IDENTIFICATION AND FUNCTIONAL ANALYSIS OF CODING GENES AND LONG NON-CODING RNAS INVOLVED IN NEURONAL DEVELOPMENT

UNIVERSITY OF NAPLES FEDERICO II

PHD IN BIOLOGY

XXXI CYCLE

PhD Student

Dr. **V**incenza **A**liperti

Tutor

Prof. **A**ldo **D**onizetti

Coordinator

Prof. **S**alvatore **C**ozzolino

INDEX

ABSTRACT (English) 4
ABSTRACT (Italian) 5
ABBREVIATIONS 6
1. INTRODUCTION 11
1.1 Human Neurodevelopment 11
1.2 Programmed cell death in neurodevelopment 13
1.3 Regulatory landscape of human neurodevelopment 14
1.4 Role of neurotrophins in neuronal development 17
1.5 Immediate Early genes in brain 19
2. AIM OF THE STUDY 21
3. MATERIALS & METHODS 23
3.1 Cell cultures 23
3.2 Cellular treatments 23
3.3 RNA isolation, retrotranscription and quantitative PCR (qPCR) analysis 24
3.4 Microarray analysis 24
3.5 lncRNAs classification and functional analysis 25
3.6 Generation of EGR1 and LINC00473 KO cell lines 25
3.7 iPSCs 26
3.8 Western Blotting 26
3.9 EGR1 iper-expression: cloning, transfection and subcellular fractionation protocol 27
3.10 Immunofluorescence analysis 27
3.11 Statistical analysis 28
TABLES 29
4. RESULTS 31
PART I 31
Identification and preliminary characterization of genes involved in Early Response to BDNF Stimulation 31
4.I.1 Identification on new players in early response to BDNF stimulation 31
PART II 40
EGR1 in neuronal differentiation 40
4.II.1 *EGR1* 40
4.II.2 *EGR1* expression level during RA-dependent differentiation 41
4.II.3 EGR1 knockout cell line generation and validation 41
4.II.4 KO cells are unable to properly differentiate under RA stimulation 43
4.II.5 Effects of EGR1-KO on the RA pathway 44
4.II.6 Dysregulation of BDNF system in EGR1-KO cells 45

4.II.7 Effect of TPA stimulation on EGR1-KO differentiation 46
4.II.8 EGR1 knockout in iPSCs 47
4.II.9 Identification of a new alternative splicing isoform for *EGR1* 48
PART III 52
Preliminary characterization of LINC00473 (C6orf176) 52
4.III.1 *LINC00473* overview 52
4.III.2 *LINC00473* expression levels under RA and BDNF stimulation 53
4.III.3 LINC00473 expression levels under FBS stimulation 54
4.III.4 LINC00473 KO 55
5. DISCUSSION 58
6. REFERENCES 65

ABSTRACT (English)

Neuronal differentiation is a complex process characterized by different cellular events. It is finely regulated by numerous transcription factors, many of which have been identified while others remain unknown. Recently, long non-coding RNAs (lncRNAs) have emerged as key regulators of gene expression, with an essential role in cell differentiation. Therefore, the identification and characterization of these transcripts and the functional relationships between mRNA and lncRNA are fundamental to the understanding the complex transcriptional processes underlying gene regulation. In this regard, an essential molecule for neuronal development, the neurotrophin BDNF (Brain-Derived Neurotrophic Factor), is known for its action in post-transcriptional regulation, while little is known about the transcriptional programs it triggers and through which it influences the development and survival of neurons. The results presented in this thesis concerns the identification of several Immediate Early Genes (IEGs), belonging to both coding and non-coding that, engaged by BDNF, can play a fundamental role in the biology of this neurotrophin. Among coding genes, the involvement of the transcription factor EGR1 (Early growth response protein 1) in neuronal differentiation was investigated. Taking advantage of genome editing techniques (CRISPR-Cas9) and the SH-SY5Y cell line as a human neuronal model, I showed that the knockout cells for EGR1 are unable to differentiate, underlining an essential role of the transcriptional programs regulated by EGR1 in the survival of neurons. Furthermore, I identified a new alternative splicing isoform for EGR1 likely involved in neuronal differentiation. The preliminary results suggested that this isoform may act as a dominant negative of the canonical isoform, contributing to a fine regulation of its effect on transcriptional regulation. In the final part of my thesis, I focused my efforts on a primary characterization of the most differentially expressed lncRNA under BFNF stimulation, the *LINC00473* gene. I provided evidence that this gene may encode for a little protein highly conserved in primate species. Functional analysis by the use of a KO cell line generated by CRISPR/Cas9 showed that the putative protein affected the transcript level of the IEGs with a relevant impact on the regulation of gene expression.

ABSTRACT (Italian)

Il differenziamento neuronale è un processo complesso caratterizzato da diversi eventi cellulari. È finemente regolato da numerosi fattori trascrizionali, molti dei quali sono stati identificati, mentre altri restano ancora sconosciuti. Recentemente, i long non-coding RNAs (lncRNAs) sono emersi come regolatori chiave dell'espressione genica, con un ruolo essenziale nel differenziamento cellulare. Pertanto, l'identificazione e la caratterizzazione di tali trascritti e delle relazioni funzionali tra mRNA e lncRNA, risultano fondamentali per la comprensione dei complessi processi trascrizionali alla base della regolazione genica. A tal proposito, una molecola essenziale per lo sviluppo neuronale, la neurotrofina BDNF (*Brain-Derived Neurotrophic Factor*) è nota per la sua azione mediante la regolazione post-trascrizionale, mentre poco si conosce sui programmi trascrizionali che innesca e tramite i quali influenza lo sviluppo e la sopravvivenza dei neuroni. Il presente lavoro di tesi ha permesso di individuare una serie di Immediate Early Genes (IEGs), sia codificanti che non codificanti, che, ingaggiati da BDNF, possono avere un ruolo fondamentale nella biologia di tale neurotrofina. Tra i codificanti, l'attenzione si è focalizzata sul fattore di trascrizione EGR1 (*Early growth response protein 1*), il cui ruolo è particolarmente documentato nell'adulto. Grazie alle tecniche di *genome editing* basate sul CRISPR-Cas9 ed alla linea cellulare SH-SY5Y, è stato possibile dimostrare una funzione chiave di EGR1 nello sviluppo dei neuroni. Le cellule knockout per tale fattore di trascrizione sono infatti risultate incapaci di differenziare, sottolineando un'importante influenza dei programmi trascrizionali regolati da EGR1 nella sopravvivenza dei neuroni. Inoltre, gli studi riportati nel presente lavoro di tesi hanno permesso di identificare un'isoforma di splicing alternativo inedita per EGR1. I risultati preliminari ottenuti fanno ipotizzare che tale isoforma possa agire come dominante negativo dell'isoforma canonica concorrendo ad una fine regolazione della funzione di EGR1. Un altro gene su cui il presente lavoro di tesi si è concentrato è invece appartenente a quelli codificanti per i lncRNA, noto come *LINC00473*. Le analisi hanno permesso di ipotizzare che tale gene possa avere una rilevante funzione nello sviluppo e funzionamento dei neuroni. In particolare, è stato interessante osservare che probabilmente tale gene codifica per una piccola proteina primate-specifica il cui ruolo può influenzare l'espressione di altri IEGs con un impatto rilevante sulla regolazione dell'espressione genica.

