International School of Advanced Studies
Area of Neuroscience

Curriculum in Functional and Structural Genomics

SINEUP long non-coding RNAs: from molecular mechanism to therapeutic application.

Thesis submitted for the degree of "Doctor Philosophiae"

CANDIDATE
Bianca Pierattini

SUPERVISOR
Prof. Stefano Gustincich

25th July 2022

Declaration

The work described in this thesis was carried out at SISSA (International School of Advanced Studies) in Trieste and at the Italian Institute of Technology in Genoa, between November 2017 and July 2022.

Part of the work described in this thesis is included in the following papers:

N6-methyladenosine modification regulates SINEUP non-coding RNA activity.
Pierattini B., D'Agostino S., Bon C., Peruzzo O., Alendar A., Espinoza S., Valentini P., Pandolfini L. and Gustincich S. *Manuscript in preparation.*

SINEUP non-coding RNAs rescue defective OPA1 expression and activity in cellular models of Dominant Optic Atrophy.
Pierattini B., Del Dotto V., Bon C., D'Agostino S., Carelli V., Zucchelli S., Espinoza S. and Gustincich S. *Manuscript in preparation*

Internal Ribosome Entry Site RNAs act in trans through antisense base-pairing in linear and circular non-coding RNAs.
D'Agostino S.*, Matey A.*, Volpe M., **Pierattini B.**, Cheung P.L.P., Bon C., Peruzzo O., Armirotti A., Scarpato M., Di Carlo V., Santoro C., Persichetti F., Espinoza S., Zucchelli S., Sanges R. and Gustincich S. *Submitted*

Towards SINEUP-based therapeutics: Design of an in vitro synthesized SINEUP RNA.
Valentini P., **Pierattini B.**, Zacco E., Mangoni D., Espinoza S., Webster N.A., Andrews B., Carninci P., Tartaglia G.G., Pandolfini L., Gustincich S., Mol Ther Nucleic Acids 2022 Feb 2;27:1092-1102. doi: 10.1016/j.omtn.2022.01.021. eCollection 2022 Mar 8

SINEUPs: a novel toolbox for RNA therapeutics.
Espinoza S., Bon C., Valentini P., **Pierattini B**., Tettey Matey A., Damiani D., Pulcrano S., Sanges R., Persichetti F., Takahashi H., Carninci P., Santoro C., Cotella D. and Gustincich S., Assays in Biochem. 2021 Oct 27;65(4):775-789. doi: 10.1042/EBC20200114.

SINEUP non-coding RNAs rescue defective frataxin expression and activity in a cellular model of Friedreich's Ataxia.

Bon C., Luffarelli R., Russo R., Fortuni S., **Pierattini B.**, Santulli C., Fimiani C., Persichetti F., Cotella D., Mallamaci A., Santoro C., Carninci P., Espinoza S., Testi R., Zucchelli S., Condò I. and Gustincich S. Nucleic Acid Research 2019 Nov18;47(20):10728-10743.doi:10.1093/nar/gkz798

Table of contents

Abstract

The post-genomic era has brought to light a previously unknown world of transcripts with the discovery of non-coding RNAs (ncRNAs). Indeed, it became evident that only as few as 1-2% of mammalian transcriptome consists of protein-coding mRNAs. Among several families of ncRNAs, long non-coding RNAs (lncRNAs) are under intense scrutiny for their heterogenicity of forms and molecular activities. A new class of antisense lncRNAs, known as SINEUPs, were previously identified for their ability to specifically enhance the translation of their target sense mRNA. LncRNAs and mRNA were transcribed from a sense/antisense pair locus with an head-to-head divergent configuration. SINEUPs activity relies on the combination of two domains: an overlapping region, or binding domain (BD), that confers specificity, and an embedded inverted SINEB2 element, or effector domain (ED), enhancing target mRNA translation. This new class of transcripts embodies the model of lncRNAs as flexible and versatile modular scaffolds enabling interactions between RNA, DNA and proteins. Furthermore, it represents a promising new RNA therapeutics platform to increase endogenous expression of a protein of interest within a physiological range.

In this work, I provided new insights on the molecular mechanism of SINEUP activity, focusing on the role of N6-methyladenosine (m^6A) modification, and on a Proof-Of-Concept therapeutic application of SINEUPs to rescue haploinsufficient OPA1 gene expression in Dominant Optic Atrophy (DOA).

m^6A is the most common RNA modification found in mRNAs and ncRNAs, where it is post-transcriptionally installed in the cell nucleus and can exert regulatory functions in many cellular processes such as nuclear export and translation. Here, I observed that both the natural SINEUP *AS Uchl1,* acting in rodent cells, and the synthetic shorter miniSINEUP-DJ1, acting in human cells, are m^6A-modified. Results indicate METTL3 enzyme as the main responsible for SINEUP RNA modification. I then applied Nanopore direct RNA sequencing to map m^6A-modified residues and a reverse transcription assay for validation. I monitored SINEUP activity upon METTL3 knock-down and in the presence of mutations on sites of m^6A deposition. Interfering with a proper m^6A modification led to a dominant negative effect of SINEUPs RNA on endogenous DJ1 protein levels in both experimental conditions. Applying ribosome fractionation analysis in conditions of inhibition of proper m^6A deposition, I observed an enrichment of the target DJ1 mRNA associated to 40S and 60S ribosome fractions and a

concomitant depletion from polysomes. These results provide a mechanistic model for its dominant negative effect on endogenous DJ1 protein. These data also suggest the presence of an m^6A-dependent step in the molecular mechanism of SINEUP activity at the ribosome and contribute to a better understanding of the role of RNA modifications in the regulation of lncRNAs function.

From a therapeutic point of view, SINEUPs are proposed as a new platform for the treatment of i. haploinsufficient diseases, where the lack of a functional allele prevents healthy phenotype formation; ii. complex multifactorial diseases, where increasing a compensatory pathway could preserve or restore physiological activities. Here, I applied SINEUP technology to increase endogenous levels of OPA1 protein to treat DOA, the most common inherited optic neuropathy caused in 75% of cases by heterozygous mutations in OPA1 gene. DOA is an early-onset autosomal dominant haploinsufficient disorder, with a prevalence ranging from 1:12000 to 1:50000 births and characterized by degeneration of the retinal ganglion cells that leads to optic nerve atrophy and blindness. OPA1 is a ubiquitously expressed dynamin-related GTPase protein with crucial functions in mitochondrial homeostasis, that localizes in the Inner Mitochondrial Membrane (IMM), reaching highest expression levels in brain, retina and heart. By *in vitro* screening, I identified OPA1-specific miniSINEUPs able to increase selectively both human and murine OPA1 proteins in a range sufficient to restore neuronal cell functions. Currently, a major limitation to the development of SINEUPs as a RNA drug is represented by their length, that should be reduced to less than 60 nts to allow cost-effective manufacturing and efficient *in vivo* delivery. Recently, encouraging data have proved that the incorporation of chemically modified ribonucleotides restores IVT SINEUP RNA activity, making an important progress for its development as a drug. Here, I successfully designed and tested shorter SINEUP RNA variants that allowed us to reduce their size from ~250 nts down to ~50 nts. Indeed, by transfecting 2'OMeA modified ASO-SINEUP-OPA1, I was able to upregulate endogenous OPA1 protein translation of around 1.8 fold, as achieved with standard plasmid-driven expression of the same nanoSINEUP-OPA1 RNA. Most importantly, I applied previously selected mini- and nanoSINEUP to prove the functional rescue of DOA patients' fibroblasts defects in mitochondrial morphology and activity. In summary, I was able to identify OPA1-specific SINEUPs promoting the recovery of disease-associated defects in patient-derived cellular model of DOA and I optimized SINEUP technology for its development as RNA therapeutic molecule for the treatment of haploinsufficient diseases.

Acronyms

AAV, adeno-associated virus;

ACT, Actin beta;

AGO2, argonaute 2;

AS Uchl1, lncRNA antisense to UchL1 mRNA;

AS, antisense;

ASOs, antisense oligonucleotides;

ATP, adenosine triphosphate;

BD, binding domain;

bp, base pair;

cDNA, complementary DNA;

CDS, coding sequence;

circRNA, circular RNA

cox7B, cytochrome c oxidase subunit 7B;

CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats;

DGCR8, DiGeorge syndrome critical region gene 8;

DICER, endoribonuclease Dicer or helicase with RNase motif;

DOA, Dominant Optic Atrophy;

ds, double-stranded

ED, effector domain;

ENCODE, Encyclopedia of DNA Elements;

eRNA, enhancer RNA;

ESC, embryonic stem cells

FANTOM, Functional Annotation of the Mammalian Genome;

FBS, fetal bovine serum;

GAPDH, Glyceraldehyde-3-phosphate dehydrogenase

GFP, green fluorescent protein;

HEK, human embryonic kidney;

HRP, horseradish peroxidase;

IL, internal loop;

IMM, inner mitochondrial membrane;

IMS, inner mitochondrial space;

invSINEB2, inverted SINE of B2 subfamily;

IRES, Internal Ribosome Entry Site;

lincRNA, long intergenic non-coding RNA;

LINE, long interspersed elements;

lncRNA, long non-coding RNA;

lncRNP, lncRNA-protein complex;

LTR, long terminal repeat;

MIRb, mammalian interspersed repetitive (MIR) element b;

miRNA, micro RNA;

MOI, multiplicity of infection;

mRNA, messenger RNA;

MTS, mitochondrial targeting sequence;

ncRNA, non-coding RNA;

NMR, nuclear magnetic resonance spectroscopy;

NMHV, nuclear localization signal – MS2 coat protein interacting domain – HA epitope – (3x) VP16 transactivating domain;

nts, nucleotides;

OMM, outer mitochondrial membrane;

ORF, open reading frame;

OXPHOS, oxidative phosphorylation;

OCR, oxygen consumption rate

PD, Parkinson's disease;

piRNA, piwi-interacting RNA;

PRC2, polycomb repressor complex 2;

pri-miRNA, primary miRNA;

PTM, post-transcriptional modification;

qRT-PCR, quantitative real time PCR;

RBP, RNA-binding protein;

RGC, Retinal Ganglion Cell;

RISC, RNA-induced silencing complex;

RIP, RNA immunoprecipitation

RNA Pol, RNA polimerase;

aRNA, activating RNA;

RNAi, RNA interference;

RNase H, Ribonuclease H;

rRNA, ribosomal RNA;

ROS, reactive oxidative species;

RT, retrotranscriptase;

S, sense;

saRNA, small activating RNA;

shRNA, short harping RNA;

SINE, short interspersed element;

SINEB2, short interspersed element of B2 subfamily;

SINEUP, AS lncRNA with embedded inverted SINE B2 element that UP-regulate target **mRNA** translation;

siRNA, short-interfering RNA;

snoRNA, small nucleolar RNA;

SSOs, spice-switching oligonucleotides;

tRNA, transfer RNA;

TE, transposable elements;

TM, transmembrane;

TSS, Transcriptional Start Site;

Uchl1, Ubiquitin carboxyl-terminal hydrolase L1;

UTR, untranslated region;

WB, Western Blot

Introduction

The first draft of the human genome has defined the beginning of the so-called "post-genomic era". Key discoveries of this revolutionary time were achieved thanks to large-scale genomic projects such as FANTOM[1] and ENCODE[2], that developed new technologies to unveil the complexity of regulatory elements in genomes and the transcriptomes of mice and humans at unprecedented depth. These consortia surprisingly found that the majority of the mammalian genome is pervasively transcribed (70-80%), with only a very small fraction of the transcripts having protein-coding potential (1-2%). The remaining major portion of transcripts is a diversified repertoire of non-coding RNAs (ncRNAs), including small ncRNAs, long non-coding RNAs (lncRNAs) and Transposable Elements (TEs). Such a deep analysis of various organisms' transcriptomes highlighted that the number of protein-coding transcripts is reasonably static, while the relative amount of non-coding transcripts positively correlates with organisms' complexity (Table 1)[3]. While in most bacterial species the 90% of the genome is represented by protein-coding DNA, the human genome contains only about five times more protein coding genes than *Escherichia coli*, three times more than eucaryotic yeast and even less than mice and *Caenorhabditis elegans*, which is very surprising considering the difference in complexity between these species[4]. This crucial finding shifted scientists' focus from DNA to RNA, suggesting that the latter plays a much more relevant role in regulatory mechanisms than previously believed.

Organism	*Kbp*	*Coding genes*	*Ratio (bp/gene)*
Escherichia coli	4.6×10^3	4.3×10^3	1.1×10^3
Saccaromyces cerevisiae	1.2×10^4	5.8×10^3	2.2×10^3
Caenorhabditis elegans	1.0×10^5	2.2×10^4	4.6×10^3
Drosophila melanogaster	1.2×10^5	1.5×10^4	8.2×10^3
Mus musculus	2.8×10^9	2.3×10^4	1.2×10^5
Homo sapiens	3.3×10^9	2.0×10^4	1.7×10^5

Table 1 Comparative genome complexity[4].

1. LncRNAs

Among ncRNAs, several families of small ncRNAs have been identified, such as transfer RNAs (tRNAs), small interfering RNAs (siRNAs), microRNAs (miRNAs), Piwi-interacting RNAs (piRNAs) and small nucleolar RNAs (snoRNAs), each one exerting specific functions. However, the largest and most heterogeneous portion of

transcriptomes is represented by long non-coding RNAs (lncRNAs), defined as transcripts longer than 200 nucleotides, often polyadenylated and without evident open reading frames (ORFs)[5]. More than 100000 lncRNAs are annotated at date, as reported by LNCipedia v5.2[6]. The majority of them shares features with canonical mRNAs as they are mostly transcribed by RNA polymerase II, they undergo splicing, 5'-capping and polyadenylation[7]. On the other hand, lncRNAs are expressed at lower levels, show an higher tissue-specificity and their localization is often restricted to the cell nucleus compartment, compared to protein-coding genes[8]. lncRNAs are often found to be part of ribonucleoprotein complexes that can regulate gene expression acting at different steps of the process[5]. Indeed, they are able to function as ligands for proteins and also to bind specific DNA or RNA molecules through a base-pairing mechanism[9], features shared with miRNAs, snRNAs and other small ncRNAs. However, lncRNAs show some peculiar, additional features: they are able to fold in secondary or higher order structures, which confers them a flexible and versatile potential as modular scaffolds[10] able to induce protein interactions[11]. It is well known that lncRNAs contribute to regulate gene expression with different molecular mechanisms depending on their physical features, subcellular localization and interaction with other molecules both in physiological conditions and in diseases[12].

Being the most heterogenous class of ncRNAs, lncRNAs can be classified according to the following criteria: anatomical properties, subcellular localization and mechanism of action.

1.1 Anatomical properties

Based on the genomic position relative to their nearby protein-coding genes, lncRNAs can be classified as intergenic (long intergenic non-coding RNAs, or lincRNAs) if they do not overlap with any other gene. When they do overlap with exons and intron of genes, they can be both sense (S) and antisense (AS) lncRNAs, as transcribed from the opposite DNA strand that overlaps protein-coding gene.

1.2 Subcellular localization

A distinctive feature of lncRNAs is their preferential subcellular compartmentalization in the cell nucleus, compared to canonical mRNAs. lncRNAs are also less conserved among species and display a general lower expression level when compared to mRNAs. Moreover, while the two transcript classes show a comparable length, lncRNAs are reported to contain fewer and longer exons, lower GC content and very short introns[4,8,13].

Nuclear

Nuclear lncRNAs can regulate transcription[14], contribute to the organization and structure of subnuclear components[15–17] and to chromatin state regulation[18–20]. The nuclear localization and fate of lncRNAs are finely regulated from transcription to nuclear export by the selective usage of sequence motifs acting in *cis* and factors acting in *trans*. They are mostly transcribed from phosphorylation-dysregulated RNA Pol II which leads them to be weakly co-transcriptionally spliced and to be terminated in a polyadenylation signal-independent manner. Altogether, these peculiar features cause an accumulation of lncRNAs on chromatin with a subsequent rapid degradation[12]. Several studies have focused on the characterization of lncRNAs localization pattern with high-throughput approaches, combining computational methods with high-resolution single-molecule imaging techniques[2,21,22], but the molecular mechanism that tethers lncRNAs in the cell nucleus is still scarcely known.

Cytoplasmic and organelles-localized

Another portion of lncRNAs reside instead in the cytoplasm, where they take action in post-transcriptional gene expression processing through different mechanisms: acting as miRNA sponges[23], regulating target mRNA metabolism, sequestering specific proteins, altering protein post-translational modifications[24] or modulating translation[25–27]. Interestingly, some lncRNAs have been reported as undergoing shuttling across subcellular compartments in response to specific stimuli[25,28]. Indeed, recent advances in imaging techniques, such as imaging of thousands of barcoded RNAs or APEX-RIP, have allowed a specific and unbiased quantification of RNAs localized in cellular compartments such as the nucleus and the cytosol, but also mitochondria[29], exosomes and *endoplasmic reticulum* (ER)[30]. It is estimated that the majority of cytoplasmic lncRNAs co-localize with polysomes[12,31] thanks to *cis* elements, such as long "pseudo" 5' untranslated (5'UTR) regions, deriving their name from their localization upstream "pseudo-open ORFs". The fate and functions of ribosome-associated lncRNAs is still scarcely known but certainly worth of further investigations.

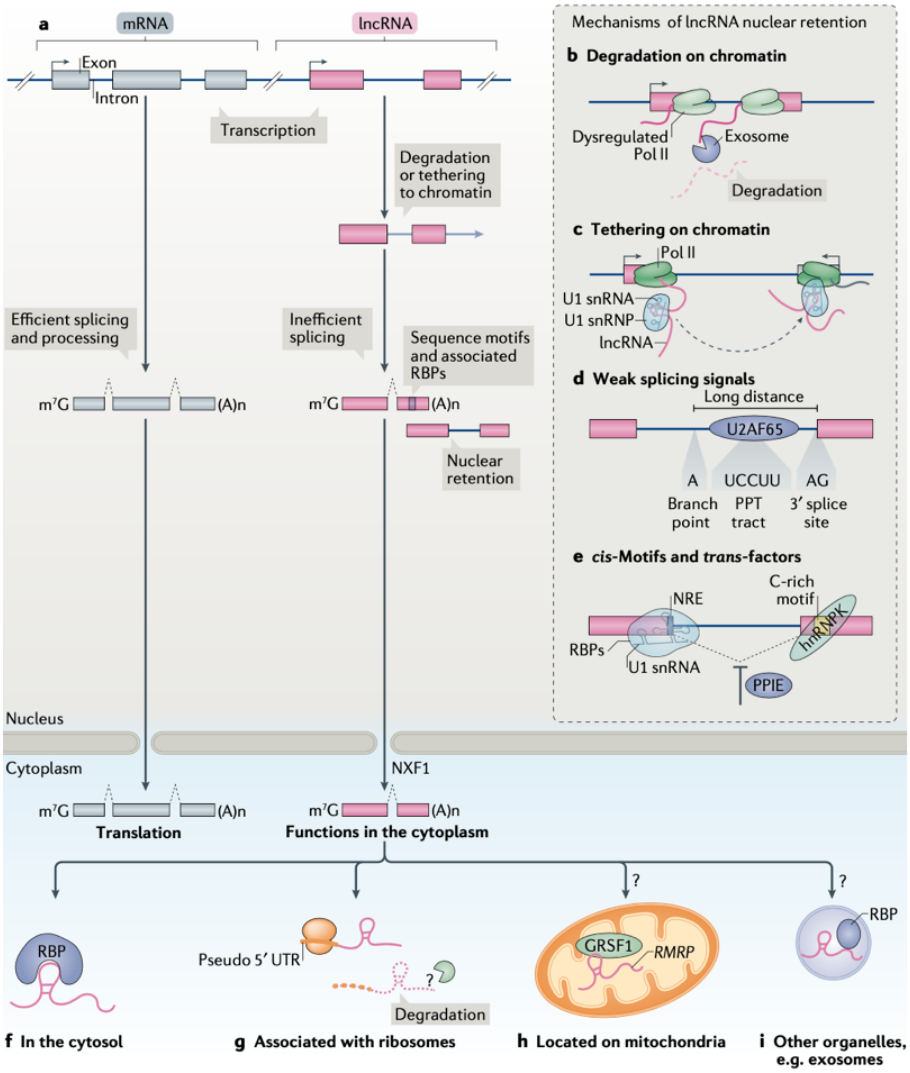


Figure 1 Biogenesis and fate of lncRNAs[12]. **a** Biogenesis of lncRNAs: differently from canonical mRNAs, most lncRNAs are transcribed by RNA pol II but inefficiently processed and therefore retained in the cell nucleus, while others are spliced and exported to the cytoplasm, mainly by nuclear RNA export factor 1 (NFX1). **b** dysregulated-RNA pol II lncRNAs are degraded by nuclear exosome. **c** lncRNAs containing U1 snRNA binding motif can recruit U1 snRNA and associate with Pol II. **d** The sequence between the 3' splice site and the branch point is longer in lncRNAs than in mRNAs and contains a shorter polypyrimidine tract. This results in inefficient splicing. **e** Sequence motifs and co-factors contribute to subcellular lncRNAs localization in cis and in trans respectively. NRE U1 snRNA-binding site C-rich motifs can recruit U1 snRNP19 and hnRNPK to enhance nuclear localization of lncRNAs. Other RBPs, such as PPIE 6, inhibit splicing of lncRNAs. **f** In the cytoplasm, lncRNAs may interact with different RBPs. **g** In the cytoplasm, many lncRNAs are associated with ribosomes via a pseudo 5'UTR. These transcripts usually have a short half-life. **h** Several lncRNAs are sorted to subcellular organelles, such as mitochondria, by unknown mechanisms. **i** Some other lncRNAs are found in other organelles, such as exosomes, probably forming complexes with RBPs.

1.3 Mechanism of action

lncRNAs can modulate gene expression with many different mechanisms that result in chromatin structure and function modulation, in the regulation of other genes' transcription, RNA splicing, stability and translation (Figure 2 and 3). lncRNAs are also

involved in the metabolism of cellular organelles and nuclear condensates. Several interdependent factors are considered as key regulators of lncRNA function: the relative location of the lncRNA and target gene, the formation of RNA-DNA and RNA-protein interactions and whether the lncRNA effect is exerted by the transcript or by its transcription[12].

Chromatin regulation

LncRNAs are able to mediate chromatin de-compaction exploiting their negative charge and producing an efficient switch of gene expression. lncRNAs mechanisms of action, both in *cis* and *trans* conformation, may involve direct or indirect DNA interactions, in the former case based on sequence-complementarity and in the latter involving protein interactions. Several lncRNAs were reported to modulate the recruitment of Polycomb repressive complex 2 (PRC2) with both in *cis* and in *trans* interactions. An example is represented by *ANRIL* lncRNA, which is able to recruit PRC1 and PRC2 to the promoter of *CDKN2A* and *CDKN2B* genes, closely positioned, modulating their expression and eventually regulating cell senescence[19]. In particular, *ANRIL* is also able to work in *trans* thanks to embedded *Alu* TE, that enable this lncRNA to recruit PRC1 and PRC2 to distant targets[18]. Another well-known example is represented by *HOTAIR* lncRNA, which acts in *trans* as negative regulator of HOXD genes, through the recruitment of PRC2 and a protein complex that de-methylases H3K4[32]. Other lncRNAs are able to recruit chromatin modifiers with positive gene regulatory functions, such as *HOTTIP* which is involved in the HOXA gene cluster regulation maintaining the chromatin organization in this *locus*[33]. Moreover, some lncRNAs have also been reported to act as decoys, able to repress multiple genes at the same time, like *lncPRESS1* that exerts its function on sirtuin 6, eventually leading to the repression of several pluripotency genes[12]. Nevertheless, an important feature of lncRNAs is their capability to directly bind DNA, thus altering chromatin state with the formation of hybrid structures also known as R-loops, that can be generated both in *cis* and in *trans* by lncRNAs[34], and RNA-DNA-DNA triplexes, that have been reported to mediate both gene silencing and activation[35].

Transcription regulation

The transcriptional regulation by lncRNAs can occur through two main mechanisms: the lncRNA itself can regulate transcription of nearby *loci,* or the transcription or splicing of the lncRNA itself can produce a chromatin state or a steric distortion that alters the expression of proximal genes. The most representative example of this kind of mechanism is *XIST* lncRNA, which inactivates in *cis* the X chromosome from which it is

transcribed in female mammalian organisms. This lncRNA is required for the initiation of the X chromosome inactivation and not for its maintenance, thus requiring a fine regulation of its time of expression. *XIST* tethers PRC2 to the inactive chromosome through the formation of repressive heterochromatin, which relies on the cooperation between the lncRNA and several proteins, including transcriptional repressor YY1. *XIST* expression is reported to be controlled both in positive and negative ways by other lncRNAs, such as *Tsix* and *Jpx*[36,37] respectively. Other lncRNAs, instead, can act in *cis* directly or indirectly interacting with chromatin near a transcription start site (TSS), promoting its inactive state; an example of this case is represented by *COOLAIR* lncRNA, from *A. Thaliana*, whose expression is induced by low temperatures and causes histone demethylation in the gene body of *FLOWERING LOCUS C (FLC)*, thus reducing its transcription[5]. In other cases, lncRNAs can interfere with the transcription machinery to alter transcription factors or RNA Pol II recruitment at the promoter, histone modification or chromatin accessibility[12]. A class of enhancer-associated lncRNAs (elncRNAs) has recently been reported. They are produced by RNA pol II binding to specific enhancers and their expression levels induce an increase of messenger RNAs (mRNAs) expression from neighboring genes. While eRNAs are relatively short, bidirectional, capped, non-polyadenilated and unspliced, elncRNAs are in most cases unidirectional, polyadenylated and spliced transcripts. elncRNAs can act on pre-existing chromatin conformation or promoting chromatin looping through the interaction with scaffold proteins. These interactions enable regulatory contacts between enhancers and promoters even when they are located very distant from each other[12].

Scaffolding and condensate formation

Several lncRNAs were reported to be essential for the assembly and activity of nuclear condensates, which are RNA-protein compartments without membrane. Two of the most studied lncRNAs can fit into this category: *nuclear paraspeckle assembly transcript 1 (NEAT1)* and *metastasis-associated lung adenocarcinoma transcript 1 (MALAT1). NEAT1* is a key player in the organization and function of paraspeckles[16,38], acting as two different isoforms originating from alternative splicing: *NEAT1 long* and *NEAT1 short*. The longest transcript is essential for paraspeckles assembly and was demonstrated to recruit core proteins to initiate the assembly through liquid-liquid phase separation[15]. *MALAT1*, one of the most highly expressed lncRNAs in cell cultures, is specifically located in nuclear speckles and exerts important functions in pre-mRNA splicing and transcription, cancer progression and metastasis formation[17,39]. Its artificial downregulation in human lung

cancer cells showed that MALAT1 controls metastatic gene expression program, but a detailed picture of its mechanism of action has not been defined yet[39].

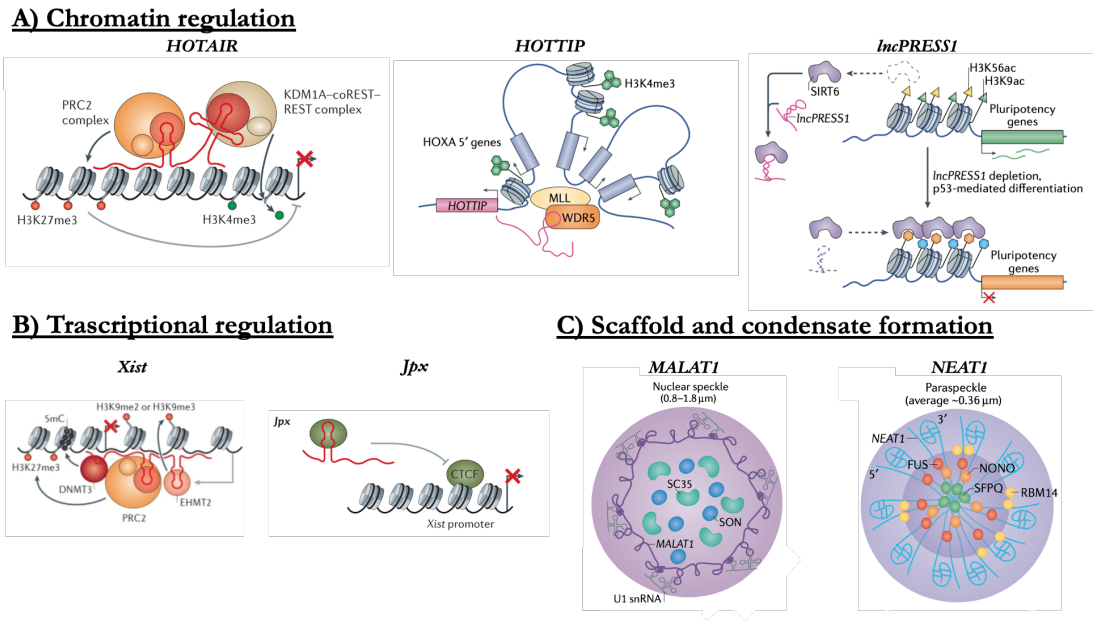


Figure 2 lncRNAs activity in the nucleus[5,12]. **A) Chromatin regulation:** lncRNAs can interact with chromatin modifiers recruiting them to target-gene promoters, activating or suppressing their transcription. HOXA transcript antisense RNA (*HOTAIR*) acts in trans as HOXD genes regulator; HOXA transcript at the distal tip (*HOTTIP*) acts at the 5' genes of HOXA gene cluster in cis through chromatin looping promoting histone H3 Lys4 trimethylation; p53-regulated and embryonic stem cell-specific lncRNA *lncPRESS1* supports human embryonic stem cells pluripotency by sequestering chromatin modifiers from target genes' promoters. **B) Transcriptional regulation:** lncRNAs can be involved in dosage compensation and genomic imprinting. Examples are *Xist, Kcnq1* and Airn lncRNAs that induce the formation of repressive chromatin by DNA methyltransferase 3 (DNMT3), PRC2 and N-methyltransferase (EHMT2). The lncRNA Jpx binds the transcriptional repressor CTCF inhibiting its binding to Xist promoter, eventually activating Xist transcription. **C) Scaffold and condensate formation:** *MALAT1* is localized at the periphery of nuclear speckles and is involved in pre-mRNA splicing regulation. At the periphery, it interacts with U1 snRNA, while proteins and splicing components are located in the center of the structure. NEAT1 lncRNA is essential for the formation of paraspeckles. It sequesters numerous proteins to form a core-shell spheroidal nuclear body. The middle region of the transcript is located in the center of paraspeckles and the distal regions are in the periphery.

Post-transcriptional regulation

LncRNAs can also alter gene expression at post-transcriptional, translational, and post-translational levels with different mechanisms. Indeed, they may be able to regulate mRNA splicing and half-life binding to RNA-binding proteins (RBPs) through specific consensus sequences or structures and forming lncRNA-protein complexes (lncRNPs). In other cases, lncRNAs may be able to modulate post-translational modifications of splicing factors, or to repress splicing through the formation of RNA-RNA hybrids, or to alter a target gene's splicing through chromatin remodelling[12]. Some other lncRNAs display miRNA-complementary sequences that allow them to sequester specific miRNAs, competing with the endogenous targets for their binding and thus reducing miRNAs

availability. For this reason, they are also known as competing endogenous RNAs (ceRNAs) and their expression is very finely regulated to counter-act, in a physiological state, miRNAs activity. Recently, a growing number of examples of this kind of mechanism was observed in circular RNAs (circRNAs). CircRNAs, in general, exhibit a longer half-life compared to linear RNAs; as a consequence, circRNAs' turnover can be controlled by a perfectly matched miRNA target site[5,23]. Post-transcriptional regulation can also be exerted by lncRNAs through a direct base-pairing with other RNAs. An example is represented by Staufen-mediated mRNA decay, that is triggered by lncRNAs with embedded Alu elements or retroelements in human and short interspersed nuclear elements (SINE) in mouse. These lncRNAs are able to recruit Staufen homolog 1 (STAU1) protein which recognizes double-stranded RNA (dsRNA) and promotes RNA decay[12]. On the other hand, both *β-site APP-cleaving enzyme 1-antisense (BACE1-AS)* and *tissue differentiation-inducing non-protein-coding RNA (TINCR)* increase the stability of their target mRNAs[5]. Another example of post-transcriptional regulation, that relies on base pairing in *trans*, is represented by a nuclear transcript antisense to ubiquitin carboxy-terminal hydrolase L1 (*AS Uchl1*), that contains an embedded inverted SINEB2 element and will be discussed more in detail in the following paragraphs.

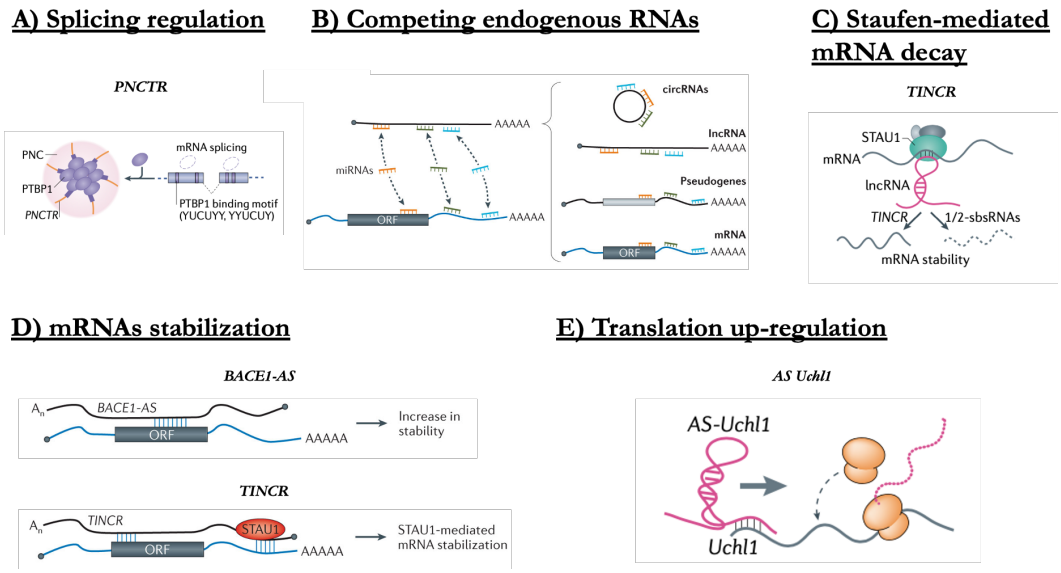


Figure 3 Post-transcriptional lncRNAs activity[5,12]. **A) Splicing regulation:** lncRNAs can interact in trans with RBPs forming structural motifs or through sequence motifs. Pyrimidine-rich non-coding transcript (*PNCTR*) sequesters PTBP1 protein to perinuclear compartment, suppressing PTBP1-mediated mRNA splicing. **B) Competing endogenous RNAs:** ncRNAs, including lncRNAs and circRNAs, can compete with endogenous mRNAs for miRNA binding, resulting in a crosstalk between these different RNA classes. **C) Staufen-mediated mRNA decay:** this process is induced by intermolecular base-pairing between Alu element (or SINE in mice) in the 3'UTR and an Alu element within a half-STAU-binding site RNA. **D) mRNA stabilization:** two examples are reported. In the first, base pairing between specific regions of the human β-site APP-cleaving enzyme 1 (BACE1) mRNA and its antisense transcript BACE1-AS induces stabilization of target mRNA and increases BACE1 protein expression. In the other example, STAU1-

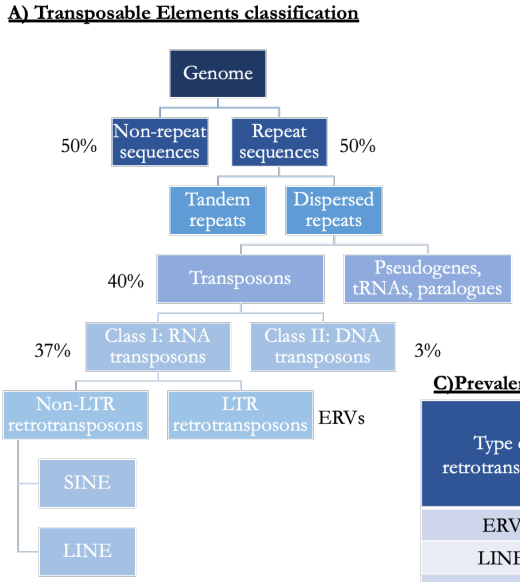
mRNA stabilization was described in tissue-differentiation process that induces ncRNA *TINCR*, recognizing its target mRNA through base-pairing. **E) Translation up-regulation:** SINEB2 element of mouse AS Uchl1 complementarily binds its sense target mRNA promoting its association to polysome and thus inducing its translation.

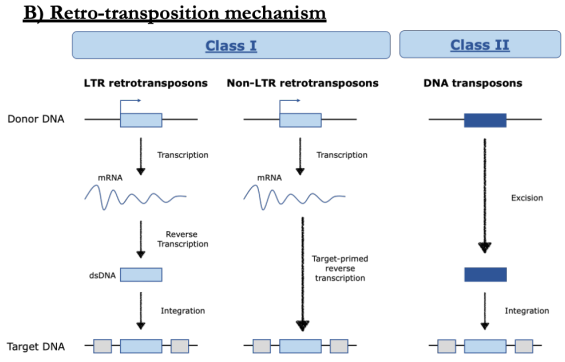
1.3 LncRNAs and TEs relationship

LincRNAs have been found as part of regulatory complexes in about 30% of cases in embryonic stem cells (ESCs), defining a potential general pattern[11]. A model was thus proposed where lncRNAs might be the crucial elements to build a "modular RNA code", pacing cell growth states and essential biological processes[9]. A peculiar feature of lncRNAs is the relevance that their secondary and tertiary structure seems to exert. Indeed, several studies reported that common structural features seem to be much more conserved in multiple lncRNAs than their primary sequences, thus suggesting a deep connection between structure and function[40]. An important contribution to the biogenesis and regulation of lncRNAs is given by TEs. This information is coherently mirrored by the discovery that TEs can be found embedded in 75% of mature lncRNAs in vertebrates, while they are scarcely present in protein-coding transcripts[5,41,42].

