; Shahab, A.; Skancke, J.; Suzuki, A. M.; Takahashi, H.; Tilgner, H.; Trout, D.; Walters, N.; Wang, H.; Hayashizaki, Y.; Reymond, A.; Antonarakis, S. E.; Hannon, G. J.; Ruan, Y.; Carninci, P.; Sloan, C. A.; Learned, K.; Malladi, V. S.; Wong, M. C.; Barber, G. P.; Cline, M. S.; Dreszer, T. R.; Heitner, S. G.; Karolchik, D.; Kirkup, V. M.; Meyer, L. R.; Long, J. C.; Maddren, M.; Raney, B. J.; Grasfeder, L. L.; Giresi, P. G.; Battenhouse, A.; Sheffield, N. C.; Showers, K. A.; London, D.; Bhinge, A. A.; Shestak, C.; Schaner, M. R.; Kim, S. K.; Zhang, Z. Z.; Mieczkowski, P. A.; Mieczkowska, J. O.; Liu, Z.; McDaniell, R. M.; Ni, Y.; Rashid, N. U.; Kim, M. J.; Adar, S.; Zhang, Z.; Wang, T.; Winter, D.; Keefe, D.; Iyer, V. R.; Zheng, M.; Wang, P.; Gertz, J.; Vielmetter, J.; Partridge, E. C.; Varley, K. E.; Gasper, C.; Bansal, A.; Pepke, S.; Jain, P.; Amrhein, H.; Bowling, K. M.; Anaya, M.; Cross, M. K.; Muratet, M. A.; Newberry, K. M.; McCue, K.; Nesmith, A. S.; Fisher-Aylor, K. I.; Pusey, B.; DeSalvo, G.; Parker, S. L.; Balasubramanian, S.; Davis, N. S.; Meadows, S. K.; Eggleston, T.; Newberry, J. S.; Levy, S. E.; Absher, D. M.; Wong, W. H.; Blow, M. J.; Visel, A.; Pennachio, L. A.; Petrykowska, H. M.; Abyzov, A.; Aken, B.; Barrell, D.; Barson, G.; Berry, A.; Bignell, A.; Boychenko, V.; Bussotti, G.; Davidson, C.; Despacio-Reyes, G.; Diekhans, M.; Ezkurdia, I.; Frankish, A.; Gilbert, J.; Gonzalez, J. M.; Griffiths, E.; Harte, R.; Hendrix, D. A.; Hunt, T.; Jungreis, I.; Kay, M.; Khurana, E.; Leng, J.; Lin, M. F.; Loveland, J.; Lu, Z.; Manthravadi, D.; Mariotti, M.; Mudge, J.; Mukherjee, G.; Notredame, C.; Pei, B.; Rodriguez, J. M.; Saunders, G.; Sboner, A.; Searle, S.; Sisu, C.; Snow, C.; Steward, C.; Tapanari, E.; Tress, M. L.; Van Baren, M. J.;

Washietl, S.; Wilming, L.; Zadissa, A.; Zhang, Z.; Brent, M.; Haussler, D.; Valencia, A.; Addleman, N.; Alexander, R. P.; Auerbach, R. K.; Balasubramanian, S.; Bettinger, K.; Bhardwaj, N.; Boyle, A. P.; Cao, A. R.; Cayting, P.; Charos, A.; Cheng, Y.; Eastman, C.; Euskirchen, G.; Fleming, J. D.; Grubert, F.; Habegger, L.; Hariharan, M.; Harmanci, A.; Iyengar, S.; Jin, V. X.; Karczewski, K. J.; Kasowski, M.; Lacroute, P.; Lam, H.; Lamarre-Vincent, N.; Lian, J.; Lindahl-Allen, M.; Min, R.; Miotto, B.; Monahan, H.; Moqtaderi, Z.; Mu, X. J.; O'Geen, H.; Ouyang, Z.; Patacsil, D.; Raha, D.; Ramirez, L.; Reed, B.; Shi, M.; Slifer, T.; Witt, H.; Wu, L.; Xu, X.; Yan, K. K.; Yang, X.; Struhl, K.; Weissman, S. M.; Penalva, L. O.; Karmakar, S.; Bhanvadia, R. R.; Choudhury, A.; Domanus, M.; Ma, L.; Moran, J.; Victorsen, A.; Auer, T.; Centanin, L.; Eichenlaub, M.; Gruhl, F.; Heermann, S.; Hoeckendorf, B.; Inoue, D.; Kellner, T.; Kirchmaier, S.; Mueller, C.; Reinhardt, R.; Schertel, L.; Schneider, S.; Sinn, R.; Wittbrodt, B.; Wittbrodt, J.; Jain, G.; Balasundaram, G.; Bates, D. L.; Byron, R.; Canfield, T. K.; Diegel, M. J.; Dunn, D.; Ebersol, A. K.; Frum, T.; Garg, K.; Gist, E.; Hansen, R. S.; Boatman, L.; Haugen, E.; Humbert, R.; Johnson, A. K.; Johnson, E. M.; Kutyavin, T. V.; Lee, K.; Lotakis, D.; Maurano, M. T.; Neph, S. J.; Neri, F. V.; Nguyen, E. D.; Qu, H.; Reynolds, A. P.; Roach, V.; Rynes, E.; Sanchez, M. E.; Sandstrom, R. S.; Shafer, A. O.; Stergachis, A. B.; Thomas, S.; Vernot, B.; Vierstra, J.; Vong, S.; Wang, H.; Weaver, M. A.; Yan, Y.; Zhang, M.; Akey, J. M.; Bender, M.; Dorschner, M. O.; Groudine, M.; MacCoss, M. J.; Navas, P.; Stamatoyannopoulos, G.; Beal, K.; Brazma, A.; Flicek, P.; Johnson, N.; Lukk, M.; Luscombe, N. M.; Sobral, D.; Vaquerizas, J. M.; Batzoglou, S.; Sidow, A.; Hussami, N.; Kyriazopoulou-Panagiotopoulou, S.; Libbrecht, M. W.; Schaub, M. A.; Miller, W.; Bickel, P. J.; Banfai, B.; Boley, N. P.; Huang, H.; Li, J. J.; Noble, W. S.; Bilmes, J. A.;

Buske, O. J.; Sahu, A. D.; Kharchenko, P. V.; Park, P. J.; Baker, D.; Taylor, J.; Lochovsky, L. An Integrated Encyclopedia of DNA Elements in the Human Genome. *Nature* **2012**, *489* (7414), 57–74. https://doi.org/10.1038/nature11247.

(19) Derrien, T.; Johnson, R.; Bussotti, G.; Tanzer, A.; Djebali, S.; Tilgner, H.; Guernec, G.; Martin, D.; Merkel, A.; Knowles, D. G.; Lagarde, J.; Veeravalli, L.; Ruan, X.; Ruan, Y.; Lassmann, T.; Carninci, P.; Brown, J. B.; Lipovich, L.; Gonzalez, J. M.; Thomas, M.; Davis, C. A.; Shiekhattar, R.; Gingeras, T. R.; Hubbard, T. J.; Notredame, C.; Harrow, J.; Guigó, R. The GENCODE v7 Catalog of Human Long Noncoding RNAs: Analysis of Their Gene Structure, Evolution, and Expression. *Genome Res.* **2012**, *22* (9), 1775–1789. https://doi.org/10.1101/gr.132159.111.

(20) Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M. C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; Kodzius, R.; Shimokawa, K.; Bajic, V. B.; Brenner, S. E.; Batalov, S.; Forrest, A. R. R.; Zavolan, M.; Davis, M. J.; Wilming, L. G.; Aidinis, V.; Allen, J. E.; Ambesi-Impiombato, A.; Apweiler, R.; Aturaliya, R. N.; Bailey, T. L.; Bansal, M.; Baxter, L.; Beisel, K. W.; Bersano, T.; Bono, H.; Chalk, A. M.; Chiu, K. P.; Choudhary, V.; Christoffels, A.; Clutterbuck, D. R.; Crowe, M. L.; Dalla, E.; Dalrymple, B. P.; Bono, B.; Gatta, G. Della; Bernardo, D. Di; Down, T.; Engstrom, P.; Fagiolini, M.; Faulkner, G.; Fletcher, C. F.; Fukushima, T.; Furuno, M.; Futaki, S.; Gariboldi, M.; Georgii-Hemming, P.; Gingeras, T. R.; Gojobori, T.; Green, R. E.; Gustincich, S.; Harbers, M.; Hayashi, Y.; Hensch, T. K.; Hirokawa, N.; Hill, D.; Huminiecki, L.; Iacono, M.; Ikeo, K.; Iwama, A.; Ishikawa, T.; Jakt, M.; Kanapin, A.; Katoh, M.; Kawasawa, Y.; Kelso, J.; Kitamura, H.; Kitano, H.; Kollias, G.; Krishnan, S. P. T.; Kruger, A.; Kummerfeld, S. K.; Kurochkin, I. V.; Lareau, L. F.; Lazarevic, D.; Lipovich, L.;

Liu, J.; Liuni, S.; McWilliam, S.; Babu, M. M.; Madera, M.; Marchionni, L.; Matsuda, H.; Matsuzawa, S.; Miki, H.; Mignone, F.; Miyake, S.; Morris, K.; Mottagui-Tabar, S.; Mulder, N.; Nakano, N.; Nakauchi, H.; Ng, P.; Nilsson, R.; Nishiguchi, S.; Nishikawa, S.; Nori, F.; Ohara, O.; Okazaki, Y.; Orlando, V.; Pang, K. C.; Pavan, W. J.; Pavesi, G.; Pesole, G.; Petrovsky, N.; Piazza, S.; Reed, J.; Reid, J. F.; Ring, B. Z.; Ringwald, M.; Rost, B.; Ruan, Y.; Salzberg, S. L.; Sandelin, A.; Schneider, C.; Schönbach, C.; Sekiguchi, K.; Semple, C. A. M.; Seno, S.; Sessa, L.; Sheng, Y.; Shibata, Y.; Shimada, H.; Shimada, K.; Silva, D.; Sinclair, B.; Sperling, S.; Stupka, E.; Sugiura, K.; Sultana, R.; Takenaka, Y.; Taki, K.; Tammoja, K.; Tan, S. L.; Tang, S.; Taylor, M. S.; Tegner, J.; Teichmann, S. A.; Ueda, H. R.; Nimwegen, E.; Verardo, R.; Wei, C. L.; Yagi, K.; Yamanishi, H.; Zabarovsky, E.; Zhu, S.; Zimmer, A.; Hide, W.; Bult, C.; Grimmond, S. M.; Teasdale, R. D.; Liu, E. T.; Brusic, V.; Quackenbush, J.; Wahlestedt, C.; Mattick, J. S.; Hume, D. A.; Kai, C.; Sasaki, D.; Tomaru, Y.; Fukuda, S.; Kanamori-Katayama, M.; Suzuki, M.; Aoki, J.; Arakawa, T.; Iida, J.; Imamura, K.; Itoh, M.; Kato, T.; Kawaji, H.; Kawagashira, N.; Kawashima, T.; Kojima, M.; Kondo, S.; Konno, H.; Nakano, K.; Ninomiya, N.; Nishio, T.; Okada, M.; Plessy, C.; Shibata, K.; Shiraki, T.; Suzuki, S.; Tagami, M.; Waki, K.; Watahiki, A.; Okamura-Oho, Y.; Suzuki, H.; Kawai, J.; Hayashizaki, Y. Molecular Biology: The Transcriptional Landscape of the Mammalian Genome. *Science (80-.).* **2005**, *309* (5740), 1559–1563. https://doi.org/10.1126/science.1112014.