ABBREVIATIONS

AD	Alzheimer disease
AKT	Serine-Threonine Protein Kinase
APAF-1	Apoptotic Peptidase Activating Factor 1
ARC	Activity Regulated Cytoskeleton Associated Protein
ATF3	Activating Transcription Factor 3
BACE1	Beta-secretase 1
BACE1-AS	Beta-secretase 1 Antisense
BAX	BCL2 Associated X, Apoptosis Regulator
BCL-2	B-cell lymphoma 2
BDNF	Brain-derived neurotrophic factor
BID	BH3 Interacting Domain Death Agonist
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
ceRNA	Competitive endogenous RNA
C-FOS	Fos Proto-Oncogene C
CLIP	cross-linking immunoprecipitation
CNS	Central Nervous System
CPAT	Coding Potential Assessment Tool
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeats
CYP26a1	Cytochrome P450 Family 26 Subfamily A Member 1
DAPI	4′, 6-diamidino-2-phenylindole
DAVID	Functional Annotation Bioinformatics Microarray Analysis
DIABLO	Diablo IAP-Binding Mitochondrial Protein
DMEM	Dulbecco’s modified Eagle’s medium
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
DUSP5	Dual Specificity Phosphatase 5
ECL	Enhanced ChemiLuminescence
EDTA	ethylendiaminetetracetic acid

EGR1	Early Growth Response 1
EGR2	Early Growth Response 2
EGR3	Early Growth Response 3
EGR4	Early Growth Response 4
EGTA	Ethylene Glycol Tetraacetic Acid
EMT	Epithelial-mesenchymal transition
EP2	E-type prostanoid receptor 2
EP4	E-type prostanoid receptor 4
ERK	Extracellular signal–regulated kinase
F3	Coagulation Factor III, Tissue Factor
FAM46A	Terminal Nucleotidyltransferase 5A
FBS	Fetal Bovine Serum
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GDNF	Glial Cell Derived Neurotrophic Factor
GREAT	Genomic Regions Enrichment of Annotations Tool
GO	Gene ontology
gDNA	genomic DNA
gRNA	RNA guide
HAND2-AS	Heart And Neural Crest Derivatives Expressed 2 Antisense
HD	Huntington disease
H1FX	H1 Histone Family Member X
HOMER1a	Homer Scaffold Protein 1a
HOXD	Homeobox D
HPRT1	Hypoxanthine Phosphoribosyl Transferase 1
IAPs	Apoptosis inhibitors
IEG	Immediate-early gene
IGF2BP1	Insulin Like Growth Factor 2 MRNA Binding Protein 1
IGF2BP2	Insulin Like Growth Factor 2 MRNA Binding Protein 2
IGF2BP3	Insulin Like Growth Factor 2 MRNA Binding Protein 3
IGFBP7	Insulin Like Growth Factor Binding Protein 7
INDEL	INsertion/DELetion

iPSCs	Induced pluripotent stem cells
KCl	potassium chloride
KLF10	Kruppel Like Factor 10
KO	knockout
lncRNA	long non-coding RNA
lncRNome	long noncoding RNA knowledgebase
LTD	long-term depression
LTP	Long term potentiation
MALAT1	Metastasis Associated Lung Adenocarcinoma Transcript 1
MAPK	Mitogen-activated protein kinase
MAP3K14	Mitogen-Activated Protein Kinase Kinase Kinase 14
$MgCl_2$	magnesium chloride
MIAT	Myocardial infarction associated transcript
miRNA	microRNA
NaCl	sodium chloride
ncRNA	non-coding RNA
NEUROD1	Neuronal Differentiation 1
NGF	Nerve growth factor
NGFR	Nerve growth factor receptor
NPAS4	Neuronal PAS Domain Protein 4
NR4A1	Nuclear Receptor Subfamily 4 Group A Member 1
NR4A3	Nuclear Receptor Subfamily 4 Group A Member 3
NRAGE	Neurotrophin receptor-interacting MAGE homologue
NRIF	Neurotrophin receptor interacting factor
NT-3	Neurotrophin-3
NT-4	Neurotrophin-4
NTRK	Neurotrophic tropomyosin receptor kinase
PCD	Programmed cell death
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PD	Parkinson disease

PFA	paraformaldehyde
PI3K	Phosphatidylinositol-3- kinase
PKA	Protein kinase A
PKC	Protein kinase C
PLC	Phospholipase C
PTB	Polypyrimidine tract-binding protein
PUM2	Pumilio RNA Binding Family Member 2
QKI	KH Domain Containing RNA Binding
RA	Retinoic acid
RAD51	RAD51 Recombinase
RARα	Retinoic Acid Receptor Alpha
RARβ	Retinoic Acid Receptor Beta
RARγ	Retinoic Acid Receptor Gamma
RBP	Retinol-binding protein
RHOA	Ras Homolog Family Member A
RHOB	Ras Homolog Family Member B
RHOF	Ras Homolog Family Member F, Filopodia Associated
RIPA	Radioimmunoprecipitation assay buffer
RPL19	Ribosomal Protein L19
RT	Room temperature
SB	StemMACS basal medium
SC-1	SPARC-like protein 1
SDS	Sodium Dodecyl Sulphate
SEM	mean standard error
SH-2	Src-Homology-2
SIK1	Salt Inducible Kinase 1
siRNA	small interfering RNA
SOX2	SRY-Box 2
SOX9	SRY-Box 9
SYBU	Syntabulin
TBS	tris-buffered saline

TNRC6A	Trinucleotide Repeat Containing 6A
TNRC6C	Trinucleotide Repeat Containing 6C
TPA	12-O-tetradecanoyl-phorbol-13-acetate
TRAF6	TNF Receptor Associated Factor 6
Tris-HCl	TRIS hydrochloride
WDR1	WD Repeat Domain 1
WT	wild type
ZSCAN10	Zinc Finger And SCAN Domain Containing 10

1. INTRODUCTION

1.1 Human Neurodevelopment

The CNS is probably the most complex of all the biological systems. The mature brain is made up of over 100 billion neurons that represent the information processing cells that can vary in their size, shape and function (Pakkenberg and Gundersen, 1997). Basically, neurons have a cell body (called the soma) and various processes that include dendrites, branching processes that extend only for a short distance away from the neuron cell body, and an axon, a separate process that is typically longer than the dendrites (Figure 1). To support the general function of the nervous system, neurons have evolved special abilities for sending electrical signals along axons. These are wrapped in a fatty substance called myelin that makes the transmission of electrochemical signals between regions efficient (Figure 1; Stiles and Jernigan, 2010). The communication between neurons is achieved at synapses by the process of neurotransmission.

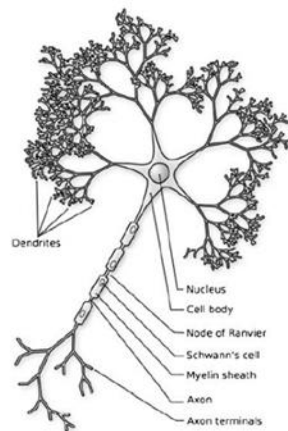

Fig 1. Schematic drawing of a neuron (Stiles and Jernigan, 2010).

During embryogenesis and adulthood, the development of nervous system is a complex and dynamic process that involves a precisely orchestrated sequence of genetic, epigenetic, environmental, biochemical and physical events. The differentiation of neural progenitor cells that will form nervous system requires complex cascades of molecular signals and the fine regulation of gene expression networks. Neural progenitors are self-renewing cells able to divide and increase their number. During embryogenesis, after several replication cycles, the number of cells increases until the cell division switches from symmetric to asymmetric. In this new phase, two different types of cells are produced: a neural progenitor cell, which continues to proliferate (self-renewal), and a neuron, which no longer proliferates. In the early stages of development, neurons undergo to a migration process referred to as somal translocation (Figure 2 A) (Nadarajah and Parnavelas, 2002; Stiles and Jernigan, 2010). As development proceeds, the brain becomes

larger and the primary mode of neuronal migration changes. Because of the greater distances, neurons require a special population of cells called "radial glial guides" to support their migration (Figure 2 B) (Rakic, 1972; Stiles and Jernigan, 2010). Although the radial glial guides were originally thought to be a special, transient population of cells, it has recently been discovered that they are neural progenitor cells (Noctor et al. 2001; Noctor et al. 2002; Parnavelas et al. 2002; Weissman et al. 2003; Stiles and Jernigan, 2010). Very recent studies have identified a second proliferative zone that will later develop into the basal ganglia. These neurons traverse long distances using a mode of migration that has been termed "tangential migration" (Figure 2 C). Tangential migration involves a variety of signalling pathways not seen in radial migration (Marin and Rubenstein 2001; Huang, 2009; Stiles and Jernigan, 2010; Valiente and Marin 2010).