TEs classification

TEs are defined as genomic sequences able to move from their original location in the genome to another position (Figure 4). Two major classes of TEs have been identified based on the mechanism adopted to mobilize: a first one consisting of retrotransposons (class I) and a second one consisting of DNA transposons (class II)[43,44]. The latter group encodes a transposase enzyme that catalyzes the transposon's own cut and insertion into another genomic site. Instead, retrotransposons activity depends on a reverse transcriptase (RT) enzyme that first transcribes them into an RNA intermediate, which is then integrated in a new genomic *locus* and can only be found in eukaryotic genomes. Retrotransposons represent around 90% of all TEs[41], representing around 37% of the human genome, and also the majority of lncRNA-embedded TEs[45]. This class can be further divided into three sub-classes: long terminal repeat (LTR) elements, long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). LTR retrotransposons are autonomous and very similar to retroviruses in both structure and mechanism of amplification; they exhibit a ORF encoding for an RT enzyme flanked by LTRs. On the contrary, LINEs and SINEs do not present terminal repeats and are therefore called non-LTR retrotransposons[44].





C)Prevalence of Transposable Elements

Type of retrotransposon	Occupation in the genome (human)	Copy number (human)	Occupation in the genome (mouse)	Copy number (mouse)
ERVs	8%	500k	10%	500k
LINEs	18%	870k	9%	660k
Alu/SINEs	11%	1500k	8%	1500k

Figure 4 Transposable Elements. **A) Transposable Elements classification:** composition of the different classes of transposable elements in the human genome. Percentages in the tree correspond to the genomic proportion of each type of element[45]. **B) Retro-transposition mechanism:** class I transposons are also referred to as "retrotransposons" as their activity is dependent on reverse transcription, or "copy and paste" transposons, as they are "copied" in RNA before being "pasted" in the target DNA. LTR transposons are first transcribed in mRNA, then retrotranscribed in dsDNA and finally integrated in the target DNA. Instead, non LTR transposons, such as SINE and LINE elements, are first transcribed in mRNA and then targed-primed reverse transcribed in target DNA. Class II DNA transposons, on the other hand, do not need to be transcribed, as they are "cut and paste" into target DNA. **C) Prevalence of Transposable Elements** in human and mouse genome[45].

LINE elements

LINE-1 (L1) are the most common class among retrotransposons, representing around 17-20% of the genome and represent the only active autonomous retrotransposon group. Transcription of L1 elements could be a risk for genome integrity, so several epigenetic strategies have evolved to suppress it, including DNA methylation, histone methylation and small RNA expression. More than 520000 copies of L1s are annotated in the human genome, among which 100-140 are potentially capable to autonomously retrotranspose. About 20 of them are particularly active, producing a ~6 kb long transcript and capable of undertaking a full life cycle finally leading to translation of ORF1 and ORF2, trafficking back to the nucleus and insertion in a new genomic position. Most annotated L1s are, instead, truncated or mutated and, therefore, they can be defined retrotranspositionally incompetent[46].

SINE elements

SINEs are short elements, up to 1kb long, evolved from RNA genes, such as tRNAs, 7SL RNAs, 5S RNAs[43]. They are non-autonomous, since they do not encode for a reverse-

transcriptase enzyme, so in most cases they take advantage of L1s retrotransposition machinery to mobilize. The most representative SINE element is the Alu repeat, that reaches over a million copies within the human genome. SINEs can be found in all mammals, reptiles, fishes and in some invertebrates and flowering plants, while they are missing in *Drosophila* and in most unicellular eukaryotes[47], suggesting its evolutionary recent origin[48]. In rodents, SINE elements account for 7.6% of the genome and can be divided into two families: B1 and B2[49]. They can act as functional elements involved in the regulation of various biological processes. An example of conserved function in human and mouse takes place during stress response. Even if the majority of SINE elements are usually silent, stress signals, such as heat shock, can cause a massive induction of their transcription, which commonly results in the inhibited expression of multiple genes[50,51].

Role of TEs in lncRNAs

Recently, many studies have focused on the high prevalence of TEs in lncRNAs compared to mRNAs: in humans. 83% of lncRNAs present in a list of 28 tissues and cell lines contains at least one embedded TE and the same prevalence was observed in mouse, even if at a lower extent[45]. Furthermore, it is estimated that 42% of human lncRNAs originates from a TE. In comparison, only 5.5% of human protein-coding transcript derive from TEs and only 39% of mRNAs contain a TE[41]. It has been found that different TEs contribute at different extent to lncRNAs sequences: 13% from LINEs, 7.7% from SINEs, 3.5% from LTRs and 2.2% from DNA TEs[52]. Furthermore, around 19% of TE-containing lncRNAs, derive more than 50% of their sequence from a TE. Interestingly, TEs may be found in the majority of cases at the last exon of lncRNAs (56%)[53]. Some TEs, such as Alu sequences, contribute to isoform variety and exon content of lncRNAs since they present splicing signals that, upon insertion in a gene, can create new splicing sites and/or exons[45]. Importantly, TEs were also found in correspondence to or near lncRNAs Transcription Start Sites (TSSs), thus suggesting their possible involvement in the regulation of lncRNA transcription. From these evidences it was then proposed that TEs may contribute to lncRNAs evolution by conferring them tissue-specific expression through transcriptional regulatory signals[5,41,42,45]. In a number of cases, Alu sequences were found in proximity of lncRNAs 3' end in sense orientation, which, instead, suggests their role in determining a polyadenylation signal[45], as reported in Kapusta et al., where 30% of the polyadenylation signals in their lncRNA dataset overlapped with TEs[42]. TEs may also function as DNA regulatory elements of lncRNA expression and to participate in the

post-transcriptional modification of the lncRNA they are embedded in. The number of roles assigned to lncRNA-embedded TEs is constantly increasing, but it is still not exactly reflecting the very frequent presence of TEs in this class of transcripts since a detailed dissection of TEs contribution in lncRNAs functions is rarely pursued[45].

2. SINEUPs: a new functional class of antisense lncRNAs.

Bidirectional transcription leads to the co-existence of RNA molecules transcribed from opposite DNA strands. Indeed, natural overlapping sense/antisense (S/AS) pairs of transcripts have been annotated[8]. S/AS pairs may present all possible combinations of protein-coding and lncRNA genes. Indeed, they can be found 5' head-to-head divergent, 3' tail-to-tail convergent and fully overlapping configurations (Figure 5).

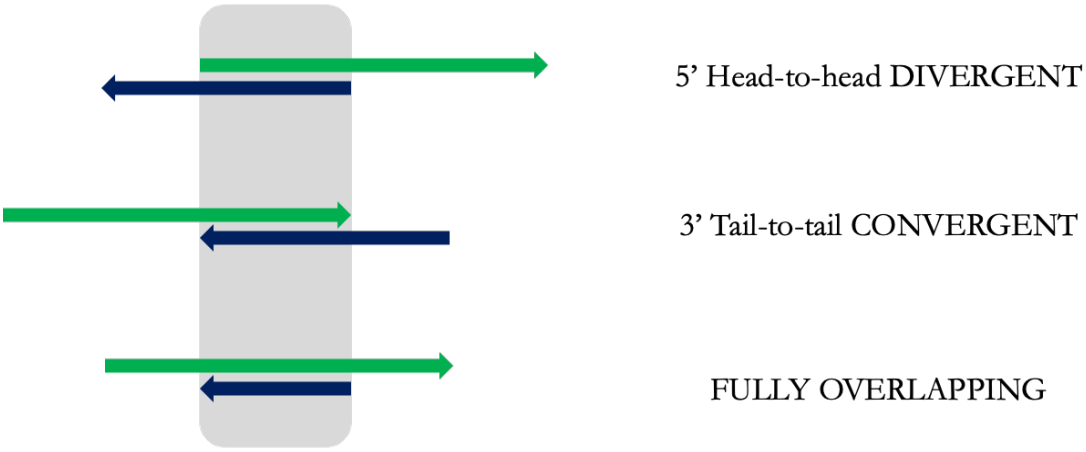


Figure 5 S/AS pairs genomic organization. Sense genes are reported in green, AS ones in blue. Arrows indicate 5'→3' direction; grey box indicates overlapping regions[54].

It has been reported that lncRNAs in AS orientation to nearby protein-coding genes constitute the 60-70% of the whole transcribed genome[55]. These natural AS transcripts (NATs) can be encoded in *cis* and transcribed from a promoter on the opposite strand of the corresponding protein-coding gene and commonly present a partial overlapping sequence with the sense transcript[55–57]. As many as 10077 and 8091 S/AS pairs were recently annotated in human and mouse genomes respectively. A significant amount was shown to derive from retrotransposition events in one of the genes or from alternative polyadenylation signals. It was also observed that this genomic configuration, with an antisense transcript partially complementary to the sense one, increases the probability of the respective intron retention[57]. NATs can modulate the corresponding sense gene expression acting in *cis* or in *trans* through several mechanisms[58].

2.1 Natural SINEUPs: AS Uchl1 and others

Focusing their studies on mouse genomic loci associated with Parkinsons's disease (PD), Carrieri and colleagues reported the presence of the aforementioned spliced lncRNA transcript in the Uchl1/PARK5 gene, named AS Uchl1, mapping in antisense orientation to the protein-coding counterpart *Uchl1* mRNA (Figure 6)[25,54]. The two transcripts were described in 5' head-to-head configuration with *AS Uchl1* overlapping to the first 73 nts of sense *Uchl1*, including the AUG starting codon (-40/+33, with +1 position corresponding to the A of the starting AUG). The remaining part of *AS Uchl1* lncRNA contains two embedded repetitive sequences: a SINEB1, with features corresponding to F1 subclass (Alu), and a SINEB2 belonging to B3 subclass[25].

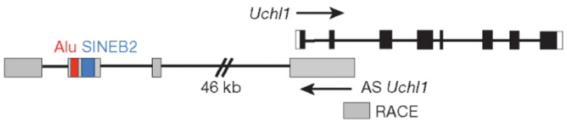


Figure 6 Natural AS Uchl1 *and* Uchl1 *genomic organization*[25].

In mouse, *AS Uch*l1 lncRNA is co-expressed with *Uchl1* mRNA in 40% of cases and is not expressed in absence of the sense transcript[25]. The two transcripts were also reported to be differentially localized in the cell: mature *Uchl1* mRNA preferentially localizes to the cytoplasm, while, in physiological conditions, *AS Uchl1* is retained in the nucleus, as the majority of reported lncRNAs[25]. Upon *AS Uchl1* overexpression, a significant increase of UCHL1 endogenous protein was detected, with no alteration in *Uchl1* mRNA levels, suggesting a potential post-transcriptional regulatory function of the lncRNA[25]. Deletion analysis of *AS Uchl1* allowed the identification of two functional domains within the lncRNA that are essential for its activity: the 5' overlapping region and the inverted SINEB2 element (invSINEB2/invB2)[25]. It was also observed that, upon CAP-dependent translation inhibition, induced with rapamycin administration, *AS Uchl1* shuttles from the nucleus to the cytoplasm, where it exerts its activity, inducing translation of the sense protein-coding *Uchl1* mRNA by increasing its association to heavy polysomes[25]. Importantly, AS Uchl1 lncRNA is able to post-transcriptionally induce UCHL1 protein translation only when containing the overlapping and the invSINEB2 sequences. In the effort to evaluate the possible presence of other AS lncRNAs with the same function and genomic configuration relative to their sense protein-coding gene, FANTOM3 dataset was interrogated, leading to the identification of 31 S/AS transcript pairs with 5'head-to-head overlapping and containing an invSINEB2 element[59]. Among the list, *AS Uxt* was

proved to retain a similar function to *AS Uchl1*: upon overexpression, UXT protein level was increased with no alteration of the corresponding mRNA expression[25]. Later on, the activity of another AS lncRNA included in the same list was successfully validated targeting Elastin[60]. Given that SINEB2 elements are not present in the human genome, it was important that an AS transcript to protein phosphatase 1 regulatory subunit 12A (PPP1R12A), named R12A-AS1, containing a short Free Right Alu Monomer repeat element (FRAM) was identified in human[59]. In this case, the invSINEB2 ability to up-regulate translation is conferred to the lncRNA by the FRAM element. Indeed, despite their lack of primary sequence homology, SINEB2 and FRAM elements both bind the dsRNA-binding protein ILF3, showing evidence of convergent evolution in different TEs[61]. Another lncRNA, AS to Integrin-Alpha FG-GAP Repeat-Containing Protein 2 (ITFG2) also displayed SINEUP activity, mediated by an inverted MIRb TE[59]. 129 potential human natural SINEUPs were computationally identified as part of S/AS pairs, retaining a region overlapping to a protein-coding gene in a head-to-head configuration and combined with an embedded TE[59]. More recently, AS lncRNAs with similar genomic organization and function to *AS Uchl1* were identified among dysregulated transcripts in Autism Spectrum Disorder (ASD). To this end, human neural progenitors with a 50% knock-down of *CHD8* mRNA expression[62] obtained with a short hairpin RNA (shRNA), recapitulated the haploinsufficient genotype in selected ASD individuals. Among the genes differentially expressed, *RAB11B-AS1* lncRNA was experimentally demonstrated to increase protein levels encoded from the corresponding sense transcript *RAB11B* through a specific overlapping antisense sequence and an invSINEB2 element[62].

Taken together, these data prove *AS Uchl1* is the representative member of a new functional class of natural AS lncRNAs that up-regulate translation of sense overlapping transcripts. In all cases reported so far, the biological activity of such lncRNAs depends on the combination of two distinctive elements: the overlapping region (Binding Domain, BD), that confers target specificity through complementary base-pairing with target mRNA, and the embedded inverted SINEB2 element (Effector Domain, ED) that is required for translation induction. This new class of lncRNAs has been named SINEUPs, since they take advantage of a SINEB2 element to UP-regulate a target mRNA translation in a selective, post-transcriptional manner[63,64]. SINEUP lncRNAs modular architecture provide sequences as binding sites for regulatory complexes and for RNA:RNA pairing. At the same time, this modular structure makes SINEUPs easy to be manipulated and synthetically designed to re-direct their activity towards a specific target mRNA of

interest, as it will be further discussed in the next section. Nevertheless, natural SINEUPs' exact mechanism of action remains to be elucidated.

2.2 Synthetic SINEUPs and miniSINEUP development

Replacing the BD sequence with one complementary to a given target mRNA, synthetic SINEUP lncRNAs can be designed, re-directing *AS Uchl1* activity towards exogenously expressed transcripts, such as Green Fluorescent Protein (GFP) (Figure 7)[25,64], or endogenous ones[54,60,65–68]. Synthetic SINEUPs have been successfully applied to enhance protein translation to a wide range of target mRNA including FLAG-tagged proteins, secreted recombinant antibodies, and cytokines. Remarkably, synthetic SINEUPs were proven to be active on endogenous mRNAs both *in vitro* and *in vivo*, with the first example represented by PARK7/DJ-1-targeting SINEUPs[63]. Other SINEUPs have been shown to be active in cell lines of mouse, hamster, monkey and human origin, proving their wide applicability in *in vitro* experimental settings along with their scalability. Unpublished results also demonstrate that synthetic SINEUPs are active in *Drosophila* cells, confirming their relying on an evolutionary conserved cell process (Matey A. et al., unpublished)[69]. Importantly, from a therapeutic point of view SINEUPs represent an ideal tool to perturb gene expression *in vivo* since they are able to induce a fold-induction from 1.5 to 3 fold, within a physiological range, avoiding the potential side effects of uncontrolled large overexpression.

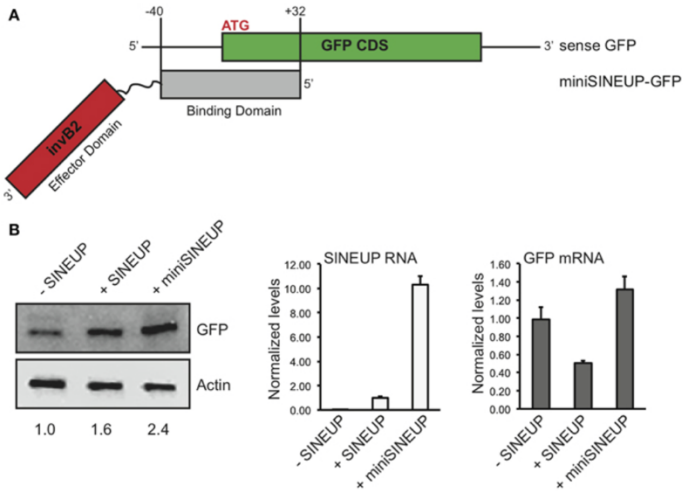


Figure 7 miniSINEUP-GFP design and activity[64]. **A)** Domain organization of miniSINEUP-GFP. BD and ED are reported in gray and green, respectively. **B)** HEK 293T/17 cells were co-transfected with sense GFP vector together with empty vector (-SINEUP), SINEUP-GFP as positive control (+SINEUP) and miniSINEUP-GFP (+miniSINEUP-GFP).

To better understand the structure/function relationship and the molecular mechanism of SINEUP activity, BDs and EDs have been extensively studied, leading to the optimization of artificial SINEUPs and the identification of minimal structural features required for activity[70]. Indeed, while the first generation of *AS Uchl1*-derived synthetic SINEUPs were about 1200 nts long, composed of a 73 nts long BD, a 172 nts long ED, a partial Alu element and a 3' tail[25], a shorter version of SINEUP RNAs, named miniSINEUPs, was obtained from the exclusive combination of a BD and an ED, reducing the RNA length to ~ 250 nts[63]. miniSINEUPs ability to increase protein levels was proved for several targets, including GFP and DJ1[60,64,70]. Although the anatomy of natural *AS Uchl1* has been taken as a model for BD design of synthetic SINEUPs, additional BDs were successfully tested for several target mRNAs: -40/+4, -40/0, -14/+4 and -14/0 have been empirically identified as most probably effective BD variants[67,68]. Intriguingly, examples of effective BDs include sequences targeting the starting AUG as well as internal in frame AUG sequences found along the ORF of the target mRNA. The exact BD design rules are not entirely defined yet. On the contrary, it is well known that a precise knowledge of the real TSS of the target mRNA is crucial for appropriate BD design. This can be achieved using FANTOM5[71] datasets and ZENBU Genome Browser Tool for data visualization[72] that allow to monitor TSS usage of a specific mRNA in the tissue of interest.

2.3 Mechanism of action: recent insights

Recent studies have focused on SINEUPs secondary structure to gain better knowledge of their molecular mechanism of action and to optimize its use as therapeutics. Applying chemical footprinting, four Internal Loops (IL) and three Stem Loop (SL) were detected within natural *AS Uchl1* invSINEB2 (Figure 8)[73]. Based on this observation, a deletion analysis revealed that, upon hairpin structure (SL1) deletion (nucleotides 68-77, ΔSL1 mutant), *AS Uchl1* completely lost its capacity to induce UCHL1 protein translation, thus proving the essential role of SL1 in SINEUP activity[73]. SL1 hairpin structure was then further refined through NMR studies performed on the fragment in solution, showing an A-type helical stem terminated by a triloop structure[73]. Using a combination of experimental data (Nuclear Overhauser Effect, NOE) and molecular dynamics simulations, a minimal set of four SL1 conformations compatible with experimental data was obtained[74]. More recently, NMR "fingerprints" allowed the identification of minimal units retaining original structure and function within the invSINEB2 element as one

dynamic domain and two discrete structured domains, named C and M domains[75]. More in detail, the 31-199 nts fragment showed an identical fold and retained 80% of SINEUP activity compared to the full length invSINEB2 element[75]. Altogether these data provide important information for the identification of minimal structural elements required for ED activity, which is a necessary step towards the miniaturization and optimization of the molecule. Furthermore, these observations represent the starting point for comparative studies on other EDs found in natural SINEUPs and for the identification of structural commonalities. Indeed, we recently reported structural and functional similarities between an Internal Ribosome Entry Site (IRES) element and the *AS Uchl1* invSINEB2[76]. In particular, a parallelism between HCV IRES IIId domain, known to be essential for IRES activity, and invSINEB2 SL1 was established trough a mutation analysis. Interestingly, it was demonstrated that invSINEB2 could work as an IRES element with a bi-cistronic assay and, the other way around, IRES elements were proven to be active as SINEUP EDs when substituting the invSINEB2 in a synthetic SINEUP RNA[76].

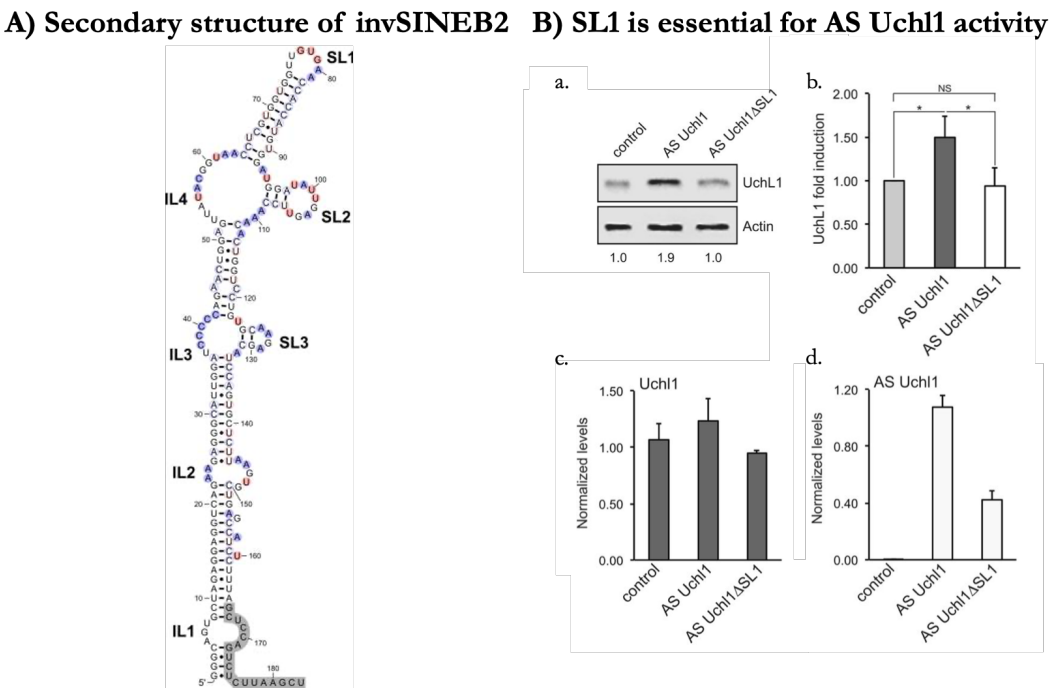


Figure 8 Secondary structure of invSINEB2 from AS Uchl1 functional characterization[73]. **A) Secondary structure of invSINEB2:** blue and red highlights tDMS and CMCT reactive nucleotides, respectively. Grey-highlighted segment indicate DNA primer hybridization site. **B) SL1 is essential for AS Uchl1 activity:** murine neuroblastoma N2A cell line was transfected with AS Uchl1 and ΔSL1 mutant constructs. Control cells were transfected with an empty control. a. representative Western Blot protein analysis. b. graphical representation of AS Uchl1 and ΔSL1 mutant activity on UCHL1 protein. c. *Uchl1* mRNA levels as measured by qRT-PCR Real-time is stable upon constructs transfection. d. AS Uchl1 and ΔSL1 mutant expression level.

Another important piece of information was acquired with studies focusing on invSINEB2 protein interactors. In a first study previously mentioned, the double-stranded RNA-binding protein ILF3 was identified as *AS Uchl1* RNA interactor[61]. Intriguingly, ILF3 was also demonstrated to bind FRAM element[61], which had been reported as embedded TE acting as ED in human natural SINEUPs[59]. This similarity is not reflected as high sequence homology between the two genomic elements, suggesting, together with no evidence of a clear consensus motif for ILF3 binding, that the RNA-protein interaction results from a similarity in the two lncRNAs secondary structure. The interaction with ILF3 influences *AS Uchl1* subcellular localization[61]. Recently, it was also showed that SINEUP RNAs interact with other RBPs, such as polypyrimidine tract binding protein-1 (PTBP1) and heterogeneous nuclear ribonucleoprotein K (HNRNPK)[77]. In this study it was also demonstrated that these proteins binding is essential for SINEUP RNA functional subcellular localization and for the assembly of translation initiation machinery. By knocking-down or over-expressing PTBP1 and HNRNPK proteins, both SINEUP shuttling from nucleus to cytoplasm and activity were altered, proving the important contribution given by RBPs in SINEUP RNA dynamics and functionality[77].

2.4 SINEUPs as a novel toolbox for RNA therapeutics

In last years, gene therapy has paved the way for nucleic acid-based therapies. With significant improvements in safety and efficacy, these technology arose hopes to target undruggable diseases, culminating with the FDA approval of the first gene modification therapy drug for genetic disease treatment in December 2017[78]. To express therapeutic molecules *in vivo,* different options are available: DNA gene delivery, most commonly with viral vectors, and RNA-based systems (Figure 9). DNA-based therapies have been recently developed to replace defective or missing proteins, as vaccines encoding specific antigens or as a treatment for genetic disorders[69,78–80]. In gene therapy, a limitation is represented by the relatively small cargo capacity of Adeno-Associated Virus (AAV) vectors, which shortens the list of genes that it is possible to deliver. Moreover, transgene expression achieved with this technology can reach levels well beyond the physiological range, which could be detrimental in terms of safety. The lack of specific promoters for each cell type could lead to ectopic gene expression in non-specific tissues.

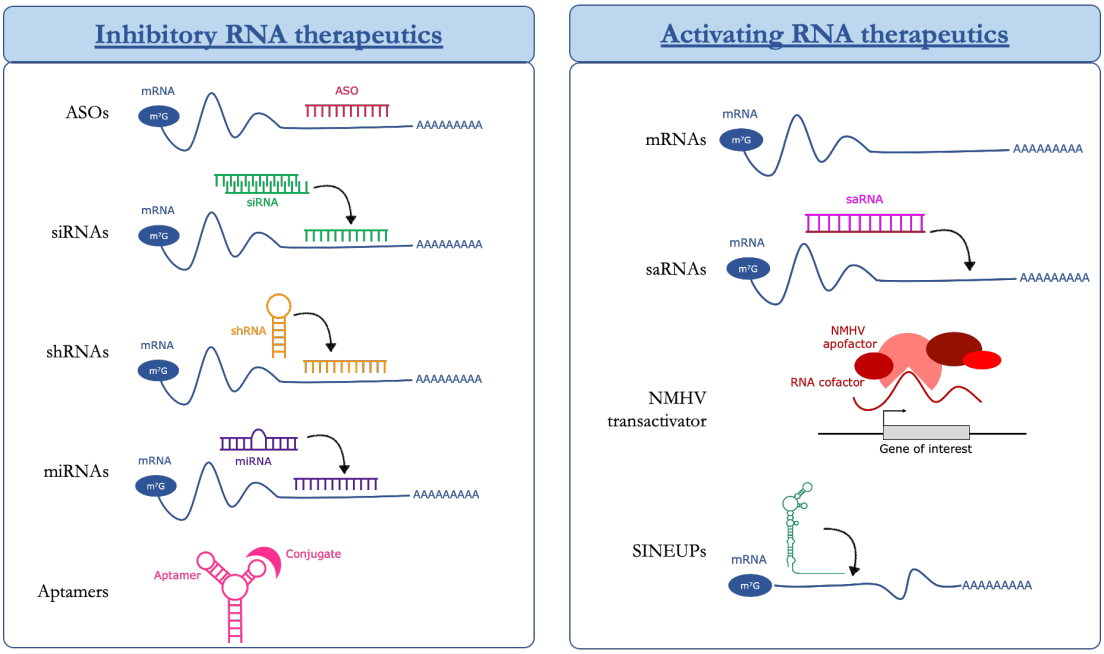


Figure 9 RNA therapeutics under development at date. Most RNA therapeutics technologies rely on complementary base pairing of synthetic RNA molecules. They are usually extensively modified to obtain a higher resistance to RNases, the lowest level of immunogenicity and the maximum binding affinity to their targets. Among these drugs, two main groups can be identified: inhibitory and activatory RNA therapeutics. The first group includes Antisense Oligonucleotides (ASOs), small interfering RNAs (siRNA) and short hairpin RNAs (shRNAs), synthetic microRNAs (miRNA), including miRNA mimics, and aptamers, short single stranded RNAs that can bind a target molecule delivering drugs, like other therapeutic RNAs or proteins or other compounds. The activatory RNA therapeutics group is much less numerous and comprises full synthetic mRNA molecules, small activating RNAs (saRNAs), NMHV transactivators, and SINEUPs, synthetic RNAs designed to redirect natural AS Uchl1 lncRNA activity against a gene of interest.

On the other hand, in the case of RNA therapeutics, major limitations are represented by the rapid degradation of exogenous RNA molecules by ubiquitous RNases, difficulties in the development of effective delivery strategies of negatively charged RNA across the hydrophobic cytoplasmic membrane, and a high immunogenicity risk caused by exogenous RNA leading to toxicity and impaired translation[79,81]. Nevertheless, major advantages that contribute to a rapid expansion and development of RNA-based drugs, especially when compared to DNA-based therapies, include the ability to target previously undruggable diseases, their convenient production in terms of both time and costs and the possibility to develop them as personalized drugs, or easily adapting them to evolving pathogens[79].

The path towards RNA therapeutics development can take two different directions depending on the gene expression dysfunction that causes a disease: inhibitory or activating RNAs (Figure 9).

Inhibitory RNAs

The category of inhibitory RNAs comprises a vast repertoire of molecules: antisense oligonucleotides (ASOs), miRNAs, miRNA sponges, small interfering RNAs (siRNA), short hairpin RNAs (shRNAs), and aptamers.

ASOs are short single-stranded molecules complimentary to a defined target RNA sequence. They can be divided in two sub-classes: the more commonly used RNase H-dependent ASOs, dependent on the enzyme to hydrolyze the RNA strand of an RNA/DNA duplex, and RNase H-independent ASOs[69,79,82,83]. RNase H-dependent ASOs were reported to be more efficient in gene-expression knock-down[79].

siRNAs can be single or double stranded RNA molecules that follow endogenous miRNA pathway to mediate specific, perfectly complementary mRNA silencing by loading them onto RNA-induced silencing complex (RISC)[79,82,84].

shRNAs exploit miRNA maturation pathway as well and are commonly delivered to cells with viral vector systems. Interestingly, two bifunctional shRNA molecules are currently under phase I clinical trial evaluation[79,82,84]. This type of shRNA show higher efficiency as they produce multiple transcripts with perfect and imperfect complementarity to drive degradation and translation inhibition at once[79,82,84].

miRNAs are small ncRNAs that reduce the expression level of multiple RNAs at once by blocking their translation or inducing their degradation. miRNA drugs can be divided into two subclasses: miRNA mimics, dsRNAs mimicking endogenous miRNAs activity, and miRNA inhibitors that are ssRNA molecules synthetically designed to interfere with endogenous miRNAs[79,82,84].

Aptamers are short single-stranded nucleic acids that can target a variety of molecules exploiting their tertiary structure[79].

Activating RNAs.

The class of activating RNA therapeutics comprises non-degradative ASOs, RNA activators (RNAa) and Nuclear localization signal – MS2 coat protein RNA interacting domain – HA epitope – (3x) VP16 trans-activating domain (NMHV) transcription factors.

Non-degradative ASOs are applied to up-regulate target genes' expression with different mechanisms, such as interfering with miRNA activity or modulating mRNAs processing acting on their splicing. Among these, the most important examples include exon-skipping ASOs or splice-switching oligonucleotides (SSOs)[85–87].

RNA activation technology applies small RNAs to enhance transcription and it was first described in 2006 by Li and colleagues[88]. More recently, several reports confirmed RNAa as a common mechanism of gene expression regulation[89–91]. These RNA molecules can be sense or antisense oriented and can target TSSs[88,92], sequences nearby polyadenylation sites[93], cis-active elements within the gene of interest or the transcribed region of the gene[94]. Some small activating RNAs (saRNAs) can bind AS transcripts interfering with their repressive functions[95] or can induce gene locus transactivation[92]. Two of their features are particularly attractive for application: they have a prolonged effect[88] and reach an mRNA up-regulation that commonly falls within the physiological range[89,91].

NMHV transcription factors is a new class of artificial trans-activators, which are RNA-programmable enzymes. They consist of a synthetic ribonucleoprotein transcription factor that stimulate transcription, combined with a ncRNA domain that drives the enzyme to a specific target gene[96]. The two domains are linked by two accessory domains: an MS2 RNA-interacting domain[97] and a hairpin interactor[98], joined to the transcription factor and to the ncRNA domain respectively.

SINEUPs as a new therapeutic platform

In current medical practice, there are several unmet therapeutic needs to increase protein levels *in vivo*. As a broad classification, we can envision the use of SINEUP technology to:

i. genetic diseases with the lack of one functional allele for a single (haploinsufficiency) or multiple (copy number variations; CNVs) genes;
ii. ii. complex diseases where the increase of compensatory pathways may preserve or restore physiological activities.

Haploinsufficiencies are a wide spectrum of diseases (more than one thousand) where the protein product of both alleles is required to ensure the normal phenotype, but one allele is inactive due to hereditary or germline mutations leading to lower expression of a functional protein. They are very heterogeneous (each of them involving a different gene) and rare (they occur in a very limited number of patients), limiting drug development by the private sector. Importantly, recent data has shown that an uncontrolled, ectopic, large expression of some of these target genes can be detrimental, phenocopying the disease or leading to life-threatening side-effects. These worrisome results strongly support the need for new technologies able to restore the physiological range of expression of the gene of interest.

Targeted enhancement of protein level would also be beneficial for diseases caused by pathogenic CNVs. In these cases, heterozygous deletion of a portion of a chromosome

(from 50 bps to several million bases) leads to multiple genes' haploinsufficiencies. Pathogenic CNVs are significantly enriched for genes involved in neurodevelopment and include at least one dosage-sensitive gene, whose duplication or deletion is usually negatively selected. Recurrent deletion and duplication syndromes can either manifest with similar characteristics or with mirror image traits indicating that duplication of the very same region may be pathogenic (reciprocal CNVs). As a consequence, overexpressing large amounts of proteins encoded by hemideleted genes can be pathogenic as phenocopies of duplications. It is therefore crucial to increase protein levels of multiple target genes within the physiological range. Since no technologies are currently available to restore the expression of multiple genes, large pathogenic CNVs remain untreatable.

In many complex, multifactorial diseases, the increase of pro-survival factors and enzymes may improve the well-being of patients. As an example, exogenous delivery of neurotrophic factors has been proposed as therapeutic treatments for neurodegenerative diseases. However, dosage and bioavailability issues hamper the therapeutic benefits with current delivery strategies. Moreover, toxicity from off-target distribution highlights the need for tissue-specific expression. Therefore, there is a crucial unmet therapeutic need to enhance the expression of compensatory pathways within the physiological range and in the appropriate brain region and body district at large. Similarly, increasing the concentrations of transcription factors and enzymes involved in pathways whose efficiency is lowered in neurodegenerative diseases, such as autophagy and mitochondrial biogenesis, can result in valuable novel therapeutic options.

SINEUP technology presents specific advantages over current technologies to increase protein expression *in vivo* for therapy:

i. the increase of protein quantities is within physiological range. This is important when the overexpression of the protein of interest is toxic as in the case of proteins with pro-oncogenic properties or of reciprocal CNVs, when duplication could elicit a mirror disease
ii. SINEUP activity occurs exclusively in cells that express the mRNA target avoiding the toxicity associated to the ectopic expression of the protein in unwanted cells and not requiring cell-type specific promoters
iii. AAV can be used to induce the expression of large proteins overcoming constrains in cDNA length

iv. it is the only technology targeting translation leaving room for combinatorial therapy

v. multi-target SINEUPs can increase expression of several proteins at the same time.

The first demonstration of SINEUP activity *in vivo* was obtained in a medakafish (*Oryzias latipes*) model of microphtalmia with linear skin lesions (MLS) syndrome[99]. It is caused by mutations in enzymes of the mitochondrial respiratory chain including the subunit 7B of cytochrome c oxidase (cox7b). Medakafish model of MLS displays down-regulation of cox7b resulting in microcephaly and microphtalmia. When a SINEUP for medakafish cox7b was microinjected into embryos, the microcephaly and microphtalmia disease phenotype was completely reverted, due to restoration of physiological levels of cox7b, with no transcriptional effects[99].

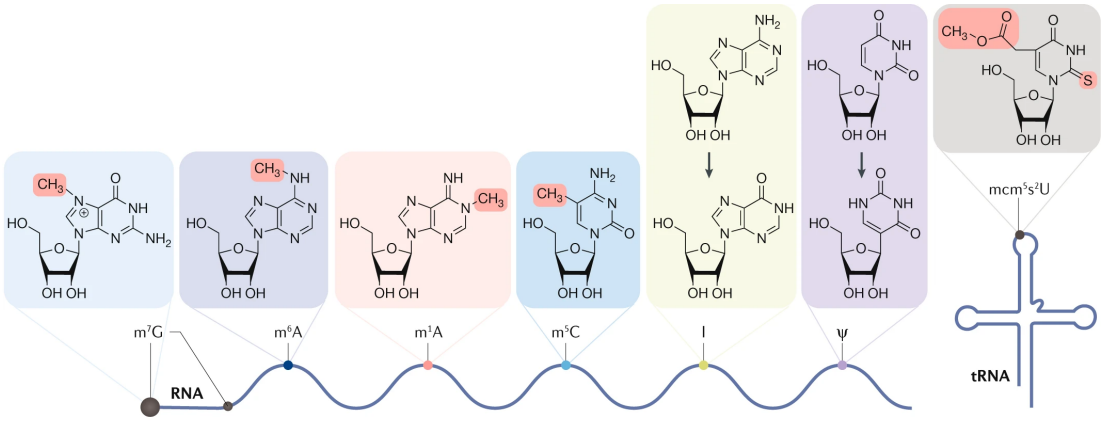
As a second Proof-of-Concept, our lab focused its attention on rescuing haploinsufficiency in patients' cells from a human disease: Friedreich's Ataxia (FRDA)[68]. FRDA is a fatal and presently untreatable genetic disease due to a decreased expression of frataxin (FXN), caused by hyperexpansion of GAA repats. The FXN gene encodes for frataxin, a protein involved in the biosynthesis of the iron-sulfur cluster (ISC). An insufficient ISC biosynthesis leads to decreased mitochondrial activity. *SINEUP-* and *miniSINEUP-FXNs* increased frataxin expression of 2-folds in FRDA-derived fibroblasts and lymphoblasts re-establishing frataxin physiological levels. This increase was sufficient to restore the physiological mitochondrial activity of patients'-derived primary cell lines, a major disease-associated phenotypic trait[68].