(21) Djebali, S.; Davis, C. A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; Xue, C.; Marinov, G. K.; Khatun, J.; Williams, B. A.; Zaleski, C.; Rozowsky, J.; Röder, M.; Kokocinski, F.; Abdelhamid, R. F.; Alioto, T.; Antoshechkin, I.; Baer, M. T.; Bar, N. S.; Batut, P.; Bell, K.;

Bell, I.; Chakrabortty, S.; Chen, X.; Chrast, J.; Curado, J.; Derrien, T.; Drenkow, J.; Dumais, E.; Dumais, J.; Duttagupta, R.; Falconnet, E.; Fastuca, M.; Fejes-Toth, K.; Ferreira, P.; Foissac, S.; Fullwood, M. J.; Gao, H.; Gonzalez, D.; Gordon, A.; Gunawardena, H.; Howald, C.; Jha, S.; Johnson, R.; Kapranov, P.; King, B.; Kingswood, C.; Luo, O. J.; Park, E.; Persaud, K.; Preall, J. B.; Ribeca, P.; Risk, B.; Robyr, D.; Sammeth, M.; Schaffer, L.; See, L. H.; Shahab, A.; Skancke, J.; Suzuki, A. M.; Takahashi, H.; Tilgner, H.; Trout, D.; Walters, N.; Wang, H.; Wrobel, J.; Yu, Y.; Ruan, X.; Hayashizaki, Y.; Harrow, J.; Gerstein, M.; Hubbard, T.; Reymond, A.; Antonarakis, S. E.; Hannon, G.; Giddings, M. C.; Ruan, Y.; Wold, B.; Carninci, P.; Guig, R.; Gingeras, T. R. Landscape of Transcription in Human Cells. *Nature* **2012**, *489* (7414), 101–108. https://doi.org/10.1038/nature11233.

(22) Carninci, P.; Kasukawa, T.; Katayama, S.; Gough, J.; Frith, M. C.; Maeda, N.; Oyama, R.; Ravasi, T.; Lenhard, B.; Wells, C.; Kodzius, R.; Shimokawa, K.; Bajic, V. B.; Brenner, S. E.; Batalov, S.; Forrest, A. R. R.; Zavolan, M.; Davis, M. J.; Wilming, L. G.; Aidinis, V.; Allen, J. E.; Ambesi-Impiombato, A.; Apweiler, R.; Aturaliya, R. N.; Bailey, T. L.; Bansal, M.; Baxter, L.; Beisel, K. W.; Bersano, T.; Bono, H.; Chalk, A. M.; Chiu, K. P.; Choudhary, V.; Christoffels, A.; Clutterbuck, D. R.; Crowe, M. L.; Dalla, E.; Dalrymple, B. P.; Bono, B.; Gatta, G. Della; Bernardo, D. Di; Down, T.; Engstrom, P.; Fagiolini, M.; Faulkner, G.; Fletcher, C. F.; Fukushima, T.; Furuno, M.; Futaki, S.; Gariboldi, M.; Georgii-Hemming, P.; Gingeras, T. R.; Gojobori, T.; Green, R. E.; Gustincich, S.; Harbers, M.; Hayashi, Y.; Hensch, T. K.; Hirokawa, N.; Hill, D.; Huminiecki, L.; Iacono, M.; Ikeo, K.; Iwama, A.; Ishikawa, T.; Jakt, M.; Kanapin, A.; Katoh, M.; Kawasawa, Y.; Kelso, J.; Kitamura, H.; Kitano, H.; Kollias, G.; Krishnan, S. P. T.; Kruger, A.; Kummerfeld, S. K.;

Kurochkin, I. V.; Lareau, L. F.; Lazarevic, D.; Lipovich, L.; Liu, J.; Liuni, S.; McWilliam, S.; Babu, M. M.; Madera, M.; Marchionni, L.; Matsuda, H.; Matsuzawa, S.; Miki, H.; Mignone, F.; Miyake, S.; Morris, K.; Mottagui-Tabar, S.; Mulder, N.; Nakano, N.; Nakauchi, H.; Ng, P.; Nilsson, R.; Nishiguchi, S.; Nishikawa, S.; Nori, F.; Ohara, O.; Okazaki, Y.; Orlando, V.; Pang, K. C.; Pavan, W. J.; Pavesi, G.; Pesole, G.; Petrovsky, N.; Piazza, S.; Reed, J.; Reid, J. F.; Ring, B. Z.; Ringwald, M.; Rost, B.; Ruan, Y.; Salzberg, S. L.; Sandelin, A.; Schneider, C.; Schönbach, C.; Sekiguchi, K.; Semple, C. A. M.; Seno, S.; Sessa, L.; Sheng, Y.; Shibata, Y.; Shimada, H.; Shimada, K.; Silva, D.; Sinclair, B.; Sperling, S.; Stupka, E.; Sugiura, K.; Sultana, R.; Takenaka, Y.; Taki, K.; Tammoja, K.; Tan, S. L.; Tang, S.; Taylor, M. S.; Tegner, J.; Teichmann, S. A.; Ueda, H. R.; Nimwegen, E.; Verardo, R.; Wei, C. L.; Yagi, K.; Yamanishi, H.; Zabarovsky, E.; Zhu, S.; Zimmer, A.; Hide, W.; Bult, C.; Grimmond, S. M.; Teasdale, R. D.; Liu, E. T.; Brusic, V.; Quackenbush, J.; Wahlestedt, C.; Mattick, J. S.; Hume, D. A.; Kai, C.; Sasaki, D.; Tomaru, Y.; Fukuda, S.; Kanamori-Katayama, M.; Suzuki, M.; Aoki, J.; Arakawa, T.; Iida, J.; Imamura, K.; Itoh, M.; Kato, T.; Kawaji, H.; Kawagashira, N.; Kawashima, T.; Kojima, M.; Kondo, S.; Konno, H.; Nakano, K.; Ninomiya, N.; Nishio, T.; Okada, M.; Plessy, C.; Shibata, K.; Shiraki, T.; Suzuki, S.; Tagami, M.; Waki, K.; Watahiki, A.; Okamura-Oho, Y.; Suzuki, H.; Kawai, J.; Hayashizaki, Y. Molecular Biology: The Transcriptional Landscape of the Mammalian Genome. *Science (80-.).* **2005**, *309* (5740), 1559–1563. https://doi.org/10.1126/science.1112014.

(23) Mattick, J. S.; Taft, R. J.; Faulkner, G. J. A Global View of Genomic Information - Moving beyond the Gene and the Master Regulator. *Trends Genet.* **2010**, *26* (1), 21–28. https://doi.org/10.1016/j.tig.2009.11.002.

(24) Shao, T.; Pan, Y. hong; Xiong, X. dong. Circular RNA: An

Important Player with Multiple Facets to Regulate Its Parental Gene Expression. *Mol. Ther. - Nucleic Acids* **2021**, *23*, 369–376. https://doi.org/10.1016/j.omtn.2020.11.008.

(25) Nagano, T.; Fraser, P. No-Nonsense Functions for Long Noncoding RNAs. *Cell* **2011**, *145* (2), 178–181. https://doi.org/10.1016/j.cell.2011.03.014.

(26) Kapranov, P.; Willingham, A. T.; Gingeras, T. R. Genome-Wide Transcription and the Implications for Genomic Organization. *Nat. Rev. Genet.* **2007**, *8* (6), 413–423. https://doi.org/10.1038/nrg2083.

(27) Hon, C. C.; Ramilowski, J. A.; Harshbarger, J.; Bertin, N.; Rackham, O. J. L.; Gough, J.; Denisenko, E.; Schmeier, S.; Poulsen, T. M.; Severin, J.; Lizio, M.; Kawaji, H.; Kasukawa, T.; Itoh, M.; Burroughs, A. M.; Noma, S.; Djebali, S.; Alam, T.; Medvedeva, Y. A.; Testa, A. C.; Lipovich, L.; Yip, C. W.; Abugessaisa, I.; Mendez, M.; Hasegawa, A.; Tang, D.; Lassmann, T.; Heutink, P.; Babina, M.; Wells, C. A.; Kojima, S.; Nakamura, Y.; Suzuki, H.; Daub, C. O.; De Hoon, M. J. L.; Arner, E.; Hayashizaki, Y.; Carninci, P.; Forrest, A. R. R. An Atlas of Human Long Non-Coding RNAs with Accurate 5′ Ends. *Nature* **2017**, *543* (7644), 199–204. https://doi.org/10.1038/nature21374.

(28) Mattick, J. S. Long Noncoding RNAs in Cell and Developmental Biology. *Semin. Cell Dev. Biol.* **2011**, *22* (4), 327. https://doi.org/10.1016/j.semcdb.2011.05.002.

(29) Chan, J. J.; Tay, Y. Noncoding RNA: RNA Regulatory Networks in Cancer. *International Journal of Molecular Sciences*. MDPI AG May 1, 2018. https://doi.org/10.3390/ijms19051310.

(30) Wang, K. C.; Chang, H. Y. Molecular Mechanisms of Long Noncoding RNAs. *Mol. Cell* **2011**, *43* (6), 904–914. https://doi.org/10.1016/j.molcel.2011.08.018.