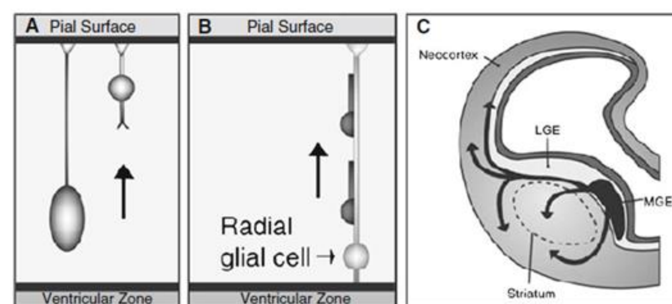

Fig 2. Different modes of neuronal migration. **(A)** Neuron migration by somal translocation: neuron extends a long basal process, which is an extension of the cell body and attaches itself to the pial surface, the outer surface of the developing brain. The nuclei of cells then move through the cytoplasm of the basal process. **(B)** Neuron migration radial glial guide: guides extend a basal process that attaches to the pial surface of the brain which forms a kind of scaffold along which neurons can migrate. Each glial scaffold can support the migration of many neurons. **(C)** Neuron migration from the second proliferative zone in ganglionic eminences by tangential migration (arrows indicate the direction of migration for different neuron populations). Neurons use several guidance molecules produced in local regions along their migratory route to direct their movement into the cortex. (Stiles and Jernigan, 2010).

The migration of neurons into the developing neocortex leads to the formation of an ordered 6-layer structure (Cooper, 2008; Stiles and Jernigan, 2010). At this point, young neurons must become part of information processing networks. They must, therefore, develop neuronal processes (axons and dendrites) that allow them to communicate with other neurons. Each cell has many dendrites and a single axon that can extend to a certain distance from the cell. Once the axon has reached its target, the connections called synapses are formed with the target cell. Synaptogenesis consists of three phases: the immature synapses are formed between the axons and the dendrites; therefore, the synapses undergo maturation and convert from to silent into an active state; finally, the synaptic number is reduced to the neuronal connections within the circuit (Craige et al., 2006; Budday et al., 2015).

Although the production and migration of neurons are largely prenatal events, proliferation and migration of glial progenitors continue for a long time after birth, and differentiation and

maturation of these cells continue during childhood. Neuron-glia interactions are not yet fully defined, but these interactions play an important role in the functional organization of neural circuits during postnatal life. In the postnatal period, neurogenesis continues to a very limited degree; however, new neurons continue to emerge and migrate into the subventricular zone and the dentate gyrus of the hippocampus (Stiles and Jernigan, 2010). These exceptional forms of neurogenesis seem to continue throughout adult life but produce only a small percentage of the neuronal population (Ming and Song, 2011; Kempermann et al., 2018).

1.2 Programmed cell death in neurodevelopment

Programmed cell death (PCD) contributes to the development of the nervous system in an evolutionarily conserved manner. PCD is the basis of the neurotrophic theory, whereby cell death derives from a surplus of neurons (Yamaguchi and Miura, 2015). Much of the PCD occurs through apoptosis, a well-characterized genetic program that exhibits specific morphological characteristics such as membrane blebbing, nuclear and cytoplasmic withdrawal, condensation of chromatin and DNA fragmentation (Kerr et al., 1972; Yeo and Gautier, 2004; Yamaguchi and Miura, 2015). The components of the cellular apoptotic machinery can be classified into pro-apoptotic regulators, such as BAX, BID, APAF-1, DIABLO and anti-apoptotic regulators, such as BCL-2, BCL-XL, apoptosis inhibitors (IAPs) and caspases, which belong to a family of cysteine proteases that are the key effectors of the process (Yamaguchi and Miura, 2015). The regulation of caspase activities is complex and derives from the interplay between the anti- and pro-apoptotic regulators (Yeo and Gautier, 2004). Executioner caspases, including CASPASE-3/7, are physiologically activated by extrinsic death ligands or intrinsic signals such as DNA damage, survival factor deprivation, ER stress, abnormal ion flow or reactive oxygen overproduction (Figure 3) (Green et al., 2014; Yamaguchi and Miura, 2015).

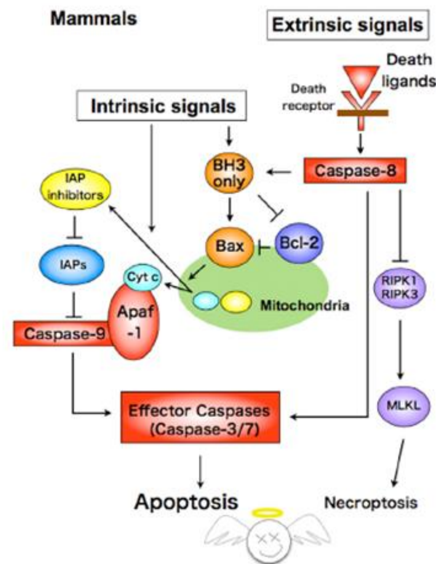

Fig 3. Schematic representation of molecular pathways leading to apoptosis by effector caspase activation (Yamaguchi and Miura, 2015).

Two important processes involve substantial loss of neural elements: naturally occurring cell death, which involves the normal loss of 50% or more of the neurons within a brain region; synaptic exuberance and pruning in which there is a massive excess production of connections followed by the systematic elimination of up to 50% of those connections. Both processes reflect non-pathological events that play an essential role in establishing the complex networks of the developing brain (Stiles and Jernigan, 2010). The uptake of neurotrophic substances protects against the apoptosis cascade (Levi-Montalcini, 1964; Oppenheim, 1989). Neurotrophic factors are produced by target neurons at synaptic sites and are taken up by the afferent neurons that make effective connections with the targets (Huang and Reichardt, 2001). During development, it is thought that neurons compete for neurotrophic resources. According to the neurotrophic hypothesis (Oppenheim, 1989), neurons that establish effective connections are able to obtain more neurotrophic factor and are more likely to survive.

1.3 Regulatory landscape of human neurodevelopment

Recent genome-wide profiling studies of the developmental transcriptome of the human brain by several groups (Johnson et al., 2009; Somel et al., 2009; Ip et al., 2010; Colantuoni et al., 2011; Kang et al., 2011; Lambert et al., 2011; Mazin et al., 2013; Miller et al., 2014; Pletikos et al., 2014; Jaffe et al., 2015) revealed that the great majority of protein-coding genes (at least 86% according to Kang et al., 2011) and an ever-expanding number of non-coding genes are used at some point in the building of the human CNS. In particular, they evidenced remarkable

dynamicity of gene expression during prenatal and early postnatal development, accounting for approximately 2/3 of the variance in global expression. Transcriptional studies also demonstrated correlations between gene expression dynamics and the morphological and functional development of brain regions and, consequently, shed light on the timing of developmental processes and the onset of specific biological functions (Johnson et al., 2009; Kang et al., 2011; Parikshak et al., 2013; Willsey et al., 2013; Miller et al., 2014). For example, gene co-expression network analyses have revealed that the developmental brain transcriptome can be segregated into distinct modules or clusters of genes with highly correlated expression. Many transcriptional processes are shared and coordinated across different regions as the brain develops. In particular, the module enriched for genes associated with the neuronal specification is most highly expressed during the embryonic and early fetal periods, while the module enriched for genes associated with synaptic function and ion channels begins to rise in late fetal period and plateaus in early childhood. Recent approaches have successfully integrated the spatiotemporal dimensions of the human brain transcriptome with gene mutation discoveries to generate testable hypotheses about when and in which regions and/or cell types in the developing human CNS the expression of disease-associated genes converge. It is intriguing to speculate that selective dysfunction of spatially and temporally regulated gene expression may in part explain differences in the age of onset and affected neural circuits in neurological and psychiatric disorders (Silbereis et al., 2016).