As a Proof-Of-Concept of SINEUP activity in a complex disease, our laboratory confronted a mouse model of Parkinson Disease (PD)[67]. PD is one of the most common neurodegenerative diseases and is caused by loss of dopaminergic (DA) neurons in the Substantia Nigra. Glial cell-derived neurotrophic factor (GDNF) is a neurotrophic factor, promoting the survival of DA neurons[25] and it has been studied as potential agent to halt neurodegeneration in PD. However, long-term delivery of GDNF resulted in toxicity. We have recently shown that the expression of *miniSINEUP-GDNF* RNA in the mouse striatum delivered by an AAV9 vector led to an increase of endogenous GDNF protein of two-fold for at least six months and the potentiation of DA system's functions. The common side effects caused by the ectopic expression of GDNF were not observed. *miniSINEUP-GDNF* was able to ameliorate motor deficits and neurodegeneration of DA neurons in a neurochemical mouse model of PD[67].

As described above, SINEUP molecules could be administered following two different strategies: as DNA molecules through AAV delivery, when SINEUP molecules are chronically expressed *in vivo*, or as chemically synthesized RNA molecules. The first approach is particularly indicated when a physiological upregulation of the target mRNA translation is crucial for phenotypic rescue of the disorder and the uncontrolled expression of the protein in given tissues could be detrimental. In other cases, SINEUPs delivery as RNA molecules could be advantageous as well. Indeed, with this approach, there would be no stable genomic alteration, reducing genotoxicity risks. The development of RNA therapeutics has been massively optimized in recent years, especially for siRNAs and ASOs and could be exploited for SINEUP technology application. Limitations in this sense are posed by SINEUP RNA molecule length, that should be reduced to less than 60 nts to allow cost-effective manufacturing and efficient *in vivo* delivery. It is reported that *in vitro* transcribed (IVT) SINEUP RNA is not active in cells, suggesting that chemical modification would be needed to preserve RNA function and stability. Recently, encouraging data have proved that the incorporation of chemically modified ribonucleotides restore IVT SINEUP RNA activity, making an important progress for its development as a RNA drug[100,101]. However, it would be extremely important to gain further knowledge about natural chemical modification of SINEUP RNA to improve its efficacy.

3. RNA modifications

The first discovered modification of an RNA base was pseudouridine (Ψ), as far back in time as 1951[102], shortly after the discovery of DNA 5-methylcytosine. Over time, the number of identified RNA post-transcriptional modifications (PTMs) has constantly increased, reaching more than 150 known modifications at date[103,104]. Only in recent times, the term "epitranscriptomics" was coined to indicate all modifications that are added on transcribed nucleotides. Such chemical modifications are, indeed, catalytically deposited and removed by specific enzymes. Being detected by specific "reader" proteins, PTMs can perturb RNAs features and functions, including stability, structure, interactions and subcellular localization[105]. Chemical modifications were reported on all four ribonucleotides consisting in the addition of a chemical group or in the modification of the original RNA sequence (Figure 10). 75% of PTMs is represented by the addition of methyl groups, causing a gain of positive electrostatic charge[105].



Writers	• METTL1 • WBSCR22 • RNMT	• METTL3–METTL14 • METTL16 • METTL5 • ZCCHC4	• TRMT10 • TRM6–TRM61	• NSUN1 to NSUN7 • DNMT2	• ADAR1 and ADAR2 • ADAT2 and ADAT3	• PUS1 to PUS10 • PUS7L • RPUSD1 to RPUSD4 • DKC1	• ELP1 and ELP3 • ALKBH8 (mcm^5U) • CTU1 and CTU2 (mcm^5s^2U)
Readers	Unknown	• YTHDF1 to YTHDF3 • YTHDC1 and YTHDC2	YTHDF2?	ALYREF	Unknown	Unknown	Unknown
Erasers	Unknown	• ALKBH5 • FTO?	ALKBH1 and ALKBH3	Unknown	Unknown	Unknown	Unknown
A–T G–C pairing	No effect	No effect	Disrupted	No effect	Altered	No effect	Altered
Hoogsteen pairs	Disrupted	No effect	No effect	No effect	Disrupted	Altered	Altered

Figure 10 Internal RNA modifications[106]. Specific modification groups, with relative main protein factors involved in deposition, removal and recognition, are reported. Effect of each modification on base-pairing is also reported.

3.1 Pseudouridine (Ψ)

Among RNA bases, uridine has the largest number of reported modifications. 5-ribosyluracil, also known as pseudouridine (Ψ), besides being the first discovered modification, is also the most overall abundant one, reported in almost all RNA classes[102] (rRNA, snRNA, snoRNA, tRNA and mRNA). Its deposition is catalyzed by a conserved family of enzymes called pseudouridine synthase. In tRNAs Ψ is conserved in position 55 within the T-arm loop, contributing to tertiary structure, but it has also been reported in numerous other positions along tRNA primary structure, where it is suggested to contribute to stability and structure[105].

3.2 7-methylguanosine (m^7G)

m^7G has been detected in 5' cap structures of eukaryotic mRNA as well as in internal positions within tRNAs and rRNAs across all species. Its deposition is catalyzed by a number of enzymes, among which the best-characterized is METTL1[107,108]. This modification dramatically perturbs the charge density of RNA, possibly altering the secondary RNA structure, but does not impair Watson-Crick G:C base pair formation. Altogether these features make m^7G neutral to base-pairing and not altering reverse transcription, leading to a difficult detection of the modification.

3.3 5-methylcytosine (m^5C)

m^5C was discovered more than 40 years ago and mapped in mRNAs and lncRNAs through a modification of bisulfite methodology used for DNA[109]. It has been reported in tRNA, rRNA, lncRNA and in mouse and human protein-coding transcripts, where they are usually found about 100 nts downstream TSSs and in the UTRs[110]. This modification and the protein responsible for its deposition and elimination resulted important to control the fate and function of RNAs, since their alteration was linked with pathological states[110]. m^5C is read by the ALYREF protein that functions as mRNA export adaptor, suggesting a possible role of this modification in nuclear export regulation[109].

3.4 2'O-Methyl-Adenosine (2'OMeA, Am)

2'O-Methyl-Adenosine is often found in the extended cap structure in mRNAs, where 2'O sites of the first and second nucleotides next to the cap are methylated. The first of them can also bear an additional N6-methylation (m^6Am), proven to regulate mRNAs and ncRNAs stability. 2'O-Methyl-Adenosine also affects translation efficiency and self-non-self recognition, being important for induction of self-tolerance or de-immunization of exogenous RNA[101].

3.5 N6-methyladenosine (m^6A)

N6-methyladenosine is the most common RNA modification in many different species among viruses[111,112], bacteria, yeast, plants and mammals[113]. Watson-Crick base pairing of an m^6A with and opposite U forces rotation of the carbon-nitrogen bond destabilizing the RNA duplex to form locally unstructured transcripts[109]. Indeed, m^6A peaks are reported to be preferentially enriched within unstructured regions of RNA, compared to loops or bulges of stem-loops[114]. This implicates a wide range of effects of m^6A modification, often relying on a deep structure-function connection[115]. In mammals, m^6A is the most common RNA modification found in mRNAs and ncRNAs, where it is post-transcriptionally deposited in the cell nucleus and can exert regulatory functions in many cellular processes such as RNA splicing, stability, nuclear export, and translation. High-throughput transcriptome-wide approaches, commonly combined with antibody enrichment, have allowed the identification of a consensus motif for m^6A sites deposition: DRACH (with D = G, A, or U; R = G or A; and H = C, A, or U)[116,117]. While consensus sequences are quite common through the transcriptome, only a few of them are actually methylated with a site and transcript specificity that is still poorly understood[118]. Another important step has been achieved with the identification of several m^6A-modifying proteins with different roles: "writers" (methyltransferases, mainly METTL3 and its

adaptors)[119,120], "erasers" (demethylases FTO and ALKBH5) and "readers" (YTHDF1, YTHDF2, YTHDF3 and others)[121]. Despite rising interest in understanding m^{6}A modification function, the majority of technologies currently used to map modified residues still rely on very long and complex protocols such as site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET)[122], m^{6}A individual-nucleotide-resolution cross-linking and immunoprecipitation (miCLIP)[123], or are based on peculiar signatures derived from reverse transcription (RT)[124,125], or rely on m^{6}A sensitive RNase digestion[126]. In this scenario, a major improvement is represented by Nanopore direct RNA sequencing (DRS), that allows to directly sequence native RNA molecules with a system that has been reported to intrinsically retain information about RNA modifications[127–130]. Indeed, the presence of nucleotide modifications can induce shifts in time and current intensity while the nucleic acid passes through the sequencing pore. These variations have been successfully used to detect DNA modifications at single-nucleotide resolution, while RNA modification detection is under optimization.

m^{6}A modification in mRNAs

m^{6}A-RIP-Seq analysis demonstrated that m^{6}A residues are enriched in specific regions along mRNA transcripts[114]: near the stop codon, evenly distributed up- and downstream of the site, in the 5'UTR, with strikingly high concentration and interesting tissue-specific differences, and in 3'UTR outside region adjacent to the stop codon and within the coding sequence, with lower concentration. No m^{6}A site was found in polyA termination. Many different effects of m^{6}A on mRNAs regulation have been reported. In some cases, m^{6}A has a role in modulating RNA-protein interactions, either blocking or inducing them[115]. It can also regulate RNA functionality altering its structure or folding, as the A•U basepair still forms but is slightly destabilized, as already mentioned. Although this effect results to be light on an RNA duplex, it could influence some interactions that are based on duplex stability, such as microRNA-mRNA ones[114]. In other cases, m^{6}A modification of mRNAs has been reported to regulate isoform diversity by alternative splicing. mRNAs that undergo alternative splicing are, indeed, more likely to contain m^{6}A and METTL3 binding domain than mRNAs that display a single isoform[114]. However, a molecular mechanism for this function is not fully understood yet. Nevertheless, the first characterized m^{6}A function was to cause mRNA instability[118], established through a comparison between m^{6}A- and non-m^{6}A-containing RNAs half-lives[131]. More recently, this function appears to be co-regulated by YTH domain-containing family[132]. m^{6}A may

also enhance mRNA export from the cell nucleus taking advantage of YTHDC1 reader protein binding[118]. Other examples refer m^{6}A function as translational enhancer and will be discussed in the following paragraphs.

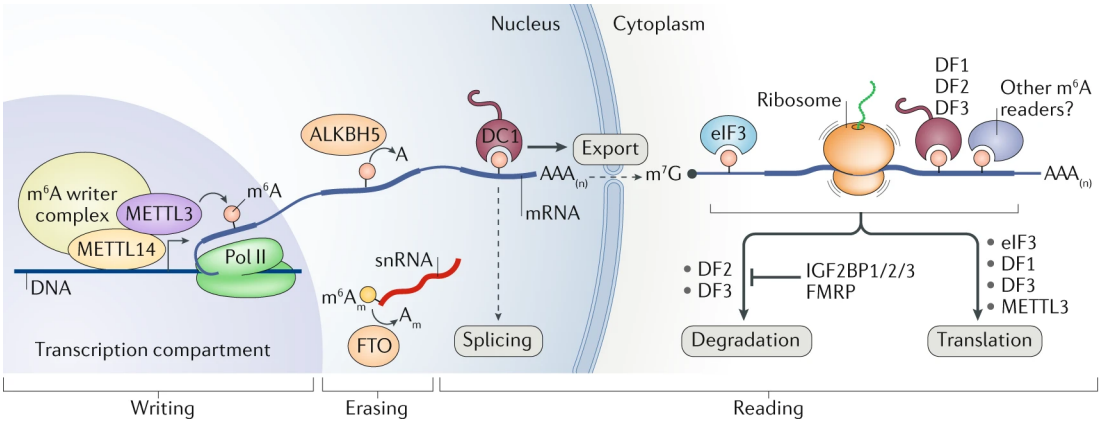


Figure 11 m^{6}A mRNA modification life cycle[118]. m^{6}A writer complex is composed by the core methyltransferase-like protein 3 (METTL3) and its adaptors. m^{6}A erasers are also mainly localized in the cell nucleus. Here, the modified nucleotide can be recognized by nuclear reader proteins, such ash YTHDC1 (DC1), that can affect splicing or other processes, such as mRNA export. Once in the cytoplasm, m^{6}A can bind other reader proteins that can act on its stability, translation efficiency or subcellular localization.

m^{6}A modification in lncRNAs

To date, most studies investigate m^{6}A modifications role in mRNAs, while much less is known about their function in lncRNAs, with major attention focusing on cancer-related transcripts[110,133]. For example, in the case of human lncRNA *MALAT1*, one m^{6}A methylation site has been identified within a hairpin stem and it has been demonstrated to destabilize the transcript structure making the DRACH sequence opposing U-tract more accessible for RNA-binding proteins[115]. A recent study reported a YTHDC1-driven pathway required for *XIST* functions involving 78 m^{6}A residues[134]. Furthermore, in the case of circRNAs, it has been reported that m^{6}A modifications may play a relevant role in their biogenesis and also in their cap-independent translation[135]. Recently, chromatin state and transcriptional regulatory function was also reported for m^{6}A modification on chromatin-associated RNAs (carRNAs)[136]. This class of RNAs comprises promoter-associated, enhancer and repeat RNAs. It was demonstrated that YTHDC1 protein induced the decay of a subset of these m^{6}A-methylated RNAs, especially of LINE-1 family, through the nuclear exosome targeting-mediated nuclear degradation[136].

m^{6}A role in translation

An interesting functional effect of m^{6}A is translation upregulation. Thanks to several recent studies, m^{6}A has been demonstrated to mediate translation upregulation through three different mechanisms. The first consists of a direct translation activation by

METTL3[137], which, according to this model, remains anchored to the methylated transcript after its export to the cytoplasm (Figure 12). Here, METTL3 would bind eIF3 that interacts with the mRNA cap-associated proteins, inducing the formation of a loop in the same mRNA, and eventually allowing ribosomes at the stop codon to reload onto 5'UTR[137]. Another m^{6}A-mediated translation upregulation mechanism involves the m^{6}A reader YTHDF1 protein. According to this model, the reader protein would bind eIF3, that, in turn, recruits the 40S ribosome unit to the mRNAs, enhancing their translation[138]. An additional example of this particular function of m^{6}A modification is a proposed mechanism involving a direct binding of a 5'UTR-contained m^{6}A to eIF3[139]. In this case of study, the modification is required to be located in the 5'UTR to exert the translation up-regulatory function. Intriguingly, this m^{6}A-dependent translation initiation mechanism does not require eIF4E, m^{7}G-containing mRNA cap-binding protein, thus defining a new model of cap-independent translation initiation, alternative to the well-established IRES model. Indeed, 5'UTR m^{6}As differ from IRES elements for their ability to recruit ribosomal preinitiation complexes to the 5' end of mRNAs instead of allowing internal ribosome entry[139]. Furthermore, the importance of 5'UTR m^{6}A-dependent translation initiation is confirmed by their selectivity for specific forms of stress, such as heat shock[139]. Despite growing knowledge, more studies are needed to determine m^{6}A relevance and mechanisms in translation up-regulation processes.

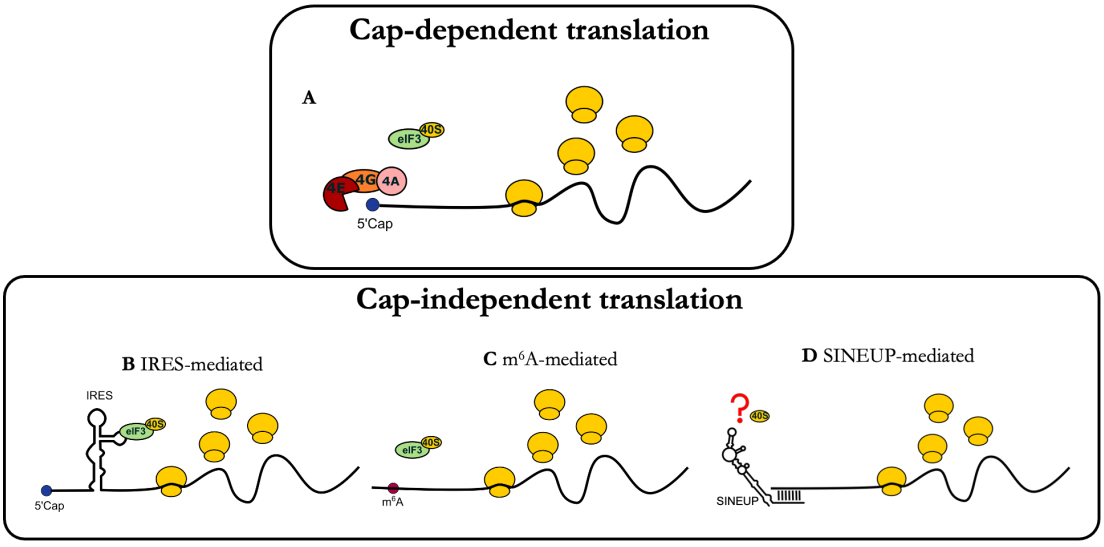


Figure 12 Translation initiation mechanisms in eukaryotes. **A) Cap-dependent translation.** In eukaryotic cells under physiological conditions, most of the mRNAs are transcribed with a cap-dependent and scanning-dependent mechanism. This process starts with eIF4F complex binding to m^{7}G cap located at the 5'end of mRNAs. The helicase eIF4A unwinds this region of the mRNA, while eIF4E helps the assembly of the other factors and eIF4G binds eIF3 recruiting the ribosome. The whole complex scans the along the mRNA in a 5'→3' direction to identify the AUG starting codon. Once this is recognized, scanning is arrested, part of the co-factors is released and recycled for translation of other mRNAs, while the remaining complex binds the 60S ribosome to form an elongation-competent 80S ribosome. **B) IRES-mediated translation**

Canonical cap-dependent translation is often inactivated as a protective mechanism against different types of cellular stress, such as viral infections. In such cases, translation of proteins can be initiated through an Internal Ribosome Entry Site (IRES). These RNA element was first discovered in viruses such as picornaviruses and Hepatitis C virus, and later were also found in cellular RNAs, reaching a number of 50 and 70 viral and cellular IRESs at date respectively[140]. IRESes can interact with translation initiation factors or directly with the 40S subunit, allowing to bypass the cellular block in cap-dependent translation initiation and leading to protein production. This leads to ribosomal positioning at or near the initiation codon and promoting translation initiation. Their activity is regulated by IRES trans-acting factors (ITAFs) that are RNA-binding proteins (RBPs). IRESes are also reported as short elements that can base-pair with ribosomal RNA (rRNA), similar to bacterial Shine-Dalgarno sequences. **C) m^{6}A-mediated translation** An alternative translation initiation mechanism was recently reported to involve m^{6}A modification. Indeed, it was demonstrated that uncapped mRNAs are translated in cell-free extracts if containing m^{6}A in their 5'UTR in absence of EIF4F complex. It was also reported that m^{6}A modification can directly bind eIF3, that, in turn, recruits the initiation complex for translation initiation. It has been proposed that m^{6}A-triggered translation initiation could be used in cellular stress conditions[139]. **D) SINEUP-mediated translation** Another cap-independent translation initiation mechanism is suggested by SINEUP lncRNAs. They have been reported to be active under mTORC pathway inhibition, which is a well-established cap-dependent repressive condition. SINEUP lncRNAs have been proved to increase their mRNA target association to heavy polysomes, but the molecular mechanism of their activity is not entirely understood at date[25,77].

3.6 RNA modifications in the SINEUP world: what is known

In a diversified and dynamic context, such as the one of epitranscriptomics, the study of the complex relationship between endogenous RNA modifications and functions is very attractive. In the case of SINEUP technology, information about nucleotide modifications could be of extreme interest for two different purposes: on one hand it would be important to characterize endogenous RNA modifications to better understand their potential involvement in natural SINEUP activity regulation, on the other hand, the results of the investigation could be advantageous for optimizing SINEUP RNA for therapy. Unfortunately, no data is available at date about endogenous SINEUP RNA modifications; instead, some recent studies focused on the identification of optimal modifications that allow to design fully functional *in vitro* transcribed (IVT) SINEUPs. The first report of IVT SINEUP application was described in a Medaka model of microphtalmia with linear skin lesions[99]. A fully synthetic IVT SINEUP was designed targeting endogenous *cox7B* mRNA and successfully achieved the protein level rescue in cox7b morphants, with functional rescue of eye and brain size. In human cells, on the other hand, IVT SINEUP RNAs targeting exogenous transfected GFP and endogenous *DJ-1* mRNA, previously validated SINEUP targets[54], were reported to be inactive[100,101]. However, encouraging data have proved that the incorporation of chemically modified ribonucleotides restore IVT SINEUP RNA activity, making an important progress for its development as a drug[100,101]. In a first study, m^{5}C, Ψ and N^{1}Ψ were used to replace all C or U nucleotides during SINEUP RNA *in vitro* transcription. The three modified IVT

(mIVT) RNAs were able to rescue SINEUP activity on *EGFP* mRNA both in a cell system and in cell-free extract[100]. A second study selected three different modifications among many others as the most promising ones: 2'Omethyladenosine (2'OMeA), m^6A and, again, Ψ. All three modifications were able to preserve SINEUP functionality when introduced in the IVT reaction at optimized ratio and combination, however their effects were not additive. It was also suggested that different combinations of modifications may stabilize a core structural domain which could be disrupted by an excess of such modified nucleotides. The percentage of modified nucleotides needed to rescue SINEUP activity was indeed way lower than what expected, since, from MS analysis, the 20% of Am in the Am+m^6A combination was sufficient to rescue activity, while a fully m^6A modified transcript was inactive. In addition, the study underlines how both an appropriate set of chemical modifications and a functional structural arrangement are essential to design a functional SINEUP RNA therapeutic molecule[101].

4. SINEUPs therapeutic application: Dominant Optic Atrophy

Dominant Optic Atrophy (DOA) (OMIM #165500) is an optic neuropathy with early onset of visual impairment that, eventually, can lead to blindness[141,142]. It was first described by Batten in 1896 and appears with variable degree of visual loss.

4.1 Clinical features

DOA is the most common hereditary optic neuropathy, with a prevalence ranging from 1:12000[142] in Denmark, due to a founder mutation, to 1:50000[143–146]. Main symptoms are progressive decrease in visual acuity, tritanopia (confusion in distinguishing blue-yellow hues), loss of sensitivity in central visual fields and optic disk pallor, as visible in Figure 13B. Unlike most common causes of blindness, DOA is not associated with age progression, as patients usually lose their vision before adulthood. In most cases, patients start to lose sight during the first or second decade of their life, they present a non-syndromic, bilateral optic neuropathy that causes a progressive, irreversible and symmetric decrease of vision, triggered degeneration of Retinal Ganglion Cells (RGCs), located in the inner layer of the retina, and optic nerve atrophy (Figure 13A)[141,142,146,147]. Clinical examinations report bilateral pallor of the temporal sector of the optic nerve head, loss of RGCs entering the optic nerve, thickness reduction of the prepapillary retinal nerve fiber layer as can be seen with high resolution optical coherence tomography[142,143,148–150]. Visual acuity typically remains better than 20/200, with most patients retaining moderately good functional vision until relatively late in the disease progression[143]. From a histopathological

point of view, DOA is found as RGCs degeneration focused in the papillomacular bundle. It is also possible to confirm DOA diagnosis with a demonstration of preferential RGCs loss in the macula and peripapillary retinal nerve fiber layer with high-resolution optical coherence tomography imaging. Electrophysiological studies can be also performed in indefinite cases of DOA[143]. In all cases of DOA, pupillary light reflex and circadian rhythms are preserved[143]. Since RGCs originate from an extension of the diencephalon, DOA can be classified as a central nervous system disease; at the same time the disease is also defined as a mitochondropathy, because the causative genes encode for ubiquitously expressed Inner Mitochondrial Membrane (IMM) proteins[144]. However, DOA penetrance is incomplete and its severity is largely variable, with patients that display visual acuity from 20/20 to light perception[151]. The disease is also reported to present, in up to 20% of cases, extra-ocular symptoms, producing a DOAplus phenotype. This form of the disease presents sensorineural hearing loss, progressive external ophthalmoplegia, peripheral neuropathy, myopathy, ataxia and progressive external ophtalmoplegia[143,144,151,152]. Hearing loss is the prevalent extraocular feature and is reported in association with a specific OPA1 mutation in exon 14 (c.1334G > A, p.R445H)[143]. An age-related decrease in OPA1 protein was reported in sedentary but not in active humans in association with muscle loss[150]. In mice, *OPA1* muscle-specific deletion was found to induce precocious senescence and premature death, while conditional inducible deletion only alters mitochondrial morphology and function with no change in mtDNA content[150]. In these cases, OPA1 loss results in ER stress that induces a catabolic program of muscle loss and systemic aging[150]. Clinical evaluation and correct diagnosis of this disorder are further hindered by a wide inter- and intra-familial variability and incomplete penetrance[153], which suggests that genetic background and environmental factors could have a significant role in the determination of clinical phenotype.

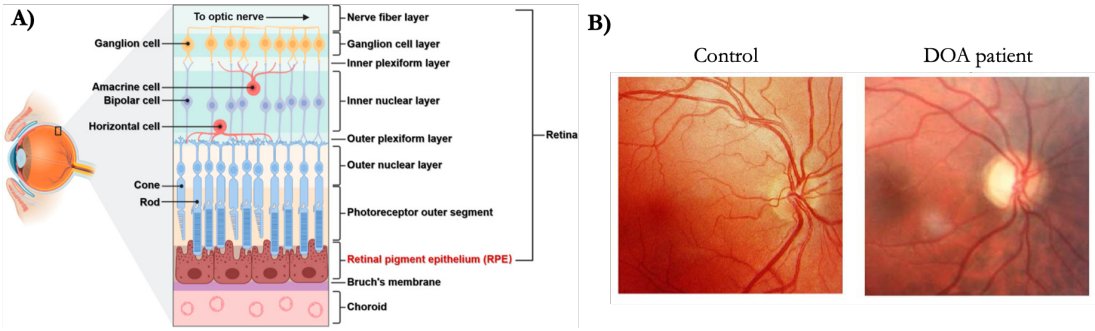


Figure 13 Retinal structure and optic disk examination. **A) Retinal structure.** The retina is composed of multiple layers and different cell types. Retinal Ganglion Cells axons project into the nerve fiber layer and form the optic nerve[154]. **B) Eye fundus examination in healthy control and DOA patient.** It appears an evident

optic nerve pallor in the DOA patient, particularly on the temporal side, while the rest of the retina appears comparable to the healthy control[144].

4.2 Etiology

Four genes have been identified as responsible of DOA: OPA1 (3q28-29), OPA3 (19q13.2-13.3), OPA4 (18q12.2-12.3) and OPA5 (22q12.1-q13) with OPA1 gene accounting for 65-90% of cases, and the other causative genes each one for less than 1% of cases [143,147,153,155]. More than 500 variants are reported at date on the locus-specific database dedicated to OPA1 (https://databases.lovd.nl/shared/genes/OPA1) 80% of which are pathogenic[156,157] (Figure 14); the large majority of variants can be found in the dynamin and in the GTPase domains and pathogenic mutations account for more than 60% of the total number of variants. In about 50% of cases pathogenic mutations cause the production of a truncated form of the protein that undergoes mRNA decay and eventually leads to a loss of function of the mutant allele, suggesting haploinsufficiency as the most common causative mechanism of DOA[158–160]. Missense mutations, generally found in the GTPase domain, account for around 27% of total number of mutations, show a dominant-negative effect, often associated with the severe syndromic disease identified as DOA "plus", characterized by multisystemic involvement and a large spectrum of clinical features, including Parkinsonism and dementia[143,161]. The remaining mutations comprise 27% causing splice variant, 23% of frame shift mutations, 16,5 % of nonsense mutations and 6% deletion or duplication errors[143]. Only a few mutations are recurrently reported, but some are frequently reported: c.2873_2876del variant in exon 29, that induces p.(Val959Glyfs*3) frameshift mutation that leads to premature truncation, reported 22 times, c.1311A>G variant in exon 14, inducing a missense mutation p.(Ile437Met) which is asymptomatic and reported 16 times, c.2635C>T variant in exon 26, inducing a nonsense mutation p.(Arg879*), reported 16 times. As for mutations that lead to premature termination, frame shift and deletion/insertion mutations also result in incomplete and significant decrease of OPA1 wild-type protein, leading to haploinsufficient phenotype. DOA plus syndromes are usually caused by missense mutations affecting the GTPase domain and suggest a possible dominant-negative effect as a contributor to pathogenic phenotype development[143]. Recently, OPA1 has been reported to be dramatically downregulated in both *in vitro* and *in vivo* models of prion diseases and its overexpression alleviated prion-induced mitochondrial dysfunctions[162].

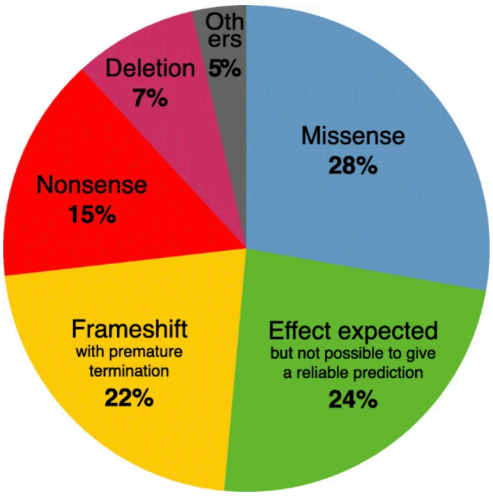


Figure 14 OPA1 protein pathogenic mutations effect. Other consequences (5%) include: synonymous, no protein production, duplication, and extension. Reported data are deposited in a public database that includes 831 patients, 697 of which have DOA, 47 DOA plus, 83 are asymptomatic or unclassified. Four other patients showed a phenotype not reported in association with OPA1[157].

4.3 OPA1 protein

Structure

OPA1 protein is part of the dynamin superfamily that reaches highest expression levels in brain, retina and heart[147]. Dynamin proteins are GTPases and include classical dynamins (1,2,4), mitofusins, Drp1, OPA1, Mx proteins, guanylate-binding proteins (GBPs) and atlastins in eukaryotic cells. Studies on protein structures show that all dynamins include a GTPase domain that binds GTP, hydrolyzing it, a α-helical bundle domain, a middle domain involved in oligomerization and a GTPase effector domain (GED), with the last two involved in stimulation of GTPase activity. At date, no crystal structure of OPA1 protein is available, probably for its high complexity. Human *OPA1* gene consists of 30 exons that give rise to at least eight alternatively spliced variants. OPA1 protein presents a mitochondrial targeting sequence (MTS) followed by a transmembrane (TM) domain, cleaved to exert mitochondrial functions. A recent study applied the threading approach to obtain a model of the three-dimensional structure of OPA1 protein without MTS and TM domains, using OPA1 isoform 8 (Figure 15)[163]. This model identified a N-terminal and C-terminal region. The N-terminal region is rich in α-helices with no specific domain and differs among the eight spliced variants for the long peptide chain length. On the other side, the C-terminal domain of the protein was found as a compact structure containing the GTPase, PH and GED domains, that differs from the one of other dynamin proteins mediating membrane fusion, but was found similar to dynamin 1 and MxA, that can self-assembly in dimers by the middle domain and GED. The evolutionary

conservation of the C-terminal domain was significantly higher than the one of the N-terminal region and, interestingly, an analysis of the localization of known pathogenic mutations reported that the sites were mostly located in the C-terminal region[163].

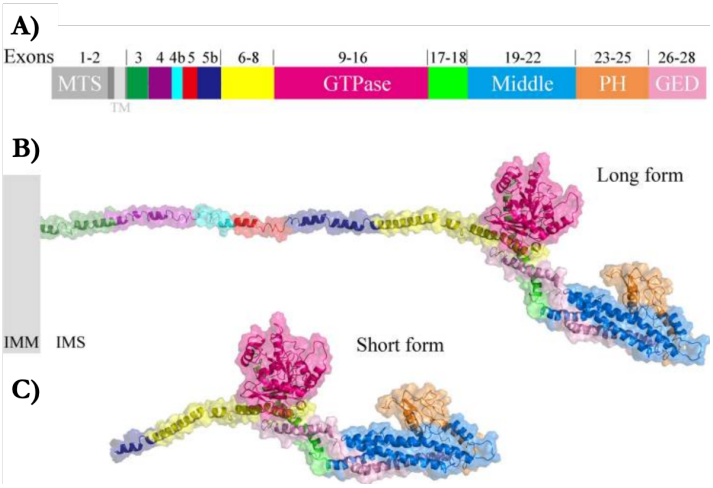


Figure 15 Three-dimensional model of OPA1 protein. **A) Domain architecture of OPA1 isoform 8.** Each domain is labeled with different colors and the exon number is also reported. **B) Structural model of long OPA1 protein.** Isoform 8 without MTS and TM (long form) highlights an α-helices rich N-terminal domain with no specific domain, and a dense structure in C-terminal domain containing GTPase, PH and GED domains. **C) Structural model of OPA1 protein short isoform.**

Processing

All OPA1 protein isoforms are ubiquitous, but each one's expression level is tissue dependent. Once the precursor protein is imported through outer mitochondrial membrane (OMM) and IMM translocases, OPA1 long forms (l-forms) are produced by MTS cleavage (Figure 16). These can be further proteolytically processed at the N-terminal domain producing short forms (s-forms), that are soluble and located in the inner mitochondrial space (IMS). Four isoforms that retain exon 4b upon alternative splicing are totally processed to form soluble s-forms by two peptidases located in the IMM: OMA1, activated upon stress conditions, cleaves exon 5 in S1 site, while YME1L, constitutively active, cleaves exon 5b in S2 site. A finely-tuned mechanism regulates OMA1 and YME1L levels depending on both $\Delta\Psi_m$ and ATP levels and regulating, as a result, OPA1 l-/s- forms' balance and network morphology[164]. An unbalance in l-/s-forms ratio with more abundant s-forms than physiological state causes fusion inhibition and mitochondrial network fragmentation. In turn, the two events trigger a mitochondrial quality control (MQC) process that marks mitochondrial dysfunctional fragments for removal by mitophagy. MQC is regulated by a group of factors that include PINK1/Parkin axis, known to be involved in Parkinson disease pathogenesis. Indeed, it

has been also reported that Parkin and OPA1 proteins expression is also linked through the ubiquitination of NF-kB essential modulator[165,166].

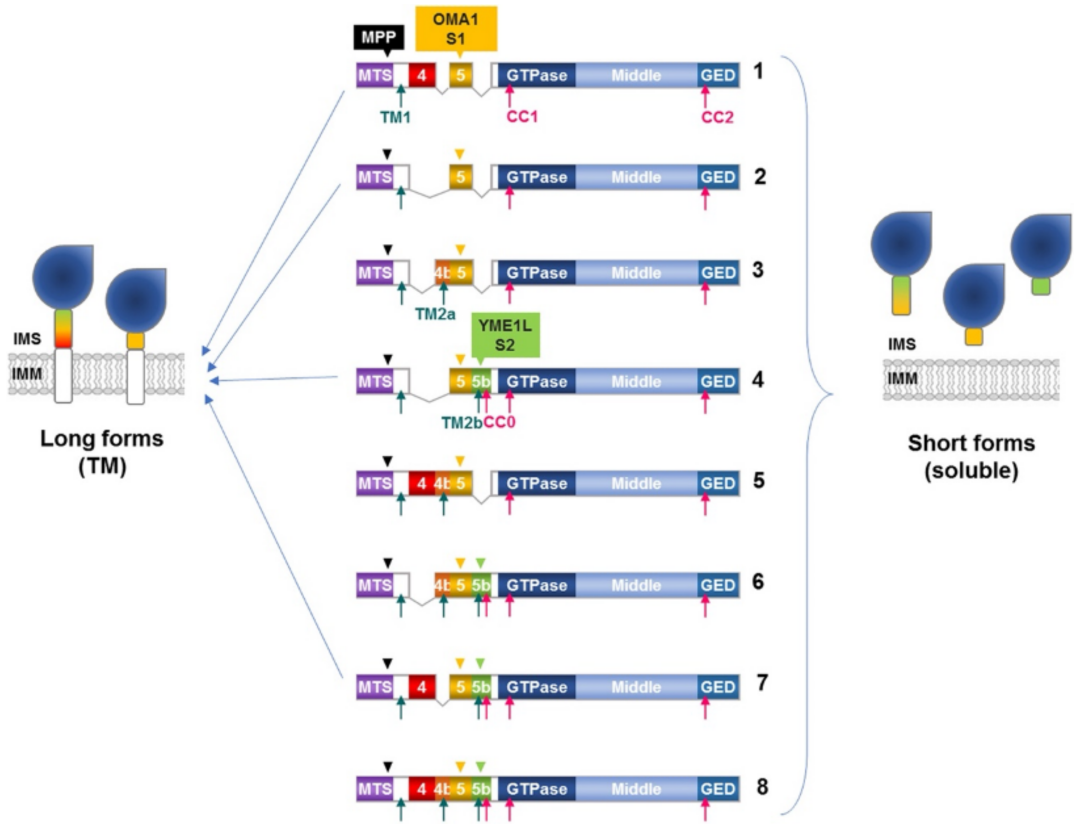


Figure 16 OPA1 protein isoforms[164]. Human *OPA1* gives rise to eight isoforms upon alternative splicing. After import in the IMM, the long forms are generated by cleavage of mitochondrial targeting sequence (MTS) by mitochondrial processing peptidase (MPP). Transmembrane domains (TM1, TM2a, TM2b) and coiled coil domains (CC0, CC1, CC2) are indicated with blue and red arrows respectively. All isoforms contain exon 5, including S1 cleavage site, while isoforms 4, 6, 7, 8 also contain S2 site. When these sites are cleaved by OMA1 and YME1L concerted activity, soluble short forms of OPA1 protein are produced.