(31) Arun, G.; Diermeier, S. D.; Spector, D. L. Therapeutic Targeting of Long Non-Coding RNAs in Cancer. *Trends Mol. Med.* **2018**, *24* (3), 257–277.

https://doi.org/10.1016/j.molmed.2018.01.001.
(32) Ulitsky, I.; Bartel, D. P. XLincRNAs: Genomics, Evolution, and Mechanisms. *Cell* **2013**, *154* (1), 26. https://doi.org/10.1016/j.cell.2013.06.020.
(33) Fatica, A.; Bozzoni, I. Long Non-Coding RNAs: New Players in Cell Differentiation and Development. *Nature Reviews Genetics*. January 2014, pp 7–21. https://doi.org/10.1038/nrg3606.
(34) Yao, R. W.; Wang, Y.; Chen, L. L. Cellular Functions of Long Noncoding RNAs. *Nat. Cell Biol.* **2019**, *21* (5), 542–551. https://doi.org/10.1038/s41556-019-0311-8.
(35) Iyer, M. K.; Niknafs, Y. S.; Malik, R.; Singhal, U.; Sahu, A.; Hosono, Y.; Barrette, T. R.; Prensner, J. R.; Evans, J. R.; Zhao, S.; Poliakov, A.; Cao, X.; Dhanasekaran, S. M.; Wu, Y. M.; Robinson, D. R.; Beer, D. G.; Feng, F. Y.; Iyer, H. K.; Chinnaiyan, A. M. The Landscape of Long Noncoding RNAs in the Human Transcriptome. *Nat. Genet.* **2015**, *47* (3), 199–208. https://doi.org/10.1038/ng.3192.
(36) Statello, L.; Guo, C. J.; Chen, L. L.; Huarte, M. Gene Regulation by Long Non-Coding RNAs and Its Biological Functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22* (2), 96–118. https://doi.org/10.1038/s41580-020-00315-9.
(37) Clark, M. B.; Johnston, R. L.; Inostroza-Ponta, M.; Fox, A. H.; Fortini, E.; Moscato, P.; Dinger, M. E.; Mattick, J. S. Genome-Wide Analysis of Long Noncoding RNA Stability. *Genome Res.* **2012**, *22* (5), 885–898. https://doi.org/10.1101/gr.131037.111.
(38) Guttman, M.; Garber, M.; Levin, J. Z.; Donaghey, J.; Robinson, J.; Adiconis, X.; Fan, L.; Koziol, M. J.; Gnirke, A.; Nusbaum, C.; Rinn, J. L.; Lander, E. S.; Regev, A. Ab Initio Reconstruction of Cell Type-Specific Transcriptomes in Mouse Reveals the Conserved Multi-Exonic Structure of LincRNAs. *Nat. Biotechnol.* **2010**, *28* (5), 503–510. https://doi.org/10.1038/nbt.1633.
(39) Quinn, J. J.; Chang, H. Y. Unique Features of Long Non-

Coding RNA Biogenesis and Function. *Nat. Rev. Genet.* **2016**, *17* (1), 47–62. https://doi.org/10.1038/nrg.2015.10.

(40) Chen, J.; Liu, Y.; Min, J.; Wang, H.; Li, F.; Xu, C.; Gong, A.; Xu, M. Alternative Splicing of LncRNAs in Human Diseases. *Am. J. Cancer Res.* **2021**, *11* (3), 624–639.

(41) Cabili, M.; Trapnell, C.; Goff, L.; Koziol, M.; Tazon-Vega, B.; Regev, A.; Rinn, J. L. Integrative Annotation of Human Large Intergenic Noncoding RNAs Reveals Global Properties and Specific Subclasses. *Genes Dev.* **2011**, *25* (18), 1915–1927. https://doi.org/10.1101/gad.17446611.

(42) Wu, H.; Yang, L.; Chen, L. L. The Diversity of Long Noncoding RNAs and Their Generation. *Trends Genet.* **2017**, *33* (8), 540–552. https://doi.org/10.1016/j.tig.2017.05.004.

(43) Yin, Q. F.; Yang, L.; Zhang, Y.; Xiang, J. F.; Wu, Y. W.; Carmichael, G. G.; Chen, L. L. Long Noncoding RNAs with SnoRNA Ends. *Mol. Cell* **2012**, *48* (2), 219–230. https://doi.org/10.1016/j.molcel.2012.07.033.

(44) Wu, H.; Yin, Q. F.; Luo, Z.; Yao, R. W.; Zheng, C. C.; Zhang, J.; Xiang, J. F.; Yang, L.; Chen, L. L. Unusual Processing Generates SPA LncRNAs That Sequester Multiple RNA Binding Proteins. *Mol. Cell* **2016**, *64* (3), 534–548. https://doi.org/10.1016/j.molcel.2016.10.007.

(45) Salzman, J.; Gawad, C.; Wang, P. L.; Lacayo, N.; Brown, P. O. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS One* **2012**, *7* (2). https://doi.org/10.1371/journal.pone.0030733.

(46) Chen, L. L. The Biogenesis and Emerging Roles of Circular RNAs. *Nat. Rev. Mol. Cell Biol.* **2016**, *17* (4), 205–211. https://doi.org/10.1038/nrm.2015.32.

(47) Chen, R.; Wang, S. K.; Belk, J. A.; Amaya, L.; Li, Z.; Cardenas, A.; Abe, B. T.; Chen, C. K.; Wender, P. A.; Chang, H. Y. Engineering Circular RNA for Enhanced Protein Production. *Nat. Biotechnol.* **2022**, 1–11.

https://doi.org/10.1038/s41587-022-01393-0.

(48) Nukala, S. B.; Jousma, J.; Cho, Y.; Lee, W. H.; Ong, S. G. Long Non-Coding RNAs and MicroRNAs as Crucial Regulators in Cardio-Oncology. *Cell Biosci.* **2022**, *12* (1), 24. https://doi.org/10.1186/s13578-022-00757-y.

(49) Isoda, T.; Moore, A. J.; He, Z.; Chandra, V.; Aida, M.; Denholtz, M.; Piet van Hamburg, J.; Fisch, K. M.; Chang, A. N.; Fahl, S. P.; Wiest, D. L.; Murre, C. Non-Coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell* **2017**, *171* (1), 103-119.e18. https://doi.org/10.1016/j.cell.2017.09.001.

(50) Mumbach, M. R.; Granja, J. M.; Flynn, R. A.; Roake, C. M.; Satpathy, A. T.; Rubin, A. J.; Qi, Y.; Jiang, Z.; Shams, S.; Louie, B. H.; Guo, J. K.; Gennert, D. G.; Corces, M. R.; Khavari, P. A.; Atianand, M. K.; Artandi, S. E.; Fitzgerald, K. A.; Greenleaf, W. J.; Chang, H. Y. HiChIRP Reveals RNA-Associated Chromosome Conformation. *Nat. Methods* **2019**, *16* (6), 489–492. https://doi.org/10.1038/s41592-019-0407-x.

(51) Paul, I. J.; Duerksen, J. D. Chromatin-Associated RNA Content of Heterochromatin and Euchromatin. *Mol. Cell. Biochem.* **1975**, *9* (1), 9–16. https://doi.org/10.1007/BF01731728.

(52) Dueva, R.; Akopyan, K.; Pederiva, C.; Trevisan, D.; Dhanjal, S.; Lindqvist, A.; Farnebo, M. Neutralization of the Positive Charges on Histone Tails by RNA Promotes an Open Chromatin Structure. *Cell Chem. Biol.* **2019**, *26* (10), 1436-1449.e5. https://doi.org/10.1016/j.chembiol.2019.08.002.

(53) Brockdorff, N.; Ashworth, A.; Kay, G. F.; McCabe, V. M.; Norris, D. P.; Cooper, P. J.; Swift, S.; Rastan, S. The Product of the Mouse Xist Gene Is a 15 Kb Inactive X-Specific Transcript Containing No Conserved ORF and Located in the Nucleus. *Cell* **1992**, *71* (3), 515–526.

https://doi.org/10.1016/0092-8674(92)90519-I.
(54) Boeren, J.; Gribnau, J. Xist-Mediated Chromatin Changes That Establish Silencing of an Entire X Chromosome in Mammals. *Curr. Opin. Cell Biol.* **2021**, *70*, 44–50. https://doi.org/10.1016/j.ceb.2020.11.004.
(55) Ji, P.; Diederichs, S.; Wang, W.; Böing, S.; Metzger, R.; Schneider, P. M.; Tidow, N.; Brandt, B.; Buerger, H.; Bulk, E.; Thomas, M.; Berdel, W. E.; Serve, H.; Müller-Tidow, C. MALAT-1, a Novel Noncoding RNA, and Thymosin B4 Predict Metastasis and Survival in Early-Stage Non-Small Cell Lung Cancer. *Oncogene* **2003**, *22* (39), 8031–8041. https://doi.org/10.1038/sj.onc.1206928.
(56) Gutschner, T.; Hämmerle, M.; Eißmann, M.; Hsu, J.; Kim, Y.; Hung, G.; Revenko, A.; Arun, G.; Stentrup, M.; Groß, M.; Zörnig, M.; MacLeod, A. R.; Spector, D. L.; Diederichs, S. The Noncoding RNA MALAT1 Is a Critical Regulator of the Metastasis Phenotype of Lung Cancer Cells. *Cancer Res.* **2013**, *73* (3), 1180–1189. https://doi.org/10.1158/0008-5472.CAN-12-2850.
(57) Yamazaki, T.; Souquere, S.; Chujo, T.; Kobelke, S.; Chong, Y. S.; Fox, A. H.; Bond, C. S.; Nakagawa, S.; Pierron, G.; Hirose, T. Functional Domains of NEAT1 Architectural LncRNA Induce Paraspeckle Assembly through Phase Separation. *Mol. Cell* **2018**, *70* (6), 1038-1053.e7. https://doi.org/10.1016/j.molcel.2018.05.019.
(58) Clemson, C. M.; Hutchinson, J. N.; Sara, S. A.; Ensminger, A. W.; Fox, A. H.; Chess, A.; Lawrence, J. B. An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles. *Mol. Cell* **2009**, *33* (6), 717–726. https://doi.org/10.1016/j.molcel.2009.01.026.
(59) Bond, C. S.; Fox, A. H. Paraspeckles: Nuclear Bodies Built on Long Noncoding RNA. *J. Cell Biol.* **2009**, *186* (5), 637–644. https://doi.org/10.1083/jcb.200906113.
(60) Pospiech, N.; Cibis, H.; Dietrich, L.; Müller, F.; Bange, T.;

Hennig, S. Identification of Novel PANDAR Protein Interaction Partners Involved in Splicing Regulation. *Sci. Rep.* **2018**, *8* (1), 1–9. https://doi.org/10.1038/s41598-018-21105-6.

(61) Bielli, P.; Bordi, M.; Di Biasio, V.; Sette, C. Regulation of BCL-X Splicing Reveals a Role for the Polypyrimidine Tract Binding Protein (PTBP1/HnRNP I) in Alternative 5' Splice Site Selection. *Nucleic Acids Res.* **2014**, *42* (19), 12070–12081. https://doi.org/10.1093/nar/gku922.

(62) Tsai, M. C.; Manor, O.; Wan, Y.; Mosammaparast, N.; Wang, J. K.; Lan, F.; Shi, Y.; Segal, E.; Chang, H. Y. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science (80-.).* **2010**, *329* (5992), 689–693. https://doi.org/10.1126/science.1192002.

(63) Xin, X.; Li, Q.; Fang, J.; Zhao, T. LncRNA HOTAIR: A Potential Prognostic Factor and Therapeutic Target in Human Cancers. *Front. Oncol.* **2021**, *11*, 2904. https://doi.org/10.3389/fonc.2021.679244.

(64) Ghafouri-Fard, S.; Hajiesmaeili, M.; Shoorei, H.; Bahroudi, Z.; Taheri, M.; Sharifi, G. The Impact of LncRNAs and MiRNAs in Regulation of Function of Cancer Stem Cells and Progression of Cancer. *Front. Cell Dev. Biol.* **2021**, *9*. https://doi.org/10.3389/fcell.2021.696820.

(65) Rinn, J. L.; Kertesz, M.; Wang, J. K.; Squazzo, S. L.; Xu, X.; Brugmann, S. A.; Goodnough, L. H.; Helms, J. A.; Farnham, P. J.; Segal, E.; Chang, H. Y. Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* **2007**, *129* (7), 1311–1323. https://doi.org/10.1016/j.cell.2007.05.022.