As an important layer of gene expression, the epigenetic mechanisms act in concert with *trans* and *cis* components of the regulatory circuitry and play a critical role in regulating spatiotemporal gene expression patterns (Maze et al., 2014; Nord et al., 2015; Shibata et al., 2015; Silbereis et al., 2016). These mechanisms, including DNA methylation, histone modifications, and non-coding RNAs, can also be affected by various extrinsic factors, thus providing a molecular link between external cues and gene expression (Silbereis et al., 2016). DNA methylation of cytosine, primarily at CpG nucleotides, plays a key role in neural development and function. DNA methylation levels change rapidly during fetal development but slow down after birth and with aging (Numata et al., 2012; Silbereis et al., 2016). The accumulation of methylation in the non-CpG context (mCH) happens during early postnatal development (first 2 years postpartum) and through adolescence, with a small decrease thereafter. Lister et al. (2013) noted that the accumulation of mCH and gene expression were negatively correlated; genes highly expressed in adult neurons lost both CpG and non-CpG methylation progressively during development (Lister et al., 2013). In addition to DNA methylation, histone modification is an essential mechanism for establishing cellular diversity and regulating the timing of developmental processes (Silbereis et al., 2016). Methylation (mono-, di-, and tri-) and acetylation are the most extensively studied modifications (Tessarz and Kouzarides, 2014). For example, Cheung and colleagues (2010) profiled histone 3 lysine 4 trimethylation (H3K4me3) across development in neuronal and non-neuronal cells of the human prefrontal cortex. They highlighted significant remodelling during

postnatal development and the aging of prefrontal neurons. In particular, they observed higher levels of H3K4me3 methylation of NEUROD1 and several members of the cadherin and semaphorin families in newborns compared to older samples. Dysregulation of epigenetic modifications are associated with various human diseases, including neurodevelopmental disorders (Bale et al., 2010; Millan, 2013). In particular, an increasing number of mutations in histone lysine methylation-related genes have been identified as intellectual disability-associated genes (Ronan et al., 2013; Vissers et al., 2016).

The correct gene expression regulation during development requires the intricate interplay between transcription factors, epigenetic modifications, and non-coding RNAs (ncRNAs). The latter represents a large, but poorly characterized component of the human transcriptome. ncRNAs play a critical role in transcriptional and post-transcriptional regulation of gene expression and affect the overall transcriptional landscape and proteomic diversity of a cell (O'Carroll and Schaefer, 2013; Morris and Mattick, 2014; Silbereis et al., 2016). ncRNAs can be classified as short or long RNAs, on the basis on their length. Short microRNAs (miRNAs) are a relatively well-known class of RNA whose mechanism of action has been unravelling. miRNAs bind their mRNA target based on sequence complementary and trigger the degradation or the translation of inhibition of the mRNA (O'Carroll and Schaefer, 2013; Silbereis et al., 2016). Long non-coding RNAs (lncRNAs) represents a relatively newly identified class of RNA. The lncRNAs can act as regulators at different levels: at the transcriptional level by means of chromatin remodeling mechanisms and histone modifications; at the post-transcriptional level through interaction with mRNAs, with consequent modulation of splicing, inhibition of protein synthesis, degradation of mRNAs and formation of endogenous siRNAs; they can interact with other biological molecules, modulating the activity of proteins, their localization and the formation of RNA-protein complexes; finally they can interact with miRNAs (Wilusz et al., 2009; Wu et al., 2013). Recent studies have shown that many miRNAs and lncRNAs are expressed in the human and non-human primate brain and are spatially and temporally regulated (Somel et al., 2011; O'Carroll and Schaefer, 2013; Silbereis et al., 2016). A growing body of evidences emphasizes the role of lncRNAs in neurodevelopment and brain function, and a functional correlation with several neurodevelopmental, neurodegenerative and psychiatric diseases (for instance, see St Laurent and Wahlestedt, 2007; van de Vondervoort et al., 2013; Wu et al., 2013; Barry, 2014; Roberts et al., 2014; Tushir and Akbarian, 2014). For example, *MIAT*, also termed as Gomafu, exhibits deregulation in multiple diseases, including up-regulation in ischemic stroke and down-regulation in schizophrenia (Sun et al., 2018); *BACE1-AS* stabilizes the *BACE1* transcript, protecting it from RNA degradation through RNA–RNA pairing, and thus contributing to the progression of Alzheimer's (Clark and Blackshaw, 2014).

1.4 Role of neurotrophins in neuronal development

Four neurotrophins have been characterized in mammals: nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3) and neurotrophin-4 (NT-4) (Hallbook, 1999). They share sequence homology and structure similarity and evolved from a common ancestral gene (Hallbook, 1999).

Neurotrophins act by binding their cognate NTRK receptors (Trk, tropomyosin-related kinase). NGF is specific for NTRK1; BDNF and NT-4 are specific for NTRK2; NT-3 activates NTRK3 and it is also able to activate less efficiently each of the other NTRK receptors (Figure 4).

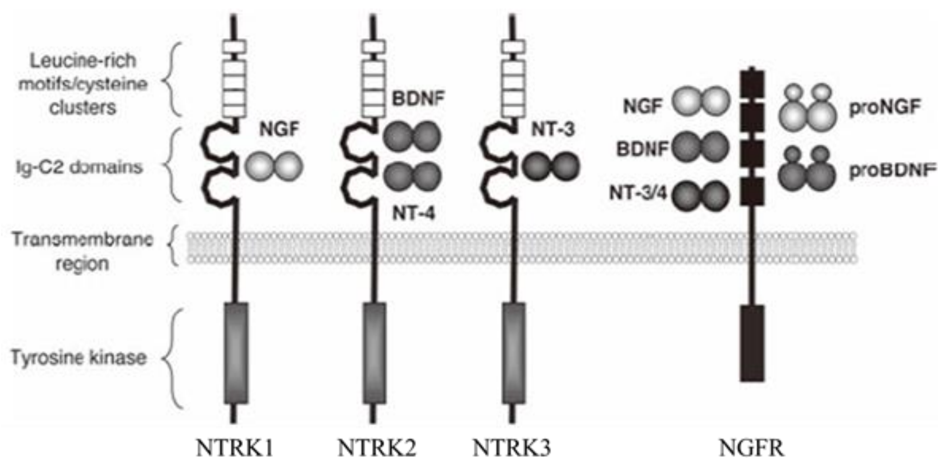

Fig 4. Neurotrophins receptors and specific ligands (https://www.sinobiological.com/neurotrophin-receptor-sinobio.html).

Ligand engagement of NTRK receptors has been shown to result in phosphorylation of cytoplasmic tyrosine residues on the cytoplasmic domains of these receptors. Phosphorylation of other tyrosine residues promotes signalling by creating docking sites for adaptive proteins containing phosphotyrosine (PTB) or src-homology-2 (SH-2) binding patterns (Huang and Reichardt, 2001; Pawson and Nash, 2000). These adapter proteins couple NTRK receptors to intracellular signaling cascades, which include the Ras/ERK metabolic pathway (signal extracellular signal-regulated kinase), the phosphatidylinositol-3- kinase (PI-3 kinase) / AKT-kinase and the phospholipase C (PLC) -γ1 (Figure 5) (Reichardt and Fariñas, 1997; Kaplan and Miller, 2000; Huang and Reichardt, 2001).

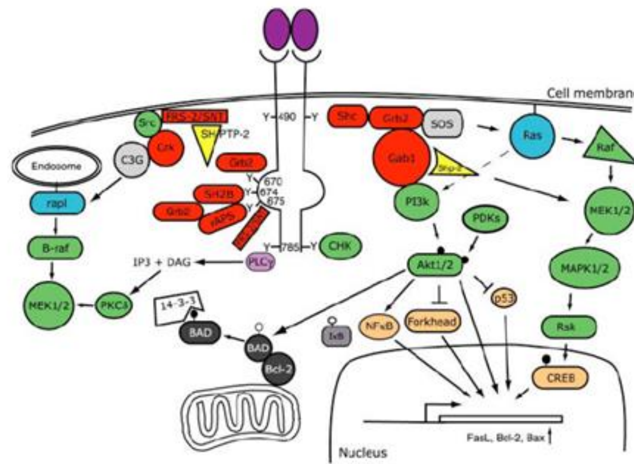

Fig 5. Schematic diagram of NTRK receptor-mediated signal transduction pathways. In this diagram, adaptor proteins are red, kinase green, small G proteins blue, and transcription factors brown. (Huang and Reichardt, 2001).