Function

OPA1 protein has primary functions in mitochondrial network dynamics: together with Mitofusins MFN1 and MFN2, promotes mitochondrial fusion, a process associated with increased respiratory efficiency, co-operates with pro-fission DRP1 and DNM2 and contributes to mtDNA maintenance. OPA1 protein polymerization also preserves *cristae* morphogenesis, facilitating the activity of respiratory super-complexes[167]. It has a main role in controlling the apoptotic process as it is fundamental for the compartmentalization of cytochrome C, whose dysregulated release leads to cell death. Any of the eight OPA1 isoforms can support its three essential functions: energetics, structural and mtDNA maintenance, but a balance between long and short isoforms is reported to be a crucial requirement for a full recovery of the mitochondrial network[164]. A much more complex

picture emerges when trying to fully characterize each OPA1 isoforms' function, in particular for the great variability among the different isoforms' expression among human tissues. To this purpose, a recent study carried out a molecular and biochemical analysis in OPA1-null cells where single splice forms were stably expressed to evaluate each one's effect[164]. It was reported that any OPA1 isoform maintains the physiological mtDNA content, restores *cristae* structure and preserves respiratory complexes organization, while mitochondrial network morphology, that appears completely fragmented in OPA1$^{-/-}$ cells, can be partially rescued by mRNA isoforms giving rise to both l- and s-forms[164]. Importantly, to achieve full recovery of mitochondrial network morphology, at least two OPA1 isoforms, with a specific balance of l- and s- isoforms were necessary, suggesting the need for a multiplicity of isoforms to flexibly shape mitochondrial dynamics as a response to diverse metabolic and stress conditions that may perturb cellular homeostasis[164].

Mitochondria play an essential role in cellular homeostasis processes such as organelle dynamics control, interaction with other organelles, apoptosis regulation, calcium homeostasis maintenance and autophagy, but, most importantly, they are key suppliers of cellular energy through oxidative phosphorylation (OXPHOS). In mitochondria, five enzymatic complexes cooperate in the process and lead to the synthesis of adenosine triphosphate (ATP), through electrons shuttling from complex I to IV. The energy is then conserved and used by complex V to catalyze ATP synthesis. Alteration of this complex multi-step process can cause a reduction of ATP synthesis and an increase of reactive oxidative species (ROS), inducing damage in the respiratory chain and activation of apoptotic pathways up to mtDNA mutations accumulation. All these molecular effects can lead to energy failure and eventually cell death. In cells, mitochondria can be found as discrete organelles or as part of a network. Their transition between these two phases depends on fusion and fission processes, that are regulated by a protein machinery guided by OPA1 protein. Transport and distribution of mitochondria require an interplay with cytoskeletal proteins whose alteration provokes metabolic imbalance. In most cases, mitochondrial dysfunctions lead to neurodegeneration, addressing particularly RGC cells. This cell type is hardly damaged by energy failure because they present narrower not myelinated axons, that imply the absence of saltatory conduction of action potentials, highly requiring energy supply from mitochondria clustering within unmyelinated retinal and prelaminar sectors and less abundant in the posterior part of lamina cribrosa.

4.4 Therapeutic approaches

At date, no therapy is available to cure DOA and there is currently no active clinical trial. Since the disorder leads to degeneration of the optic nerve and loss of retinal functionality, visual aids such as glasses or contacts would not help to improve vision impairment. The most promising therapeutic approaches aim to complement OPA1 defective function through gene therapy or mitochondrial dysfunctions-targeting drugs. DOA is a good candidate for gene therapy for several reasons: it is a monogenic disorder, and the eye is an excellent organ to target since it is small, compartmentalized and easily accessible. Furthermore, the eyes can be monitored by noninvasive approaches (electroretinography (ERG) and optical coherence tomography (OCT))[78,168]. A first pre-clinical Adeno-Associated Viral (AAV) vector-based gene therapy trial targeting OPA1 has been recently reported to be successful in a mouse model of the disease carrying the delTTAG mutation thanks to the introduction a functional copy of OPA1 gene[169]. Nevertheless, looking at possible limitations of this type of approach for DOA, a first issue could be represented by the limited capacity of AAV vectors, given that OPA1 gene is quite large in size. Also, a crucial point, that has not yet been fully elucidated, is which isoform of the OPA1 gene should be selected and used in humans to gain maximal phenotypic rescue in affected RGCs. Besides isoform 1 that has been used in mice[148,169,170], all eight isoforms, when overexpressed within physiological levels, have similar ability to rescue OPA1 essential functions (mtDNA, cristae and energetics), but not network dynamics functionality[148,164,168,170]. Therefore, the ideal therapy should not only finely restore OPA1 physiological protein amount, but it should also fulfill tissue-specific requisites such as the physiological balance between l- and s-forms. Other therapeutic options are represented by strategies aimed at increasing neuronal survival by using growth factors such as ciliary neurotrophic factor (CNTF), glia-cell derived neurotrophic factor (GDNF), brain-derived neurotrophic factor (BDNF) or antioxidant transcription factors such as NRF2[171]. Furthermore, the reduction of ATP synthesis peculiar to DOA patients' fibroblasts[172] and patient-derived iPSCs[173] suggests the use of electron donors and acceptors like coenzyme Q10 (CoQ10) and riboflavin as potential drugs, but CoQ10 efficacy has resulted limited to disorders characterized by its primary deficiency[168]. Idebenone and EPI-743, CoQ10 analogues, were demonstrated to be potentially more effective than CoQ10, but a randomized, placebo controlled trial on a mouse model of DOA harboring OPA1^{Q285STOP} mutation, reported limited therapeutic effect of idebenone on RGCs dendropathy[174]. Moreover, the increased ROS levels reported in OPA1 mutated cells[170] could suggest

using antioxidants such as vitamin C, E, B2, B3, B12, lipoic and folic acids to reduce toxic effects of OPA1 mutations, but these supplements administration is not supported by any evidence of efficacy.

Disease models

Cellular models

Patients' fibroblasts and lymphoblasts are extensively studied as OPA1 mutations pathophysiology model. The majority of clinical observations reported in affected patients has been confirmed in these cellular models: defective mitochondrial network dynamics, energetic metabolism, cristae structure maintenance and increased sensitivity to apoptosis stimuli, while depletion of mtDNA copy number has been reported in fibroblasts in a few cases of missense or compound heterozygous mutations[175]. Along with these alterations, an increase of ROS production, low levels of antioxidant enzymes and alteration of calcium uptake have been reported. Moreover, basal mitophagy resulted increased in OPA1 dominant-negative mutations, while haploinsufficiency seems to correlate with reduced mitochondrial turnover and autophagy[175].

Other cellular models take advantage of induced pluripotent stem cell (iPSC) technology to generate *in vitro* human models of the disorder from patient-derived primary cells. iPSC can be differentiated in specific cell types, allowing the study of pathophysiological effects of OPA1 mutations in specific tissues. Recent data report the generation of dopaminergic neurons carrying OPA1 haploinsufficient mutation from iPSCs derived from two patients of the same family that developed different phenotypes: DOA and DOA with syndromic Parkinsonism[149]. The two models showed reduced oxygen consumption rate (OCR), complex I levels and activity, while the only Parkinsonism model also presented mitochondrial fragmentation and increased OPA1 s-forms[149].

Mouse models

The most well established DOA animal models are three mouse lines with a heterozygous germline mutant *OPA1* allele[176–178] that show a mild, age-dependent RGCs dysfunction and loss, optic nerve degeneration and mild neuromuscular impairment. The first, B6;C3-*OPA1*$^{329\text{-}355del}$ mouse, has an in-frame deletion of 27 aminoacidic residues in the dynamin GTPase domain[176]. The second, B6;C3-*OPA1*^{Q285STOP} mouse, instead, has a truncated form of the protein[177]. Both models show a 50% OPA1 expression reduction and the homozygous condition is embryonically lethal. The third model is a knock-in mouse carrying *OPA1 c.2708_2711delTTAG* mutation, common in humans, on a C57BL16/J

background[178]. This model reports a 25% OPA1 protein reduction in brain, retina, and optic nerve and a 50% reduction in oxidative fibers and heart, with embryonic lethality when in homozygous condition, in line with the other two models. All three models show mild, age-dependent RGCs dysfunction and loss, and optic nerves degeneration. Autophagic elimination of fusion-impaired mitochondria was reported in B6;C3-*OPA1*^{Q285STOP} and *OPA1 c.2708_2711delTTAG* mice. Increased mitophagy was also reported in B6;C3-*OPA1*^{Q285STOP} mice and confirmed in mouse RGCs overexpressing mutated OPA1. B6;C3-*OPA1*$^{329\text{-}355del}$ also showed an unbalanced redox state, probably increasing mitochondrial ROS. All these disfunctions may lead to most RGCs death, while melanopsin-expressing RGCs are reported to survive, according to what observed in humans affected from mitochondrial optic neuropathies[175]. Aging impairment of cardiac function was also reported in all considered models, while in humans, it was reported for the first time in two patients with homozygous recessive *OPA1* mutation leading to encephalopathy and hypertrophic cardiomyopathy[179]. Altogether the three models show a phenotype in accord with the human disease. Their characterization reported mild neuromuscular impairment mirroring the clinical spectrum of the human disorder, ranging from DOA to DOA "plus". All these models are useful for drugs and therapies testing: *OPA1*^{delTTAG} mouse, in fact, has already been used to test OPA1 isoform 1 gene therapy, showing a reduction of RGCs degeneration[169].

Materials and methods

Oligonucleotides

The complete list of oligonucleotides used for quantitative real-time PCR experiments and RT assays is included in Supplementary Table 1.

Plasmids

Complete list of plasmids is reported in Supplementary Table 2. For *EGFP* mRNA and miniSINEUP-GFP WT expression I used a plasmid vector derived from pEGFP-C2 vector (Clontech) previously described in[68], substituting miniSINEUP-FXN with miniSINEUP-GFP. miniSINEUP-DJ1 plasmid was generated based on SINEUP-DJ1 and miniSINEUP-GDNF, previously described[63,67]. ASUchl1 expressing plasmid was previously described[25]. m^6A sites miniSINEUP mutants were all synthetized by commercial preparation service (GeneScript).

miniSINEUP-OPA1 plasmids were generated using miniSINEUP DJ1 as backbone, replacing DJ1 BD with BDs designed in antisense orientation to target a common region between all eight OPA1 isoforms, targeting both the first and the second in frame AUG with longer (-40/+4) or shorter (-14/+4) overlapping regions[67,68]. The constructs were were all synthesized by commercial preparation service (GeneScript). nanoSINEUP-OPA1 plasmid was generated replacing the invSINEB2 ED with nanoED (64-92) from nanoSINEUP-GFP. For *EGFP* mRNA and miniSINEUP-GFP expression I used a plasmid vector derived from pEGFP-C2 vector (Clontech) previously described in [68] , substituting miniSINEUP-FXN with miniSINEUP-GFP. Shorter EDs were gene synthesised by commercial preparation service (GeneScript) and cloned in the same plasmid backbone.

Lentiviral backbone plasmids

Lentiviral plasmids used in this study have been optimized using pLV[Exp]-CMV>mCherry from Vector Builder. Starting from this, expression cassettes conformation was modified to have optimal conformation for lentiviral particles packaging. Ctrl plasmid harboring ΔBD, miniSINEUP-OPA1 and nanoSINEUP-OPA1, containing -14/+4 M1-targeting BD selected as most effective from HEK293T/17 screening, were generated by Gibson Cloning from pLKO-based vectors previously produced by cloning. In the optimized LV backbone plasmid hPGK promoter drives

TurboRFP reporter gene expression, while CAG promoter drives SINEUP RNAs expression.

ASO-SINEUPs

nanoSINEUP-OPA1, nanoSINEUP-GFP and femtoSINEUP-GFP RNA oligo were obtained by commercial preparation service with 2'OMe-A modification to achieve maximum purity and efficiency of modification incorporation (IDT). 2'OMe-A RNA modification was previously reported to guarantee optimal SINEUP RNA functionality[101].

Cell lines

MN9D cells were obtained from M. J. Zigmond and maintained in culture with High Glucose Dulbecco's Modified Eagle Medium (Thermo Fisher, Cat. No. 41965069) supplemented with 10% fetal bovine serum (Thermo Fisher Cat. No. 10270106) and 1% antibiotics (penicillin/streptomycin).

HEK 293T/17 were obtained from ATCC® (Cat. No. CRL-11268™) and maintained in culture with High Glucose Dulbecco's Modified Eagle Medium (Thermo Fisher, Cat. No. 41965069) supplemented with 10% fetal bovine serum (Thermo Fisher Cat. No. 10270106) and 1% antibiotics (penicillin/streptomycin), as suggested by the vendor.

Sh-Ctrl and Sh-METTL3 stably transfected A549 cells were maintained in culture with High Glucose DMEM (Thermo Fisher, Cat. No. 41965069), supplemented with 10% fetal bovine serum (Thermo Fisher Cat. No. 10270106) and 1% antibiotics (penicillin/streptomycin). To induce ShRNA expression, cells were treated with 2 ng/mL doxycycline (Sigma, Cat. No. D9891) every other day and transfected on day 7 of induction in all experiments.

All human fibroblasts lines (Table 2) were maintained in culture with High Glucose or Galactose supplemented Dulbecco's modified Eagle's medium (DMEM) (Gibco by Life Technologies, Cat. No. 41090-028) supplemented with 10% fetal bovine serum (Sigma, Cat. No. F2442) and 1% antibiotics (penicillin/streptomycin). All lines were obtained from the laboratory of Prof. Valerio Carelli (University of Bologna).

Cell line	*Sex and age*	*Clinical state*	*Mutation*
Ctrl 6	Male, 50 yo	Unaffected	-
Ctrl 9	Female, 40 yo	Unaffected	-
F40D	Male, 55 yo	DOA affected	3q28 c.703 C>T p.R235X Ex.7
F171	Female, 27 yo	DOA affected	3q28 c.2823_26delAGTTp.K941fsX966

Table 2 List of patients' derived fibroblasts lines. Complete list of fibroblasts lines used in this study with detailed description of sex, age, clinical state and OPA1 mutation.

Plasmid DNA and RNA oligo transfections

MN9D, A549, HEK293T/17, Neuro2a and C8-D1A astrocyte tipe I cells were plated in 6 well-plates or 150 mm dishes (for m^6A-RIP experiments and nucleo-cytoplasm fractionation) and transfected respectively with 1 ug or 16 ug of control or miniSINEUP encoding plasmids or 7 pmol RNA oligo using Lipofectamine 2000 (Thermo Fisher Scientific, Cat. No. 11668019) and following manufacturer's instructions. Cells were harvested 48 hours after transfections. For SINEUP activity experiments, RNA and proteins were obtained from the same transfection in each biological replicate.

Stable transduction of DOA patients' fibroblasts

Patients' fibroblasts were used to establish constitutively expressing Ctrl, miniSINEUP-OPA1 and nanoSINEUP-OPA1 cell lines. $2x10^5$ were transduced with constitutive lentiviral vectors previously described at MOI 10. 72h after transduction cells were trypsinized to detach and washed 2X with PBS. Cells were then FACS analyzed and TurboRFP positive cells were sorted and put back in culture. A second round of transduction was performed 7 days after the first on sorted cells. 72h after transduction a second sorting for TurboRFP positive cells was performed to ensure isolation of stably transduced cells. Cell lines were then expanded to perform functional assays.

Western Blot

Cells were lysed in radioimmunoprecipitation assay (RIPA) buffer with the addition of protease inhibitor cocktail (Sigma-Aldrich, Cat. No. P83490), briefly sonicated, and boiled with 1X Laemmli Buffer for 5 min at 95°C. 5 µg of total lysate were resolved by 10% or 4-20% SDS-PAGE TGX pre-cast gels (Bio-Rad) and transferred to nitrocellulose membrane using Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were blocked with 5% non-fat dry milk in TBS/0.1% Tween 20 and incubated with the following

indicated primary and secondary antibodies: anti-β-actin 1:15000 (Sigma Aldrich, Cat. No. A2066), anti-Uchl1 1:1000 (Cell Signaling, Cat. No. 3524S), anti-DJ1 1:8000 (Enzo Lifesciences, Cat. No. ADI-KAM-SA100-E), anti-OPA1 1:1000 (BD Bioscience, Cat 612606). The antibody against OPA1 detects at least 5 different isoforms of the protein, with an apparent molecular weight ranging from approximately 80 to 100 kDa. Proteins of interest were visualized with the SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific, Cat. No. 34579). Western blotting images were acquired with ChemiDoc MP Imaging System (Bio-Rad), and band intensity was calculated using ImageJ Software.

RNA extraction, Retro-transcription and qRT-PCR Real-time

Total RNA was extracted from cells using RNeasy Mini Kit (Qiagen, Cat. No. 74106) following the manufacturer's instructions. RNA samples were subjected to DNase treatment (Qiagen, Cat. No. 79254) to avoid plasmid DNA contamination. A total of 500 ng RNA was subjected to retro-transcription using iScript cDNA Synthesis Kit (Bio-Rad, Cat. No. 1708890), according to the manufacturer's instructions. qRT-PCR was carried out using SYBR green fluorescent dye (SsoAdvanced Universal SYBR Green Supermix, Bio-Rad, Cat. No. 1725271) and CFX96 Real time PCR System (Bio-Rad). The reactions were performed on diluted cDNA (10 ng). Human and mouse glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were used as the normalizing controls respectively for HEK293T/17 and patients' fibroblasts or for N2A and astrocytes in all qRT-PCR experiments. The amplified transcripts were quantified using the comparative Ct method, and the differences in gene expression were presented as normalized fold expression with the ΔΔCt method.

Methyl-RNA immunoprecipitation (m^6A-RIP)

m^6A-RIP was performed as previously described with some modifications[116]. Briefly, cells were harvested 48h post-transfection and total RNA was extracted with QIAzol reagent or RNeasy Mini Kit (Qiagen, Cat. No. 74106) and DNA contamination was removed by treatment with DNaseI following manufacturer's instructions. Total RNA extract, diluted with IP buffer (10 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1 % NP-40, RNase inhibitor supplemented) was then incubated with m^6A (SySy, Cat. No. 202111) with 1:10 ug ratio between antibodies and RNA for 2 hours at 4°C on a rotating wheel. The mixture was

then immunoprecipitated with G-coupled dynabeads (Thermo Fisher, Cat. No. 10003D) at 4°C for additional 2 hours. Beads were washed 5 times with IP buffer and resuspendend in QIAzol reagent and RNA was extracted according to manufacturer's protocol and analysed by RT-qPCR Real-time. Normal mouse IgG antibody (SantaCruz, Cat. No. 2025) and beads-only samples were used as negative controls.

In vitro transcription

For production of unmodified miniSINEUP RNA, synthetic double stranded DNA miniSINEUP template was cloned downstream a T7 or SP6 promoter. 1.4 ug of linearized DNA template was used to transcribe and purify miniSINEUP RNAs with MEGAscript T7 or SP6 Transcription Kit (Thermo Fisher Scientific, Cat. No. AM1333 or AM1330) following manufacturer's protocol.

Nanopore targeted direct RNA sequencing

RNA sequencing was performed following Oxford Nanopore Technologies (Oxford, UK) instruction, on FLO-MIN106D flowcells (R9.4.1 chemistry) and direct-RNA sequencing kit (SQK-RNA002). For library preparation 2 ug of total RNA from each replicate were used with custom reverse transcription adapters complementary to the 3' end of miniSINEUP-DJ1 RNA following ONT sequence specific DRS protocol. (DSS_9081_v2_revM_14Aug2019). Total RNA with 0.01 % unmodified IVT miniSINEUP-DJ1 RNA spike-in was used as negative control. The amount of IVT spiked-in RNA to be added was determined by comparison with expression levels of transfected miniSINEUP-DJ1 through qRT-PCR Real-time in order to maintain the amount of reads within the same range for each sample and replicate.

I compared the differences in electrical signal between 1) transfected *versus* IVT-spiked samples and 2) transfected WT *versus* METTL3 Knock-Down Cells using xPore (https://doi.org/10.1038/s41587-021-00949-w). The putative m^6A sites were extracted in a stringent way by intersecting the statistically significant positions identified by both comparisons and retaining those contained within a DRACH motif.

Subcellular fractionation

For subcellular fractionation experiments cells were transfected as previously described. Nucleo-cytoplasmic fractionation was performed as previously described[180]. Nucleus and

cytoplasmic RNAs were extracted using RNeasy Mini Kit (Qiagen, Cat. No. 74106) with DNase I treatment to remove DNA contamination.

The purity of nuclear and cytoplasmic fractions was confirmed by qRT-PCR as on pre-ribosomal RNA 45S and GAPDH and/or CytB respectively. qRT-PCR reactions were performed as previously described.

Polysome fractionation

Polysome fractionation was performed as previously described[181]. Briefly, A549 ShCtrl and ShMETTL3 were plated in 15 cm plates and transfected as previously reported after 7 days of doxycycline induction with control, miniSINEUP-DJ1 WT or miniSINEUP-DJ1 A46U;AAA109-111UUU vectors. 48 hours after transfection, cells were incubated with 0.1 mg/mL cycloheximide (CHX) for 10 minutes at 37°C. Cells were then washed with 0.1 mg/mL CHX-supplemented PBS and harvested by scraping. Collected cells were centrifuged at 400 xg for 5 minutes at 4°C. Cell pellets were resuspended in 400 µL of ice-cold lysis buffer (20 mM Tris-HCl pH 7.5, 100mM KCl, 5 mM $MgCl_2$, 0.5% Nonidet P-40) supplemented with 0.1 mg/mL CHX and RNase inhibitors. The cell lysate was incubated for 15 min on ice followed by centrifugation at 16000 × *g* at 4°C for 10 min to separate the nuclei. The cytoplasmic lysate was layered onto a 15–50% sucrose gradient and centrifuged in an SW41Ti Beckman rotor at 41 000 × *g* at 4°C for 3 h. The sucrose gradient was separated into 29 fractions calculated by Triax flow cell (Biocomp). 527µL of each fraction was then used to extract and analyze by Sodium Acetate overnight precipitation. Briefly, 170 µL 3M Sodium Acetate pH 5.3 was added to each fraction together with 1 mL of 100% EtOH, 3 µL Glycoblue co-precipitant (Thermo Fisher Scientific, Ct. No. AM9515) and 300 µg IVT GFP spike-in, used for qRT-PCR Real-time normalization of target RNA expression levels. The solution was incubated overnight at -20°C for precipitation. The following day samples were centrifuged at 13000 xg for 30 minutes at 4°C. Pellets for Free, 40S and 60S and three fractions each for 80S, light and heavy polysomes were pulled together at this point, washing with 1 mL 100% EtOH. Samples were then centrifuged 13000 xg for 15 minutes at 4°C and EtOH was removed. Pellets were then resuspended in H_2O with subsequent cleanup using RNeasy Mini Kit (Qiagen, Cat. No. 74106) and DNase I treatment to remove DNA contamination. As a control, *GAPDH* mRNA analysis was performed, according to previous publications.

IVT GFP RNA spiked-in was used to normalize for RNA precipitation efficiency. qRT-PCR reactions were performed as previously described.

Functional role of m^{6}A-modification in SINEUP RNA activity.

1. Results

1.1 Natural *AS Uchl1* lncRNA and synthetic miniSINEUP-DJ1 RNAs are m^{6}A-methylated.

We previously reported that the natural SINEUP *AS Uchl1* activity is triggered by inhibition of cap-dependent translation, which can be induced by several stress conditions, such as depletion of mTOR pathway activity. This causes *AS Uchl1* RNA shuttling from the nucleus to the cytoplasm, with consequent increase of *Uchl1* mRNA association to heavy polysomes and UCHL1 protein up-regulation[25,64]. ncRNAs, as well as mRNAs, were previously reported to be extensively m^{6}A-methylated with a wide range of functional effects derived from this modification. Among them, many evidences support a relevant involvement of m^{6}A methylation of RNAs in transcription and translation regulation[135,139] as well as in RNA subcellular localization[182]. Since SINEUPs' mechanism of action is still not fully known, I hypothesized that m^{6}A modification could play a functional role in SINEUP activity. I therefore identified natural *AS Uchl1* and synthetic miniSINEUP-DJ1 as representative SINEUP RNAs for further studies. Indeed, while *AS Uchl1* is a natural murine SINEUP, miniSINEUP-DJ1 is a synthetically designed shorter RNA targeting endogenous human *DJ1/Park7*. miniSINEUPs are exclusively composed of an invSINEB2 element, acting as ED, combined with an overlapping sequence specific for each target mRNA, the BD. miniSINEUPs have been successfully applied to cellular and animal disease models[67,68,99], as they show a comparable efficiency with a reduced size, compared to full-length *AS Uchl1* and synthetic SINEUPs, which is a major advantage towards the development of an RNA-based therapeutic molecule.

By performing a methyl-RNA immunoprecipitation (m^{6}A-RIP) in untreated MN9D cells, I found that endogenous *AS Uchl1* was modified in physiological conditions (Figure 17A), compared to an unmodified *in vitro* transcribed (IVT) spiked-in RNA encoding a portion of *EGFP* mRNA, used as negative control. I also performed an m^{6}A-RIP on overexpressed *AS Uchl1*, when post-transcriptional protein translation up-regulation was active, as confirmed by Western Blot (WB) and qRT-PCR analysis of UCHL1 protein level and *Uchl1* mRNA expression level analysis (Supplementary Figure 1A-D), and reported to be preferentially localized in the cytoplasm[61]. I found an enrichment of overexpressed *AS Uchl1* lncRNA in m^{6}A antibody-immunoprecipitated RNA comparable to what was observed for endogenous *AS Uchl1* (Figure 17A). To verify the presence of

m^{6}A modification on synthetic miniSINEUP RNA, I performed an m^{6}A-RIP on miniSINEUP-DJ1-transfected A549 cells. By qRT-PCR analysis I detected a significant enrichment level of miniSINEUP-DJ1 compared to the negative control proving the presence of m^{6}A sites (Figure 17B).

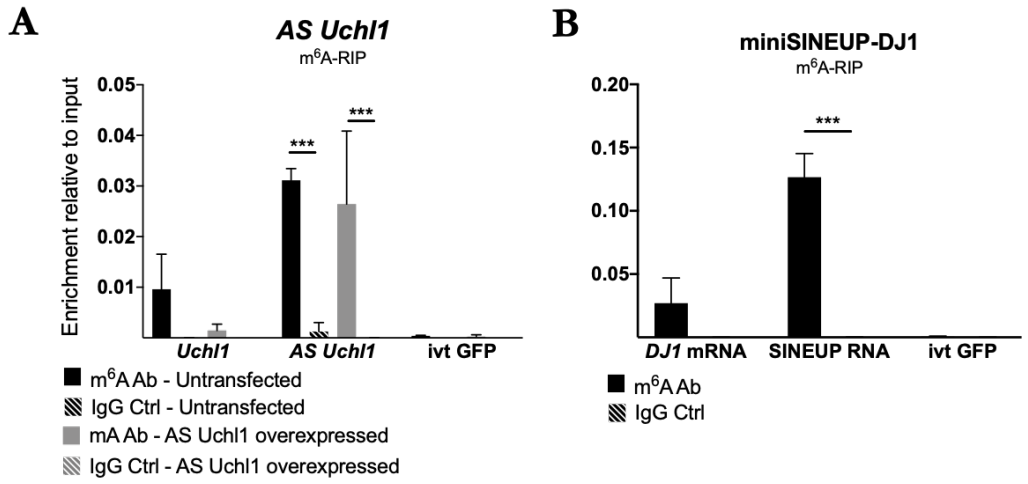


Figure 17 m^{6}A-RIP qRT-PCR. **A)** *Uchl1* mRNA and AS Uchl1 RNA relative enrichment in m^{6}A immunoprecipitated RNA from untransfected and AS Uchl1-overexpressing MN9D cells. ivt GFP spiked-in and IgG Ctrl samples (striped columns) were used as negative controls. **B)** *DJ1* mRNA and miniSINEUP-DJ1 RNA relative enrichment in m^{6}A immunoprecipitated RNA from untransfected and AS miniSINEUP-DJ1 transfected A549 cells. ivt GFP spiked-in and IgG Ctrl samples (striped columns) were used as negative controls. Data are expressed as enrichment relative to input and indicate mean ± SEM from three independent experiments. p values are calculated by Two-way ANOVA follows by Sidak multiple comparison (***, p<0,001).

To evaluate whether m^{6}A modification of synthetic miniSINEUPs was not exclusive to A549 cells, I also performed the same RNA immunoprecipitations on HEK293T miniSINEUPs-transfected cells, confirming miniSINEUP RNA methylation is not a cell-type dependent feature of synthetic miniSINEUPs (Supplementary Figure 1I). This result also confirmed that m^{6}A-methylation is not exclusive to *AS Uchl1* lncRNA and is conserved in miniSINEUPs shorter ncRNAs as a common feature between a natural lncRNA found in a murine system and the synthetic one directed against human *DJ1* mRNA. To confirm miniSINEUP-DJ1 activity in A549 cell line, cells were transfected with miniSINEUP-DJ1 or a miniSINEUP lacking the BD (miniSINEUP-ΔBD) as a negative control, as previously reported[68]. As expected, miniSINEUP-DJ1 was able to induce around 1.5 fold increase of DJ1 protein compared to the control, as assessed by Western Blot analysis (Supplementary Figure 1E). Both *DJ1* mRNA (Supplementary Figure 1G) and miniSINEUP RNA levels (Supplementary Figure 1H) were analyzed by

qRT-PCR Real-time, confirming no change in *DJ1* mRNA level upon miniSINEUP-DJ1 expression and therefore miniSINEUP's post-transcriptional activity.

1.2 Identification of SINEUP lncRNAs m^{6}A methylation sites.

I applied the m^{6}A prediction score algorithm (SRAMP)[183] to identify m^{6}A DRACH consensus sites within natural *AS Uchl1* transcript. m^{6}A consensus site has been deeply investigated and is now well established as D = G, A, or U; R = G or A; and H = C, A, or U. The analysis identified nine putative consensus sites with different degrees of confidence along *AS Uchl1* full-length sequence: two in the overlapping BD region, one adjacent to a partial Alu sequence, one in the invSINEB2 element, and other five in the downstream region. I then manually annotated other three weaker putative consensus sites: 407, 425, 455, within the invSINEB2 element (Figure 18 and Supplementary Figure 1J).

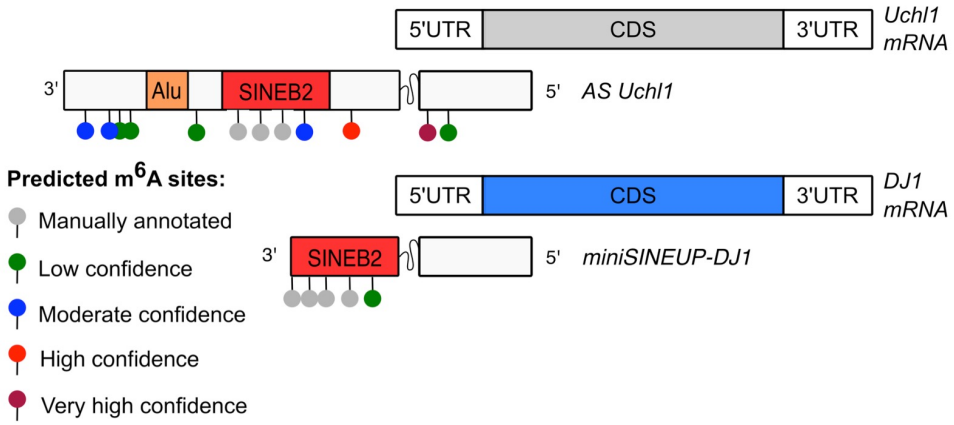


Figure 18 Annotation of predicted m^{6}A sites in AS Uchl1 and miniSINEUP-DJ1 RNAs. SRAMP algorithm was interrogated (https://www.cuilab.cn/sramp) resulting DRACH consensus sites are annotated with confidence level reported in the legend. Additional consensus sites were manually annotated (grey).

By applying SRAMP algorithm to the miniSINEUP-DJ1, one putative site was identified in position A46. I then manually annotated other three weaker putative consensus sites: A61, A81 and A111 (Figure 18 and Supplementary Figure 1K). Most of the common methods used to map post-transcriptional modifications are currently based on RNA immunoprecipitation[117], chemoenzymatic substitution of the modified base[122,184] or detection through specific reverse transcriptase (RT) enzyme efficiency alteration in correspondence of m^{6}A modification sites[124,125]. All these methods suffer from low sensitivity and specificity and errors introduced by complex and long protocols. In the effort to overcome these issues and to be able to map m^{6}A methylation sites on miniSINEUP RNA, I used a combined approach relying on Nanopore targeted direct

RNA sequencing and a recently developed RT-qPCR based method that involves the use of *BstI* enzyme for reverse transcription of putative m^{6}A residues[125]. Indeed, it is well established that modified nucleotides induce relevant signal deviations during the sequencing of a nucleic acid molecule with this technology, allowing the mapping of modification sites on both DNA and RNA molecules with a resolution that is very close to single nucleotide[127,130,185]. Since invSINEB2 is essential for SINEUP activity, the m^{6}A sites it contained were more likely to play a crucial role in SINEUP activity regulation. Indeed, deletion of other portions within *AS Uchl1* sequence have been proven to reach unaltered activity levels[25,63,77]. Here, I took advantage of A549 cells stably transduced with inducible ShCtrl or ShMETTL3 viral vector. To get reliable results, I compared the sequencing of IVT miniSINEUP-DJ1 RNA transfected in A549 ShCtrl cells to the same IVT RNA transfected in A549 ShMETTL3 derived from 3 biological replicates. As a non-modified control, I also used the very same IVT miniSINEUP-DJ1 RNA spiked in total RNA extract from A549 ShCtrl cells (Figure 19A). In this way, I were able to detect two relevant modification sites: A46 and A111 of the invSINEB2 element (Figure 19B). To validate Nanopore sequencing results, I took advantage of *BstI* enzyme reverse transcriptase (RT), whose efficiency was reported to be markedly reduced when used in combination with a primer adjacent to an m^{6}A residue. I performed the *BstI* RT followed by qPCR of the product to the very same RNA samples used for Nanopore sequencing, using four reverse transcription primers (A46+, A63+, A81+ and A111+) located adjacent to m^{6}A putative sites, and one primer with no m^{6}A consensus site nearby (-). After qRT-PCR Real-time of the RT product, I compared *BstI* reverse transcription efficiency with primers (+) to the one with primer (-) and I also compared it to RT products from MRT enzyme reactions, whose efficiency is not affected by the proximity of m^{6}A residues, performed with the same primers. I then applied the following formula to calculate relative m^{6}A level with qRT-PCR, as previously reported: Relative m^{6}A = $2^{-(Ctprimer(-)BstI-Ctprimer(-)MRT/Ctprimer(+)BstI-Ctprimer(+)MRT)}$. I confirmed the absence of m^{6}A sites in IVT miniSINEUP-DJ1 spiked-in RNA and IVT miniSINEUP-DJ1 transfected in ShMETTL3 cells (Figure 19C, E), while I observed a significant relative m^{6}A level at A46 and A111 sites in IVT miniSINEUP-DJ1 ShCtrl cells (Figure 19D). I also assessed METTL3 knock-down by qRT-PCR analysis of *METTL3* mRNA levels in the same samples (Figure 19F).

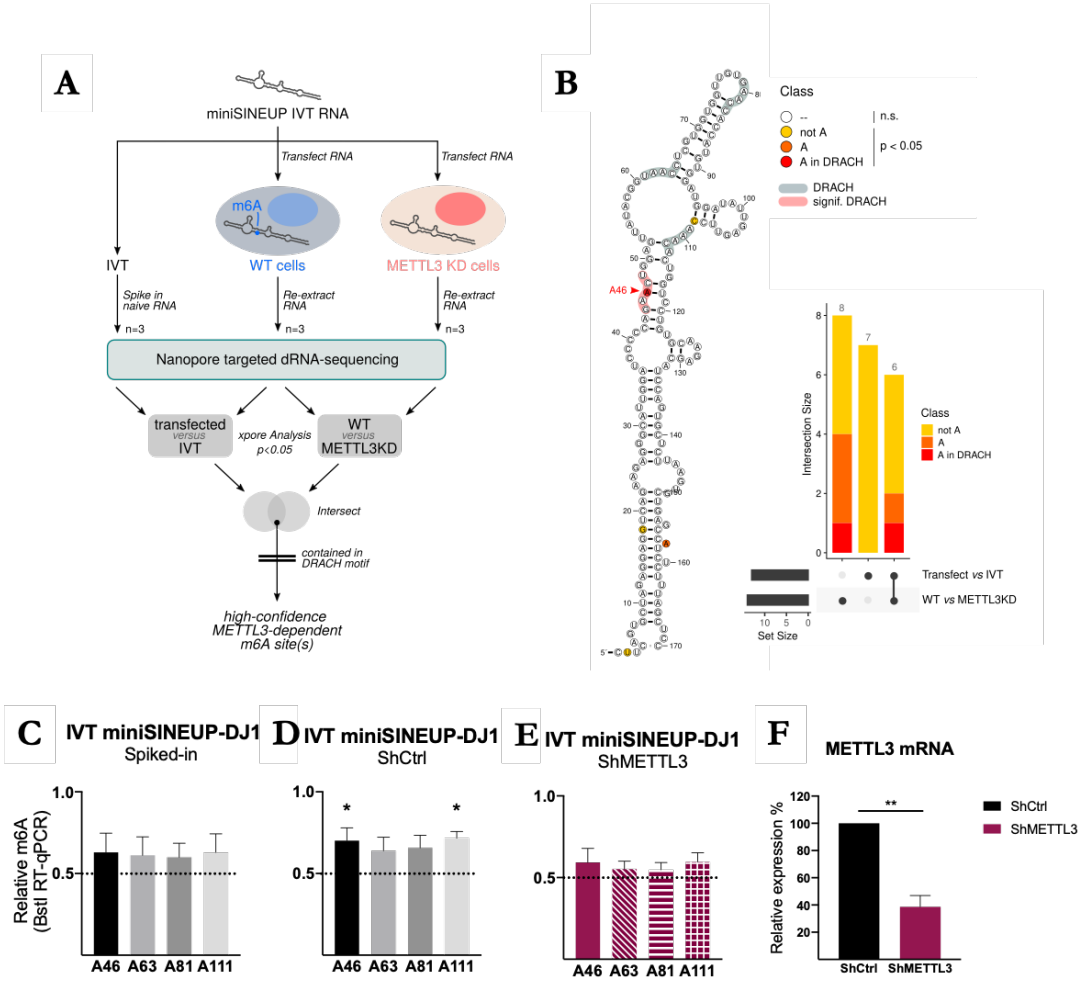


Figure 19 Nanopore direct RNA sequencing mapping of m⁶A sites in transfected IVT miniSINEUP-DJ1 and validation with RT assay. **A)** Experimental design **B)** Nanopore direct RNA sequencing analysis with resulting m⁶A sites annotated as colored nucleotides and DRACH consensus sites highlighted, as reported in the legend. **C-E)** m⁶A sites validation through *BstI* RT assay on Nanopore-analysed samples. Columns represent each putative site probability of harboring m⁶A modification as difference in retro-transcription efficiencies of *BstI* and *MRT* reverse transcriptase using the same primer adjacent to putative consensus site (Relative m⁶A <= 0.5 indicates absence of modification, Relative m⁶A > 0.5 indicates presence of m⁶A site). Data indicate mean ± SEM from three independent experiments. p values are calculated by One Sample t and Wilcoxon test, comparing to 0.5 control (*, p<0,05,). **C)** miniSINEUP-DJ1 RNA relative m⁶A quantification in IVT miniSINEUP-DJ1 RNA spiked in A549 ShCtrl total RNA extract. **D)** miniSINEUP-DJ1 RNA relative m⁶A quantification in IVT miniSINEUP-DJ1 transfected A549 ShMETTL3 cells. **E)** miniSINEUP-DJ1 RNA relative m⁶A quantification in IVT miniSINEUP-DJ1 transfected A549 ShMETTL3 cells. **F)** *METTL3* mRNA expression in ShCtrl and ShMETTL3 cells used for m⁶A mapping Nanopore direct RNA sequencing and BstI RT-qPCR analysis. Data indicate mean ± SEM from three independent experiments. p values are calculated by One Sample t and Wilcoxon test, comparing to 100% control (**, p<0,01).