(66) Luo, M.; Jeong, M.; Sun, D.; Park, H. J.; Rodriguez, B. A. T.; Xia, Z.; Yang, L.; Zhang, X.; Sheng, K.; Darlington, G. J.; Li, W.; Goodell, M. A. Long Non-Coding RNAs Control Hematopoietic Stem Cell Function. *Cell Stem Cell* **2015**, *16* (4), 426–438. https://doi.org/10.1016/j.stem.2015.02.002.

(67) Noh, J. H.; Kim, K. M.; McClusky, W. G.; Abdelmohsen,

K.; Gorospe, M. Cytoplasmic Functions of Long Noncoding RNAs. *Wiley Interdisciplinary Reviews: RNA*. Blackwell Publishing Ltd May 1, 2018. https://doi.org/10.1002/wrna.1471.

(68) Wickens, M.; Bernstein, D. S.; Kimble, J.; Parker, R. A PUF Family Portrait: 3′UTR Regulation as a Way of Life. *Trends Genet.* **2002**, *18* (3), 150–157. https://doi.org/10.1016/S0168-9525(01)02616-6.

(69) Lee, S.; Kopp, F.; Chang, T. C.; Sataluri, A.; Chen, B.; Sivakumar, S.; Yu, H.; Xie, Y.; Mendell, J. T. Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **2016**, *164* (1–2), 69–80. https://doi.org/10.1016/j.cell.2015.12.017.

(70) Grelet, S.; Link, L. A.; Howley, B.; Obellianne, C.; Palanisamy, V.; Gangaraju, V. K.; Diehl, J. A.; Howe, P. H. A Regulated PNUTS MRNA to LncRNA Splice Switch Mediates EMT and Tumour Progression. *Nat. Cell Biol.* **2017**, *19* (9), 1105–1115. https://doi.org/10.1038/ncb3595.

(71) Statello, L.; Guo, C. J.; Chen, L. L.; Huarte, M. Gene Regulation by Long Non-Coding RNAs and Its Biological Functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22* (2), 96–118. https://doi.org/10.1038/s41580-020-00315-9.

(72) Omote, N.; Sakamoto, K.; Li, Q.; Schupp, J. C.; Adams, T.; Ahangari, F.; Chioccioli, M.; DeIuliis, G.; Hashimoto, N.; Hasegawa, Y.; Kaminski, N. Long Noncoding RNA TINCR Is a Novel Regulator of Human Bronchial Epithelial Cell Differentiation State. *Physiol. Rep.* **2021**, *9* (3). https://doi.org/10.14814/phy2.14727.

(73) Gong, C.; Maquat, L. E. LncRNAs Transactivate STAU1-Mediated MRNA Decay by Duplexing with 39 UTRs via Alu Eleme. *Nature* **2011**, *470* (7333), 284–290. https://doi.org/10.1038/nature09701.

(74) Yoon, J. H.; Abdelmohsen, K.; Srikantan, S.; Yang, X.; Martindale, J. L.; De, S.; Huarte, M.; Zhan, M.; Becker, K. G.; Gorospe, M. LincRNA-P21 Suppresses Target MRNA

Translation. *Mol. Cell* **2012**, *47* (4), 648–655. https://doi.org/10.1016/j.molcel.2012.06.027.

(75) Liu, B.; Sun, L.; Liu, Q.; Gong, C.; Yao, Y.; Lv, X.; Lin, L.; Yao, H.; Su, F.; Li, D.; Zeng, M.; Song, E. A Cytoplasmic NF-KB Interacting Long Noncoding RNA Blocks IκB Phosphorylation and Suppresses Breast Cancer Metastasis. *Cancer Cell* **2015**, *27* (3), 370–381. https://doi.org/10.1016/j.ccell.2015.02.004.

(76) Anderson, D. M.; Makarewich, C. A.; Anderson, K. M.; Shelton, J. M.; Bezprozvannaya, S.; Bassel-Duby, R.; Olson, E. N. Widespread Control of Calcium Signaling by a Family of SERCA-Inhibiting Micropeptides. *Sci. Signal.* **2016**, *9* (457), ra119. https://doi.org/10.1126/scisignal.aaj1460.

(77) Salvatori, B.; Biscarini, S.; Morlando, M. Non-Coding RNAs in Nervous System Development and Disease. *Frontiers in Cell and Developmental Biology*. Frontiers Media S.A. May 6, 2020. https://doi.org/10.3389/fcell.2020.00273.

(78) Clark, B. S.; Blackshaw, S. Understanding the Role of LncRNAs in Nervous System Development. *Adv. Exp. Med. Biol.* **2017**, *1008*, 253–282. https://doi.org/10.1007/978-981-10-5203-3_9.

(79) Rusconi, F.; Battaglioli, E.; Venturin, M. Psychiatric Disorders and Lncrnas: A Synaptic Match. *Int. J. Mol. Sci.* **2020**, *21* (9). https://doi.org/10.3390/ijms21093030.

(80) Chodroff, R. A.; Goodstadt, L.; Sirey, T. M.; Oliver, P. L.; Davies, K. E.; Green, E. D.; Molnár, Z.; Ponting, C. P. Long Noncoding RNA Genes: Caaonservation of Sequence and Brain Expression among Diverse Amniotes. *Genome Biol.* **2010**, *11* (7), R72. https://doi.org/10.1186/gb-2010-11-7-r72.

(81) Zayia, L. C.; Tadi, P. Neuroanatomy, Motor Neuron. *StatPearls* **2020**.

(82) Vangoor, V. R.; Gomes-Duarte, A.; Pasterkamp, R. J. Long

Non-Coding RNAs in Motor Neuron Development and Disease. *J. Neurochem.* **2021**, *156* (6), 777–801. https://doi.org/10.1111/jnc.15198.

(83) Stifani, N. Motor Neurons and the Generation of Spinal Motor Neuron Diversity. *Front. Cell. Neurosci.* **2014**, *8* (OCT), 293. https://doi.org/10.3389/fncel.2014.00293.

(84) De Santis, R.; Garone, M. G.; Pagani, F.; de Turris, V.; Di Angelantonio, S.; Rosa, A. Direct Conversion of Human Pluripotent Stem Cells into Cranial Motor Neurons Using a PiggyBac Vector. *Stem Cell Res.* **2018**, *29*, 189–196. https://doi.org/10.1016/j.scr.2018.04.012.

(85) Mehler, M. F.; Mattick, J. S. Noncoding RNAs and RNA Editing in Brain Development, Functional Diversification, and Neurological Disease. *Physiol. Rev.* **2007**, *87* (3), 799–823. https://doi.org/10.1152/physrev.00036.2006.

(86) Mercer, T. R.; Dinger, M. E.; Sunkin, S. M.; Mehler, M. F.; Mattick, J. S. Specific Expression of Long Noncoding RNAs in the Mouse Brain. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (2), 716–721. https://doi.org/10.1073/pnas.0706729105.

(87) Fernandes, D. P.; Bitar, M.; Jacobs, F. M. J.; Barry, G. Long Non-Coding RNAs in Neuronal Aging. *Non-coding RNA* **2018**, *4* (2), 12. https://doi.org/10.3390/ncrna4020012.

(88) Shi, C.; Zhang, L.; Qin, C. Long Non-Coding RNAs in Brain Development, Synaptic Biology, and Alzheimer's Disease. *Brain Res. Bull.* **2017**, *132*, 160–169. https://doi.org/10.1016/j.brainresbull.2017.03.010.

(89) Yen, Y. P.; Hsieh, W. F.; Tsai, Y. Y.; Lu, Y. L.; Liau, E. S.; Hsu, H. C.; Chen, Y. C.; Liu, T. C.; Chang, M.; Li, J.; Lin, S. P.; Hung, J. H.; Chen, J. A. Dlk1-Dio3 Locus-Derived LncRNAs Perpetuate Postmitotic Motor Neuron Cell Fate and Subtype Identity. *Elife* **2018**, *7*. https://doi.org/10.7554/eLife.38080.

(90) Ray, M. K.; Wiskow, O.; King, M. J.; Ismail, N.; Ergun, A.; Wang, Y.; Plys, A. J.; Davis, C. P.; Kathrein, K.; Sadreyev,

R.; Borowsky, M. L.; Eggan, K.; Zon, L.; Galloway, J. L.; Kingston, R. E. CAT7 and Cat7l Long Non-Coding Rnas Tune Polycomb Repressive Complex 1 Function during Human and Zebrafish Development. *J. Biol. Chem.* **2016**, *291* (37), 19558–19572. https://doi.org/10.1074/jbc.M116.730853.

(91) Gao, T.; Li, J.; Li, N.; Gao, Y.; Yu, L.; Zhuang, S.; Zhao, Y.; Dong, X. Lncrps25 Play an Essential Role in Motor Neuron Development Through Controlling the Expression of Olig2 in Zebrafish. *J. Cell. Physiol.* **2020**, *235* (4), 3485–3496. https://doi.org/10.1002/jcp.29237.

(92) Lu, Q. R.; Sun, T.; Zhu, Z.; Ma, N.; Garcia, M.; Stiles, C. D.; Rowitch, D. H. Common Developmental Requirement for Olig Function Indicates a Motor Neuron/Oligodendrocyte Connection. *Cell* **2002**, *109* (1), 75–86. https://doi.org/10.1016/S0092-8674(02)00678-5.

(93) Park, H. C.; Mehta, A.; Richardson, J. S.; Appel, B. Olig2 Is Required for Zebrafish Primary Motor Neuron and Oligodendrocyte Development. *Dev. Biol.* **2002**, *248* (2), 356–368. https://doi.org/10.1006/dbio.2002.0738.

(94) Briese, M.; Saal, L.; Appenzeller, S.; Moradi, M.; Baluapuri, A.; Sendtner, M. Whole Transcriptome Profiling Reveals the RNA Content of Motor Axons. *Nucleic Acids Res.* **2015**, *44* (4), e33–e33. https://doi.org/10.1093/nar/gkv1027.

(95) Ng, S. Y.; Bogu, G. K.; Soh, B. S.; Stanton, L. W. The Long Noncoding RNA RMST Interacts with SOX2 to Regulate Neurogenesis. *Mol. Cell* **2013**, *51* (3), 349–359. https://doi.org/10.1016/j.molcel.2013.07.017.

(96) Carvelli, A.; Setti, A.; Desideri, F.; Galfrè, S. G.; Biscarini, S.; Santini, T.; Colantoni, A.; Peruzzi, G.; Marzi, M. J.; Capauto, D.; Di Angelantonio, S.; Ballarino, M.; Nicassio, F.; Laneve, P.; Bozzoni, I. A Multifunctional Locus Controls Motor Neuron Differentiation through Short and Long Noncoding RNAs. *EMBO J.* **2022**, *41* (13), e108918. https://doi.org/10.15252/embj.2021108918.