Each neurotrophin also binds to the low-affinity p75NTR (NGFR) receptor, which is a member of the tumor necrosis factor receptor superfamily (Frade and Barde, 1999; Huang and Reichardt, 2001). NGFR interacts with proteins, including TRAF6, RhoA, NRAGE (neurotrophin receptor-interacting MAGE homologue), SC-1, and NRIF, and regulates gene expression, cell cycle, and growth cone motility. Binding of neurotrophins to NGFR has also been shown to activate the Jun kinase pathway, which can be inhibited by activation of the NTRK receptors-dependent Ras-phosphatidylinositol-3-kinase (PI3K) pathway by NTRK receptors. Most intriguing, both *in vitro* and *in vivo* evidences indicates that ligand engagement of NGFR can directly induce neuronal death via apoptosis (Frade and Barde 1999; Friedman, 2000; Huang and Reichardt, 2001) (Figure 6).

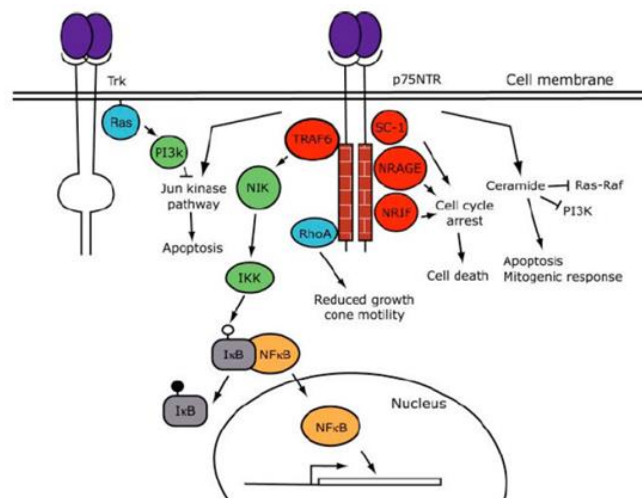

Fig 6. Schematic diagram of p75NTR-mediated signal transduction pathways. In this diagram, adaptor proteins are red, kinase green, small G proteins blue, and transcription factors brown. (Huang and Reichardt, 2001).

Neurotrophins regulate development, maintenance, and function of vertebrate nervous systems (Huang and Reichardt, 2001). They are initially synthesized as precursor proteins (pro-neurotrophins), which are processed intracellularly to be secreted mostly in a mature, biologically active form (Mowla et al., 1999; Mowla et al., 2001; Lu et al., 2005; Matsumoto et al., 2008). During development, neurotrophins stimulate the formation of appropriate synaptic connections, controlling the direction and rate of axon growth (Wang and Poo, 2005; Li et al., 2005), as well as the shape of dendritic arbores and spines (Ji et al., 2005; Melo et al., 2013).

The neurotrophin BDNF plays a central role both in neuronal development and in the adult nervous system, and dysfunction in its signalling may contribute to several neurodegenerative disorders (Pruunsild et al., 2007). BDNF interaction with NTRK2 receptor activates three signaling pathways: PI3K-Akt (PI3K, phosphatidyl inositol-3 kinase), Ras-MAPK (MAPK, mitogen-activated protein kinase), and PLCγ-Ca^{+} (PLC, phospholipase C) (Duman and Voleti, 2012). It is also known to regulate a large spectrum of processes of the nervous system, including cell survival, growth and differentiation (Casaccia-Bonnefil et al., 1999; Bibel and Barde, 2000; Huang and Reichardt, 2003; Park and Poo, 2013; Suliman et al., 2013; Zagrebelsky and Korte, 2014), synaptic plasticity of neurons and LTP (Xu et al., 2000; Bramham and Messaoudi, 2005; Gottmann et al., 2009; Minichiello, 2009; Mizui et al., 2014; Zagrebelsky and Korte, 2014; Leal et al., 2015). The essential role of BDNF in neuronal processes during development and adulthood, support its potential therapeutic use in the treatment of both neurological and psychiatric disorders (Pruunsild et al., 2007; Nagahara and Tuszynski, 2011; Weissmiller and Wu, 2012). In fact, BDNF exerts potent pro-survival effects in models of neurological diseases such as Parkinson's (PD) (Howells et al., 2000; van der Kolk et al., 2015), Huntington's (HD) (Zuccato et al., 2008; Jiang et al., 2013), and Alzheimer's (AD) diseases (Holsinger et al., 2000; Michalski and Fahnestock, 2003; Peng et al., 2005; Faria et al., 2014), as well as depression and other psychiatric disorders (Karege et al., 2002; Aydemir et al., 2005; Gonul et al., 2005; Cunha et al., 2006).

1.5 Immediate Early genes in brain

Immediate-early genes (IEGs) are genes that are induced rapidly and transiently in response to a wide range of cellular stimuli. They act in many biological processes: regulation growth, differentiation and cell cycle. Because of the wide variety of extrinsic stimuli to which they respond, their mechanism of action can vary depending on the stimulus received, its duration and the type of cell (Bahrami and Drabløs, 2016).

IEGs play a key role in brain development; in particular their expression is selectively and rapidly increased in certain neurons of specific brain regions, associated to the formation of memory and

learning (Minatohara et al., 2015). They are also involved in important processes such as neuronal plasticity, which persists even in adulthood and therefore is not limited only to the early stages of development (Pérez-Cadahìa et al., 2011).

Approximately 30-40 neuronal IEGs have been identified; 10-15 are regulators, whose protein products are classified as inducible transcription factors, such as C-FOS and EGR1 (Davis et al., 2003). C-FOS is involved in cell proliferation and differentiation after extracellular stimuli, and its deregulation has been often associated to oncogenic progression (Velazquez et al., 2015). Only 40% of knockout c-Fos mouse embryos survive until birth showing the importance of this gene for development (Velazquez et al., 2015). EGR1 is a transcription factor widely studied in adults, where it plays a clear role in synaptic plasticity (Adams et al., 2017). Several evidences in different cell lines led to hypothesize that EGR1 can play a role in the development of neurons (Milbrandt, 1987; Aliperti and Donizetti 2016). ARC (also knowns as Arg3.1), NPAS4 and HOMER1a (a shorter variant of Homer1) are all induced by neuronal activity and they are therefore considered IEGs (Xiao et al., 1998; Korb and Finkbeiner, 2011; Kim et al., 2018). ARC is not a transcription factor but acts as an effector involved in various neuronal signaling pathways. *ARC* mRNA is rapidly transcribed in response to neuronal activity, and precisely targeted to activated synapses in neuronal dendrites (Farris et al., 2014; Kim et al., 2018). NPAS4 is a brain-specific transcription factor, which plays a role in the development of inhibitory synapses by regulating the activity-dependent gene programs in cultured neurons (Kim et al., 2018). HOMER1a is primarily located at the postsynaptic density where it competes with other constitutively expressed Homer proteins for the interaction with Homer-binding proteins. As such, HOMER1a negatively regulates excitatory synapse structure and function (Sala et al., 2003; Kim et al., 2018).

Alterations in the expression of IEGs have been associated with various cognitive disorders, such as autism, schizophrenia and dementia (Thiel and Cibelli, 2002).

2. AIM OF THE STUDY

The central nervous system (CNS) is a large network of interconnecting and intercommunicating cells that form functional circuits and is responsible for intellectual and cognitive functions. Generation of the CNS occurs during embryonic development from a pool of immature progenitors that will give rise to all the neurons in the brain and spinal cord. During neuronal development transcriptional regulation plays fundamental role in each step of the process from neural fate determination to complete neuronal maturation. At a molecular level, this process is coordinated by transcription factors involved in all the cellular processes, including cell cycle exit, migration, survival, and acquisition of neuronal features, such as dendrites and axon maturation, and functional synapses. Dysfunction of gene expression regulation during critical developing periods lead to neurodevelopmental abnormalities and mental disorders. Despite the progress on the molecular mechanisms controlling neuronal development, a lot remains to be clarified about regulation of gene expression and regulation of transcriptional activity underlying this phenomenon. This is especially true considering that beyond classical transcription factors and miRNA, it is now emerging another recently identified class of molecules, the long non-coding RNA (lncRNA), as a fundamental element of gene expression regulation underlying neuronal development.