The same method was then used to map plasmid-encoded miniSINEUP-DJ1 m⁶A sites in ShCtrl and ShMETTL3 cells (Figure 20A) and *AS Uchl1* m⁶A sites in MN9D cells (Figure 20B). For miniSINEUP-DJ1, I were able to confirm A46 and A111 as m⁶A sites previously identified with Nanopore direct RNA sequencing in ShCtrl cells (Figure 20A, 19B) while no putative m⁶A site was found modified in ShMETTL3 knock-down cells

(Figure 20A). In the case of *AS Uchl1* I observed a marked relative m^{6}A level in sites A275, A390 and A455, with the last two contained in the invSINEB2 element and corresponding to A46 and A111 (Figure 20B).

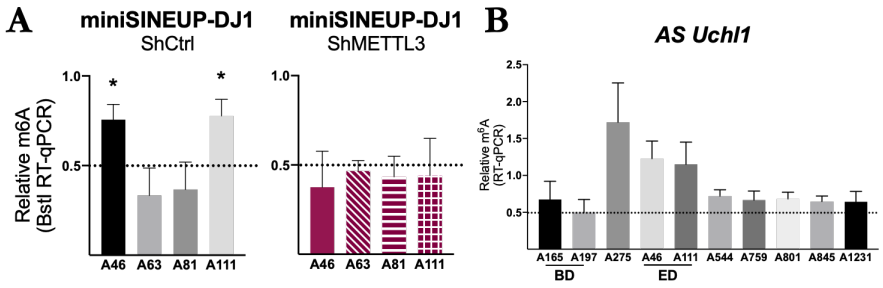


Figure 20 m^{6}A sites mapping in plasmid encoded miniSINEUP-DJ1 and AS Uchl1 RNA. m^{6}A sites mapping through *BstI* RT assay. Columns represent each putative site probability of harboring m^{6}A modification as difference in retro-transcription efficiencies of *BstI* and *MRT* reverse transcriptase using the same primer adjacent to putative modification site (Relative m^{6}A <= 0.5 indicates absence of modification, Relative m^{6}A > 0.5 indicates presence of m^{6}A site). **A)** miniSINEUP-DJ1 RNA relative m^{6}A quantification in plasmid-encoded miniSINEUP-DJ1 transfected in A549 ShCtrl (left, scale of grey) and ShMETTL3 (right, purple) cells. **D)** AS Uchl1 RNA relative m^{6}A quantification in annotated DRACH consensus sites. Data indicate mean ± SEM from at least three independent experiments. p values are calculated by One Sample t and Wilcoxon test, comparing to 0.5 control (*, p<0,05,).

1.3 METTL3 expression regulates synthetic miniSINEUPs activity without altering RNA subcellular distribution.

Given the effects of decreased METTL3 expression on m^{6}A sites detection by Nanopore sequencing, I carried out m^{6}A-RIP experiments on METTL3-depleted cells transfected with miniSINEUP-DJ1. A significantly lower level of enrichment of miniSINEUP-DJ1 RNA was observed upon ShMETTL3 expression induction, confirming METTL3 enzyme as the main writer of miniSINEUP RNA m^{6}A modification (Figure 21A). To evaluate the functional consequences of METTL3-dependent m^{6}A modification on SINEUP activity, I tested miniSINEUP-DJ1 activity in A549 METTL3 knock-down cells, compared to controls. Interestingly, while miniSINEUP-DJ1 activity was confirmed to reach around 1.6 fold increase in control cells, in ShMETTL3 cells DJ1 protein level decreased to around 0.5 fold compared to the negative control (Figure 21B), showing a dominant negative effect on endogenous DJ1 expression. qRT-PCR on *DJ1* mRNA confirmed an equal level in both ShCtrl and ShMETTL3 cells (Supplementary Figure 2A), and a comparable transfection efficiency of SINEUP plasmid (Supplementary Figure 2B). *METTL3* mRNA knock-down was also confirmed by qRT-PCR analysis (Figure 21C).

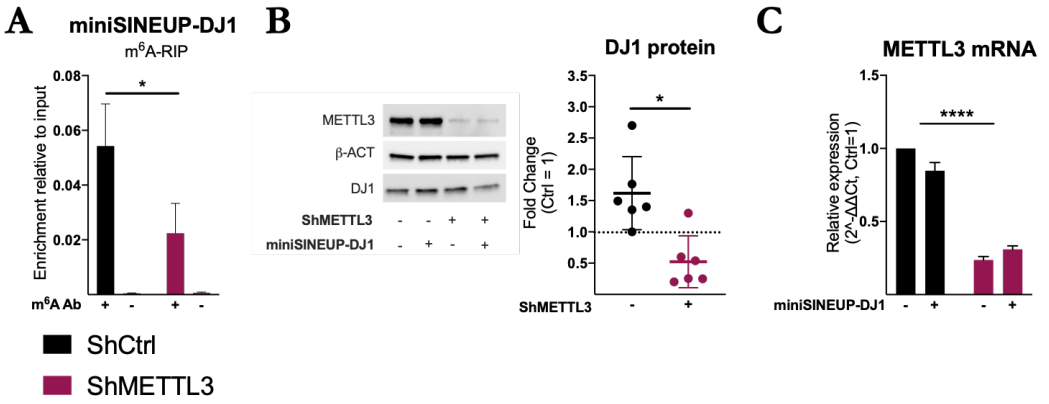


Figure 21 METTL3 enzyme is responsible for miniSINEUP-DJ1 m6A modification and regulates its activity. **A)** miniSINEUP-DJ1 enrichment level in m6A immunoprecipitated total RNA from ShCtrl (black) and ShMETTL3 (purple). IgG ctrl (-m6A Ab) and IVT GFP spike-in (Supplementary Figure x) were used as negative controls. Data are expressed as enrichment relative to input. **B)** miniSINEUP-DJ1 activity in Ctrl and METTL3 KD cells. Left: representative western blot analysis with METTL3 and DJ1 antibodies show protein levels in ShCtrl and ShMETTL3 cells. Right: summary of DJ1 protein levels in ShCtrl (black) and ShMETTL3 (purple) cells. Fold changes in DJ1 protein expression are relative to Ctrl plasmid transfected cells. **C)** *METTL3* mRNA levels relative quantification with qRT-PCR Real-time in samples described in Ctrl and miniSINEUP-DJ1 transfected ShCtrl and ShMETTL3 cells. Data indicate mean ± SEM from at least three independent experiments. p values are calculated by Two-way ANOVA follows by Sidak multiple comparison (*, p<0,05, ****, p<0,0001).

Since SINEUP RNA subcellular distribution was previously reported to be a key factor in the regulation of target mRNA translation[25,77] and m6A modification has been reported to influence subcellular RNA localization in several cases, I investigated whether a METTL3-dependent differential distribution of SINEUP RNA was responsible for the reduction in activity upon METTL3 depletion, analyzing nuclear and cytoplasmic RNA levels upon subcellular fractionation followed by qRT-PCR analysis (Figure 22). I observed no difference in the subcellular distribution between ShCtrl and ShMETTL3 knockdown cells, with ~ 20% of miniSINEUP-DJ1 RNA localized in the cell nucleus and ~80% in the cytoplasm, in accordance with what has been reported for other SINEUP RNAs[25,61,77]. No variation in *DJ1* mRNA subcellular distribution was also observed (~40% in the nucleus and ~60% in the cytoplasm).

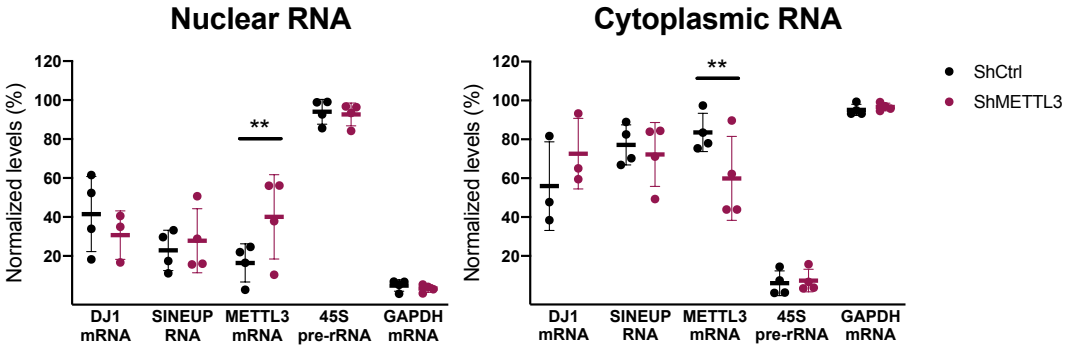


Figure 22 METTL3 enzyme depletion does not alter miniSINEUP-DJ1 nor target mRNA subcellular localization. Subcellular distribution of DJ1 mRNA, miniSINEUP-DJ1 RNA, METTL3 mRNA was analysed in ShCtrl (black) and ShMETTL3 (purple). Nucleocytoplasmic fractionation was performed and RNA levels in nuclear (left) and cytoplasmic (right) fractions were quantified by qRT-PCR Real-time. Purity of cellular fractions was checked by monitoring GAPDH and 45S pre-rRNA levels. Data are expressed as percentages of total RNA and derive from three independent experiments. Data indicate mean ± SEM from four independent experiments. p values are calculated by Two-way ANOVA follows by Sidak multiple comparison (**, p<0,01).

Taken together, these results show that m^{6}A modification of miniSINEUP-DJ1 RNA is deposed by METTL3 m^{6}A writer whose down-regulation negatively regulates SINEUP activity without altering its subcellular distribution.

1.4 m^{6}A methylation sites regulate miniSINEUP-DJ1 activity.

With the aim of gaining a more detailed insight on the role of m^{6}A modification in SINEUP activity, I mutated m^{6}A sites in the invSINEB2 element of synthetic SINEUPs as identified through Nanopore targeted direct RNA sequencing and *BstI*-RT-qPCR. Since I previously reported that there is a strong correlation between structure and functionality of SINEUP invSINEB2 ED and that this is a highly structured molecule[73,75], I engineered a point mutation in A46 m^{6}A site, by substitution of the A with a U, in the effort to perturb RNA secondary structure as little as possible. In the case of A111 m^{6}A site, instead, I had to perform a 3-nucleotide mutation to avoid possible formation of any cryptic consensus sites, substituting AAA109-111 with UUU sequence. I then transfected miniSINEUP-DJ1 WT and its mutants A46U, AAA109-111UUU and A46U;AAA109-111UUU in A549 cells and I assessed RNA m^{6}A-modification level through m^{6}A-RIP and target protein expression through WB. Following m^{6}A-RIP all SINEUP RNA were significantly less enriched in the immunoprecipitated RNA, which indicates a successful killing of all m^{6}A modified sites (Figure 23A). Interestingly, while miniSINEUP-DJ1 WT confirmed its capability to induce DJ1 protein expression increase to around 1.5 fold, upon A46 and AAA109-111UUU single m^{6}A sites mutation, I did not observe any

translation upregulation. Moreover, when the double m^{6}A sites mutant miniSINEUP-DJ1 A46;AAA109-111UUU was transfected, a decrease in DJ1 protein levels to around 0.5 fold compared to the control, was observed which is in line with miniSINEUP-DJ1 WT activity in ShMETTL3 knock-down cells (Figure 23B) and with a dominant negative effects on endogenous DJ1 protein levels. *DJ1* mRNA and miniSINEUP-DJ1s RNA expression were analyzed by qRT-PCR to confirm miniSINEUP's post-transcriptional activity and unaltered expression of m^{6}A sites mutants (Supplementary Figure 2F-G).

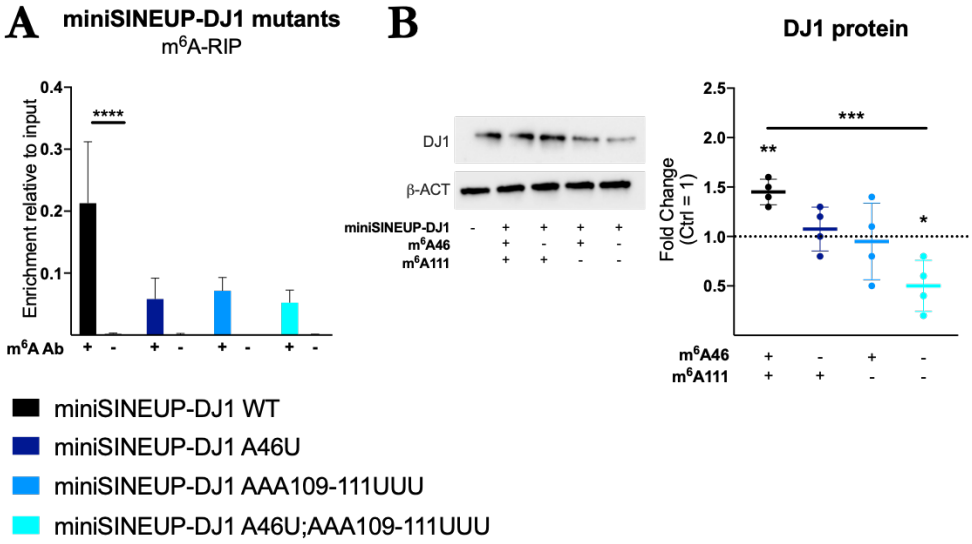


Figure 23 m^{6}A methylation sites regulate miniSINEUP-DJ1 activity. Mutation analysis was performed by selective removal of m^{6}A consensus sites A46 and A111, previously identified as modified residues, from miniSINEUP ED sequence. miniSINEUP-DJ1 WT, single m^{6}A site mutants A46U and AAA109-111UUU and double m^{6}A sites mutant A46U;AAA109-111UUU mutants were transfected in A549 cells. **A)** Total RNA was then m^{6}A immunoprecipitated and miniSINEUP RNAs' enrichment was analysed with qRT-PCR Real-time. Data are expressed as enrichment relative to input and indicate mean ± SEM from three independent experiments. p values are calculated by Two-way ANOVA follows by Sidak multiple comparison (****, p<0,0001). **B)** miniSINEUP-DJ1 mutants' activity. Fold changes in DJ1 protein expression are relative to Ctrl plasmid transfected cells. Left: representative western blot analysis with DJ1 antibody. Right: summary of DJ1 protein levels in miniSINEUP-DJ1 WT and mutants transfected cells. Data indicate mean ± SEM from four independent experiments. p values are calculated by One sample t and Wilcoxon test for comparison with Control plasmid and One-way ANOVA followed by Dunnett's multiple comparison test (*, p<0,05, **, p<0,01, ***, p<0,001).

The reduction of activity was not due to an alteration in *DJ1* target mRNA nor in miniSINEUP-DJ1 RNA subcellular distribution as proved by qRT-PCR analysis of subcellular fractions of nuclear and cytosolic miniSINEUP-DJ1 mutants RNAs (Figure 24). In summary, the requirement of m^{6}A methylation for miniSINEUP-DJ1 activity was proved by two experimental strategies: upon depletion of the methyltransferase enzyme responsible for their deposition and by mutating the plasmid-encoded RNA in both the mapped modified sites.

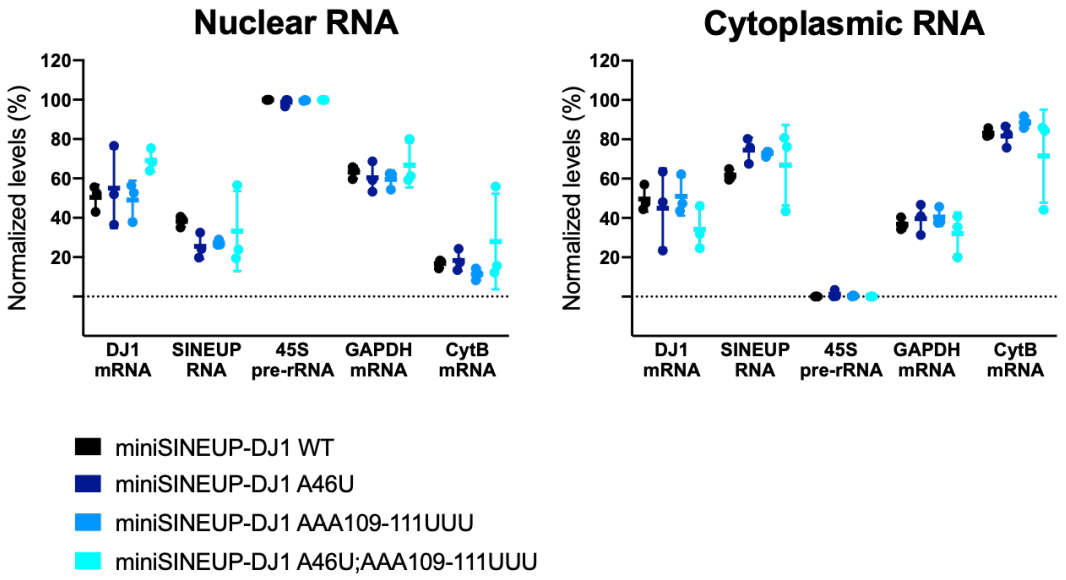


Figure 24 m⁶A sites mutation does not alter miniSINEUP-DJ1 nor target mRNA subcellular localization. Subcellular distribution of DJ1 mRNA and miniSINEUP-DJ1 WT and mutants RNA. Nucleocytoplasmic fractionation was performed and RNA levels in nuclear (left) and cytoplasmic (right) fractions were quantified by qRT-PCR Real-time. Purity of cellular fractions was checked by monitoring 45S pre-rRNA (nuclear) and GAPDH and Cytochrome B (cytoplasmic) levels. Data are expressed as percentages of total RNA and derive from three independent experiments. Data indicate mean ± SEM from three independent experiments. p values are calculated by Two-way ANOVA follows by Sidak multiple comparison.

1.5 miniSINEUP-DJ1 translation enhancing activity is impaired upon loss of m^{6}A modification.

Natural *AS Uchl1* activity has been previously described as post-transcriptional, reporting the translation enhancement of the target *Uchl1* mRNA as the result of its increased association to heavy polysomes[25]. More recently, the same mechanism was observed for a synthetic SINEUP-GFP, reporting a shift of *EGFP* mRNA to heavy polysome and a SINEUP RNA distribution gradually decreasing from free RNA fractions to heavy polysome ones[77].

I therefore applied ribosome fractionation analysis to further dissect the mechanism causing the reduction of DJ1 protein expression upon miniSINEUP-DJ1 m^{6}A sites removal by both METTL3 depletion and synthetic mutation. To this purpose, *DJ1* target mRNA and miniSINEUP RNA distribution were analyzed using as a normalizing factor the IVT EGFP fragment previously used as m^{6}A-RIP negative control, as well as *GAPDH* mRNA and as a reference. Ribosome fractionation was first performed on A549 ShCtrl and ShMETTL3 untransfected cells using a 15-50% sucrose gradient separated into 13 fractions to compare endogenous *DJ1* and *GAPDH* mRNAs distribution in physiological condition (Supplementary Figure 3D-F). In these two conditions, a comparable profile

and a similar distribution for *DJ1* and *GAPDH* mRNA was observed. Indeed, the majority of *DJ1* and *GAPDH* mRNAs was found associated to heavy polysomes to comparable extent in ShCtrl and ShMETTL3 cells.

The same experiment was performed on A549 ShCtrl cells transfected with the control, miniSINEUP-DJ1 WT or miniSINEUP-DJ1 A46U;AAA109-111UUU plasmid and ShMETTL3 cells transfected with the control or miniSINEUP-DJ1 WT plasmid (Figure 25). In this experiment, to get a higher resolution of *DJ1* mRNA and SINEUP RNA distribution I analyzed all the fractions separately. In the control plasmid transfected samples, both in ShCtrl and ShMETTL3 cells, *DJ1* mRNA was found to be mostly co-localized with light polysomes (15% to 20% in each light polysome fraction, Figure 25C), with a minor shift of localization compared to untranfected cells (Supplementary Figure 3E). Importantly, upon miniSINEUP-DJ1 WT transfection in ShCtrl cells a significant increase of *DJ1* mRNA % in light polysomes fractions was detected, as expected for SINEUP activity (Figure 22C). On the other hand, in ShMETTL3 cells a marked shift of *DJ1* mRNA to non-actively translated fractions was observed: from ≈6% in ShCtrl miniSINEUP-DJ1 transfected cells to ≈13%, confirming what previously observed through western blot as DJ1 protein level decrease below control levels (Figure 25C). These results confirmed WB analysis results (Figure 22B). By analyzing SINEUP RNA a similar distribution was found across all samples, regardless the presence of DJ1 BD or *METTL3* expression, with a clear enrichment of SINEUP RNA in 40S and 60S fractions and depletion from polysomes (Figure 25D), in line with what previously reported[186].

I then analyzed *DJ1* mRNA and SINEUP RNA distribution upon miniSINEUP-DJ1 A46U;AAA109-111UUU transfection and, surprisingly, a significant accumulation of *DJ1* mRNA in 40S and 60S fractions compared to control and miniSINEUP-DJ1 transfected cells (Figure 25D) was detected, with a concomitant depletion from polysomes fractions (Figure 25B). A slightly different SINEUP RNA distribution upon m^6A sites mutation was observed, although not significant, in 40S fraction (Figure 25D). *DJ1* mRNA and SINEUP RNAs expressions were analyzed in total RNA to confirm comparable expression level (Supplementary Figure 3A). All together, these results unveil a m6A-dependent step in the molecular mechanism of SINEUP activity during translation. When SINEUP RNA is not appropriately m6A-modified, the sequestration of endogenous *DJ1 mRNA* to 40S and 60S ribosomal fractions by miniSINEUP-DJ1-A46U;AAA109-

111UUU, causes the inhibition of target mRNA translation and the consequent downregulation of endogenous DJ1 protein levels.

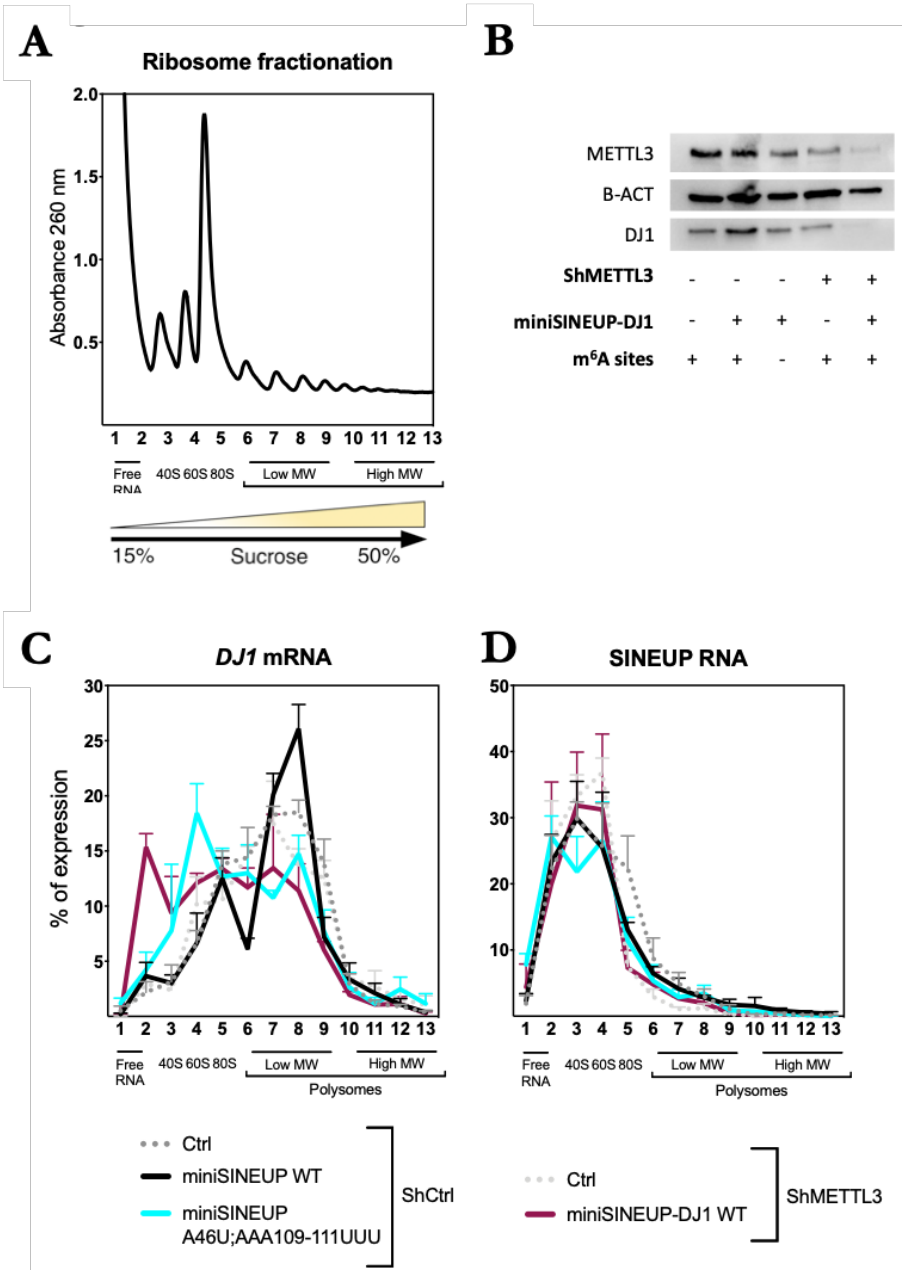


Figure 25 miniSINEUP-DJ1 translation enhancing activity is impaired upon loss of m⁶A modification. **A)** Representative ribosome fractionation profile with optical density 260 nm, obtained with a 15% to 50% sucrose gradient. RNA was extracted from single fractions. **B)** Representative western blot analysis of DJ1 and METTL3 protein level in total lysate of samples used for ribosome fractionation analysis. **C-D)** RNA distribution in polysome fractions from ShCtrl cells overexpressing Ctrl (dotted, dark grey), miniSINEUP-DJ1 WT (bold, black) and miniSINEUP-DJ1 double m⁶A sites mutants A46U;AAA109-111UUU (bold, light blue), and ShMETTL3 cells overexpressing Ctrl (dotted, light gray) and miniSINEUP-DJ1 WT (bold, purple)cells. **C)** DJ1 mRNA polysome profiling. **D)** SINEUP RNA polysome profiling. Data are expressed as percentages of total RNA in each fraction and derive from three independent experiments. Data indicate mean ± SEM from three independent experiments.

2. Discussion

We previously reported that natural *AS Uchl1* is enriched in the nucleus in physiological conditions, while, upon mTOR pathway inhibition with rapamycin administration, it is exported in the cytoplasm, where it triggers the cap-independent translation up-regulation of the overlapping *Uchl1* mRNA promoting ribosomal machinery recruitment and inducing *Uchl1* mRNA shift to heavy polysomes[25].

Recently, a new type of function for m^{6}A has been associated to mRNA translation through the direct binding of 5'UTR-contained m^{6}A to eIF3[139]. In this case, the modification was required to be located in the 5'UTR. Intriguingly, this m^{6}A-dependent translation initiation mechanism did not require eIF4E, the m^{7}G-containing mRNA cap-binding protein, thus defining a new model of cap-independent translation initiation, alternative to the well-established IRES model[139]. Another study recently reported m^{6}A modification as an important regulator of IRES-mediated translation in HCV virus life cycle[111]. While most studies investigate m^{6}A modification role in mRNAs, much less is known about their function in lncRNAs, with major attention focusing on cancer-related transcripts[110]. For example, in the case of human lncRNA MALAT1 (Metastasis Associated Lung Adenocarcinoma Transcript), one m^{6}A methylation site has been identified within a hairpin stem and it has been demonstrated to have a destabilizing effect on the transcript structure that makes the DRACH sequence opposing U-tract more accessible for RNA-binding proteins [187]. Furthermore, in the case of circRNAs hosting an IRES and a open reading frame, it has been recently reported that m^{6}A modifications may play a relevant role in their biogenesis and in their cap-independent translation [135].

Since the molecular mechanisms of SINEUP activity are still not fully understood, I hypothesized an involvement of m^{6}A modification acting in *trans*. Performing m^{6}A RNA immunoprecipitation on i. endogenous and overexpressed *AS Uchl1* targeting endogenous *Uchl1* mRNA in MN9D murine cells and on ii. synthetic miniSINEUP-DJ1 targeting endogenous *DJ1* mRNA in human A549 and HEK293T cells, I found that in all cases SINEUP RNAs were enriched in the immunoprecipitated fractions, suggesting m^{6}A modifications are a constant feature of active SINEUPs[67,68,99].

In the effort to further elucidate SINEUP activity, I used miniSINEUPs, as they are exclusively composed by the invSINEB2 as ED and the target-specific antisense sequence as BD. Using Nanopore targeted direct RNA sequencing, I found that m^{6}A modification sites are present along the invSINEB2 element in position A46 and A111 and validated these results with a reverse transcription approach. Using the latter, an additional

methylation site in position A275 was detected in *AS Uchl1* lncRNA, in the region between BD and ED that was previously demonstrated not to be essential for SINEUP activity[63,67,68]. Moreover, METTL3 was identified as the main responsible enzyme for m^6A deposition using an inducible knock-down cell system that allowed us to evaluate the effect of indirect m^6A depletion from miniSINEUP RNA. Interestingly, upon METTL3 knock-down induction, SINEUP RNA activity was abolished and, most importantly, it had a dominant negative effect on target mRNA translation. In parallel, by directly removing m^6A sites through sequence mutagenesis, a similar effect was observed, proving that m^6A methylation is playing a crucial role in SINEUP activity regulation.

Interestingly, SINEUP RNA is associated to 40S and 60S ribosomal subunit independently of m^6A since no changes were detected in SINEUP RNA ability to associate to ribosomal subunits upon removal of m^6A sites modifications with both METTL3 knock-down and mutagenesis. On the contrary, the specific increase of target mRNA translation required the presence of a selective BD and of the m^6A methylation of both A46 and A111 sites. The low association of SINEUP RNA to 80S and polysome fractions indicated that its activity took place only in the initial steps of translation, while, upon formation of 80S-target mRNA complex, SINEUP ncRNA was released, possibly for recycling or decay. Indeed, when SINEUP RNA-target mRNA complex is formed, as previously reported[186], it would not be possible for the mRNA to be translated as its 5'UTR and starting codon are paired with SINEUP RNA. This observation implicates the potential activity of an unknown RNA helicase (possibly m6A-dependent) to unwind and separate the two RNA species. Mutant SINEUP RNA indeed sequestered endogenous target mRNAs to the 40S and 60S fractions lowering their association to polysomes for active translation.

These results have two important and original implications: i. an m6A-dependent step is required for SINEUP activity at the ribosome; ii. in the absence of m6A, SINEUP RNAs possess dominant-negative activity reducing endogenous protein quantity of the target mRNA. In summary, this work has provided new cues on the molecular mechanism of SINEUP activity and for the development of new RNA-based therapeutics.

SINEUP non-coding RNAs rescue defective OPA1 expression and activity in a cellular model of Dominant Optic Atrophy.

1. Results

1.1 Synthetic miniSINEUP OPA1 design

Given that SINEUPs target specificity is based on the antisense pairing with the target mRNA sequence around the translation initiation codon AUG (Figure 26A), a deep and detailed analysis of the surrounding region of transcription initiation (TSS) is essential, particularly in tissues and cell types involved in the DOA. To have a complete picture of the TSS usage and of the 5' UTR region of endogenous human and murine *OPA1* mRNAs, we used the Cap Analysis of Gene Expression (CAGE) collection, known as FANTOM5 (Functional ANnoTation Of the Mammalian Genome), and Zenbu Genome Browser Tool[71,72]. We noticed that the annotated reference sequences did not fully recapitulate the complex scenario of OPA1 TSS usage, especially in human samples. Indeed, we observed a relevant discrepancy between UCSC annotated in UCSC and Gencode catalogue of transcripts (Figure 26B). In addition, various TSSs were also annotated for mouse sample (Figure 26D). Since a balance between long and short OPA1 protein isoforms has been proven to be an essential requirement for a full recovery of the mitochondrial network[164], we chose to design OPA1-specific mini-SINEUPs (miniSINEUP-OPA1) in antisense orientation to the common region shared between all *OPA1* mRNAs for both human and mouse, following the antisense-pairing rules of the natural Uchl1 locus[25,64], with minor modifications (Figure 26B). Although the natural BD of *AS Uchl1* was described to have a -40/+32 anatomy, with a 72 nts length, recent results from synthetic SINEUPs targeting a wide range of mRNAs, both endogenous and exogenously transfected in cells, have been shown to retain a comparable level of activity with a reduced BD size[63,67,68]. An overall reduction of SINEUPs' size would be extremely advantageous, especially for their development as RNA therapeutics. To this purpose, we generated a -40/+4 and a shorter -14/+4 BD miniSINEUP-OPA1 around the starting AUG for both human and mouse *OPA1* mRNAs (Figure 26C, E). Then, since another methionine (M125) is present, in-frame, in the third exon, we designed an additional -40/+4 BD around this sequence for human and a -41/+4 BD for mouse (Figure 26C, E). Finally, we designed an additional -14/+4 BD targeting the same second in-frame methionine (M125) in human *OPA1* mRNAs (Figure 26C). Each BD was combined with the very same ED, consisting of the invSINEB2 sequence derived from the natural AS

Uchl1 (Figure 26C, E). With this strategy, our constructs would synthesize miniSINEUP-OPA1 RNAs of around 200 nts in length.

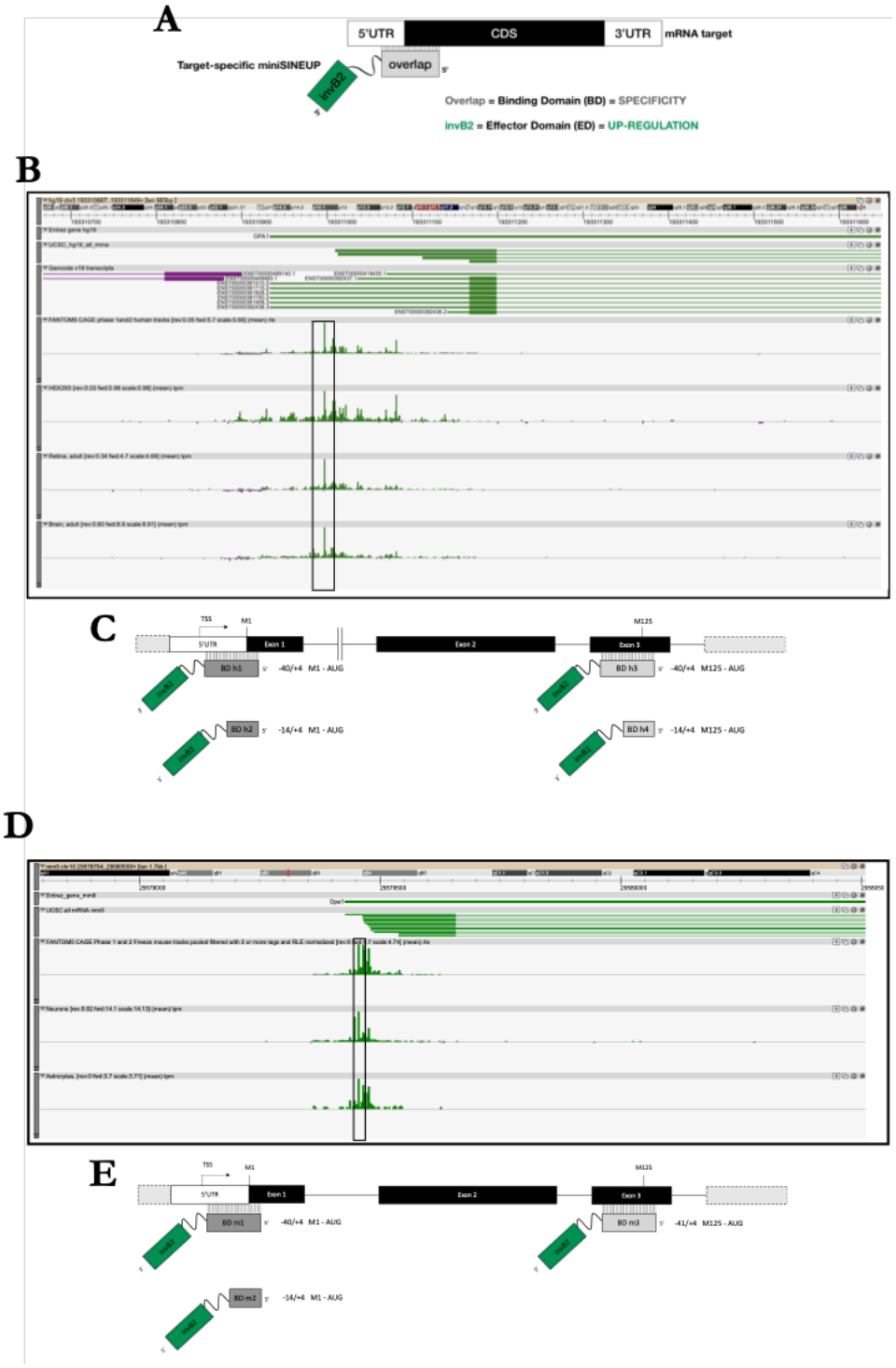


Figure 26 Design of synthetic miniSINEUPs to target OPA1 *mRNA.* **A)** Schematic representation of SINEUPs functional domains. The binding domain (BD, gray) provides SINEUP specificity and it is in antisense orientation to the sense protein coding mRNA (mRNA target). The inverted SINEB2 element (invB2) is the effector domain (ED, green) and confers enhancement of protein synthesis. 5' to 3' orientation of sense and antisense RNA molecules is indicated. Structural elements of target mRNA are shown: 5' untranslated region (5'UTR, white), coding sequence (CDS, black), and 3' untranslated region (3'UTR, white). The scheme is not drawn in scale. **B)** ZENBU genome browser view of human *OPA1* locus, showing *OPA1*

TSS in HEK293T, brain and retina. *OPA1* reference sequences and Gencode annotated transcripts are indicated. **C)** Scheme of human *OPA1* mRNA and BDs design (dark grey, M1 targeting BDs; light grey, M125 targeting BDs). The numbering refers to the position according to the methionine (i.e. -40/+4, from 40 nucleotides upstream and to 4 nucleotides downstream M1-AUG). The scheme is not drawn in scale. **D)** ZENBU genome browser view of mouse *OPA1* locus, showing *OPA1* TSS in neurons and astrocytes. *OPA1* reference sequences and Gencode annotated transcripts are indicated. **E)** Scheme of mouse *OPA1* mRNA and BDs design (dark grey, M1 targeting BDs; light grey, M125 targeting BDs). The numbering refers to the position according to the methionine (i.e. -40/+4, from 40 nucleotides upstream and to 4 nucleotides downstream M1-AUG). The scheme is not drawn in scale.