(97) Pellegrini, F.; Padovano, V.; Biscarini, S.; Santini, T.; Setti, A.; Galfrè, S. G.; Silenzi, V.; Vitiello, E.; Mariani, D.; Nicoletti, C.; Torromino, G.; De Leonibus, E.; Martone, J.; Bozzoni, I. A KO Mouse Model for the LncRNA Lhx1os Produces Motor Neuron Alterations and Locomotor Impairment. *iScience* **2023**, *26* (1), 105891. https://doi.org/10.1016/j.isci.2022.105891.

(98) Fox, A. H.; Nakagawa, S.; Hirose, T.; Bond, C. S. Paraspeckles: Where Long Noncoding RNA Meets Phase Separation. *Trends Biochem. Sci.* **2018**, *43* (2), 124–135. https://doi.org/10.1016/j.tibs.2017.12.001.

(99) Naganuma, T.; Nakagawa, S.; Tanigawa, A.; Sasaki, Y. F.; Goshima, N.; Hirose, T. Alternative 3′-End Processing of Long Noncoding RNA Initiates Construction of Nuclear Paraspeckles. *EMBO J.* **2012**, *31* (20), 4020–4034. https://doi.org/10.1038/emboj.2012.251.

(100) Chujo, T.; Yamazaki, T.; Hirose, T. Architectural RNAs (ArcRNAs): A Class of Long Noncoding RNAs That Function as the Scaffold of Nuclear Bodies. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2016**, *1859* (1), 139–146. https://doi.org/10.1016/j.bbagrm.2015.05.007.

(101) Mao, Y. S.; Sunwoo, H.; Zhang, B.; Spector, D. L. Direct Visualization of the Co-Transcriptional Assembly of a Nuclear Body by Noncoding RNAs. *Nat. Cell Biol.* **2011**, *13* (1), 95–101. https://doi.org/10.1038/ncb2140.

(102) Hirose, T. The Building Process of the Functional Paraspeckle with Long Non-Coding RNAs. *Front. Biosci.* **2015**, *7* (1), 1–47. https://doi.org/10.2741/715.

(103) Nishimoto, Y.; Nakagawa, S.; Hirose, T.; Okano, H. J.; Takao, M.; Shibata, S.; Suyama, S.; Kuwako, K. I.; Imai, T.; Murayama, S.; Suzuki, N.; Okano, H. The Long Non-Coding RNA Nuclear-Enriched Abundant Transcript 1-2 Induces Paraspeckle Formation in the Motor Neuron during the Early Phase of Amyotrophic Lateral Sclerosis. *Mol. Brain* **2013**, *6* (1), 1–18. https://doi.org/10.1186/1756-6606-

6-31.
(104) Suzuki, H.; Shibagaki, Y.; Hattori, S.; Matsuoka, M. C9-ALS/FTD-Linked Proline–Arginine Dipeptide Repeat Protein Associates with Paraspeckle Components and Increases Paraspeckle Formation. *Cell Death Dis.* **2019**, *10* (10), 1–16. https://doi.org/10.1038/s41419-019-1983-5.
(105) DeJesus-Hernandez, M.; Mackenzie, I. R.; Boeve, B. F.; Boxer, A. L.; Baker, M.; Rutherford, N. J.; Nicholson, A. M.; Finch, N. C. A.; Flynn, H.; Adamson, J.; Kouri, N.; Wojtas, A.; Sengdy, P.; Hsiung, G. Y. R.; Karydas, A.; Seeley, W. W.; Josephs, K. A.; Coppola, G.; Geschwind, D. H.; Wszolek, Z. K.; Feldman, H.; Knopman, D. S.; Petersen, R. C.; Miller, B. L.; Dickson, D. W.; Boylan, K. B.; Graff-Radford, N. R.; Rademakers, R. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72* (2), 245–256. https://doi.org/10.1016/j.neuron.2011.09.011.
(106) Renton, A. E.; Majounie, E.; Waite, A.; Simón-Sánchez, J.; Rollinson, S.; Gibbs, J. R.; Schymick, J. C.; Laaksovirta, H.; van Swieten, J. C.; Myllykangas, L.; Kalimo, H.; Paetau, A.; Abramzon, Y.; Remes, A. M.; Kaganovich, A.; Scholz, S. W.; Duckworth, J.; Ding, J.; Harmer, D. W.; Hernandez, D. G.; Johnson, J. O.; Mok, K.; Ryten, M.; Trabzuni, D.; Guerreiro, R. J.; Orrell, R. W.; Neal, J.; Murray, A.; Pearson, J.; Jansen, I. E.; Sondervan, D.; Seelaar, H.; Blake, D.; Young, K.; Halliwell, N.; Callister, J. B.; Toulson, G.; Richardson, A.; Gerhard, A.; Snowden, J.; Mann, D.; Neary, D.; Nalls, M. A.; Peuralinna, T.; Jansson, L.; Isoviita, V. M.; Kaivorinne, A. L.; Hölttä-Vuori, M.; Ikonen, E.; Sulkava, R.; Benatar, M.; Wuu, J.; Chiò, A.; Restagno, G.; Borghero, G.; Sabatelli, M.; Heckerman, D.; Rogaeva, E.; Zinman, L.; Rothstein, J. D.; Sendtner, M.; Drepper, C.; Eichler, E. E.; Alkan, C.; Abdullaev, Z.; Pack, S. D.; Dutra, A.; Pak, E.; Hardy, J.; Singleton, A.; Williams, N. M.; Heutink, P.;

Pickering-Brown, S.; Morris, H. R.; Tienari, P. J.; Traynor, B. J. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **2011**, *72* (2), 257–268. https://doi.org/10.1016/j.neuron.2011.09.010.

(107) Zu, T.; Liu, Y.; Bañez-Coronel, M.; Reid, T.; Pletnikova, O.; Lewis, J.; Miller, T. M.; Harms, M. B.; Falchook, A. E.; Subramony, S. H.; Ostrow, L. W.; Rothstein, J. D.; Troncoso, J. C.; Ranum, L. P. W. RAN Proteins and RNA Foci from Antisense Transcripts in C9ORF72 ALS and Frontotemporal Dementia. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (51), E4968–E4977. https://doi.org/10.1073/pnas.1315438110.

(108) Huynh, D. P.; Del Bigio, M. R.; Ho, D. H.; Pulst, S. M. Expression of Ataxin-2 in Brains from Normal Individuals and Patients with Alzheimer's Disease and Spinocerebellar Ataxia 2. *Ann. Neurol.* **1999**, *45* (2), 232–241. https://doi.org/10.1002/1531-8249(199902)45:2<232::AID-ANA14>3.0.CO;2-7.

(109) Satterfield, T. F.; Pallanck, L. J. Ataxin-2 and Its Drosophila Homolog, ATX2, Physically Assemble with Polyribosomes. *Hum. Mol. Genet.* **2006**, *15* (16), 2523–2532. https://doi.org/10.1093/hmg/ddl173.

(110) Elden, A. C.; Kim, H. J.; Hart, M. P.; Chen-Plotkin, A. S.; Johnson, B. S.; Fang, X.; Armakola, M.; Geser, F.; Greene, R.; Lu, M. M.; Padmanabhan, A.; Clay-Falcone, D.; McCluskey, L.; Elman, L.; Juhr, D.; Gruber, P. J.; Rüb, U.; Auburger, G.; Trojanowski, J. Q.; Lee, V. M. Y.; Van Deerlin, V. M.; Bonini, N. M.; Gitler, A. D. Ataxin-2 Intermediate-Length Polyglutamine Expansions Are Associated with Increased Risk for ALS. *Nature* **2010**, *466* (7310), 1069–1075. https://doi.org/10.1038/nature09320.

(111) Li, P. P.; Sun, X.; Xia, G.; Arbez, N.; Paul, S.; Zhu, S.; Peng, H. B.; Ross, C. A.; Koeppen, A. H.; Margolis, R. L.; Pulst, S. M.; Ashizawa, T.; Rudnicki, D. D. ATXN2-AS, a

Gene Antisense to ATXN2, Is Associated with Spinocerebellar Ataxia Type 2 and Amyotrophic Lateral Sclerosis. *Ann. Neurol.* **2016**, *80* (4), 600–615. https://doi.org/10.1002/ana.24761.

(112) Swinnen, B.; Robberecht, W.; Van Den Bosch, L. RNA Toxicity in Non-coding Repeat Expansion Disorders. *EMBO J.* **2020**, *39* (1), e101112. https://doi.org/10.15252/embj.2018101112.

(113) Sessa, L.; Breiling, A.; Lavorgna, G.; Silvestri, L.; Casari, G.; Orlando, V. Noncoding RNA Synthesis and Loss of Polycomb Group Repression Accompanies the Colinear Activation of the Human HOXA Cluster. *Rna* **2007**, *13* (2), 223–239. https://doi.org/10.1261/rna.266707.

(114) Zhang, X.; Lian, Z.; Padden, C.; Gerstein, M. B.; Rozowsky, J.; Snyder, M.; Gingeras, T. R.; Kapranov, P.; Weissman, S. M.; Newburger, P. E. A Myelopoiesis-Associated Regulatory Intergenic Noncoding RNA Transcript within the Human HOXA Cluster. *Blood* **2009**, *113* (11), 2526–2534. https://doi.org/10.1182/blood-2008-06-162164.

(115) Samarut, E.; Rochette-Egly, C. Nuclear Retinoic Acid Receptors: Conductors of the Retinoic Acid Symphony during Development. *Mol. Cell. Endocrinol.* **2012**, *348* (2), 348–360. https://doi.org/10.1016/j.mce.2011.03.025.

(116) Lalevée, S.; Anno, Y. N.; Chatagnon, A.; Samarut, E.; Poch, O.; Laudet, V.; Benoit, G.; Lecompte, O.; Rochette-Egly, C. Genome-Wide in Silico Identification of New Conserved and Functional Retinoic Acid Receptor Response Elements (Direct Repeats Separated by 5 Bp). *J. Biol. Chem.* **2011**, *286* (38), 33322–33334. https://doi.org/10.1074/jbc.M111.263681.