Each stage of neuronal maturation requires an integration of cell-intrinsic genetic programs and the extrinsic influences such as the effects of neurotrophic factors. Among neurotrophins, BDNF has been shown to be essential for the correct development of the brain, playing an important role in numerous cellular processes, such as axon migration, regulation of the size and number of dendritic spines, synaptogenesis and cellular survival.

In an effort to identify long non-coding RNA involved in the BDNF function during neuronal development, we carried out high-throughput experiments reported in the part I of the present thesis. In particular, we focused on those transcriptional events occurring immediately after BDNF stimulus, considering that they play critical roles in long-lasting neuronal changes by regulating a plethora of target genes. We took advantages from the use of a popular human neuroblastoma cell line (SH-SY5Y) as an experimental model of neuronal development.

The results of the part I provided the identification of coding and long non-coding RNA involved in the BDNF biology and thus likely important in the transcriptional programs occurring during neuronal development. In the second part of this thesis, we focused our attention on *EGR1*, one the coding genes engaged by BDNF. This gene encodes for a transcription factor whose role in the embryogenesis remains elusive. By using the recent and powerful molecular approach for genome editing, the CRISPR/Cas9, we provided preliminary investigation on the role of EGR1 in the human neuronal development.

The high-throughput analysis reported in the part I allowed us to identify a lncRNA never associated to neuronal development, the *LINC00473*. With the aim to provide a characterization of this gene, in the part III of this thesis, we analyzed the evolutionary conservation of the sequence and the expression pattern regulation. In addition, to gain preliminary insights on the molecular mechanism of action, we generated a KO model for the *LINC00473* gene.

3. MATERIALS & METHODS

3.1 Cell cultures

The human neuroblastoma cell line SH-SY5Y was used as a neuronal cell model. In particular, the N-enriched population of SH-SY5Y was obtained from the parental cell line by a procedure reported elsewhere (Piacentini et al., 1996; Bell et al., 2013; Aliperti and Donizetti 2016). The cells were grown and propagated in Dulbecco's modified Eagle's medium (DMEM, Euroclone®), supplemented with 2mM L-glutamine (Euroclone®), and a solution of 1% penicillin/streptomycin (Euroclone®) and 15% Fetal Bovine Serum (FBS, Euroclone®).

The HEK293T cell line was grown and propagated in Dulbecco's modified Eagle's medium (DMEM, Euroclone®), supplemented with 2mM L-glutamine (Euroclone®), and a solution of 1% penicillin/streptomycin (Euroclone®) and 10% Fetal Bovine Serum (FBS, Euroclone®).

3.2 Cellular treatments

The N-enriched SH-SY5Y cells were differentiated by incubation in a low serum (1.5%) medium containing RA (retinoic acid, SIGMA–Aldrich®) 10μM. In particular, $8x10^5$ cells were seeded in 35mm plates and starved for 24h by reducing FBS. Then RA was added, and the medium was refreshed every 2 days. Untreated cells were grown in the presence of only dimethyl sulfoxide (DMSO) as vehicle control. After 6 days of differentiation, the medium containing RA was removed and substituted with a medium without FBS and with BDNF 10ng mL^{-1} (PeproTech®) for a specific time.

In order to evaluate the effect of RA concentration, WT and EGR1-KO cells were treated with increasing doses of morphogen (3μM, 6μM and 10μM). $5x10^4$ cells were seeded in 24-well plates and starved for 24h by reducing FBS from 15% to 1.5%. RA was then added to the different concentrations, and the cells were followed up to 6 days with observations under an optical microscope.

For TPA differentiation, $8x10^5$ N-enriched SH-SY5Y WT and EGR1-KO cells were seeded in 35mm plates. The day after, the cells were treated with TPA (16nM) in a medium where the percentage of FBS was reduced from 15% to 1.5%. Untreated cells were grown in the presence of only dimethyl sulfoxide (DMSO). The medium was refreshed every 2 days.

For FBS stimulation, $6x10^5$ HEK293T cells were seeded in 35mm plates and starved for 24h by reducing FBS from 10% to 1.5%. After 24h of starvation, cells were stimulated with 15% FBS and collected after different time intervals as reported in the Results section.

3.3 RNA isolation, retrotranscription and quantitative PCR (qPCR) analysis

For microarray analysis, total cellular RNA was isolated using an RNeasy® Mini Kit (QIAGEN) according to the RNeasy® Mini Handbook (QIAGEN). DNA contamination was efficiently removed by on-column DNAse digestion (QIAGEN).

For the analysis of transcript levels, total cellular RNA was isolated using TRI-Reagent (SIGMA–Aldrich®) according to the manufacturer's instructions.

The concentration and purity of the RNA sample were assessed using NanoDrop® 1000 (Thermo Scientific). cDNA was synthesized from 1μg of total RNA using an Invitrogen SuperScriptIII® reverse transcriptase kit (Invitrogen). qPCR validation was performed on independent biological replicates in triplicate. Real-time PCR was performed using the SYBR green method and an Applied Biosystems 7500 System. The reaction mixture contained 50ng of cDNA template and 400nM of each forward and reverse primer in a final volume of 15μL. PCR conditions included a denaturation step (95°C for 10min) followed by 40 cycles of amplification and quantification (95°C for 35s, 60°C for 1min). Relative gene expression levels were normalized to the reference gene Hypoxanthine Phosphoribosyl Transferase 1 (HPRT1) for microarray validation, and to the gene Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) for the other analyses. It was calculated by the $2^{-\Delta\Delta Ct}$ method. The sequences of the primers used are reported in Table 1.

3.4 Microarray analysis

Microarray experiments were performed on biological triplicate samples. Total RNA quality was assessed by an Agilent 2100 Bioanalyzer (Agilent Technologies). Microarray hybridizations were performed by the Transcriptomics and Genomics core facility of the Department of Emergency and Organ Transplants (DETO) in the Nephrology Unit at the University of Bari Aldo Moro in Italy. The labeled cRNA was produced using a Low Input Quick Amp Labeling (LIQA) kit (Agilent Technologies) and hybridized for 17 hours at 65°C on an Agilent SurePrint G3 8x60K custom lncRNA expression array (Agilent Technologies). This array contains two probes for 22,001 lncRNAs targeting the Gencode v15 human lncRNA annotation, together with one probe for 17,535 randomly chosen protein-coding transcripts. After hybridization, the slide was washed according to Agilent protocols and scanned using a High-Resolution Microarray C Scanner (Agilent Technologies). The image file was processed using Agilent Feature Extraction software (v10.7.3). The microarray grid was correctly placed, and outlier pixels (which were rejected) and inlier pixels were identified. Normalization was performed according to the Quantile method. The differentially expressed probes were selected using a moderated t-test with a p-value cut-off of 0.05.