1.2 Synthetic miniSINEUP OPA1 are active *in vitro*

Human miniSINEUP-OPA1 RNAs increase OPA1 protein expression in human cells.

To screen human miniSINEUP-OPA1 activity, I took advantage of HEK 293T/17 cells, since they endogenously express *OPA1* mRNA and have already been proven to support SINEUP activity on a variety of endogenous genes[54,60,65–68]. HEK293T/17 were transfected with miniSINEUP-OPA1 (+miniSINEUP) or with a control vector harboring the only ED, without any BD (ΔBD). miniSINEUP activity was assessed as fold change in protein expression levels by western blotting, using β-actin for normalization (Figure 27). SINEUPs' post-transcriptional activity was monitored and confirmed by qRT-PCR quantification of *OPA1* mRNA (Supplementary Figure 4A). I also confirmed that all miniSINEUPs, included the ΔBD control, reached comparable level of expression upon transfection (Supplementary Figure 4B). SINEUP activity does not seem to be significantly influenced by the overlapping region: no significative activity difference was observed between M1-AUG and M125-AUG targeting miniSINEUP-OPA1. In particular, all tested human miniSINEUP-OPA1s were able to induce a significative increase of OPA1 protein levels (from ~ 1.4- to ~ 2.4-fold), with the -14/+4 M1-AUG miniSINEUP-OPA1 reaching the highest efficacy level. In summary, I successfully designed synthetic miniSINEUPs able to increase the production of human OPA1 protein without altering its mRNA levels nor the expression pattern of OPA1 isoforms.

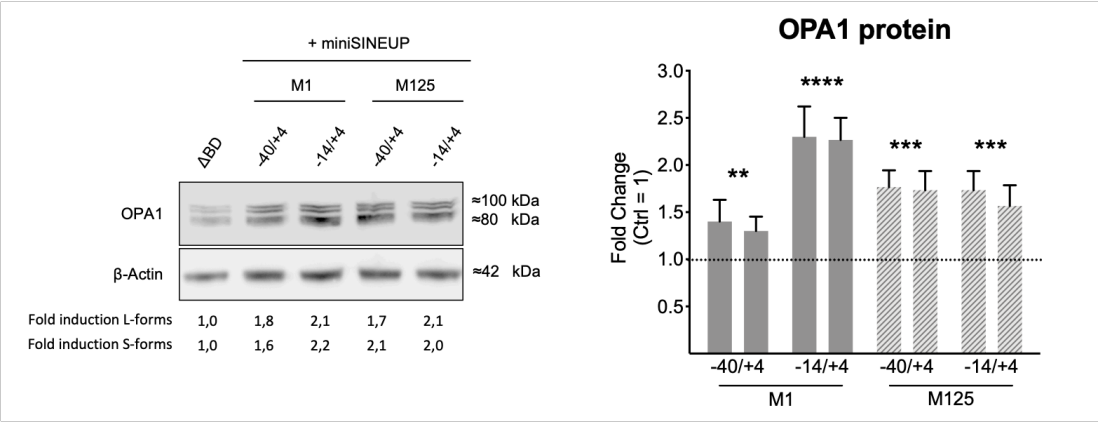


Figure 27 Synthetic miniSINEUPs increase quantities of endogenous OPA1 in vitro in human cells. HEK293T/17 cells were transfected with miniSINEUP-OPA1 variants and control vector (ΔBD) and harvested 48 hours post-transfection. Whole cell lysates were analyzed by western blotting with anti-OPA1 and anti-β-actin antibodies. Left: one representative experiment is shown. First, OPA1 band intensity was normalized to the relative β-actin band. Then, fold change values were calculated normalizing to control cells (Ctrl, ΔBD). miniSINEUP-OPA1-transfected cells showed increased levels of endogenous OPA1 protein. Right: average fold change of OPA1 protein levels. Columns represent mean ± SEM from three independent experiments. p values are calculated by One-way ANOVA followed by Dunnett's post-test.

Murine miniSINEUP-OPA1 RNAs increase OPA1 protein expression in murine cell lines.

To test mouse miniSINEUP-OPA1 activity, I used two different murine cell lines endogenously expressing *OPA1* mRNA: neuroblastoma Neuro2A (N2A) and astrocytes C8-D1A cell lines, already proven to support SINEUP activity[67]. Both cell lines were transfected with miniSINEUP-OPA1 (+miniSINEUP) or with a control vector harboring the only ED, without any BD (ΔBD). miniSINEUP activity was then assessed as fold change in protein amount by western blotting, using β-actin for normalization (Figure 28A, B). SINEUPs' post-transcriptional activity was monitored and confirmed by qRT-PCR quantification of *OPA1* mRNA (Supplementary Figure 5A). A comparable expression level among miniSINEUP and ΔBD control was confirmed by qRT-PCR Real-time (Supplementary Figure 5B). All tested murine miniSINEUP-OPA1s showed statistically significant increase of OPA1 protein (from ~ 1.5- to 2-fold in N2A cell line and ~ 1.7-fold in astrocytes). Interestingly, murine miniSINEUP-OPA1 -14/+4 M1-AUG, harboring the shorter BD, was proven to reach the highest efficiency: ~ 1.7-fold in astrocytes and ~ 2-fold in N2A. qRT-PCR Real-time analysis confirmed no variation in *OPA1* mRNA levels proving that all the miniSINEUP-OPA1s are acting at post-transcriptional level. Hence, I managed to produce synthetic miniSINEUP-OPA1s able to increase the production of murine OPA1 protein without altering its mRNA levels. Most importantly, miniSINEUPs-OPA1 were proven to be able to increase the amount

of OPA1 protein without altering the balance among different isoforms, as inferred by western blot images (Figure 28A, B).

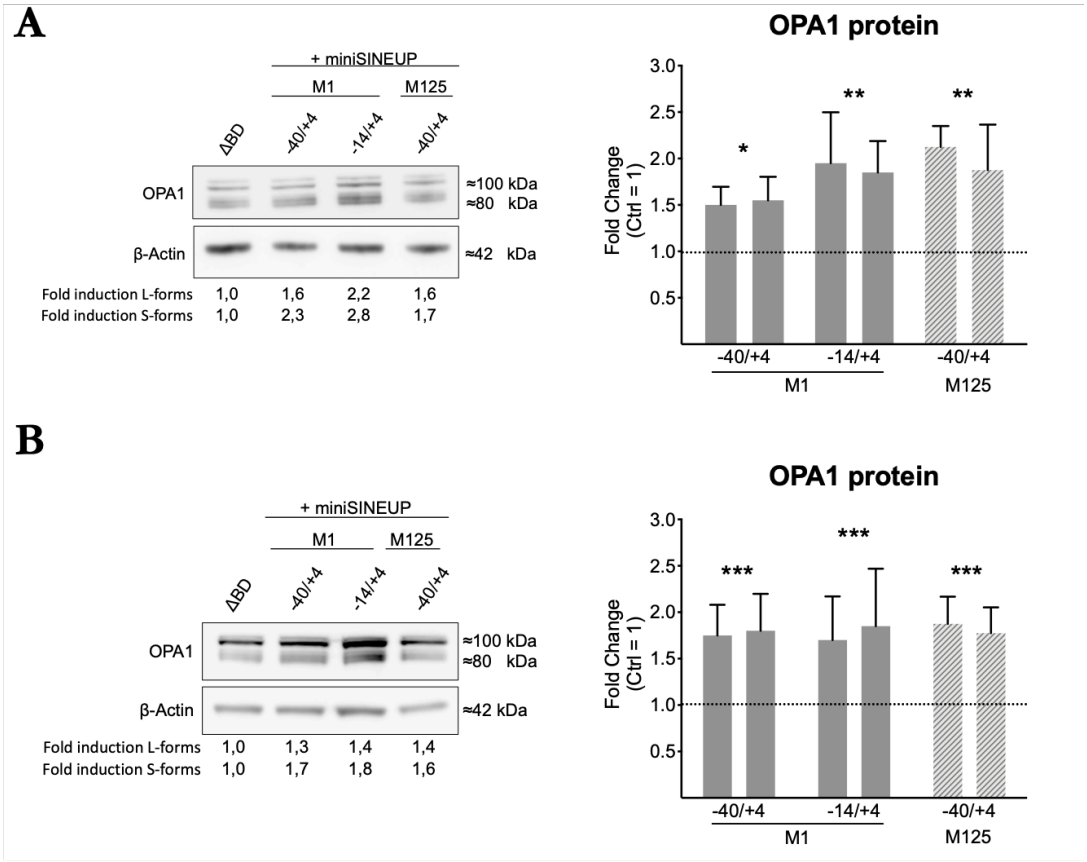


Figure 28 Synthetic miniSINEUPs increase quantities of endogenous OPA1 in vitro in mouse cells. Neuro2A cells **(A)** and C8-D1A astrocytes **(B)** were transfected with miniSINEUP-OPA1 variants and control vector (ΔBD) and harvested 48 hours post-transfection. Whole cell lysates were analyzed by western blotting with anti-OPA1 and anti-β-actin antibodies. Left: one representative experiment is shown. First, OPA1 band intensity was normalized to the relative β-actin band. Then, fold change values were calculated normalizing to control cells (Ctrl, ΔBD). miniSINEUP-OPA1-transfected cells showed increased levels of endogenous OPA1 protein. Right: average fold change of OPA1 protein levels. Columns represent mean ± SEM from three independent experiments. p values are calculated by One-way ANOVA followed by Dunnett's post-test.

1.3 SINEUP optimization for RNA therapeutics development

Recent results from our lab and others show that *in vitro* transcribed SINEUP and miniSINEUP RNAs are active in cells when appropriately modified[100,101]. Based on this information and keeping in mind that RNA molecules length and stability are strict limitation to the application of naked RNA both *in vitro* and *in vivo*, I aimed to identify a minimal active SINEUP (ASO-SINEUP) that could be successfully administered as RNA therapeutic. First, a series of shorter plasmid-encoded SINEUP EDs was designed and combined them with a GFP-targeting BD already in use in the lab[70], to be tested as a proof of principle. To reduce SINEUP RNAs length, I designed progressively shorter

ED domains with the aim to maintain the SL1 region unaltered, previously reported to be the key structural and functional domain of AS Uchl1 invSINEB2[188]. As a result, I generated a microED (44-120), two nanoEDs (59-96 and 64-92) and two femtoEDs (64-82 and 64-81) (Figure 29A). miniSINEUP-GFP and its shorter variants were transfected in HEK293T cells and EGFP protein expression level was assessed with both Western Blot and fluorescence measurement at 488 nm (Figure 29B, C). Compared to the control, all constructs were able to induce a GFP protein increase of >2.5 fold. Expression levels of *EGFP* mRNA and SINEUP RNA were monitored by qRT-PCR Real-time confirming no variation in target mRNA's expression and a comparable level of expression among all truncated SINEUP forms (Supplementary Figure 6).

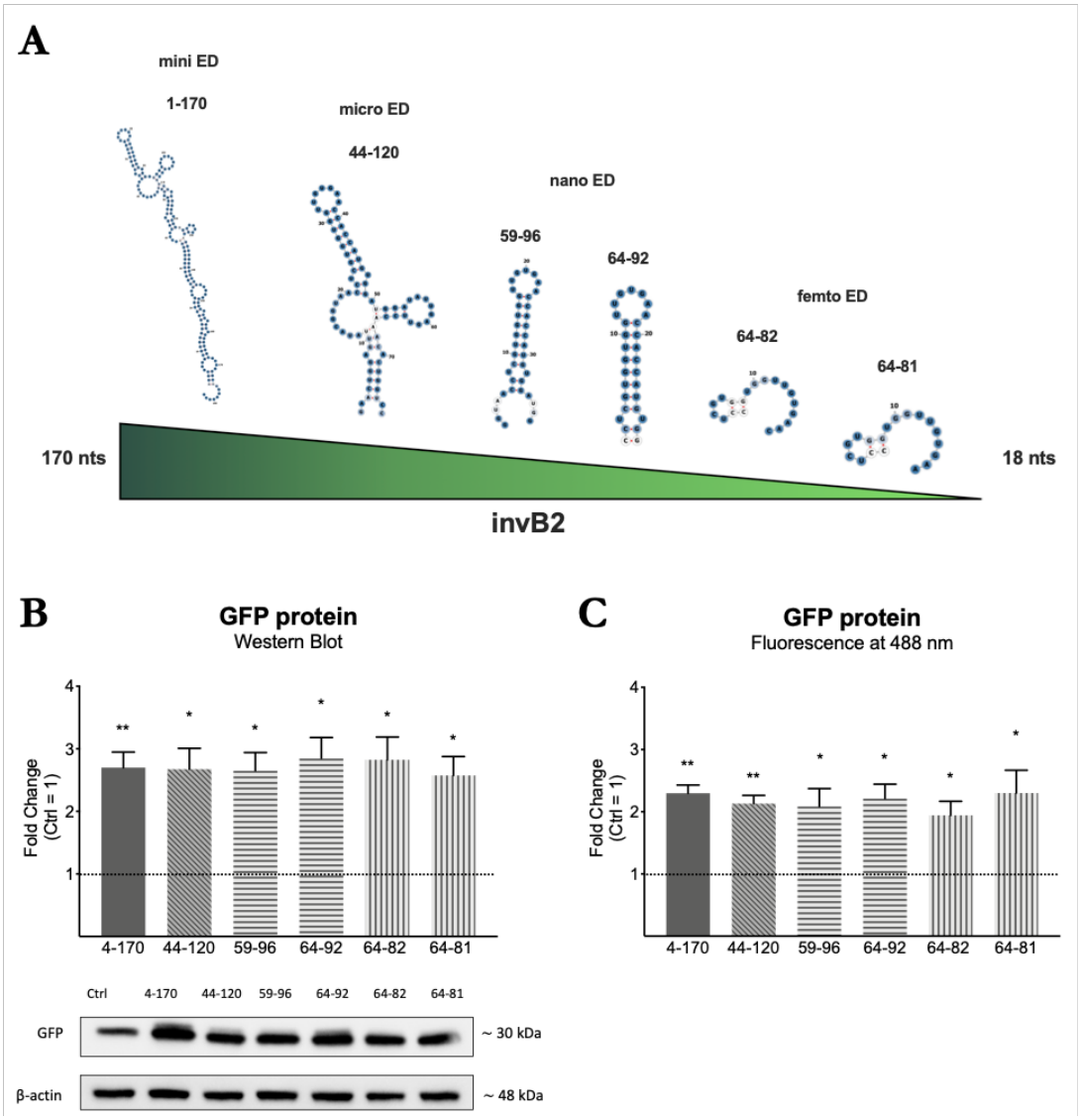


Figure 29 SINEUP miniaturization design and testing. **A)** Secondary structure prediction of mini, micro, nano and femto EDs. **B-C)** HEK293T/17 cells were transfected with pDUAL plasmid co-expressing *EGFP* target mRNA and SINEUP variants and the same plasmid expressing the only *EGFP* mRNA was transfected as negative control. Cells were harvested 48 hours post-transfection and processed for protein

and RNA analysis (Supplementary Figure 6). **B)** Whole cell lysates were analyzed by western blotting with anti-EGFP and anti-β-actin antibodies. Up: one representative experiment is shown. First, EGFP band intensity was normalized to the relative β-actin band. Then, fold change values were calculated normalizing to control cells (Ctrl). SINEUP-GFP-transfected cells showed increased levels of EGFP protein. Down: one representative experiment is shown. **C)** Harvested cells were resuspended in PBS 1X and incubated with NucBlue probe for fluorescence normalization. Cell suspension was then transferred to 96-well plate and EGFP (488 nm) and NucBlue (460 nm) fluorescence were measured. First, EGFP fluorescence intensity was normalized to the relative NucBlue intensity. Then, fold change values were calculated normalizing to control cells (Ctrl). RNA analysis is reported in Supplementary Figure 6. Columns represent mean ± SEM from three independent experiments. p values are calculated by One sample t and Wilcoxon test.

To investigate whether nano and femtoSINEUP-GFP were also active when transfected as RNA molecule, as previously shown for miniSINEUP-DJ1 in our lab[101], I transfected an *in vitro* synthesized RNA oligo carrying 2'OMe-Adenosine modification in all available positions and evaluate SINEUP activity through fluorescence measurement at 488 nm (Figure 30). Remarkably, ASO-SINEUP-GFP was able to increase GFP protein production reaching around 1.5 fold induction, even if at lower extent compared to plasmid-driven SINEUP expression. This could be due to a sub-optimal ratio between *EGFP* target mRNA and nanoSINEUP expression levels. qRT-PCR analysis confirmed plasmid-expressed SINEUPs and ASO-SINEUP-GFPs post-transcriptional activity and comparable expression levels (Supplementary Figure 6C,D).

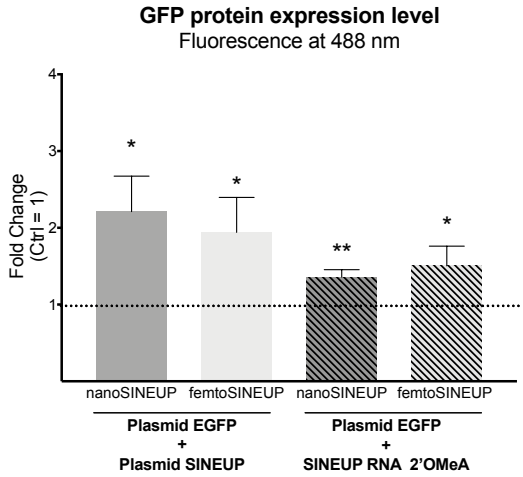


Figure 30 ASO-SINEUP-GFP are active in vitro. HEK293T/17 cells were transfected with pDUAL expressing the only *EGFP* mRNA alone and in combination with nanoED 2'OMeA RNA oligo as negative controls. pDUAL plasmids co-expressing *EGFP* mRNA with nano- and femtoSINEUP-GFP were transfected as positive controls. To test ASO-SINEUP-GFP activity pDUAL plasmid expressing the only *EGFP* mRNA was transfected in combination with ASO-SINEUP-GFP 2'OMeA RNA oligo variants. Cells were harvested 48 hours post-transfection and processed for protein and RNA analysis (Supplementary Figure 6). Harvested cells were resuspended in PBS 1X and incubated with NucBlue probe for fluorescence normalization. Cell suspension was then transferred to 96-well plate and EGFP (488 nm) and NucBlue (460 nm) fluorescence were measured. First, EGFP fluorescence intensity was normalized to the relative NucBlue intensity. Then, fold change values were calculated normalizing to control cells (Ctrl). RNA analysis is reported in Supplementary Figure 6. Columns represent mean ± SEM from at least three independent experiments. p values are calculated by One sample t and Wilcoxon test.

I then chose the shorter nanoED (64-92) to be tested in combination with OPA1 BD -14/+4 targeting M1 giving rise to nanoSINEUP-OPA1 plasmid-encoded RNA. In parallel with this construct, an *in vitro* synthesized RNA molecule with 2'OMe-Adenosine modification was tested in HEK293T cells. It is formed by the combination of OPA1 BD -14/+4 targeting M1 and the nanoED 64-92 with a 5 nts spacer in between, (Figure 31A). Interestingly, both plasmid-encoded nanoSINEUP-OPA1 RNA and ASO-SINEUP-OPA1 RNA were proved to be active, reaching a fold-induction level of around 1.7 (Figure 31B, C). qRT-PCR Real-time analysis confirmed no variation in *OPA1* mRNA levels proving that both nanoSINEUP-OPA1 and ASO-SINEUP-OPA1 are acting at post-transcriptional level (Supplementary Figure 7). This result represents a major advancement in the development of a SINEUP RNA for therapy.

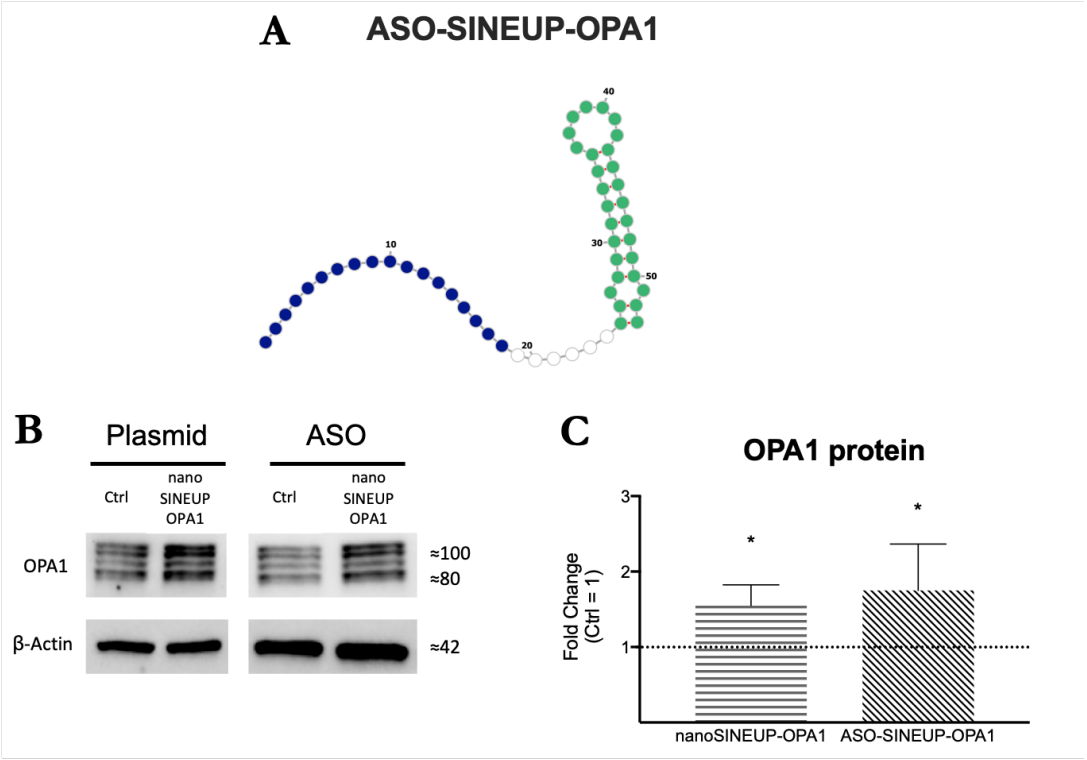


Figure 31 nanoSINEUP-OPA1 and ASO-SINEUP-OPA1 are active in vitro. HEK293T/17 cells were transfected with control or nanoSINEUP-OPA1 expressing plasmids or with nanoED or ASO-SINEUP-OPA1 with nanoED 2'OMeA RNA oligo. Cells were harvested 48 hours post-transfection and processed for protein and RNA analysis (Supplementary Figure 7). Whole cell lysates were analyzed by western blotting with anti-OPA1 and anti-β-actin antibodies. **B)** one representative experiment is shown. First, OPA1 band intensity was normalized to the relative β-actin band. Then, fold change values were calculated normalizing to control cells (Ctrl, ΔBD). miniSINEUP-OPA1-transfected cells showed increased levels of endogenous OPA1 protein. **C:** average fold change of OPA1 protein levels. Columns represent mean ± SEM from at least three independent experiments. p values are calculated by One sample t and Wilcoxon test.

1.4 OPA1 protein rescue in DOA patients' fibroblasts

DOA patients-derived fibroblasts represent the most relevant cellular model to evaluate potential therapeutic strategies. The most efficient miniSINEUP-OPA1 and the nanoSINEUP-OPA1 were cloned into a lentiviral constitutive expression vector harboring TurboRFP as reporter gene along with the previously described ΔBD as a negative control. At first, lentiviral constructs were transfected as plasmids in HEK293T cells, to ensure appropriate activity and expression level of SINEUPs and TurboRFP reporter gene (Figure 32, Supplementary Figure 8). Upon both miniSINEUP and nanoSINEUP transfection, OPA1 endogenous protein was increased of around 2-fold, confirming both RNA's activity.

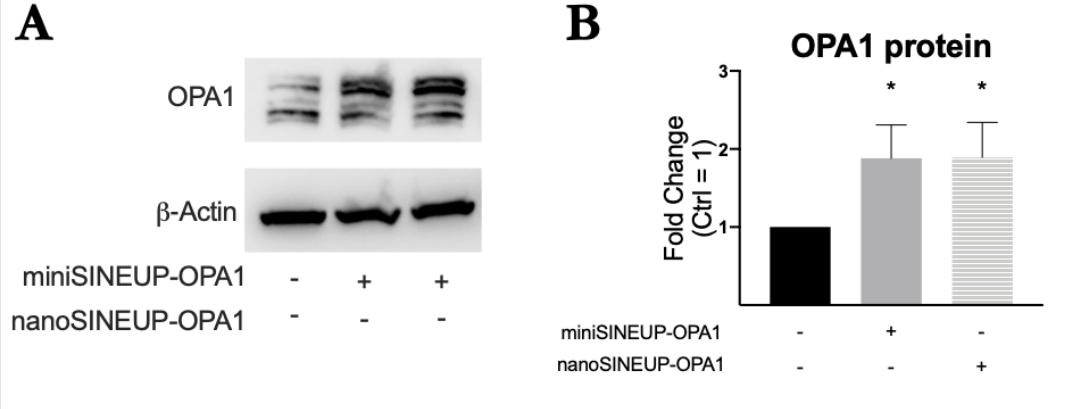


Figure 32 Lentiviral plasmids encoded mini and nanoSINEUP-OPA1 are active in vitro. HEK293T/17 cells were transfected with SINEUP-OPA1 variants and control vector and harvested 48 hours post-transfection. Whole cell lysates were analyzed by western blotting with anti-OPA1 and anti-β-actin antibodies. **A)** one representative experiment is shown. First, OPA1 band intensity was normalized to the relative β-actin band. Then, fold change values were calculated normalizing to control cells (Ctrl, ΔBD). Mini and nanoSINEUP-OPA1-transfected cells showed increased levels of endogenous OPA1 protein. **B)** Average fold change of OPA1 protein levels. Columns represent mean ± SEM from four independent experiments. p values are calculated by One sample t and Wilcoxon test.

SINEUP-OPA1 lentiviral particles were then used to infect patients' primary dermal fibroblasts. Each cell line was FACS-sorted twice for TurboRFP expression to select stably transduced cells. Among DOA patients' fibroblasts lines, we selected one derived from a male 55 years old donor carrying the mutation c.703 C>Tp.R235 in exon 7 (F40D) and one derived from a 27 years old female donor, carrying c.2823_26delAGTTpK941fsX966 mutation (F171). As controls, we used two healthy donor-derived fibroblasts lines, age and sex-matched. As a first step, I performed Western Blotting to assess OPA1 protein dosage rescue upon SINEUP expression. A marked reduction in OPA1 protein expression was detected in both DOA patients-derived cell

lines, when compared with the respective healthy control. Indeed, F40D and F171 cells were found to express only 50% and 25% OPA1 protein compared to the respective healthy control. qRT-PCR Real-time also confirmed a reduced expression of OPA1 mRNA (Supplementary Figure 9). Remarkably, upon SINEUP expression, through lentiviral infection, I was able to detect a rescue of OPA1 protein levels ranging from 65% to 100% (Figure 33).

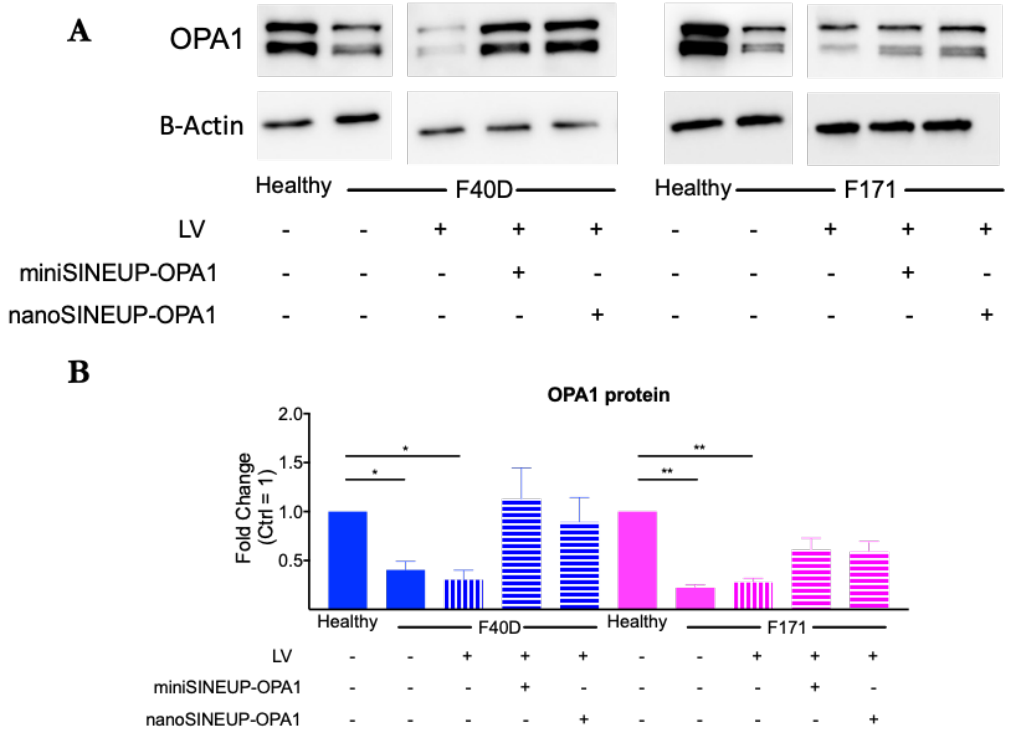


Figure 33 OPA1 protein rescue in DOA patients' fibroblasts. Human primary fibroblast lines were transduced with control, mini- and nanoSINEUP-OPA1 expressing lentiviral. Cells were FACS sorted for highest and most stable reporter expression to establish stably expressing cell lines. Each cell line was expanded and harvested for protein level assessment. Whole-cell lysates were analyzed by western blotting with anti-OPA1 and anti-β-actin antibodies. **A:** one representative experiment is shown. First, OPA1 band intensity was normalized to the relative β-actin band. Then, fold change values were calculated normalizing to healthy control cells. miniSINEUP-OPA1 and nanoSINEUP-OPA1-transduced F40D and F171 cells showed increased levels of endogenous OPA1 protein. **B:** average fold change of OPA1 protein levels. Columns represent mean ± SEM from three independent experiments. p values are calculated by One sample t and Wilcoxon test.

Discussion

Human OPA1 gene started to draw attention in 2000, when its mutations were found associated with DOA. This rare genetic disorder was originally characterized back in 1959 from the Danish ophthalmologist Paul Kjer as a progressive decrease in visual acuity, tritanopia, loss of sensitivity in central visual field and optic disk pallor. DOA is generally considered as part of diversified groups of optic neuropathies and mitochondrial diseases and, with a prevalence of 1:25000, is the most common inherited mitochondrial optic neuropathy. OPA1 protein is a ubiquitous dynamin-related GTPase residing in the inner

mitochondrial membrane (IMM) and presenting 8 differently processed isoforms whose expression level is tissue-dependent. An unbalance in long and short OPA1 forms ratio with more abundant s-forms than in physiological state causes fusion inhibition and mitochondrial network fragmentation. Importantly, to achieve full recovery of mitochondrial network morphology, at least two OPA1 isoforms with a specific balance of l- and s- isoforms are necessary, suggesting the need for a multiplicity of isoforms to flexibly shape mitochondrial dynamics as a response to different metabolic and stress conditions perturbing cellular homeostasis[164]. OPA1 protein plays a key role in mitochondrial network dynamics: it promotes mitochondrial fusion with MFN1 and MFN2, acts in concert with pro-fission proteins DRP1 and DNM2 and contributes to mtDNA maintenance. OPA1 protein polymerization also preserves *cristae* morphogenesis, facilitating the activity of respiratory chain super-complexes[167]. It has a main role in controlling the apoptotic process as it is fundamental for the compartmentalization of cytochrome C. Indeed, mitochondria play an essential role in cellular homeostasis processes such as organelle dynamics control, interaction with other organelles, apoptosis regulation, calcium homeostasis maintenance and autophagy, but, most importantly, they are key suppliers of cellular energy through oxidative phosphorylation (OXPHOS). Alteration of this complex multi-step process can cause a reduction of ATP synthesis and an increase of reactive oxidative species (ROS), inducing damage in the respiratory chain and activation of apoptotic pathways up to mtDNA mutations accumulation. All these molecular effects can lead to energy failure and eventually cell death. In most cases, mitochondrial dysfunctions lead to neurodegeneration, addressing particularly RGC cells. This cell type is severely damaged by energy failure because they present narrower not myelinated axons, that imply the absence of saltatory conduction of action potentials, highly requiring energy supply from mitochondria clustering within unmyelinated retinal and prelaminar sectors and less abundant in the posterior part of lamina cribrosa.

The essential role of OPA1 in the regulation of mitochondrial metabolism is nowadays accepted. It represents a key crossroad for mitochondrial homeostasis whose mutations reflect into a progressively enlarging repertoire of clinical phenotypes, including DOA plus syndromes, multiple sclerosis, Parkinsonism and dementia, infantile Leigh syndrome and cardiomyopathy.

Therapeutic treatments for DOA are still in early stages development, with major focus on idebenone administration, an ubiquinone analog, currently under evaluation for

treatment of Leber hereditary optic neuropathy (LHON), also characterized by RGC loss[189]. Off-label clinical trials in OPA1-caused DOA patients show some recovery of visual parameters with idebenone[190]. Unfortunately, benefits from idebenone treatment were reported to be temporary in a placebo-controlled trial on OPA1 mouse model[191]. Interestingly, recent studies report first attempts to apply gene therapy as DOA treatment. Indeed, as the majority of OPA1-related DOA cases derives from haploinsufficiency, this approach could be extremely advantageous. Recent results from clinical studies for LHON treatment reported significant bilateral improvements, indicating feasibility of gene therapy treatment of optic neuropathies[192,193]. Nevertheless, three major drawbacks can be envisioned when evaluating the development of gene therapy approaches to restore OPA1 protein levels: i. OPA1 coding sequence, being quite long, could be hard to be delivered by commonly used AAV gene therapy vectors, ii. given the presence of 8 different isoform and the requirement of a correct balance between long and short OPA1 forms, the delivered sequence should be carefully chosen, iii. overexpression of OPA1 beyond physiological levels is reported to be detrimental[194]. Furthermore, the lack of specific promoters for every cell type can give rise to ectopic expression.

In this scenario of unmet clinal needs for DOA treatment, I chose to evaluate the application of SINEUP technology as new RNA-based therapeutic approach to target OPA1 defective gene expression, allowing the rescue of protein levels within physiological range with a post-transcriptional regulation. For RNA-based therapies, chemically modified *in vitro* synthesized mRNAs can give rise to ectopic overexpression occurring in a short timeframe. Similarly, in the case of small activating RNAs (saRNAs) major limitations reside in the triggering of gene targets transcription upregulation. To date, two strategies can be pursued for the delivery of SINEUP therapeutics to patients: the first takes advantage of AAV delivery system, with chronical *in vivo* expression of a SINEUP molecule, while the second strategy could use non-viral systems, such as lipid complexes or other nanoparticles commonly used for RNA therapeutics delivery, to administrate SINEUPs as RNA molecules. A first Proof-Of-Concept of miniSINEUP AAV delivery was successfully shown in our laboratory to increase endogenous GDNF protein levels *in vivo*[67]. On the other hand, in the last years, the demonstration that the incorporation of selected chemically-modified ribonucleotides during IVT restores SINEUP activity, which is absent in non-modified IVT RNA molecules[100,101], provided promising new data on the way to a chemically synthesized, active SINEUP.

In this study, miniSINEUP-OPA1 are designed in antisense orientation 5' head-to-head to common regions among all OPA1 mRNA splice variants. Different BDs for human and murine OPA1 have been tested, each one combined with the invSINEB2 from *AS Uchl1* as ED. Our results confirm BD's flexibility previously reported for other targets[67,68] designing effective BDs around both the starting AUG and the first internal in-frame methionine. Interestingly, in both human and mouse systems, the most efficient BD were the short -14/+4 sequences, designed around the starting AUG. A major limitation of RNA-based therapeutics development is represented by the requirement of a sequence length <=50 nts for delivery purposes. Given that a miniSINEUP presents an average length of 250 nts, a further miniaturization is required. Here, for the first time, I designed and optimize nanoSINEUPs by reducing the ED length from 172 to 29 nts, with a rational that takes into account recent indication on the essential function of invSINEB2 SL1 for SINEUP activity[73]. I combined nanoED with canonical exogenous mRNA targeting GFP BD and endogenous mRNA targeting OPA1 BDs and evaluate each one's activity. Most remarkably, based on recent evidence of IVT miniSINEUP-DJ1 RNA activity when 2'O-Me-Adenosine modified[101], I was able to apply nanoSINEUP-GFP and nanoSINEUP-OPA1 in the form of chemically synthesized RNA oligo which were able to reach a comparable level of activity to their plasmid-encoded counterparts. This is, to date, the first prove that SINEUP RNA synthetic molecules as short as ≈50 nts are able to increase exogenous and endogenous target mRNA translation, shortening the distance to be covered for SINEUP technology application as RNA therapeutics.