(117) Fu, L.; Peng, S.; Wu, W.; Ouyang, Y.; Tan, D.; Fu, X. LncRNA HOTAIRM1 Promotes Osteogenesis by Controlling JNK/AP-1 Signalling-Mediated RUNX2 Expression. *J. Cell. Mol. Med.* **2019**, *23* (11), 7517–7524.

https://doi.org/10.1111/jcmm.14620.
(118) Zhang, X.; Weissman, S. M.; Newburger, P. E. Long Intergenic Non-Coding RNA HOTAIRM1 Regulates Cell Cycle Progression during Myeloid Maturation in NB4 Human Promyelocytic Leukemia Cells. *RNA Biol.* **2014**, *11* (6), 777–787. https://doi.org/10.4161/rna.28828.
(119) Lin, M.; Pedrosa, E.; Shah, A.; Hrabovsky, A.; Maqbool, S.; Zheng, D.; Lachman, H. M. RNA-Seq of Human Neurons Derived from IPS Cells Reveals Candidate Long Non-Coding RNAs Involved in Neurogenesis and Neuropsychiatric Disorders. *PLoS One* **2011**, *6* (9), e23356. https://doi.org/10.1371/journal.pone.0023356.
(120) Wan, L.; Kong, J.; Tang, J.; Wu, Y.; Xu, E.; Lai, M.; Zhang, H. HOTAIRM1 as a Potential Biomarker for Diagnosis of Colorectal Cancer Functions the Role in the Tumour Suppressor. *J. Cell. Mol. Med.* **2016**, *20* (11), 2036–2044. https://doi.org/10.1111/jcmm.12892.
(121) Chen, Y.; Wu, J.-J.; Lin, X.-B.; Bao, Y.; Chen, Z.-H.; Zhang, C.-R.; Cai, Z.; Zhou, J.-Y.; Ding, M.-H.; Wu, X.-J.; Sun, W.; Qian, J.; Zhang, L.; Jiang, L.; Hu, G.-H. Differential LncRNA Expression Profiles in Recurrent Gliomas Compared with Primary Gliomas Identified by Microarray Analysis. *Int. J. Clin. Exp. Med.* **2015**, *8* (4), 5033–5043.
(122) Chen, Z. H.; Wang, W. T.; Huang, W.; Fang, K.; Sun, Y. M.; Liu, S. R.; Luo, X. Q.; Chen, Y. Q. The LncRNA HOTAIRM1 Regulates the Degradation of PML-RARA Oncoprotein and Myeloid Cell Differentiation by Enhancing the Autophagy Pathway. *Cell Death Differ.* **2017**, *24* (2), 212–224. https://doi.org/10.1038/cdd.2016.111.
(123) Moazzam-Jazi, M.; Lanjanian, H.; Maleknia, S.; Hedayati, M.; Daneshpour, M. S. Interplay between SARS-CoV-2 and Human Long Non-Coding RNAs. *J. Cell. Mol. Med.* **2021**, *25* (12), 5823–5827. https://doi.org/10.1111/jcmm.16596.
(124) Wei, S.; Zhao, M.; Wang, X.; Li, Y.; Wang, K. PU.1

Controls the Expression of Long Noncoding RNA HOTAIRM1 during Granulocytic Differentiation. *J. Hematol. Oncol.* **2016**, *9* (1), 44. https://doi.org/10.1186/s13045-016-0274-1.

(125) Zhang, X.; Weissman, S. M.; Newburger, P. E. Long Intergenic Non-Coding RNA HOTAIRM1 Regulates Cell Cycle Progression during Myeloid Maturation in NB4 Human Promyelocytic Leukemia Cells. *RNA Biol.* **2014**, *11* (6), 777–787. https://doi.org/10.4161/rna.28828.

(126) Xin, J.; Li, J.; Feng, Y.; Wang, L.; Zhang, Y.; Yang, R. Downregulation of Long Noncoding RNA HOTAIRMI Promotes Monocyte/Dendritic Cell Differentiation through Competitively Binding to Endogenous MiR-3960. *Onco. Targets. Ther.* **2017**, *10*, 1307–1315. https://doi.org/10.2147/OTT.S124201.

(127) GTEx Project. GTEx portal https://gtexportal.org/home/ (accessed Jan 3, 2023).

(128) Mazzoni, E. O.; Mahony, S.; Closser, M.; Morrison, C. A.; Nedelec, S.; Williams, D. J.; An, D.; Gifford, D. K.; Wichterle, H. Synergistic Binding of Transcription Factors to Cell-Specific Enhancers Programs Motor Neuron Identity. *Nat. Neurosci.* **2013**, *16* (9), 1219–1227. https://doi.org/10.1038/nn.3467.

(129) Sepehrimanesh, M.; Ding, B. Generation and Optimization of Highly Pure Motor Neurons from Human Induced Pluripotent Stem Cells via Lentiviral Delivery of Transcription Factors. *Am. J. Physiol. - Cell Physiol.* **2020**, *319* (4), C771–C780. https://doi.org/10.1152/AJPCELL.00279.2020.

(130) Solomon, E.; Davis-Anderson, K.; Hovde, B.; Micheva-Viteva, S.; Harris, J. F.; Twary, S.; Iyer, R. Global Transcriptome Profile of the Developmental Principles of in Vitro IPSC-to-Motor Neuron Differentiation. *BMC Mol. Cell Biol.* **2021**, *22* (1), 1–21. https://doi.org/10.1186/s12860-021-00343-z.

(131) Smith, N. C.; Wilkinson-White, L. E.; Kwan, A. H. Y.; Trewhella, J.; Matthews, J. M. Contrasting DNA-Binding Behaviour by ISL1 and LHX3 Underpins Differential Gene Targeting in Neuronal Cell Specification. *J. Struct. Biol. X* **2021**, *5*, 100043. https://doi.org/10.1016/j.yjsbx.2020.100043.
(132) Garone, M. G.; De Turris, V.; Soloperto, A.; Brighi, C.; De Santis, R.; Pagani, F.; Di Angelantonio, S.; Rosa, A. Conversion of Human Induced Pluripotent Stem Cells (IPSCs) into Functional Spinal and Cranial Motor Neurons Using PiggyBac Vectors. *J. Vis. Exp.* **2019**, *2019* (147). https://doi.org/10.3791/59321.
(133) Ludwik, K. A.; von Kuegelgen, N.; Chekulaeva, M. Genome-Wide Analysis of RNA and Protein Localization and Local Translation in MESC-Derived Neurons. *Methods* **2019**, *162–163*, 31–41. https://doi.org/10.1016/j.ymeth.2019.02.002.
(134) Harrison, S. J.; Nishinakamura, R.; Jones, K. R.; Monaghan, A. P. Sall1 Regulates Cortical Neurogenesis and Laminar Fate Specification in Mice: Implications for Neural Abnormalities in Townes-Brocks Syndrome. *DMM Dis. Model. Mech.* **2012**, *5* (3), 351–365. https://doi.org/10.1242/dmm.002873.
(135) Asprer, J. S. T.; Lee, B.; Wu, C. S.; Vadakkan, T.; Dickinson, M. E.; Lu, H. C.; Lee, S. K. LMO4 Functions as a Co-Activator of Neurogenin 2 in the Developing Cortex. *Development* **2011**, *138* (13), 2823–2832. https://doi.org/10.1242/dev.061879.
(136) Boukhtouche, F.; Doulazmi, M.; Frederic, F.; Dusart, I.; Brugg, B.; Mariani, J. RORα, a Pivotal Nuclear Receptor for Purkinje Neuron Survival and Differentiation: From Development to Ageing. *Cerebellum* **2006**, *5* (2), 97–104. https://doi.org/10.1080/14734220600750184.
(137) Bourane, S.; Grossmann, K. S.; Britz, O.; Dalet, A.; Del Barrio, M. G.; Stam, F. J.; Garcia-Campmany, L.; Koch, S.;

Goulding, M. Identification of a Spinal Circuit for Light Touch and Fine Motor Control. *Cell* **2015**, *160* (3), 503–515. https://doi.org/10.1016/j.cell.2015.01.011.

(138) Richards, A. B.; Scheel, T. A.; Wang, K.; Henkemeyer, M.; Kromer, L. F. EphB1 Null Mice Exhibit Neuronal Loss in Substantia Nigra Pars Reticulata and Spontaneous Locomotor Hyperactivity. *Eur. J. Neurosci.* **2007**, *25* (9), 2619–2628. https://doi.org/10.1111/j.1460-9568.2007.05523.x.

(139) Jevince, A. R.; Kadison, S. R.; Pittman, A. J.; Chien, C. Bin; Kaprielian, Z. Distribution of EphB Receptors and Ephrin-B1 in the Developing Vertebrate Spinal Cord. *J. Comp. Neurol.* **2006**, *497* (5), 734–750. https://doi.org/10.1002/cne.21001.

(140) Zou, M.; Luo, H.; Xiang, M. Selective Neuronal Lineages Derived from Dll4-Expressing Progenitors/Precursors in the Retina and Spinal Cord. *Dev. Dyn.* **2015**, *244* (1), 86–97. https://doi.org/10.1002/dvdy.24185.

(141) Kaltezioti, V.; Antoniou, D.; Stergiopoulos, A.; Rozani, I.; Rohrer, H.; Politis, P. K. Prox1 Regulates Olig2 Expression to Modulate Binary Fate Decisions in Spinal Cord Neurons. *J. Neurosci.* **2014**, *34* (47), 15816–15831. https://doi.org/10.1523/JNEUROSCI.1865-14.2014.

(142) Sharma, K.; Sheng, H. Z.; Lettieri, K.; Li, H.; Karavanov, A.; Potter, S.; Westphal, H.; Pfaff, S. L. LIM Homeodomain Factors Lhx3 and Lhx4 Assign Subtype Identities for Motor Neurons. *Cell* **1998**, *95* (6), 817–828. https://doi.org/10.1016/S0092-8674(00)81704-3.

(143) Kim, M.; Fontelonga, T.; Roesener, A. P.; Lee, H.; Gurung, S.; Mendonca, P. R. F.; Mastick, G. S. Motor Neuron Cell Bodies Are Actively Positioned by Slit/Robo Repulsion and Netrin/DCC Attraction. *Dev. Biol.* **2015**, *399* (1), 68–79. https://doi.org/10.1016/j.ydbio.2014.12.014.

(144) Arber, S.; Han, B.; Mendelsohn, M.; Smith, M.; Jessell, T. M.; Sockanathan, S. Requirement for the Homeobox Gene

Hb9 in the Consolidation of Motor Neuron Identity. *Neuron* **1999**, *23* (4), 659–674. https://doi.org/10.1016/S0896-6273(01)80026-X.

(145) Fertuzinhos, S.; Krsnik, E.; Kawasawa, Y. I.; Rain, M. R.; Kwan, K. Y.; Chen, J. G.; Juda, M.; Hayashi, M.; Sestan, N. Selective Depletion of Molecularly Defined Cortical Interneurons in Human Holoprosencephaly with Severe Striatal Hypoplasia. *Cereb. Cortex* **2009**, *19* (9), 2196–2207. https://doi.org/10.1093/cercor/bhp009.

(146) Wu, S. J.; Sevier, E.; Saldi, G. A.; Yu, S.; Abbott, L.; Choi, D. H.; Sherer, M.; Qiu, Y.; Shinde, A.; Rizzo, D.; Xu, Q.; Barrera, I.; Kumar, V.; Marrero, G.; Prönneke, A.; Huang, S.; Rudy, B.; Stafford, D. A.; Macosko, E.; Chen, F.; Fishell, G. Cortical Somatostatin Interneuron Subtypes Form Cell-Type Specific Circuits. *SSRN Electron. J.* **2022**, 2022.09.29.510081. https://doi.org/10.2139/ssrn.4227261.

(147) Lim, L.; Mi, D.; Llorca, A.; Marín, O. Development and Functional Diversification of Cortical Interneurons. *Neuron* **2018**, *100* (2), 294–313. https://doi.org/10.1016/j.neuron.2018.10.009.