3.5 lncRNAs classification and functional analysis

The LNCipedia database (http://www.lncipedia.org/) (Volders et al., 2015) was used for retrieving the transcript ID, gene ID, and alternative gene name of the differentially expressed lncRNAs. Differentially expressed lncRNAs were classified by considering their position relative to adjacent protein-coding genes as reported by Mattick and Rinn (2015). The differentially expressed lncRNAs were analysed by CPAT software (http://lilab.research.bcm.edu/cpat/) (Wang et al., 2013) to assess their protein-coding potential. For the functional analysis of the lncRNAs, a list of nearby potentially regulated genes was retrieved using the computational tool GREAT (Genomic Regions Enrichment of Annotations Tool; http://bejerano.stanford.edu/great/public/html/) (McLean et al., 2010). Functional enrichment analysis for the predicted target genes and differentially expressed coding genes was performed using the DAVID system (Database for Annotation, Visualization and Integrated Discovery; http://david.abcc.ncifcrf.gov/) (Dennis et al., 2003), which uses Gene Ontology (GO) to identify the molecular function represented in the gene profile (Dennis et al., 2003). For all DAVID analyses, the significance of enrichment of each GO term was assessed by a p-value of <0.05 and ranked by the number of differentially expressed genes (count). The lncRNAs were analyzed with the lncRNome software (http://genome.igib.res.in/lncRNome) (Bhartiya et al., 2013) to identify possible interactions with proteins. The database contains information about 6,800 binding sites for proteins obtained from PAR-CLIP and CLIP-seq analyses. I obtained a list of potentially regulated miRNAs from lnCeDB (http://gyanxet-beta.com/lncedb/) (Das et al., 2014), a database that provides human lncRNA that can potentially act as competitive endogenous RNAs (ceRNAs) and interfere with the pathway of miRNAs. These miRNAs were analyzed by miR2Disease (http://www.mir2disease.org/) (Jiang et al., 2009) to find miRNAs deregulated in human diseases. The potentially regulated miRNAs were also analyzed by miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/) (Chou et al., 2018) to find their experimentally validated mRNA targets.

3.6 Generation of EGR1 and LINC00473 KO cell lines

The knockouts for the *EGR1* gene in an N-enriched SH-SY5Y cell line and the *LINC00473* gene in the HEK293T cell line were obtained using the CRISPR/Cas9 procedure as reported by Ran et al. (2013). Briefly, forward and reverse oligonucleotides (Table 2) for the gRNA were designed from the online CRISPR design tool (http://crispr.mit.edu/) and inserted in the all-in-one vector pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene plasmid #62988) for *EGR1* knockout and in the all-in-one vector AIO-PURO (Addgene plasmid #74630) for the *LINC00473* gene. The first strategy was designed to obtain an INDEL mutation in the ORF; the second strategy was based on the use of an all-in-one vector with the Cas9 Nickase and two gRNAs in order to remove a

wider region, avoiding off-targets. Subsequently, the vectors were transfected into the cells using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. After 24h of transfection, the cells were cultured under a puromycin selection (1µg mL^{-1}) for 48h. The surviving cells were left to propagate in the plate and then transferred into a 96-well plate for single clone selection by serial dilution. The monoclonal population carrying *EGR1* and *LINC00473* mutated sequences was selected by sequence analysis of genomic region of interest. The genomic DNA was isolated by Quick-gDNA™ Miniprep kit (ZYMO RESEARCH) and the region of interest was amplified using specific primers pair (Table 2). The PCR products were then subcloned in the pGEM®-T Easy Vector (Promega) and several independent clones were subjected to Sanger sequencing.

3.7 iPSCs

iPSCs were grown and propagated under feeder-free conditions in StemMACS™ IPS-Brew XF Basal medium (SB, MACS® Media), supplemented with StemMACSTM IPS-Brew XF 50X supplement (MACS® Media). The medium was changed every day.

Cells were transfected with all-in-one vector pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene plasmid #62988) containing the validated EGR1 gRNA. Amaxa™ Human Stem Cell Nucleofector™ Kit 1 (LONZA) was used for transfection according to the manufacturer's instructions. After 48h of transfection, the cells were cultured under a puromycin selection (0.5µg mL^{-1}) for 48h. Finally, single clones were picked and transferred in a 48-well plate. The genomic DNA was isolated using QIAamp DNA Blood Mini kit (QIAGEN) and the region of interest was amplified by using specific primers pairs (Table 2). The PCR products were sequenced to check the generation of mutation.

3.8 Western Blotting

N-enriched SH-SY5Y WT and EGR1-KO cell pellets were lysed in RIPA buffer (50mM Tris-HCl pH 8.8, 150mM NaCl, 1mM EDTA, 0.1% SDS, 1% Triton X-100) containing protease inhibitors (Roche), incubated on ice for 30min and centrifuged at 14000rpm for 10min at 4°C. The supernatant was collected and used for protein quantification by Bradford assay (BIO-RAD). 30µg of protein lysate for each sample was electrophoresed in SDS gel (10% acrylamide) and blotted on to a nitrocellulose membrane. The transferred membranes were blocked with 3% non-fat milk (BIO-RAD) in TBST buffer (100mM Tris-HCl pH 8, 1.5M NaCl, 1% Tween) for 1h at room temperature (RT) and incubated with specific primary antibodies in TBST with 3% non-fat milk (BIO-RAD) overnight at 4°C. After several washes with TBST, the membranes were incubated with the corresponding secondary antibodies in the same buffer. After several washes,

immunoreactive bands were visualized using ECL detection kit (Euroclone®) according to the manufacturer's instructions. Primary and secondary antibodies are listed in Table 3.

3.9 EGR1 iper-expression experiments: cloning, transfection and subcellular fractionation protocol

The CDS of the *EGR1*alternative isoform (*EGR1 Δ141-278*) was amplified by PCR using cDNA samples obtained by RNA extracted from N-type SH-SY5Y cells treated with RA for 6 days as a template. The reaction mixture contained 50ng of cDNA template and 400nM of each forward and reverse primer spanning from the start codon to the stop codon, in a final volume of 20µL. The PCR conditions included a denaturation step (95°C for 2min) followed by 38 cycles of denaturation, annealing and elongation (95°C for 30s, 60°C for 40s and 68°C for 1min). The PCR product was isolated and cloned into pCMV3 expression vector by using restriction enzymes KpnI and XbaI.

The pCMV6-EGR1 vector (SC128132, ORIGENE) and the pCMV3 with CDS of the alternative isoform of *EGR1* were used for iper-expression experiments. Both expression vectors were transfected in HEK293T cells using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. After 24h of transfection, cell pellets were lysed in a fractionation buffer (20mM HEPES pH 7.4, 10mM KCl, 2mM $MgCl_2$, 1mM EDTA, 1mM EGTA, 1mM DTT) containing protease inhibitors (Roche), incubated on ice for 20min, and centrifuged at 3000rpm for 5min at 4°C. The pellet contained nuclei and the supernatant contained cytoplasm proteins. Supernatant containing cytoplasmic proteins was transferred into a fresh tube, while the pellet of nuclei was washed with fractionation buffer and centrifuged at 3000rpm for 10min at 4°C. Finally, the pellet of nuclei was suspended in TBS buffer (100mM Tris-HCl pH 8, 1.5M NaCl) with 0.1% SDS and sonicated.

3.10 Immunofluorescence analysis

HEK293T cells were grown on coverslips, fixed in 4% paraformaldehyde (PFA), permeabilized in 0.3% PBS-Triton X-100 for 20min at RT, and then blocked in 3% PBS-BSA for 30min at RT. Cells were subsequently incubated with primary antibodies for 1.5h at RT in 3% PBS-BSA and then rinsed with PBS 1X. Specific secondary antibodies conjugated with fluorophores were used for 1h at RT in 3% PBS-BSA. Finally, the cells were washed in PBS 1X. Nuclei were stained with DAPI and coverslips were mounted using glycerol 50% in PBS 1X. Primary and secondary antibodies are listed in Table 4.

3.11 Statistical analysis

The results from independent biological replicates in triplicate are expressed as mean ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test (Prism 6 software) with a p-value cut-off of 0.05.

TABLES

Table 1. Sequences of primers used for qPCR analysis.