Most importantly, using patients' derived human dermal primary fibroblasts carrying various mutations as a model of the complex clinical features of DOA, I proved the ability of both mini- and nanoSINEUP to rescue OPA1 protein amount to healthy control levels. These cells' preparations will be now used to assess the functional rescue of DOA phenotypes, evaluated for Oxygen consumption rate assay (Seahorse), mitochondrial morphology analysis and mitochondrial DNA content assessment. AAV will be then produced to investigate the effects of SINEUP-OPA1 overexpression on the retina of a mouse model of DOA.

In conclusion, with this study, I provided strong evidence that synthetically designed SINEUP molecules are able to increase OPA1 endogenous protein expression to physiological level in a cellular model of DOA. Their ability to rescue pathological phenotypes in human patients' cells ensures the pre-clinical evaluation of a SINEUP-

based therapy to treat DOA and supports SINEUP technology as scalable platform to treat haploinsufficient diseases.

Conference proceedings

1. **Pierattini B.**, D'Agostino S., Bon C., Espinoza S., Valentini P., Pandolfini L., Gustincich S. *N^6-methyladenosine modification regulates SINEUP lncRNAs activity.* **2022 SIBBM Seminar "The RNA world"**, 20-22 June 2021, Rome, **Oral presentation**
2. D'Agostino S.*, Matey A.*, Volpe M., **Pierattini B.**, Lau P., Bon C., Peruzzo O., Braccia C., Armirotti A., Scarpato M., Di Carlo V., Santoro C., Persichetti F., Espinoza S., Zucchelli S., Sanges R., and Gustincich S. *Internal Ribosome Entry Site RNAs act in trans through an antisense sequence in linear and circular non-coding RNAs.* **2022 SIBBM Seminar "The RNA world"**, 20-22 June 2021, Rome
3. **Pierattini B.,** Valentini P., Bon C., Espinoza S., Pandolfini L., and Gustincich S. *N6-methyladenosine modification regulates SINEUP lncRNAs activity.* **EMBL in Italy 2021**, 20-21 May 2021, **Poster presentation**
4. D'Agostino S.*, Matey A.*, Volpe M., **Pierattini B.**, Lau P., Bon C., Peruzzo O., Braccia C., Armirotti A., Scarpato M., Di Carlo V., Santoro C., Persichetti F., Espinoza S., Zucchelli S., Sanges R., and Gustincich S. *Internal Ribosome Entry Site RNAs act in trans through an antisense sequence in linear and circular non-coding RNAs.* **EMBL in Italy 2021**, 20-21 May 2021
5. **Pierattini B.,** Valentini P., Bon C., Espinoza S., Pandolfini L., and Gustincich S. *N6-methyladenosine modification regulates SINEUP lncRNAs activity.* **Non-Coding RNAs: Biology and Applications, Keystone eSymposia,** 11-14th May 2021, **Poster presentation**
6. **Pierattini B.,** Bon C., Espinoza S., D'Agostino S., Carelli V., Zucchelli S., Gustincich S., *SINEUP non-coding RNAs selectively up-regulate OPA1 expression in rodent and human cells.* **The non-coding genome, EMBO/EMBL Symposia**, October 16-19th 2019, **Poster presentation**
7. Espinoza S, Bon C, **Pierattini B**, Jones MH, Luffarelli R, D'Agostino S, Valentini P, Matey A, Condò I, Cotella D, Santoro C, Zucchelli S, and Gustincich S *SINEUPs: a new antisense, long non-coding RNA-based platform to increase endogenous protein levels for therapy.* UMass Medical School, **RNA therapeutics Conference 2019** , From Base Paires to Bedsides, Worcester MA, June 26-28, 2019
8. **RNA therapeutics Conference, SMi**, London February 19-20th 2019

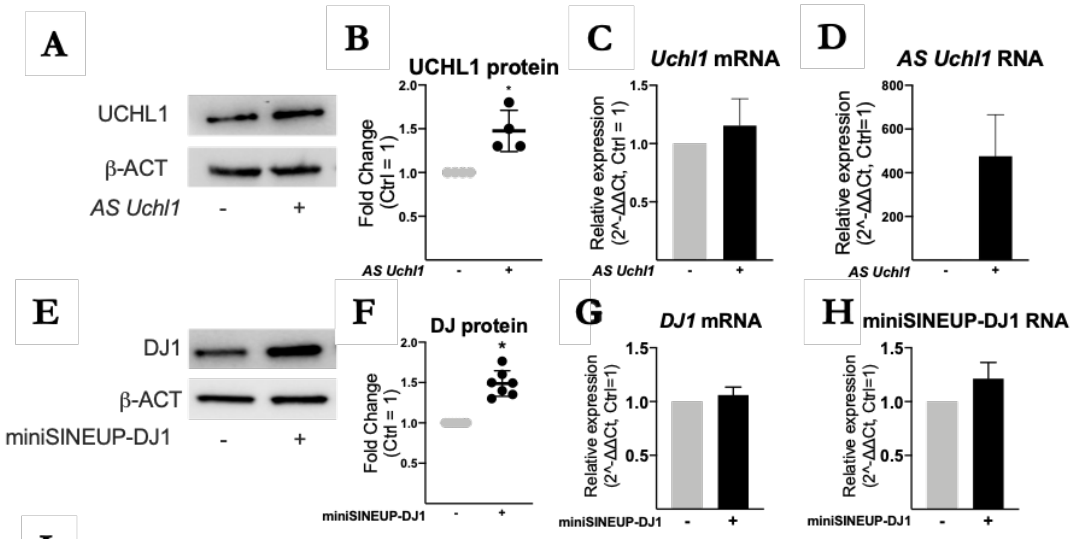
List of publications

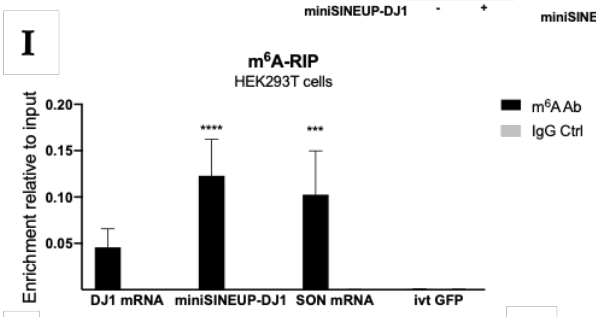
1. **Pierattini B.,** D'Agostino S., Bon C., Peruzzo O., Alendar A., Espinoza S., Valentini P., Pandolfini L. and Gustincich S. *N6-methyladenosine modification regulates SINEUP non-coding RNA activity.* Manuscript in preparation.
2. **Pierattini B.,** Del Dotto V., Bon C., D'Agostino S., Carelli V., Zucchelli S., Espinoza S. and Gustincich S. *SINEUP non-coding RNAs rescue defective OPA1 expression and activity in cellular models of Dominant Optic Atrophy.* Manuscript in preparation
3. D'Agostino S.*, Matey A.*, Volpe M., **Pierattini B.**, Cheung P.L.P., Bon C., Peruzzo O., Armirotti A., Scarpato M., Di Carlo V., Santoro C., Persichetti F., Espinoza S., Zucchelli S., Sanges R. and Gustincich S. *Internal Ribosome Entry Site RNAs act in trans through antisense base-pairing in linear and circular non-coding RNAs.* Submitted
4. Valentini P., **Pierattini B.**, Zacco E., Mangoni D., Espinoza S., Webster N.A., Andrews B., Carninci P., Tartaglia G.G., Pandolfini L., Gustincich S. *Towards SINEUP-based therapeutics: Design of an* in vitro *synthesized SINEUP RNA.* Mol Ther Nucleic Acids 2022 Feb 2; 27:1092-1102. doi: 10.1016/j.omtn.2022.01.021.eCollection 2022 Mar 8
5. Espinoza S., Bon C., Valentini P., **Pierattini B**., Tettey Matey A., Damiani D., Pulcrano S., Sanges R., Persichetti F., Takahashi H., Carninci P., Santoro C., Cotella D. and Gustincich S. *SINEUPs: a novel toolbox for RNA therapeutics.* Assays in Biochem. 2021 Oct 27;65(4):775-789. doi: 10.1042/EBC20200114.
6. Bon C., Luffarelli R., Russo R., Fortuni S., **Pierattini B.,** Santulli C., Fimiani C., Persichetti F., Cotella D., Mallamaci A., Santoro C., Carninci P., Espinoza S., Testi R., Zucchelli S., Condò I. and Gustincich S. *SINEUP non-coding RNAs rescue defective frataxin expression and activity in a cellular model of Friedreich's Ataxia.* Nucleic Acid Research 2019 Nov18;47(20):10728-10743.doi:10.1093/nar/gkz798

Patents

1. "FUNCTIONAL NUCLEIC ACID MOLECULES" to increase OPA1 expression, TSI-C-P2811GBp
2. "MODIFIED FUNCTIONAL NUCLEIC ACID MOLECULES" on chemical modifications of SINEUP RNAs, PT200555
3. "FUNCTIONAL NUCLEIC ACID MOLECULES" on circRNA function, P123747GB

Supplementary Figures





J ***AS Uchl1* putative m^{6}A sites**

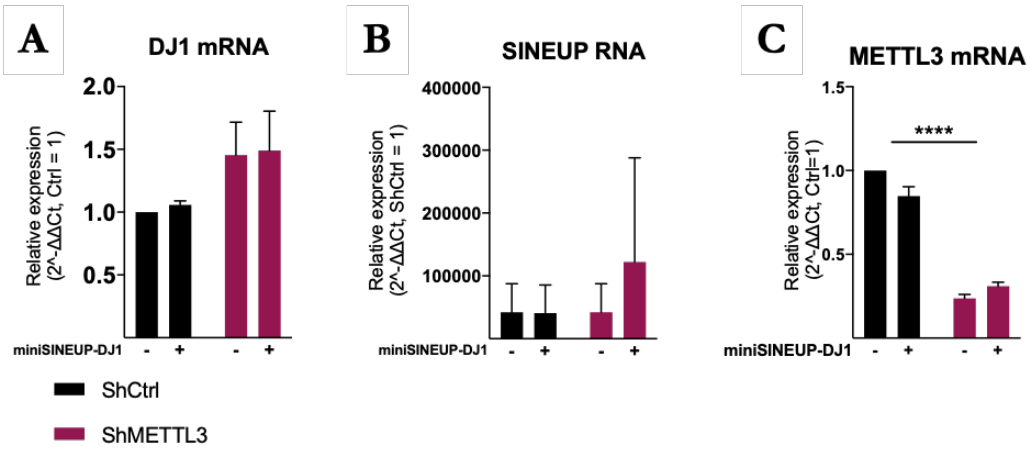
Position	Sequence	Confidence	SRAMP
165	AAACA	Low	✓
197	GAACU	High	✓
275	GGACU	Very High	✓
390	GAACU	Moderate	✓
407	UAACC	-	-
425	GAACU	-	-
455	AAACU	-	-
544	GAACA	Low	✓
759	AGACU	Low	✓
801	GGACA	Low	✓
845	AGACU	Moderate	✓
1231	AGACU	Moderate	✓

K **invSINEB2 putative m^{6}A sites**

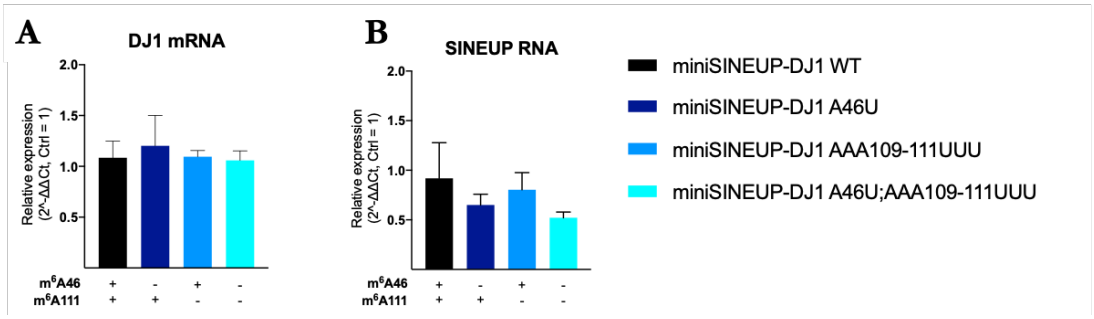
Position	Sequence	Confidence	SRAMP
46	GAACU	Low	✓
61	UAACC	-	-
81	GAACU	-	-
111	AAACU	-	-

Supplementary Figure 1 **A-D) AS Uchl1 activity in MN9D cells.** MN9D cells were transfected with control or AS Uchl1 plasmid and harvested 48 hours post-transfection. Whole cell lysates were analyzed by western blotting with anti-Uchl1 and anti-β-actin antibodies. **A)** One representative experiment is shown. AS Uchl1-transfected cells showed increased levels of endogenous UCHL1 protein. **B)** Average fold change of UCHL1 protein levels. Data indicate mean ± SEM from four independent experiments. p values are calculated by One sample t and Wilcoxon (*, p<0,05). **C)** Uchl1 mRNA levels were analyzed by qRT-PCR with specific primers. Uchl1 mRNA expression was stable (ns, p>0.05) **D)** AS Uchl1 RNA levels were analyzed by qRT-PCR with specific primers confirming overexpression. **E-H) miniSINEUP-DJ1 activity in A549 cells.** A549 cells were transfected with control (ΔBD) or miniSINEUP-DJ1 plasmid and harvested 48 hours post-transfection. Whole cell lysates were analyzed by western blotting with anti-DJ1 and anti-β-actin antibodies. **E)** One representative experiment is shown. miniSINEUP-DJ1-transfected cells showed increased levels of endogenous DJ1 protein. **F)** Average fold change of DJ1 protein levels. Data indicate mean ± SEM from four independent experiments. p values are calculated by One sample t and Wilcoxon (*, p<0,05). **G)** *DJ1* mRNA levels were analyzed by qRT-PCR with specific primers. *DJ1* mRNA expression

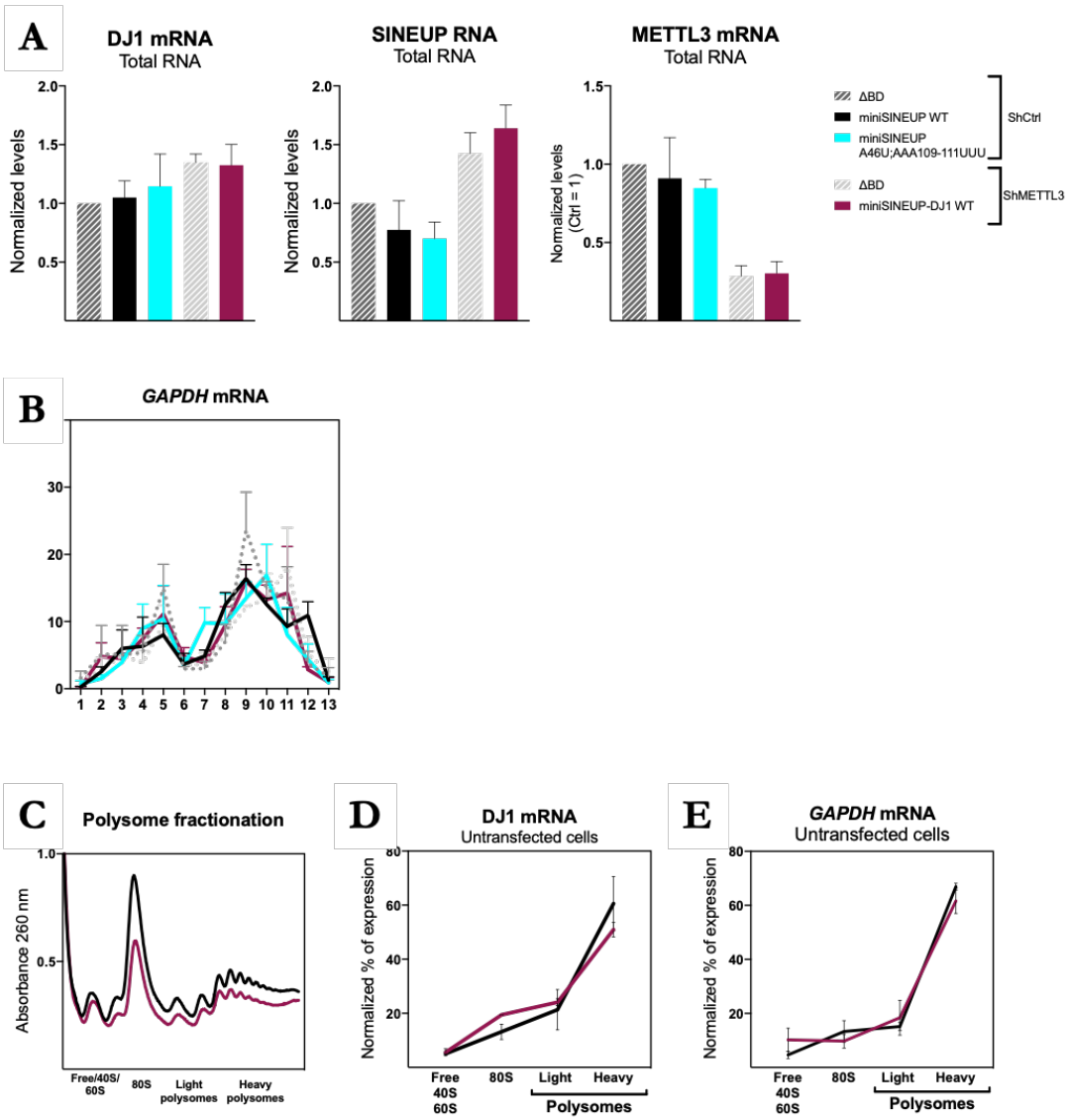
was stable (ns, p>0.05) **H)** Control and miniSINEUP-DJ1 RNA levels were analyzed by qRT-PCR with specific primers confirming overexpression. **I) m^{6}A RNA immunoprecipitation in miniSINEUP-DJ1 transfected HEK293T/17 cells.** *DJ1* mRNA and miniSINEUP-DJ1 RNA relative enrichment in m^{6}A immunoprecipitated RNA from untransfected and AS miniSINEUP-DJ1 transfected A549 cells. ivt GFP spiked-in and IgG Ctrl samples (striped columns) were used as negative controls. *hSON* mRNA was analyzed as positive control. Data are expressed as enrichment relative to input and indicate mean ± SEM from three independent experiments. p values are calculated by Two-way ANOVA follows by Sidak multiple comparison (***, p<0,001). **J)** m^{6}A consensus sequences in AS Uchl1 detailed annotation. **K)** m^{6}A consensus sequences in invSINEB2 detailed annotation.



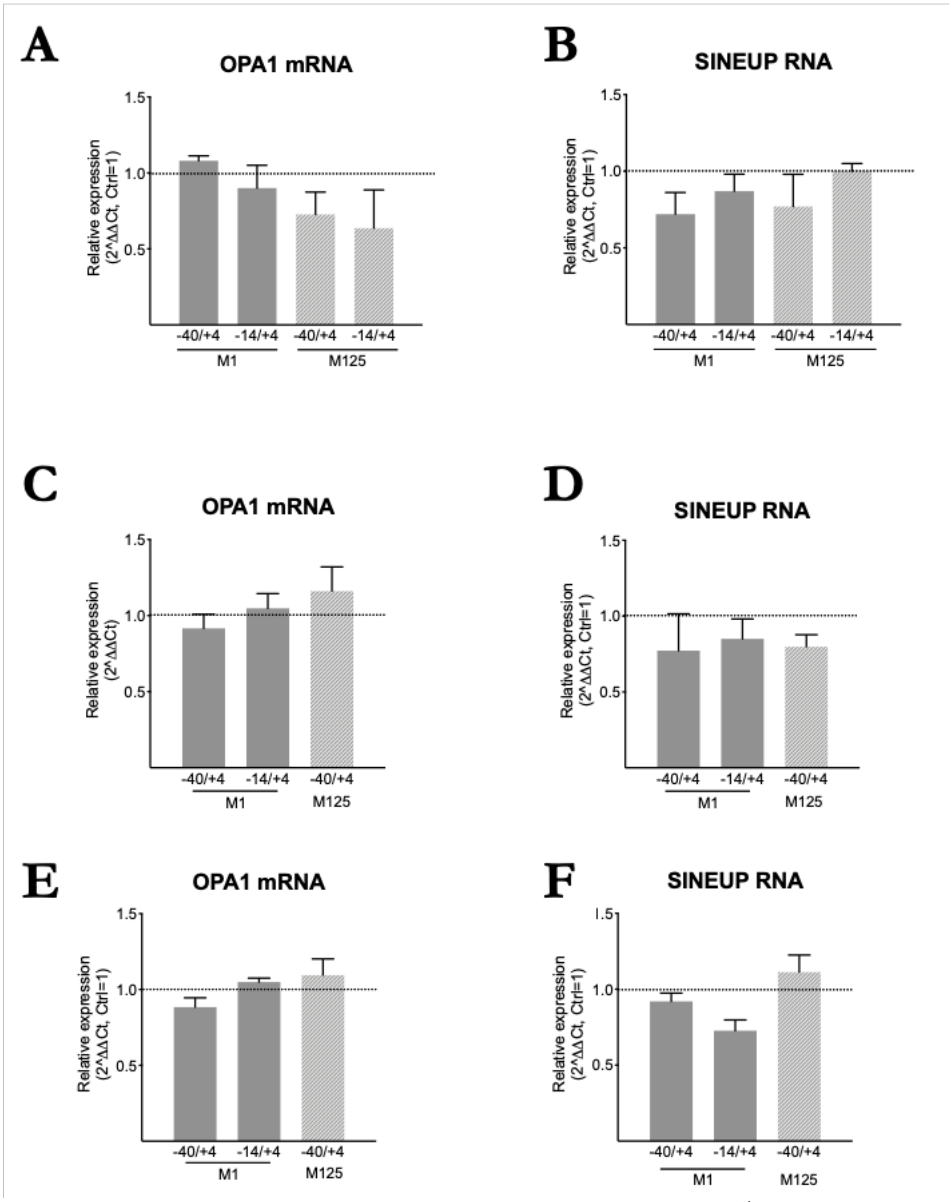
Supplementary Figure 2 Total RNA expression analysis. **A)** *DJ1* mRNA levels were analyzed by qRT-PCR with specific primers. *DJ1* mRNA expression was stable (ns, p>0.05). **B)** miniSINEUP-RNA levels were analyzed by qRT-PCR with specific primers. miniSINEUP-DJ1 expression was stable (ns, p>0.05) **C)** *METTL3* mRNA levels were analyzed by qRT-PCR with specific primers. mRNA expression confirmed 0.7 fold knock-down (****, p>0.0001)



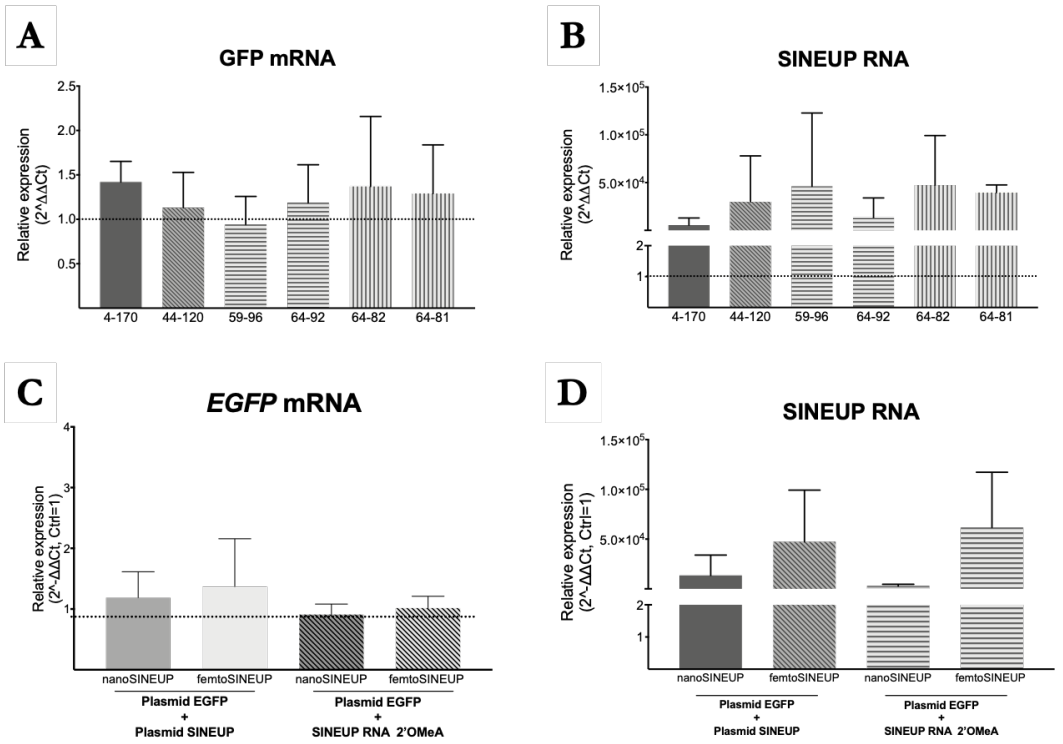
Supplementary Figure 3 Total RNA expression analysis. **A)** *DJ1* mRNA levels were analyzed by qRT-PCR with specific primers. *DJ1* mRNA expression was stable (ns, p>0.05). **B)** miniSINEUP-RNA varaints levels were analyzed by qRT-PCR with specific primers. miniSINEUP-DJ1 expression was stable (ns, p>0.05)



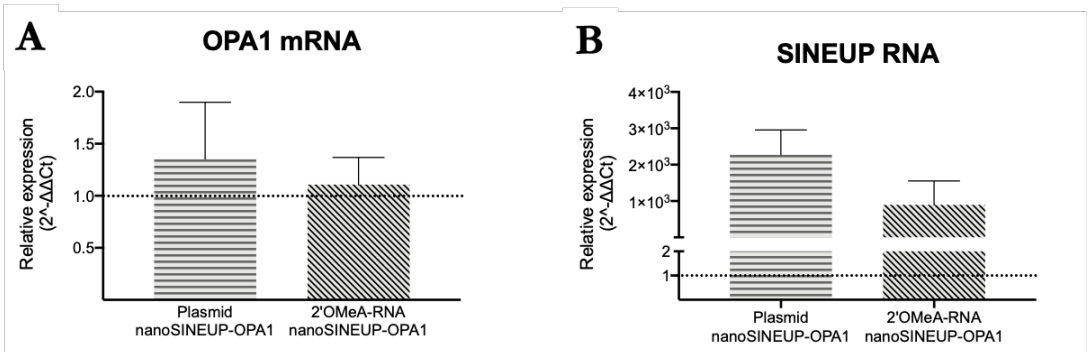
Supplementary Figure 4 **A)** Left: *DJ1* mRNA levels were analyzed by qRT-PCR with specific primers. *DJ1* mRNA expression was stable (ns, p>0.05). Center: SINEUP-RNA levels were analyzed by qRT-PCR with specific primers. miniSINEUP-DJ1 expression was stable (ns, p>0.05) Right: *METTL3* mRNA levels were analyzed by qRT-PCR with specific primers. mRNA expression confirmed -0.7 fold knock-down (****, p>0.0001). Data are normalized on *GAPDH* mRNA expression in total RNA and derive from three independent experiments. Data indicate mean ± SEM from three independent experiments. **B)** *GAPDH* mRNA polysome profiling. Data are expressed as percentages of total RNA in each fraction and derive from three independent experiments. Data indicate mean ± SEM from three independent experiments. **C-E)** Polysome profile of untransfected A549 ShCtrl (black) and ShMETTL3 (purple) knock-down cells. **C)** Representative ribosome fractionation profile with optical density 260 nm, obtained with a 15% to 50% sucrose gradient. Equal volumes from 3 fractions for each Free RNA-40S-60S, 80S, Light Polysomes and Heavy Polysomes were pulled for RNA extraction and qRT-PCR Real-time analysis. **C)** *DJ1* mRNA polysome profiling. **D)** *GAPDH* mRNA polysome profiling. Data are expressed as percentages of total RNA in each fraction and derive from three independent experiments. Data indicate mean ± SEM from two independent experiments.



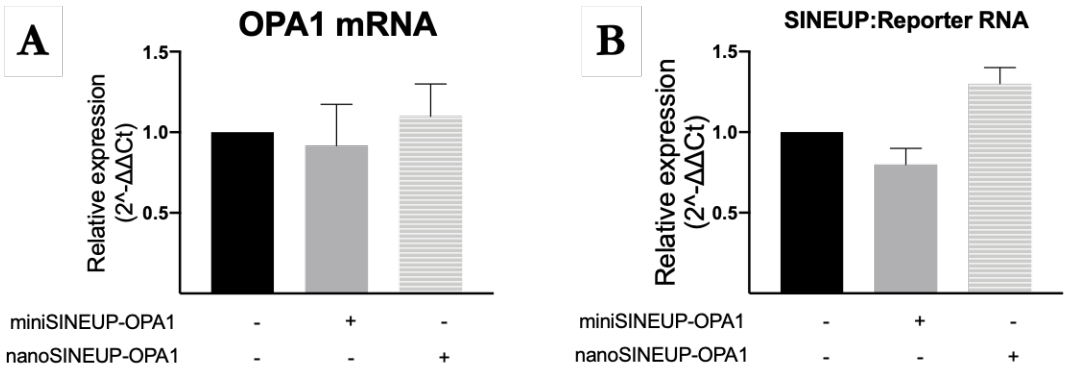
Supplementary Figure 5 Total RNA expression analysis. **A-B)** HEK293T/17 cells were transfected with miniSINEUP-OPA1 variants and control vector (ΔBD) and harvested 48 hours post-transfection. **A)** *OPA1* mRNA levels were analyzed by qRT-PCR with specific primers. *OPA1* mRNA expression was stable (ns, p>0.05). **B)** miniSINEUP-RNA variants expression level was analyzed by qRT-PCR with specific primers. miniSINEUP variants expression was stable (ns, p>0.05) **C-D)** N2A cells were transfected with miniSINEUP-OPA1 variants and control vector (ΔBD) and harvested 48 hours post-transfection. **C)** *OPA1* mRNA levels were analyzed by qRT-PCR with specific primers. *OPA1* mRNA expression was stable (ns, p>0.05). **D)** miniSINEUP-RNA variants expression level was analyzed by qRT-PCR with specific primers. miniSINEUP variants expression was stable (ns, p>0.05). **E-F)** C8-D1A astrocytes were transfected with miniSINEUP-OPA1 variants and control vector (ΔBD) and harvested 48 hours post-transfection. **E)** *OPA1* mRNA levels were analyzed by qRT-PCR with specific primers. *OPA1* mRNA expression was stable (ns, p>0.05). **F)** miniSINEUP-RNA variants expression level was analyzed by qRT-PCR with specific primers. miniSINEUP variants expression was stable (ns, p>0.05)



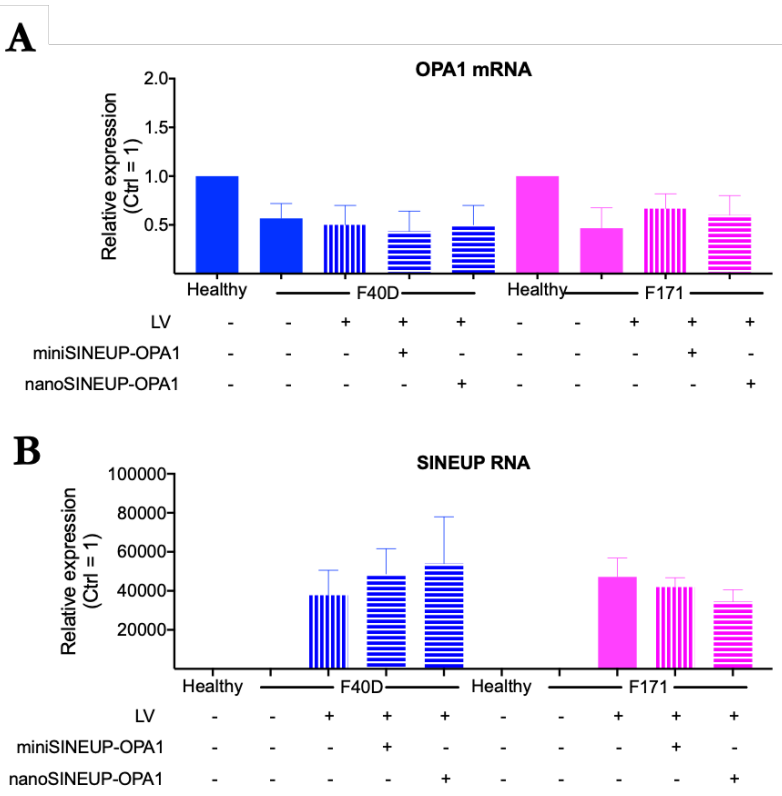
Supplementary Figure 6 Total RNA expression analysis. **A-B)** HEK293T/17 cells were transfected with SINEUP-GFP variants and control vector and harvested 48 hours post-transfection. **A)** *EGFP* mRNA levels were analyzed by qRT-PCR with specific primers. *EGFP* mRNA expression was stable (ns, p>0.05). **B)** SINEUP-RNA variants expression level was analyzed by qRT-PCR with specific primers. SINEUP variants expression was stable (ns, p>0.05). **C-D)** HEK293T/17 cells were transfected with SINEUP-GFP variants and control plasmid vector, expressing the only EGFP, or with control and ASO-SINEUP-GFP. ASO-SINEUP-GFP negative control consists of the only nanoED sequence. Cells were harvested 48 hours post-transfection. **C)** *EGFP* mRNA levels were analyzed by qRT-PCR with specific primers. *EGFP* mRNA expression was stable (ns, p>0.05). **D)** SINEUP-RNA variants expression level was analyzed by qRT-PCR with specific primers. SINEUP variants expression was stable (ns, p>0.05)



Supplementary Figure 7 Total RNA expression analysis. HEK293T/17 cells were transfected with SINEUP-OPA1 variants and control plasmid vector, or with control and ASO-SINEUP-OPA1 variants. ASO-SINEUP-OPA1 negative control consists of the only nanoED sequence. Cells were harvested 48 hours post-transfection. **A)** *OPA1* mRNA levels were analyzed by qRT-PCR with specific primers. *OPA1* mRNA expression was stable (ns, p>0.05). **D)** SINEUP-OPA1 RNA variants expression level was analyzed by qRT-PCR with specific primers. SINEUP variants expression was stable (ns, p>0.05)



Supplementary Figure 8 Total RNA expression analysis. **A-B)** HEK293T/17 cells were transfected with SINEUP-OPA1 variants and control vector and harvested 48 hours post-transfection. **A)** *OPA1* mRNA levels were analyzed by qRT-PCR with specific primers. *OPA1* mRNA expression was stable (ns, p>0.05). **B)** SINEUP-RNA variants:reporter TurboRFP mRNA expression level was analyzed by qRT-PCR with specific primers. SINEUP variants expression was stable (ns, p>0.05)



Supplementary Figure 8 Total RNA expression analysis. Human primary fibroblast lines were transduced with control, mini- and nanoSINEUP-OPA1 expressing lentiviral. Cells were FACS sorted for highest and most stable reporter expression to establish stably expressing cell lines. **A)** *OPA1* mRNA levels were analyzed by qRT-PCR with specific primers. *OPA1* mRNA expression was stable (ns, p>0.05). **D)** SINEUP-OPA1 RNA variants expression level was analyzed by qRT-PCR with specific primers. SINEUP variants expression was stable (ns, p>0.05)

Supplementary tables

SybrGreen qRT-PCR Oligo Name:	*Forward (5'-->3')*	*Reverse (5'-->3')*
hGAPDH	TCTCTGCTCCTCCTGTTC	GCCCAATACGACCAAATCC
mACT	CACACCCGCCACCAGTTC	CCCATTCCCACCATCACACC
45S rRNA	GAACGGTGGTGTGTCGTT	GCGTCTCGTCTCGTCTCACT
hMETTL3	CTGAGGCAGGAGAATTGCTT	GGCAGCCATACACGTTAAGA
hHPRT1	TGACACTGGCAAAACAATGC	GGTCCTTTTCACCAGCAAGCT
hSON	TGACAGATTTGGATAAGGCTCA	GCTCCTCCTGACTTTTTAGCAA
hCREBBP	CATGGCCAAGATGGGAATAA	TGCATCTGAGACATATTTGGC
hRFP	AAGCTGTACATGGAGGGCAC	CCGGGCATCTTGAGGTTCTT
AS Uchl1	CTGGTGTGTATCTCTTATGC	CTCCCGAGTCTCTGTAGC
mUchl1	CCCGCCGATAGAGCCAAG	ATGGTTCACTGGAAAGGG
pTS invB2	CAGTGCTAGAGGAGGTCAGAAGA	GGAGCTAAAGAGATGGCTCAGCACT
Overlap	CTCGGGGTTAATCTCCATCGGC	TCTGCTCCCGTCTCCC
m6A amp	ATATGTTTACAAGCCCCACACCA	TCTGACCTCCTCTAGCACTGA
hOPA1	GACAAAGAGAAAATTGACCAACTTCAGGAAG	CTTAAGCTTTCTATGATGAATGCCTTTGTCA
mOPA1	GCTTCAAAGACCCCAACTAAGGACAC	CCTTGCTGGCCAAAAGTTCCTGCG
nanoSINEUP hOPA1.2	GCCGGCGGGGAATCTG	TGGTGGTTCACAACCACCACG
hDJ1	GAGACGGTCATCCCTGTAG	CATCTTCAAGGCTGGCATC
EGFP	GCCCGACAACCACTACCTGAG	CGGCGGTCACGAACTCCAG
ivt GFP	AGGAGCGCACCATCTTC	GATGCCCTTCAGCTCGAT
BstI RT Assay Oligo Name:	Reverse (5'-->3')	
m6A -	GAGCTAAAGAGATGGCT	
m6A 44-48 +	GGTTACCGTATAACTCC	
m6A 61-65 +	TTCACAACCACCACGA	
m6A 79-83 +	CTCAATATCCATCCACATG	
m6A 109-113 +	TCTTGCACAGGACCAG	
A165+	CCTCCTCTGCTTGT	
A197+	GTGCATGGGGGAG	
A275+	GCTACCATGCCCAG	
A390+	GGTTACCGTATAACTCCAG	
A801+	CTCCCTCTCTGCTTG	
A845+	CAGTTTGCTAAGGAACATAG	
A1231+	CATCGGTTCAATGGAAG	
A759+	AAAGGGCCTTATTACAAAG	
A544+	AGCTCCCTTGCTG	
m6A ASUchl1 -	GGAGCTAAAGAGATGGC	

Supplementary Table 1 Complete list of primers used in this study.