(148) Symmank, J.; Gölling, V.; Gerstmann, K.; Zimmer, G. The Transcription Factor LHX1 Regulates the Survival and Directed Migration of POA-Derived Cortical Interneurons. *Cereb. Cortex* **2019**, *29* (4), 1644–1658. https://doi.org/10.1093/cercor/bhy063.

(149) Apicella, A. junior; Marchionni, I. VIP-Expressing GABAergic Neurons: Disinhibitory vs. Inhibitory Motif and Its Role in Communication Across Neocortical Areas. *Front. Cell. Neurosci.* **2022**, *16*, 1. https://doi.org/10.3389/fncel.2022.811484.

(150) Deng, Y.; Bi, M.; Delerue, F.; Forrest, S. L.; Chan, G.; van der Hoven, J.; van Hummel, A.; Feiten, A. F.; Lee, S.; Martinez-Valbuena, I.; Karl, T.; Kovacs, G. G.; Morahan, G.; Ke, Y. D.; Ittner, L. M. Loss of LAMP5 Interneurons Drives Neuronal Network Dysfunction in Alzheimer's

Disease. *Acta Neuropathol.* **2022**, *144* (4), 637–650. https://doi.org/10.1007/s00401-022-02457-w.

(151) Que, L.; Lukacsovich, D.; Luo, W.; Földy, C. Transcriptional and Morphological Profiling of Parvalbumin Interneuron Subpopulations in the Mouse Hippocampus. *Nat. Commun.* **2021**, *12* (1), 108. https://doi.org/10.1038/s41467-020-20328-4.

(152) Le, T. N.; Zhou, Q. P.; Cobos, I.; Zhang, S.; Zagozewski, J.; Japoni, S.; Vriend, J.; Parkinson, T.; Du, G.; Rubenstein, J. L.; Eisenstat, D. D. GABAergic Interneuron Differentiation in the Basal Forebrain Is Mediated through Direct Regulation of Glutamic Acid Decarboxylase Isoforms by Dlx Homeobox Transcription Factors. *J. Neurosci.* **2017**, *37* (36), 8816–8829. https://doi.org/10.1523/JNEUROSCI.2125-16.2017.

(153) Hoshino, C.; Konno, A.; Hosoi, N.; Kaneko, R.; Mukai, R.; Nakai, J.; Hirai, H. GABAergic Neuron-Specific Whole-Brain Transduction by AAV-PHP.B Incorporated with a New GAD65 Promoter. *Mol. Brain* **2021**, *14* (1), 33. https://doi.org/10.1186/s13041-021-00746-1.

(154) Wu, D.; Hersh, L. B. Choline Acetyltransferase: Celebrating Its Fiftieth Year. *J. Neurochem.* **1994**, *62* (5), 1653–1663. https://doi.org/10.1046/j.1471-4159.1994.62051653.x.

(155) Petralia, R. S.; Sans, N.; Wang, Y. X.; Wenthold, R. J. Ontogeny of Postsynaptic Density Proteins at Glutamatergic Synapses. *Mol. Cell. Neurosci.* **2005**, *29* (3), 436–452. https://doi.org/10.1016/j.mcn.2005.03.013.

(156) Brümmendorf, T.; Rathjen, F. G. Structure/Function Relationships of Axon-Associated Adhesion Receptors of the Immunoglobulin Superfamily. *Curr. Opin. Neurobiol.* **1996**, *6* (5), 584–593. https://doi.org/10.1016/S0959-4388(96)80089-4.

(157) Brand, Y.; Sung, M.; Pak, K.; Chavez, E.; Wei, E.; Radojevic, V.; Bodmer, D.; Ryan, A. F. Neural Cell Adhesion Molecule NrCAM Is Expressed in the Mammalian

Inner Ear and Modulates Spiral Ganglion Neurite Outgrowth in an In Vitro Alternate Choice Assay. *J. Mol. Neurosci.* **2015**, *55* (4), 836–844. https://doi.org/10.1007/s12031-014-0436-y.

(158) Huang, F. F.; Ben Aissa, M.; Lévesque, G.; Carreau, M. FANCC Localizes with UNC5A at Neurite Outgrowth and Promotes Neuritogenesis. *BMC Res. Notes* **2018**, *11* (1), 662. https://doi.org/10.1186/s13104-018-3763-1.

(159) Sun, K. L. W.; Correia, J. P.; Kennedy, T. E. Netrins: Versatile Extracellular Cues with Diverse Functions. *Development* **2011**, *138* (11), 2153–2169. https://doi.org/10.1242/dev.044529.

(160) Brose, K.; Bland, K. S.; Kuan, H. W.; Arnott, D.; Henzel, W.; Goodman, C. S.; Tessier-Lavigne, M.; Kidd, T. Slit Proteins Bind Robo Receptors and Have an Evolutionarily Conserved Role in Repulsive Axon Guidance. *Cell* **1999**, *96* (6), 795–806. https://doi.org/10.1016/S0092-8674(00)80590-5.

(161) Blockus, H.; Chédotal, A. Slit-Robo Signaling. *Dev.* **2016**, *143* (17), 3037–3044. https://doi.org/10.1242/dev.132829.

(162) Kim, M.; Lee, C. H.; Barnum, S. J.; Watson, R. C.; Li, J.; Mastick, G. S. Slit/Robo Signals Prevent Spinal Motor Neuron Emigration by Organizing the Spinal Cord Basement Membrane. *Dev. Biol.* **2019**, *455* (2), 449–457. https://doi.org/10.1016/j.ydbio.2019.07.017.

(163) Dillon, A. K.; Fujita, S. C.; Matise, M. P.; Jarjour, A. A.; Kennedy, T. E.; Kollmus, H.; Arnold, H. H.; Weiner, J. A.; Sanes, J. R.; Kaprielian, Z. Molecular Control of Spinal Accessory Motor Neuron/Axon Development in the Mouse Spinal Cord. *J. Neurosci.* **2005**, *25* (44), 10119–10130. https://doi.org/10.1523/JNEUROSCI.3455-05.2005.

(164) Bai, G.; Chivatakarn, O.; Bonanomi, D.; Lettieri, K.; Franco, L.; Xia, C.; Stein, E.; Ma, L.; Lewcock, J. W.; Pfaff, S. L. Presenilin-Dependent Receptor Processing Is Required for Axon Guidance. *Cell* **2011**, *144* (1), 106–118.

https://doi.org/10.1016/j.cell.2010.11.053.
(165) Baier, H.; Bonhoeffer, F. Axon Guidance by Gradients of a Target-Derived Component. *Science (80-.).* **1992**, *255* (5043), 472–475. https://doi.org/10.1126/science.1734526.
(166) Kolodkin, A. L.; Matthes, D. J.; Goodman, C. S. The Semaphorin Genes Encode a Family of Transmembrane and Secreted Growth Cone Guidance Molecules. *Cell* **1993**, *75* (7), 1389–1399. https://doi.org/10.1016/0092-8674(93)90625-Z.
(167) Pasterkamp, R. J. Getting Neural Circuits into Shape with Semaphorins. *Nat. Rev. Neurosci.* **2012**, *13* (9), 605–618. https://doi.org/10.1038/nrn3302.
(168) Pani, G.; De Vos, W. H.; Samari, N.; de Saint-Georges, L.; Baatout, S.; Van Oostveldt, P.; Benotmane, M. A. MorphoNeuroNet: An Automated Method for Dense Neurite Network Analysis. *Cytom. Part A* **2014**, *85* (2), 188–199. https://doi.org/10.1002/cyto.a.22408.
(169) Pinggera, A.; Striessnig, J. Cav1.3 (CACNA1D) L-Type Ca2+ Channel Dysfunction in CNS Disorders. *J. Physiol.* **2016**, *594* (20), 5839–5849. https://doi.org/10.1113/JP270672.
(170) Hirtz, J. J.; Boesen, M.; Braun, N.; Deitmer, J. W.; Kramer, F.; Lohr, C.; Müller, B.; Nothwang, H. G.; Striessnig, J.; Löhrke, S.; Friauf, E. Cav1.3 Calcium Channels Are Required for Normal Development of the Auditory Brainstem. *J. Neurosci.* **2011**, *31* (22), 8280–8294. https://doi.org/10.1523/JNEUROSCI.5098-10.2011.
(171) Tao, R.; Li, C.; Jaffe, A. E.; Shin, J. H.; Deep-Soboslay, A.; Yamin, R.; Weinberger, D. R.; Hyde, T. M.; Kleinman, J. E. Cannabinoid Receptor CNR1 Expression and DNA Methylation in Human Prefrontal Cortex, Hippocampus and Caudate in Brain Development and Schizophrenia. *Transl. Psychiatry* **2020**, *10* (1). https://doi.org/10.1038/s41398-020-0832-8.
(172) Jakowec, M. W.; Yen, L.; Kalb, R. G. In Situ Hybridization

Analysis of AMPA Receptor Subunit Gene Expression in the Developing Rat Spinal Cord. *Neuroscience* **1995**, *67* (4), 909–920. https://doi.org/10.1016/0306-4522(95)00094-Y.

(173) Inglis, F. M.; Crockett, R.; Korada, S.; Abraham, W. C.; Hollmann, M.; Kalb, R. G. The AMPA Receptor Subunit GluR1 Regulates Dendritic Architecture of Motor Neurons. *J. Neurosci.* **2002**, *22* (18), 8042–8051. https://doi.org/10.1523/jneurosci.22-18-08042.2002.

(174) Takenouchi, T.; Hashida, N.; Torii, C.; Kosaki, R.; Takahashi, T.; Kosaki, K. 1p34.3 Deletion Involving GRIK3: Further Clinical Implication of GRIK Family Glutamate Receptors in the Pathogenesis of Developmental Delay. *Am. J. Med. Genet. Part A* **2014**, *164* (2), 456–460. https://doi.org/10.1002/ajmg.a.36240.

(175) Pinheiro, P. S.; Perrais, D.; Coussen, F.; Barhanin, J.; Bettler, B.; Mann, J. R.; Malva, J. O.; Heinemann, S. F.; Mulle, C. GluR7 Is an Essential Subunit of Presynaptic Kainate Autoreceptors at Hippocampal Mossy Fiber Synapses. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (29), 12181–12186. https://doi.org/10.1073/pnas.0608891104.

(176) Zaslavsky, K.; Zhang, W. B.; McCready, F. P.; Rodrigues, D. C.; Deneault, E.; Loo, C.; Zhao, M.; Ross, P. J.; El Hajjar, J.; Romm, A.; Thompson, T.; Piekna, A.; Wei, W.; Wang, Z.; Khattak, S.; Mufteev, M.; Pasceri, P.; Scherer, S. W.; Salter, M. W.; Ellis, J. SHANK2 Mutations Associated with Autism Spectrum Disorder Cause Hyperconnectivity of Human Neurons. *Nat. Neurosci.* **2019**, *22* (4), 556–564. https://doi.org/10.1038/s41593-019-0365-8.

(177) Maretina, M. A.; Valetdinova, K. R.; Tsyganova, N. A.; Egorova, A. A.; Ovechkina, V. S.; Schiöth, H. B.; Zakian, S. M.; Baranov, V. S.; Kiselev, A. V. Identification of Specific Gene Methylation Patterns during Motor Neuron Differentiation from Spinal Muscular Atrophy Patient-Derived IPSC. *Gene* **2022**, *811*, 146109. https://doi.org/10.1016/j.gene.2021.146109.

(178) Agostini, F.; Zanzoni, A.; Klus, P.; Marchese, D.; Cirillo, D.; Tartaglia, G. G. CatRAPID Omics: A Web Server for Large-Scale Prediction of Protein-RNA Interactions. *Bioinformatics* **2013**, *29* (22), 2928–2930. https://doi.org/10.1093/bioinformatics/btt495.
(179) Cirillo, D.; Blanco, M.; Armaos, A.; Buness, A.; Avner, P.; Guttman, M.; Cerase, A.; Tartaglia, G. G. Quantitative Predictions of Protein Interactions with Long Noncoding RNAs: To the Editor. *Nat. Methods* **2016**, *14* (1), 5–6. https://doi.org/10.1038/nmeth.4100.
(180) Zang, Y.; Chaudhari, K.; Bashaw, G. J. New Insights into the Molecular Mechanisms of Axon Guidance Receptor Regulation and Signaling. In *Current Topics in Developmental Biology*; NIH Public Access, 2021; Vol. 142, pp 147–196. https://doi.org/10.1016/bs.ctdb.2020.11.008.
(181) Mann, M.; Wright, P. R.; Backofen, R. IntaRNA 2.0: Enhanced and Customizable Prediction of RNA-RNA Interactions. *Nucleic Acids Res.* **2017**, *45* (W1), W435–W439. https://doi.org/10.1093/nar/gkx279.
(182) Kikin, O.; D'Antonio, L.; Bagga, P. S. QGRS Mapper: A Web-Based Server for Predicting G-Quadruplexes in Nucleotide Sequences. *Nucleic Acids Res.* **2006**, *34* (WEB. SERV. ISS.), W676–W682. https://doi.org/10.1093/nar/gkl253.
(183) Hinnebusch, A. G.; Ivanov, I. P.; Sonenberg, N. Translational Control by 5′-Untranslated Regions of Eukaryotic MRNAs. *Science (80-.).* **2016**, *352* (6292), 1413–1416. https://doi.org/10.1126/science.aad9868.
(184) Thaler, J.; Harrison, K.; Sharma, K.; Lettieri, K.; Kehrl, J.; Pfaff, S. L. Active Suppression of Interneuron Programs within Developing Motor Neurons Revealed by Analysis of Homeodomain Factor HB9. *Neuron* **1999**, *23* (4), 675–687. https://doi.org/10.1016/S0896-6273(01)80027-1.
(185) Sakurai, T.; Ramoz, N.; Reichert, J. G.; Corwin, T. E.;

Kryzak, L.; Smith, C. J.; Silverman, J. M.; Hollander, E.; Buxbaum, J. D. Association Analysis of the NrCAM Gene in Autism and in Subsets of Families with Severe Obsessive-Compulsive or Self-Stimulatory Behaviors. *Psychiatr. Genet.* **2006**, *16* (6), 251–257. https://doi.org/10.1097/01.ypg.0000242196.81891.c9.

(186) Sakurai, T. The Role of NrCAM in Neural Development and Disorders-Beyond a Simple Glue in the Brain. *Mol. Cell. Neurosci.* **2012**, *49* (3), 351–363. https://doi.org/10.1016/j.mcn.2011.12.002.

(187) Gu, Z.; Kalamboglas, J.; Yoshioka, S.; Han, W.; Li, Z.; Kawasawa, Y. I.; Pochareddy, S.; Li, Z.; Liu, F.; Xu, X.; Wijeratne, S.; Ueno, M.; Blatz, E.; Salomone, J.; Kumanogoh, A.; Rasin, M. R.; Gebelein, B.; Weirauch, M. T.; Sestan, N.; Martin, J. H.; Yoshida, Y. Control of Species-Dependent Cortico-Motoneuronal Connections Underlying Manual Dexterity. *Science (80-.).* **2017**, *357* (6349), 400–404. https://doi.org/10.1126/science.aan3721.

(188) Sheng, J.; Xu, J.; Geng, K.; Liu, D. Sema6D Regulates Zebrafish Vascular Patterning and Motor Neuronal Axon Growth in Spinal Cord. *Front. Mol. Neurosci.* **2022**, *15*. https://doi.org/10.3389/fnmol.2022.854556.

(189) Oda, Y. Choline Acetyltransferase: The Structure, Distribution and Pathologic Changes in the Central Nervous System. *Pathol. Int.* **1999**, *49* (11), 921–937. https://doi.org/10.1046/j.1440-1827.1999.00977.x.

(190) Garone, M. G.; Salerno, D.; Rosa, A. Digital Color-Coded Molecular Barcoding Reveals Dysregulation of Common FUS and FMRP Targets in Soma and Neurites of ALS Mutant Motoneurons. *bioRxiv Mol. Biol.* **2022**, 2022.08.02.502510. https://doi.org/10.1101/2022.08.02.502510.

(191) Bolger, A. M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30* (15), 2114–2120.

https://doi.org/10.1093/bioinformatics/btu170.
(192) Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29* (1), 15–21. https://doi.org/10.1093/bioinformatics/bts635.
(193) Li, B.; Dewey, C. N. RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinformatics* **2011**, *12* (1), 323. https://doi.org/10.1186/1471-2105-12-323.
(194) Love, M. I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15* (12), 550. https://doi.org/10.1186/s13059-014-0550-8.
(195) Zhu, A.; Ibrahim, J. G.; Love, M. I. Heavy-Tailed Prior Distributions for Sequence Count Data: Removing the Noise and Preserving Large Differences. *Bioinformatics* **2019**, *35* (12), 2084–2092. https://doi.org/10.1093/bioinformatics/bty895.
(196) Liao, Y.; Wang, J.; Jaehnig, E. J.; Shi, Z.; Zhang, B. WebGestalt 2019: Gene Set Analysis Toolkit with Revamped UIs and APIs. *Nucleic Acids Res.* **2019**, *47* (W1), W199–W205. https://doi.org/10.1093/nar/gkz401.
(197) Santini, T.; Martone, J.; Ballarino, M. Visualization of Nuclear and Cytoplasmic Long Noncoding Rnas at Single-Cell Level by Rna-Fish. *Methods Mol. Biol.* **2021**, *2157*, 251–280. https://doi.org/10.1007/978-1-0716-0664-3_15.
(198) Vautrot, V.; Aigueperse, C.; Branlant, C.; Behm-Ansmant, I. Fluorescence in Situ Hybridization of Small Non-Coding RNAs. *Methods Mol. Biol.* **2015**, *1296*, 73–83. https://doi.org/10.1007/978-1-4939-2547-6_8.
(199) Cicchetti, F.; Lacroix, S.; Cisbani, G.; Vallières, N.; Saint-Pierre, M.; St-Amour, I.; Tolouei, R.; Skepper, J. N.; Hauser, R. A.; Mantovani, D.; Barker, R. A.; Freeman, T. B. Mutant Huntingtin Is Present in Neuronal Grafts in

Huntington Disease Patients. *Ann. Neurol.* **2014**, *76* (1), 31–42. https://doi.org/10.1002/ana.24174.

(200) Jeon, I.; Cicchetti, F.; Cisbani, G.; Lee, S.; Li, E.; Bae, J.; Lee, N.; Li, L.; Im, W.; Kim, M.; Kim, H. S.; Oh, S. H.; Kim, T. A.; Ko, J. J.; Aubé, B.; Oueslati, A.; Kim, Y. J.; Song, J. Human-to-Mouse Prion-like Propagation of Mutant Huntingtin Protein. *Acta Neuropathol.* **2016**, *132* (4), 577–592. https://doi.org/10.1007/s00401-016-1582-9.

(201) Song, N.; Fang, Y.; Zhu, H.; Liu, J.; Jiang, S.; Sun, S.; Xu, R.; Ding, J.; Hu, G.; Lu, M. Kir6.2 Is Essential to Maintain Neurite Features by Modulating PM20D1-Reduced Mitochondrial ATP Generation. *Redox Biol.* **2021**, *47*, 102168. https://doi.org/10.1016/j.redox.2021.102168.

(202) Rossi, F.; Beltran, M.; Damizia, M.; Grelloni, C.; Colantoni, A.; Setti, A.; Di Timoteo, G.; Dattilo, D.; Centrón-Broco, A.; Nicoletti, C.; Fanciulli, M.; Lavia, P.; Bozzoni, I. Circular RNA ZNF609/CKAP5 MRNA Interaction Regulates Microtubule Dynamics and Tumorigenicity. *Mol. Cell* **2022**, *82* (1), 75-89.e9. https://doi.org/10.1016/j.molcel.2021.11.032.

7. List of publications

1. J. Rea, V. Menci, **P. Tollis**, T. Santini, A. Armaos, M. G. Garone, F. Iberite, A. Cipriano, G. G. Tartaglia, A. Rosa, M. Ballarino, P. Laneve, E. Caffarelli. *"HOTAIRM1 regulates neuronal differentiation by modulating NEUROGENIN 2 and the downstream neurogenic cascade."* Cell Death Dis. 2020 Jul 13;11(7):527. doi:10.1038/s41419-020-02738-w.

2. P. Laneve, **P. Tollis**, E. Caffarelli. "*RNA Deregulation in Amyotrophic Lateral Sclerosis: The Noncoding Perspective*" Int J Mol Sci. 2021 Sep 24;22(19):10285. doi:10.3390/ijms221910285.

3. **P. Tollis**, A. Rocchegiani, F. Migliaccio, A. Carissimo, M. G. Garone, A. Rosa, E. D'Ambra, E. Vitiello, I. Bozzoni, P. Laneve, E. Caffarelli. "*HOTAIRM1 shapes differentiation and activity of iPSC-derived functional spinal motor neurons*". **IN PREPARATION**

Abstracts:

1. **Decoding the role of the lncRNA HOTAIRM1 in human motoneurons.**
   Tollis P, Rocchegiani A, Migliaccio F, Carissimo A, Santini T, D'Ambra E, Vitiello E, Laneve P, Caffarelli E.
   Poster presentation at "Regulatory and Non-coding RNAs 2022" (May 17th – 21st 2022), Cold Spring Harbor Laboratory, (New York, USA)

2. **Decoding the role of the lncRNA HOTAIRM1 in human motoneurons.**
   Tollis P, Rocchegiani A, Migliaccio F, Carissimo A, Santini T, D'Ambra E, Vitiello E, Laneve P, Caffarelli E.

Poster presentation at SIBBM 2022 “Frontiers in Molecular Biology, The RNA World 3.0” (June 20th – 22nd 2022), Sapienza University of Rome, (Rome, Italy)

3. **The long noncoding RNA HOTAIRM1 contributes to neuronal differentiation by regulating Neurogenin 2 expression**. Rea J, V Menci, Santini T, Rosa A, **Tollis P**, Ballarino M, Laneve P, Caffarelli E.