NAME	*BACKBONE*	*BD*	*ED (invB2)*
DJ1 O/L (BD)	pCS2+	DJ1 BD (-40/+4)	-
EGFP	pDUAL	-	-
Empty pCS2+	pCS2+	-	-
femto-DJ1.1	pCS2+	DJ1 BD (-40/+4)	64-81
femto-DJ1.2	pCS2+	DJ1 BD (-40/+4)	64-82
femto-GFP.1	pCS2+	DJ1 BD (-40/+4)	64-81
femto-GFP.1	pDUAL	GFP BD (-40/+32)	64-82
femto-GFP.2	pCS2+	GFP BD (-40/+32)	64-81
femto-GFP.2	pDUAL	GFP BD (-40/+32)	64-82
femto-OPA1H2.1	pCS2+	Human OPA1 BD2 (-14/+4)	64-81
femto-OPA1H2.2	pCS2+	Human OPA1 BD2 (-14/+4)	64-82
GFP O/L (BD)	pCS2+	GFP BD (-40/+32)	-
mini-hOPA1.1	pCS2+	Human OPA1 BD1 (-40/+4)	1-170
mini-hOPA1.2	pCS2+	Human OPA1 BD2 (-14/+4)	1-170
mini-hOPA1.2	pDUAL	Human OPA1 BD2 (-14/+4)	1-170
mini-hOPA1.3	pDUAL	Human OPA1 BD3 (-40/+4)	1-170
mini-hOPA1.3	pCS2+	Human OPA1 BD3 (-40/+4)	1-170
mini-hOPA1.4	pCS2+	Human OPA1 BD3 (-14/+4)	1-170
mini-mOPA1.1	pCS2+	Mouse OPA1 BD1 (-40/+4)	1-170
mini-mOPA1.2	pCS2+	Mouse OPA1 BD2 (-14/+4)	1-170
mini-mOPA1.3	pCS2+	Mouse OPA1 BD3 (-41/+4)	1-170
miniED/⊿BD	pCS2+	-	1-170
mini-SINEUP DJ1	pCS2+	DJ1 BD (-40/+4)	1-170
mini-SINEUP GFP	pDUAL	DJ1 BD (-40/+4)	1-170
nano-DJ1.1	pCS2+	DJ1 BD (-40/+4)	54-96
nano-DJ1.2	pCS2+	DJ1 BD (-40/+4)	64-92
nano-GFP.1	pCS2+	GFP BD (-40/+32)	54-96
nano-GFP.1	pDUAL	GFP BD (-40/+32)	54-96
nano-GFP.2	pCS2+	GFP BD (-40/+32)	64-92
nano-GFP.2	pDUAL	GFP BD (-40/+32)	64-92
nano-OPA1H2.1	pCS2+	Human OPA1 BD2 (-14/+4)	54-96
nano-OPA1H2.2	pCS2+	Human OPA1 BD2 (-14/+4)h2	64-92
pEGFP	pCMV	GFP	-
Plasmid #26046	pMSCV	human OPA1 isoform 1	-
Plasmid #62845	pclbw	myc-tagged Opa1 mouse isoform1	-
pLKOmini-hOPA1.2	pLKO LV	Human OPA1 BD2 (-14/+4)	1-170
pLKOnano-hOPA1.2	pLKO LV	Human OPA1 BD2 (-14/+4)	64-92
pLKO ⊿BD	pLKO LV	-	1-170
pLVmini-hOPA1.2	VB LV	Human OPA1 BD2 (-14/+4)	1-170

pLV nano-hOPA1.2	VB LV	Human OPA1 BD2 (-14/+4)	64-92
pLV ⊿BD	VB LV	-	1-170
SINEUP 005	pCS2+	GFP BD (-40/+32)	FL
micro-hOPA1.2	pDUAL	Human OPA1 BD2 (-14/+4)	44-120
micro-hOPA1.3	pDUAL	Human OPA1 BD3 (-40/+4)	44-120
micro-DJ1	pCS2+	DJ1 BD (-40/+4)	44-120
micro-GFP	pDUAL	GFP BD (-40/+32)	44-120
micro-GFP.1	pCS2+	GFP BD (-40/+32)	44-120
micro-hOPA1.2	pCS2+	Human OPA1 BD2 (-14/+4)	44-120
micro-hOPA1.3	pCS2+	Human OPA1 BD3 (-40/+4)	44-120
miniDJ1-A46U	pCS2+	DJ1 BD (-40/+4)	1-170 A46U
miniDJ1-AAA109-111UUU	pCS2+	DJ1 BD (-40/+4)	1-170 AAA109-111UUU
miniDJ1-A46U;AAA109-111UUU	pCS2+	DJ1 BD (-40/+4)	1-170 A46U;AAA109-111UUU
AS Uchl1	pCDNA3.1(-)	Uchl1 BD (-70/+32)	FL
AS Uchl1 A275U	pCDNA3.1(-)	Uchl1 BD (-70/+32)	FL A275U
AS Uchl1 A46U	pCDNA3.1(-)	Uchl1 BD (-70/+32)	FL A46U
AS Uchl1 AAA109-111UUU	pCDNA3.1(-)	Uchl1 BD (-70/+32)	FL AAA109-111UUU
AS Uchl1 A275U;A46U;AAA109-111UUU	pCDNA3.1(-)	Uchl1 BD (-70/+32)	FL A275U;A46U;AAA109-111UUU

Supplementary Table 2 Complete list of plasmids used in this study

Acknowledgments

This thesis is the result of more than four years of professionally and personally challenging work that wouldn't have been possible without the critical thinking, mentoring, and constructive dialogue with my supervisor Prof. Stefano Gustincich.

I am most grateful to all the other lab members and friends that never missed offering their help and advice: Rosa, Omar, Damiano, Diego, Liliana, Margherita, and all the IIT RNA lab.

I am indebted to all the members of the SINEUP network - SISSA, Università del Piemonte Orientale, IIT, and RIKEN - for thought-provoking discussions. A special mention goes to Prof. Silvia Zucchelli, whom I was lucky enough to meet but whom I wish I had the possibility to work with much longer.

My sincere thanks also go to the administrative and technical staff, especially to Eva (IIT) and Riccardo (SISSA), and to Bozzoni's and Tartaglia's lab members for the friendship and scientific support shared in the last years.

Special thanks must go to my greatest and constant supporters in the lab Luca, Sabrina, Stefano and Carlotta that not only offered their scientific experience but also their friendship and confidence when mine was shaking.

Bibliography

1. Carninci, P. *et al.* The transcriptional landscape of the mammalian genome. *Science (80-.).* **309**, 1559–1563 (2005).
2. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
3. Taft, R. J., Pheasant, M. & Mattick, J. S. The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* **29**, 288–299 (2007).
4. Lee, H., Zhang, Z. & Krause, H. M. Long Noncoding RNAs and Repetitive Elements: Junk or Intimate Evolutionary Partners? *Trends Genet.* **35**, 892–902 (2019).
5. Fatica, A. & Bozzoni, I. Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet. 2013 151* **15**, 7–21 (2013).
6. Volders, P. J. *et al.* An update on LNCipedia: a database for annotated human lncRNA sequences. *Nucleic Acids Res.* **43**, D174–D180 (2015).
7. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223 (2009).
8. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22**, 1775–1789 (2012).
9. Guttman, M. & Rinn, J. L. Modular regulatory principles of large non-coding RNAs. *Nature* **482**, 339–346 (2012).
10. Tsai, M. C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science (80-.).* **329**, 689–693 (2010).
11. Guttman, M. *et al.* lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature* **477**, 295 (2011).
12. 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. 2020 222* **22**, 96–118 (2020).
13. George, T. P. & Thomas, T. Exon Mapping in Long Noncoding RNAs Using Digital Filters. *Genomics Insights* **10**, (2017).
14. Vance, K. W. & Ponting, C. P. Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.* **30**, 348 (2014).
15. Yamazaki, T. *et al.* Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. *Mol. Cell* **70**, 1038-1053.e7

(2018).

16. Clemson, C. M. *et al.* An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **33**, 717–726 (2009).
17. Bernard, D. *et al.* A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J.* **29**, 3082–3093 (2010).
18. Alfeghaly, C. *et al.* Implication of repeat insertion domains in the trans-activity of the long non-coding RNA ANRIL. *Nucleic Acids Res.* **49**, 4954 (2021).
19. Yap, K. L. *et al.* Molecular Interplay of the Non-coding RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of INK4a. *Mol. Cell* **38**, 662 (2010).
20. Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).
21. Cabili, M. N. *et al.* Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol.* **16**, 1–16 (2015).
22. Shukla, C. J. *et al.* High-throughput identification of RNA nuclear enrichment sequences. *EMBO J.* **37**, (2018).
23. Hansen, T. B. *et al.* Natural RNA circles function as efficient microRNA sponges. *Nature* **495**, 384–388 (2013).
24. Lee, S. *et al.* Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **164**, 69–80 (2016).
25. Carrieri, C. *et al.* Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491**, 454–457 (2012).
26. Tichon, A. *et al.* A conserved abundant cytoplasmic long noncoding RNA modulates repression by Pumilio proteins in human cells. *Nat. Commun.* **7**, (2016).
27. J, M., D, M., F, D. & M, B. Non-coding RNAs Shaping Muscle. *Front. cell Dev. Biol.* **7**, (2020).
28. Kino, T., Hurt, D. E., Ichijo, T., Nader, N. & Chrousos, G. P. Noncoding RNA Gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal.* **3**, (2010).
29. Li, Y., Li, W., Hoffman, A. R., Cui, J. & Hu, J. F. The Nucleus/Mitochondria-Shuttling LncRNAs Function as New Epigenetic Regulators of Mitophagy in Cancer. *Front. Cell Dev. Biol.* **9**, 2519 (2021).
30. Bridges, M. C., Daulagala, A. C. & Kourtidis, A. LNCcation: lncRNA localization and function. *J. Cell Biol.* **220**, (2021).

31. Van Heesch, S. *et al.* Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. *Genome Biol.* **15**, (2014).
32. Rinn, J. L. *et al.* Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Non-Coding RNAs. *Cell* **129**, 1311 (2007).
33. Wang, K. C. *et al.* A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. (2011) doi:10.1038/nature09819.
34. Chen, J. Y., Zhang, X., Fu, X. D. & Chen, L. R-ChIP for genome-wide mapping of R-loops by using catalytically inactive RNASEH1. *Nat. Protoc.* **14**, 1661 (2019).
35. Jalali, S., Singh, A., Maiti, S. & Scaria, V. Genome-wide computational analysis of potential long noncoding RNA mediated DNA:DNA:RNA triplexes in the human genome. *J. Transl. Med.* **15**, (2017).
36. Lee, J. T. & Bartolomei, M. S. X-Inactivation, Imprinting, and Long Noncoding RNAs in Health and Disease. *Cell* **152**, 1308–1323 (2013).
37. Tian, D., Sun, S. & Lee, J. T. The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* **143**, 390–403 (2010).
38. 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.* **13**, 95–101 (2011).
39. Gutschner, T. *et al.* The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* **73**, 1180–1189 (2013).
40. Li, R., Zhu, H. & Luo, Y. Understanding the Functions of Long Non-Coding RNAs through Their Higher-Order Structures. *Int. J. Mol. Sci.* **17**, (2016).
41. D, K. & J, R. Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* **13**, R107 (2012).
42. Kapusta, A. *et al.* Transposable Elements Are Major Contributors to the Origin, Diversification, and Regulation of Vertebrate Long Noncoding RNAs. *PLoS Genet.* **9**, 1003470 (2013).
43. Wicker, T. *et al.* A unified classification system for eukaryotic transposable elements. (2007).
44. Huang, C. R. L., Burns, K. H. & Boeke, J. D. Active Transposition in Genomes. *Annu. Rev. Genet.* **46**, 651 (2012).
45. Fort, V., Khelifi, G. & Hussein, S. M. I. Long non-coding RNAs and transposable elements: A functional relationship. *Biochim. Biophys. Acta - Mol. Cell Res.* **1868**, 118837 (2021).

46. Hwang, S. Y. *et al.* L1 retrotransposons exploit RNA m6A modification as an evolutionary driving force. *Nat. Commun.* **12**, (2021).
47. Kramerov, D. A. & Vassetzky, N. S. Origin and evolution of SINEs in eukaryotic genomes. *Heredity (Edinb).* **107**, 487 (2011).
48. Makałowski, W., Gotea, V., Pande, A. & Makałowska, I. Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics. *Methods Mol. Biol.* **1910**, 177–207 (2019).
49. Krayev, A. S. *et al.* Ubiquitous transposon-like repeats B1 and B2 of the mouse genome: B2 sequencing. *Nucleic Acids Res.* **10**, 7461 (1982).
50. Allen, T. A., Von Kaenel, S., Goodrich, J. A. & Kugel, J. F. The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nat. Struct. Mol. Biol.* **11**, 816–821 (2004).
51. Mariner, P. D. *et al.* Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. *Mol. Cell* **29**, 499–509 (2008).
52. Ali, A., Han, K. & Liang, P. Role of Transposable Elements in Gene Regulation in the Human Genome. *Life* **11**, 1–23 (2021).
53. Kapusta, A. *et al.* Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet* **9**, e1003470 (2013).
54. Carrieri, C. *et al.* Expression analysis of the long non-coding RNA antisense to Uchl1 (AS Uchl1) during dopaminergic cells' differentiation in vitro and in neurochemical models of Parkinson's disease. *Front. Cell. Neurosci.* **9**, 1–11 (2015).
55. Katayama, S. *et al.* Antisense transcription in the mammalian transcriptome. *Science (80-.).* **309**, 1564–1566 (2005).
56. Chen, J., Sun, M., Hurst, L. D., Carmichael, G. G. & Rowley, J. D. Genome-wide analysis of coordinate expression and evolution of human cis-encoded sense-antisense transcripts. *Trends Genet.* **21**, 326–329 (2005).
57. Galante, P. A. F., Vidal, D. O., de Souza, J. E., Camargo, A. A. & de Souza, S. J. Sense-antisense pairs in mammals: functional and evolutionary considerations. *Genome Biol.* **8**, R40 (2007).
58. Pelechano, V. & Steinmetz, L. M. Gene regulation by antisense transcription. *Nat. Rev. Genet.* **14**, 880–893 (2013).
59. Schein, A., Zucchelli, S., Kauppinen, S., Gustincich, S. & Carninci, P. Identification of antisense long noncoding RNAs that function as SINEUPs in human cells. *Sci.*

Rep. **6**, 1–8 (2016).

60. Patrucco, L. *et al.* Engineering mammalian cell factories with SINEUP noncoding RNAs to improve translation of secreted proteins. *Gene* **569**, 287–293 (2015).
61. Fasolo, F. *et al.* The RNA-binding protein ILF3 binds to transposable element sequences in SINEUP lncRNAs. *FASEB J* **33**, 13572–13589 (2019).
62. Zarantonello, G. *et al.* Natural SINEUP RNAs in Autism Spectrum Disorders: RAB11B-AS1 Dysregulation in a Neuronal CHD8 Suppression Model Leads to RAB11B Protein Increase. *Front. Genet.* **12**, (2021).
63. Zucchelli, S. *et al.* SINEUPs are modular antisense long non-coding RNAs that increase synthesis of target proteins in cells. *Front Cell Neurosci* **9**, 174 (2015).
64. Zucchelli, S. *et al.* SINEUPs: A new class of natural and synthetic antisense long non-coding RNAs that activate translation. *RNA Biol* **12**, 771–779 (2015).
65. Sasso, E. *et al.* A long non-coding SINEUP RNA boosts semi-stable production of fully human monoclonal antibodies in HEK293E cells. *MAbs* **10**, 730–737 (2018).
66. Hoseinpoor, R., Kazemi, B., Rajabibazl, M. & Rahimpour, A. Improving the expression of anti-IL-2Rα monoclonal antibody in the CHO cells through optimization of the expression vector and translation efficiency. *J. Biotechnol.* **324**, 112–120 (2020).
67. Espinoza, S. *et al.* SINEUP Non-coding RNA Targeting GDNF Rescues Motor Deficits and Neurodegeneration in a Mouse Model of Parkinson's Disease. *Mol Ther* **28**, 642–652 (2020).
68. Bon, C. *et al.* SINEUP non-coding RNAs rescue defective frataxin expression and activity in a cellular model of Friedreich's Ataxia. *Nucleic Acids Res* **47**, 10728–10743 (2019).
69. Espinoza, S. *et al.* SINEUPs: a novel toolbox for RNA therapeutics. *Essays Biochem.* **65**, 775–789 (2021).
70. Takahashi, H. *et al.* Identification of functional features of synthetic SINEUPs, antisense lncRNAs that specifically enhance protein translation. *PLoS One* **13**, e0183229 (2018).
71. Forrest, A. R. *et al.* A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).
72. Severin, J. *et al.* Interactive visualization and analysis of large-scale sequencing datasets using ZENBU. *Nat Biotechnol* **32**, 217–219 (2014).

73. Podbevšek, P. *et al.* Structural determinants of the SINE B2 element embedded in the long non-coding RNA activator of translation AS Uchl1. *Sci Rep* **8**, 3189 (2018).
74. Reißer, S., Zucchelli, S., Gustincich, S. & Bussi, G. Conformational ensembles of an RNA hairpin using molecular dynamics and sparse NMR data. *Nucleic Acids Res.* **48**, 1164–1174 (2020).
75. Ohyama, T. *et al.* An NMR-based approach reveals the core structure of the functional domain of SINEUP lncRNAs. *Nucleic Acids Res* **48**, 9346–9360 (2020).
76. D'Agostino, S. *et al.* Internal Ribosome Entry Site RNAs act in trans through an antisense sequence in linear and circular non-coding RNAs.
77. Toki, N. *et al.* SINEUP long non-coding RNA acts via PTBP1 and HNRNPK to promote translational initiation assemblies. *Nucleic Acids Res* **48**, 11626–11644 (2020).
78. Russell, S. *et al.* Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet* **390**, 849–860 (2017).
79. Damase, T. R. *et al.* The Limitless Future of RNA Therapeutics. *Front. Bioeng. Biotechnol.* **9**, 161 (2021).
80. Gustincich, S., Zucchelli, S. & Mallamaci, A. The Yin and Yang of nucleic acid-based therapy in the brain. *Prog Neurobiol* **155**, 194–211 (2017).
81. Springer, A. D. & Dowdy, S. F. GalNAc-siRNA Conjugates: Leading the Way for Delivery of RNAi Therapeutics. *Nucleic Acid Ther* **28**, 109–118 (2018).
82. Winkle, M., El-Daly, S. M., Fabbri, M. & Calin, G. A. Noncoding RNA therapeutics — challenges and potential solutions. *Nat. Rev. Drug Discov. 2021 208* **20**, 629–651 (2021).
83. Kaczmarek, J. C., Kowalski, P. S. & Anderson, D. G. Advances in the delivery of RNA therapeutics: from concept to clinical reality. *Genome Med. 2017 91* **9**, 1–16 (2017).
84. Bajan, S. & Hutvagner, G. RNA-Based Therapeutics: From Antisense Oligonucleotides to miRNAs. *Cells* **9**, (2020).
85. Kole, R. & Krieg, A. M. Exon skipping therapy for Duchenne muscular dystrophy. *Adv. Drug Deliv. Rev.* **87**, 104–107 (2015).
86. Du, L. *et al.* Arginine-rich cell-penetrating peptide dramatically enhances AMO-mediated ATM aberrant splicing correction and enables delivery to brain and cerebellum. *Hum. Mol. Genet.* **20**, 3151 (2011).

87. Peacey, E., Rodriguez, L., Liu, Y. & Wolfe, M. S. Targeting a pre-mRNA structure with bipartite antisense molecules modulates tau alternative splicing. *Nucleic Acids Res.* **40**, 9836–9849 (2012).

88. Li, L.-C. C. *et al.* Small dsRNAs induce transcriptional activation in human cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17337–17342 (2006).

89. Janowski, B. A. *et al.* Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat. Chem. Biol. 2007 33* **3**, 166–173 (2007).

90. Portnoy, V., Huang, V., Place, R. F. & Li, L. C. Small RNA and transcriptional upregulation. *Wiley Interdiscip. Rev. RNA* **2**, 748 (2011).

91. Diodato, A., Pinzan, M., Granzotto, M. & Mallamaci, A. Promotion of cortico-cerebral precursors expansion by artificial pri-miRNAs targeted against the Emx2 locus. *Curr. Gene Ther.* **13**, 152–161 (2013).

92. Schwartz, J. C. *et al.* Antisense transcripts are targets for activating small RNAs. *Nat. Struct. Mol. Biol.* **15**, 842 (2008).

93. Yue, X. *et al.* Transcriptional regulation by small RNAs at sequences downstream from 3' gene termini. *Nat. Chem. Biol.* **6**, 621–629 (2010).

94. Liu, M. *et al.* The IGF2 intronic miR-483 selectively enhances transcription from IGF2 fetal promoters and enhances tumorigenesis. *Genes Dev.* **27**, 2543 (2013).

95. Modarresi, F. *et al.* Natural Antisense Inhibition Results in Transcriptional De-Repression and Gene Upregulation. *Nat. Biotechnol.* **30**, 453 (2012).

96. Fimiani, C., Goina, E. & Mallamaci, A. Upregulating endogenous genes by an RNA-programmable artificial transactivator. *Nucleic Acids Res.* **43**, 7850 (2015).

97. LeCuyer, K. A., Behlen, L. S. & Uhlenbeck, O. C. Mutants of the Bacteriophage MS2 Coat Protein That Alter Its Cooperative Binding to RNA. *Biochemistry* **34**, 10600–10606 (2002).

98. Carey, J., Uhlenbeck, O. C., Cameron, V. & de Haseth, P. L. Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site. *Biochemistry* **22**, 2601–2610 (1983).

99. Indrieri, A. *et al.* Synthetic long non-coding RNAs [SINEUPs] rescue defective gene expression in vivo. *Sci Rep* **6**, 27315 (2016).

100. Toki, N., Takahashi, H., Zucchelli, S., Gustincich, S. & Carninci, P. Synthetic in vitro transcribed lncRNAs (SINEUPs) with chemical modifications enhance target mRNA translation. *FEBS Lett.* **594**, 4357–4369 (2020).

101. Valentini, P. *et al.* Towards SINEUP-based therapeutics: Design of an in vitro

synthesized SINEUP RNA. *Mol. Ther. - Nucleic Acids* **27**, 1092–1102 (2022).

102. Cohn, W. E. & Volkin, E. Nucleoside-5'-Phosphates from Ribonucleic Acid. *Nat. 1951 1674247* **167**, 483–484 (1951).
103. Cantara, W. A. *et al.* The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res.* **39**, (2011).
104. Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* **46**, D303 (2018).
105. Mathlin, J., Le Pera, L. & Colombo, T. A Census and Categorization Method of Epitranscriptomic Marks. *Int. J. Mol. Sci.* **21**, 1–31 (2020).
106. Barbieri, I. & Kouzarides, T. Role of RNA modifications in cancer. *Nat. Rev. Cancer* **20**, 303–322 (2020).
107. Zhang, L. S. *et al.* Transcriptome-wide Mapping of Internal N 7-Methylguanosine Methylome in Mammalian mRNA. *Mol. Cell* **74**, 1304-1316.e8 (2019).
108. Pandolfini, L. *et al.* METTL1 Promotes let-7 MicroRNA Processing via m7G Methylation. *Mol. Cell* **74**, 1278 (2019).
109. Roundtree, I. A., Evans, M. E., Pan, T. & He, C. Dynamic RNA modifications in gene expression regulation. *Cell* **169**, 1187 (2017).
110. Jacob, R., Zander, S. & Gutschner, T. The Dark Side of the Epitranscriptome: Chemical Modifications in Long Non-Coding RNAs. *Int J Mol Sci* **18**, (2017).
111. Kim, G. W. & Siddiqui, A. N6-methyladenosine modification of HCV RNA genome regulates cap-independent IRES-mediated translation via YTHDC2 recognition. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
112. Li, N. & Rana, T. M. Detection of N6-methyladenosine in SARS-CoV-2 RNA by methylated RNA immunoprecipitation sequencing. *STAR Protoc.* **3**, 101067 (2022).
113. Wang, X. *et al.* Structural basis of N6-adenosine methylation by the METTL3–METTL14 complex. *Nat. 2016 5347608* **534**, 575–578 (2016).
114. Meyer, K. D. & Jaffrey, S. R. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat. Rev. Mol. Cell Biol.* **15**, 313 (2014).
115. Liu, N. *et al.* N6 -methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).
116. Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N. & Rechavi, G. Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. *Nat Protoc* **8**, 176–189 (2013).

117. Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
118. Zaccara, S., Ries, R. J. & Jaffrey, S. R. Reading, writing and erasing mRNA methylation. *Nat Rev Mol Cell Biol* **20**, 608–624 (2019).
119. Liu, J. *et al.* A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* **10**, 93–95 (2014).
120. Śledź, P. & Jinek, M. Structural insights into the molecular mechanism of the m(6)A writer complex. *Elife* **5**, (2016).
121. Zhao, B. S., Roundtree, I. A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **18**, 31–42 (2017).
122. Liu, N. *et al.* Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* **19**, 1848–1856 (2013).
123. Linder, B. *et al.* Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods* **12**, 767–772 (2015).
124. Harcourt, E. M., Ehrenschwender, T., Batista, P. J., Chang, H. Y. & Kool, E. T. Identification of a selective polymerase enables detection of N(6)-methyladenosine in RNA. *J Am Chem Soc* **135**, 19079–19082 (2013).
125. Castellanos-Rubio, A. *et al.* A novel RT-QPCR-based assay for the relative quantification of residue specific m6A RNA methylation. *Sci Rep* **9**, 4220 (2019).
126. Garcia-Campos, M. A. *et al.* Deciphering the "m. *Cell* **178**, 731-747.e16 (2019).
127. Price, A. M. *et al.* Direct RNA sequencing reveals m. *Nat Commun* **11**, 6016 (2020).
128. Xu, L. & Seki, M. Recent advances in the detection of base modifications using the Nanopore sequencer. *J Hum Genet* **65**, 25–33 (2020).
129. Ueda, H. nanoDoc: RNA modification detection using Nanopore raw reads with Deep One-Class Classification. *bioRxiv* 2020.09.13.295089 (2020) doi:10.1101/2020.09.13.295089.
130. Leger, A. *et al.* RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat. Commun.* **12**, 7198 (2021).
131. Sommer, S., Lavi, U. & Darnell, J. E. The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. *J. Mol. Biol.* **124**, 487–499 (1978).
132. Wang, X. *et al.* N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2014).

133. Miano, V., Codino, A., Pandolfini, L. & Barbieri, I. The non-coding epitranscriptome in cancer. *Brief. Funct. Genomics* **20**, 94–105 (2021).
134. Patil, D. P. *et al.* m6A RNA methylation promotes XIST-mediated transcriptional repression. *Nat. 2016 5377620* **537**, 369–373 (2016).
135. Di Timoteo, G. *et al.* Modulation of circRNA Metabolism by m. *Cell Rep* **31**, 107641 (2020).
136. Liu, J. *et al.* N6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580 (2020).
137. Choe, J. *et al.* mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* **561**, 556–560 (2018).
138. Wang, X. *et al.* N6-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388–1399 (2015).
139. Meyer, K. D. *et al.* 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* **163**, 999–1010 (2015).
140. Aitken, C. E. & Lorsch, J. R. A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.* **19**, 568–576 (2012).
141. JAEGER, W. Hereditary optic atrophy with dominant transmission; with special reference to the associated color-sense disorder. *Albrecht Von Graefes. Arch. Ophthalmol.* **155**, 457–484 (1954).
142. KJER, P. Infantile optic atrophy with dominant mode of inheritance: a clinical and genetic study of 19 Danish families. *Acta Ophthalmol Suppl* **164**, 1–147 (1959).
143. Chun, B. Y. & Rizzo, J. F. Dominant Optic Atrophy and Leber's Hereditary Optic Neuropathy: Update on Clinical Features and Current Therapeutic Approaches. *Semin Pediatr Neurol* **24**, 129–134 (2017).
144. Lenaers, G. *et al.* Dominant optic atrophy. *Orphanet J Rare Dis* **7**, 46 (2012).
145. Yu-Wai-Man, P. & Chinnery, P. F. Dominant optic atrophy: novel OPA1 mutations and revised prevalence estimates. *Ophthalmology* **120**, 1712-1712.e1 (2013).
146. Alexander, C. *et al.* OPA1, encoding a dynamin-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28. *Nat. Genet.* **26**, 211–215 (2000).
147. Delettre, C. *et al.* Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy. *Nat Genet* **26**, 207–210 (2000).
148. Civiletto, G. *et al.* Opa1 overexpression ameliorates the phenotype of two

mitochondrial disease mouse models. *Cell Metab.* **21**, 845–854 (2015).

149. Jonikas, M. *et al.* Stem cell modeling of mitochondrial parkinsonism reveals key functions of OPA1. *Ann Neurol* **83**, 915–925 (2018).
150. Tezze, C. *et al.* Age-Associated Loss of OPA1 in Muscle Impacts Muscle Mass, Metabolic Homeostasis, Systemic Inflammation, and Epithelial Senescence. *Cell Metab* **25**, 1374-1389.e6 (2017).
151. Weisschuh, N. *et al.* Mutation spectrum of the OPA1 gene in a large cohort of patients with suspected dominant optic atrophy: Identification and classification of 48 novel variants. *PLoS One* **16**, (2021).
152. Maloney, D. M. *et al.* Optimized OPA1 Isoforms 1 and 7 Provide Therapeutic Benefit in Models of Mitochondrial Dysfunction. *Front. Neurosci.* **14**, 1213 (2020).
153. Cohn, A. C. *et al.* Autosomal dominant optic atrophy: penetrance and expressivity in patients with OPA1 mutations. *Am J Ophthalmol* **143**, 656–662 (2007).
154. Yang, S., Zhou, J. & Li, D. Functions and Diseases of the Retinal Pigment Epithelium. *Front. Pharmacol.* **12**, 1976 (2021).
155. Almind, G. J. *et al.* Dominant optic atrophy in Denmark - report of 15 novel mutations in OPA1, using a strategy with a detection rate of 90%. *BMC Med Genet* **13**, 65 (2012).
156. Ferré, M. *et al.* Improved locus-specific database for OPA1 mutations allows inclusion of advanced clinical data. *Hum Mutat* **36**, 20–25 (2015).
157. Le Roux, B. *et al.* OPA1: 516 unique variants and 831 patients registered in an updated centralized Variome database. *Orphanet J. Rare Dis.* **14**, 1–9 (2019).
158. Ferré, M. *et al.* Molecular screening of 980 cases of suspected hereditary optic neuropathy with a report on 77 novel OPA1 mutations. *Hum Mutat* **30**, E692-705 (2009).
159. Yu-Wai-Man, P. Therapeutic Approaches to Inherited Optic Neuropathies. *Semin Neurol* **35**, 578–586 (2015).
160. Yu-Wai-Man, P., Griffiths, P. G. & Chinnery, P. F. Mitochondrial optic neuropathies - disease mechanisms and therapeutic strategies. *Prog Retin Eye Res* **30**, 81–114 (2011).
161. Carelli, V. *et al.* Syndromic parkinsonism and dementia associated with OPA1 missense mutations. *Ann Neurol* **78**, 21–38 (2015).
162. Wu, W. *et al.* OPA1 overexpression ameliorates mitochondrial cristae remodeling, mitochondrial dysfunction, and neuronal apoptosis in prion diseases. *Cell Death Dis*

10, 710 (2019).

163. Li, D., Wang, J., Jin, Z. & Zhang, Z. Structural and evolutionary characteristics of dynamin-related GTPase OPA1. *PeerJ* **2019**, (2019).
164. Del Dotto, V., Fogazza, M., Carelli, V., Rugolo, M. & Zanna, C. Eight human OPA1 isoforms, long and short: What are they for? *Biochim Biophys Acta* **1859**, 263–269 (2018).
165. Laforge, M. *et al.* NF-κB pathway controls mitochondrial dynamics. *Cell Death Differ.* **23**, 89–98 (2016).
166. Hu, X., Dai, Y., Zhang, R., Shang, K. & Sun, X. Overexpression of Optic Atrophy Type 1 Protects Retinal Ganglion Cells and Upregulates Parkin Expression in Experimental Glaucoma. *Front. Mol. Neurosci.* **11**, 1–12 (2018).
167. Cogliati, S. *et al.* Mitochondrial cristae shape determines respiratory chain supercomplexes assembly and respiratory efficiency. *Cell* **155**, 160–171 (2013).
168. Del Dotto, V. *et al.* OPA1: How much do we know to approach therapy? *Pharmacol Res* **131**, 199–210 (2018).
169. Sarzi, E. *et al.* OPA1 gene therapy prevents retinal ganglion cell loss in a Dominant Optic Atrophy mouse model. *Sci Rep* **8**, 2468 (2018).
170. Varanita, T. *et al.* The OPA1-dependent mitochondrial cristae remodeling pathway controls atrophic, apoptotic, and ischemic tissue damage. *Cell Metab* **21**, 834–844 (2015).
171. Yu-Wai-Man, P. *et al.* A neurodegenerative perspective on mitochondrial optic neuropathies. *Acta Neuropathol* **132**, 789–806 (2016).
172. Zanna, C. *et al.* OPA1 mutations associated with dominant optic atrophy impair oxidative phosphorylation and mitochondrial fusion. *Brain* **131**, 352–367 (2008).
173. Chen, J., Riazifar, H., Guan, M. X. & Huang, T. Modeling autosomal dominant optic atrophy using induced pluripotent stem cells and identifying potential therapeutic targets. *Stem Cell Res Ther* **7**, 2 (2016).
174. Smith, T. G., Seto, S., Ganne, P. & Votruba, M. A randomized, placebo-controlled trial of the benzoquinone idebenone in a mouse model of OPA1-related dominant optic atrophy reveals a limited therapeutic effect on retinal ganglion cell dendropathy and visual function. *Neuroscience* **319**, 92–106 (2016).
175. Dotto, V. Del & Carelli, V. Dominant Optic Atrophy (DOA): Modeling the Kaleidoscopic Roles of OPA1 in Mitochondrial Homeostasis. *Front. Neurol.* **12**, 681326 (2021).

176. Alavi, M. V. *et al.* A splice site mutation in the murine Opa1 gene features pathology of autosomal dominant optic atrophy. *Brain* **130**, 1029–1042 (2007).
177. Davies, V. J. *et al.* Opa1 deficiency in a mouse model of autosomal dominant optic atrophy impairs mitochondrial morphology, optic nerve structure and visual function. *Hum. Mol. Genet.* **16**, 1307–1318 (2007).
178. Sarzi, E. *et al.* The human OPA1delTTAG mutation induces premature age-related systemic neurodegeneration in mouse. *Brain* **135**, 3599–3613 (2012).
179. Del Dotto, V. *et al.* OPA1 Isoforms in the Hierarchical Organization of Mitochondrial Functions. *Cell Rep* **19**, 2557–2571 (2017).
180. Mayer, A. & Churchman, L. S. A Detailed Protocol for Subcellular RNA Sequencing (subRNA-seq). *Curr Protoc Mol Biol* **120**, 4.29.1-4.29.18 (2017).
181. Panda, A., Martindale, J. & Gorospe, M. Polysome Fractionation to Analyze mRNA Distribution Profiles. *Bio-protocol* **7**, (2017).
182. Meyer, K. D. & Jaffrey, S. R. Rethinking m 6 A Readers, Writers, and Erasers. doi:10.1146/annurev-cellbio-100616-060758.
183. Zhou, Y., Zeng, P., Li, Y. H., Zhang, Z. & Cui, Q. SRAMP: prediction of mammalian N6-methyladenosine (m6A) sites based on sequence-derived features. *Nucleic Acids Res* **44**, e91 (2016).
184. Shu, X. *et al.* N 6-Allyladenosine: A New Small Molecule for RNA Labeling Identified by Mutation Assay. *J Am Chem Soc* **139**, 17213–17216 (2017).
185. Zhang, Z. *et al.* Single-base mapping of m. *Sci Adv* **5**, eaax0250 (2019).
186. Toki, N. *et al.* SINEUP long non-coding RNA acts via PTBP1 and HNRNPK to promote translational initiation assemblies. *Nucleic Acids Res* **48**, 11626–11644 (2020).
187. Liu, N. *et al.* N6-methyladenosine-dependent RNA structural switches regulate RNA–protein interactions.
188. Podbevsek, P. *et al.* Structural determinants of the SINE B2 element embedded in the long non-coding RNA activator of translation AS Uchl1. *Sci Rep* **8**, 3189 (2018).
189. Klopstock, T. *et al.* Persistence of the treatment effect of idebenone in Leber's hereditary optic neuropathy. *Brain* **136**, (2013).
190. Romagnoli, M. *et al.* Idebenone increases chance of stabilization/recovery of visual acuity in OPA1-dominant optic atrophy. *Ann. Clin. Transl. Neurol.* **7**, 590 (2020).
191. Smith, T. G., Seto, S., Ganne, P. & Votruba, M. A randomized, placebo-controlled trial of the benzoquinone idebenone in a mouse model of OPA1-related dominant

optic atrophy reveals a limited therapeutic effect on retinal ganglion cell dendropathy and visual function. *Neuroscience* **319**, 92–106 (2016).

192. Yu-Wai-Man, P. *et al.* Bilateral visual improvement with unilateral gene therapy injection for Leber hereditary optic neuropathy. *Sci. Transl. Med.* **12**, (2020).

193. Newman, N. J. *et al.* Efficacy and Safety of Intravitreal Gene Therapy for Leber Hereditary Optic Neuropathy Treated within 6 Months of Disease Onset. *Ophthalmology* **128**, 649–660 (2021).

194. Cipolat, S., De Brito, O. M., Dal Zilio, B. & Scorrano, L. OPA1 requires mitofusin 1 to promote mitochondrial fusion. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15927 (2004).