Gene symbol	Ensemble code	Primer sequence
HPRT1	ENSG00000165704	F_5'- CGGGGACTTTGGGATGTC -3' R_5'- CGCTTTCCGTCGTGAATTTC -3'
GAPDH	ENSG00000111640	F_5'- CGGGGACTTTGGGATGTC -3' R_5'- CGCTTTCCGTCGTGAATTTC -3'
NTRK2	ENSG00000148053	F_5'- CGGGGACTTTGGGATGTC -3' R_5'- CGCTTTCCGTCGTGAATTTC -3'
ARC	ENSG00000198576	F_5'- GAGTCCTCAAATCCGGCTGA -3' R_5'- GCACAGCAGCAAAGACTTT -3'
SIK1	ENSG00000142178	F_5'- AAGACCGAGAACCTCCTGCT -3' R_5'- GTGGACAGAGGCTCTCCTGA -3'
NR4A1	ENSG00000123358	F_5'- CACAGCTTGCTTGTCGATGT -3' R_5'- GGTTCTGCAGCTCCTCCAC -3'
ATF3	ENSG00000162772	F_5'- ATCACAAAAGCCGAGGTAGC -3' R_5'- TCCTTCTTCTTGTTTCGGCAC -3'
EGR1	ENSG00000120738	F_5'- GAGCAGCCCTACGAGCAC -3' R_5'- GGCCACAAGGTGTTGCCA -3'
NR4A3	ENSG00000119508	F_5'- TTTGGAGCTGTTTGTCCTCA -3' R_5'- CACTCCCCAAATCCACGAAG -3'
RHOB	ENSG00000143878	F_5'- TTCGAGAACTATGTGGCCGA -3' R_5'- GCACATGAGAATGACGTCGG -3'
SYBU	ENSG00000147642	F_5'- AGAGCAGAGAGTTCAACCCC -3' R_5'- AGCATCTGAGGGTGTCTTCA -3'
LINC00473-202	ENST00000455853	F_5'- GTCAGCATACTTTGGCGGAC -3' R_5'- GTTGGTGCACGTGGGAGT -3'
RP11-182L21.2	ENST00000431300	F_5'- CACTTTGAGGTTCCCACTGC -3' R_5'- TTGATGGCAGGGAGACGC -3'
LINC01089	ENST00000545885	F_5'- AGCAGAACGTGAGGGTGTAA -3' R_5'- AGAGTCAAACTAGGCCTGCC -3'
HAND2-AS	ENST00000505032	F_5'- TACGAAGACCTTGGGCGATT -3' R_5'- GCGTTTAATGGTTCCCCTCC -3'
lnc-NPAS4	ENST00000526186	F_5'- TAGACCACCTGAGGATGACC -3' R_5'- AGAGGGCTGTCAAAGTGTGA -3'
MIAT	ENST00000423278	F_5'- CATGTGGTTCCAGACACGTT -3' R_5'- CCTTCTGTCTCCTCTGTCCC -3'
CYP26a1	ENSG00000095596	F_5'- GCAGCCACATCTCTGATCACT -3' R_5'- TGTTGTCTTGATTGCTCTTGC -3'
RARα	ENSG00000131759	F_5'- CGTGTCTCTCTGGACATTGA -3' R_5'- CCGAAGCCAGCGTTGTGCAT -3'
RARβ	ENSG00000077092	F_5'- ATCCGAAAAGCTCACCAGGA -3' R_5'- CTGAATTTGTCCCAGAGGCC -3'
RARγ	ENSG00000172819	F_5'- AGGAATCGCTGCCAGTACTG -3' R_5'- GCTTTGCTGACCTTGGTGAT -3'
BDNF	ENSG00000176697	F_5'- ACACAAAAGAAGGCTGCAGG -3' R_5'- TGCTATCCATGGTAAGGGCC -3'
NGFR	ENSG00000064300	F_5'- GAGCCTGCATGACCAGCA -3' R_5'- GCAGAGCCGTTGAGAAGC -3'
EGR1 Δ141-278	ENSG00000120738	F_5'- GAGCAGCCCTACGAGCAC -3' R_5'- GTGCGGCTCTCCAGGGAA -3'

Table 2. Sequences of gRNAs and genotyping primers.

Name	Sequence
EGR1 gRNA	TOP_5'- CACCGCTGCAGATCTCTGACCCGTT -3' BOTTOM_5'- AAACAACGGGTCAGAGATCTGCAGC -3'
EGR1 sequencing primers	F_5'- CCGACACCAGCTCTCCAG -3' R_5'- CTGCGGTCAGGTGCTCGTAG -3'
LINC00473-202 gRNA_A	TOP_5'- ACCGTGGGGGTGCTTCCTCGTTCC -3' BOTTOM_5'- AAACGGAACGAGGAAGCACCCCCA -3'
LINC00473-202 gRNA_B	TOP_5'- ACCGTGAGTCTACGTGCTATAGCC -3' BOTTOM_5'- AAACGGCTATAGCACGTAGACTCA -3'
LINC00473-202 sequencing primers	F_5'- GAACTGTCGGCTGCGGC -3' R_5'- TTCTCCAGTTACCACCCACC -3'

Table 3. Primary and secondary antibodies for Western Blot.

Antibody	Dilution
Anti-EGR1 Polyclonal Antibody (Immunobiological Sciences), rabbit (AB-83620)	1:1000
Anti-GAPDH Monoclonal Antibody (Immunological Sciences), mouse (MAB-91903)	1:5000
Donkey Anti-rabbit IgG Secondary Antibody, HRP Conjugate (A120-108P)	1:2000
Goat Anti-mouse IgM Secondary Antibody, HRP Conjugate (BA1075)	1:5000

Table 4. Primary and secondary antibodies for Immunofluorescence.

Antibody	Dilution
Anti-EGR1 Polyclonal Antibody (Boster Bio), rabbit (PA2177)	1:100
Anti-Tubulin beta-III Monoclonal Antibody (Elabscience®), mouse (E-AB-20095)	1:200
Goat Anti-Rabbit IgG Secondary Antibody, CY5 Conjugate (Bethyl, A120-201C5)	1:500
Sheep Anti-Mouse IgG Secondary Antibody, FITC Conjugate (Bethyl, A90-146F)	1:500

4. RESULTS

PART I

Identification and preliminary characterization of genes involved in Early Response to BDNF Stimulation

4.I.1 Identification on new players in early response to BDNF stimulation

The original SH-SY5Y cell line comprises at least two morphologically and biochemically distinct phenotypes: neuroblastic (N-type) and a low proportion of epithelial-like (S-type; Encinas et al., 2000) phenotypes. Taking into account that the expression level of long non-coding RNA is much lower compared to coding RNA, and that even little variation in their amount could have a great functional impact, we decided to work on a more homogenous population by performing an enrichment procedure to isolate the N-subtype cells (see Materials and Methods). To check the efficacy of the enrichment procedure, we analysed the N-type enriched population before and after RA treatment. In fact, during RA treatment, N-type cells undergo morphological changes, while S-type cells do not change their shape and progressively increase their percentage in the population. As shown in Figure 7 (left panel), in their undifferentiated state, N-type cells have small, rounded cell bodies with short neuritic processes, while S-type cells are larger and flatter with a large cytoplasmic/nuclear ratio (Figure 7, left panel). After the enrichment procedure, the N-type cells appeared clearly isolated in the cell plate by microscopy observation of the cell shape (Figure 7, middle panel). The enrichment was better evaluated and confirmed after RA-induced differentiation, when the N-type cells became morphologically more like primary neurons (Figure 7, right panel).

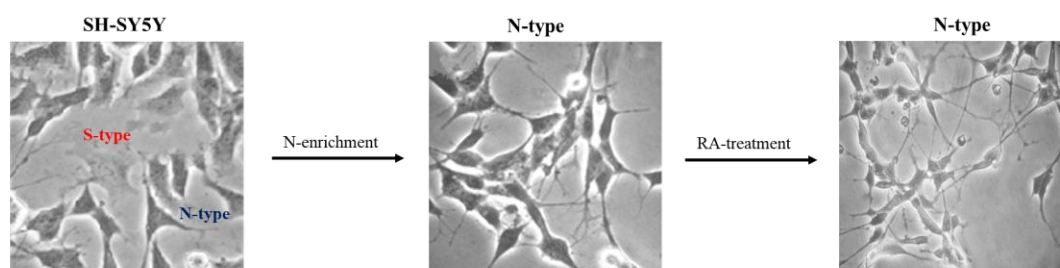

Fig 7. N-enrichment procedure for SH-SY5Y cells. Left panel: photomicrograph showing the original SH-SY5Y population including both N-type and S-type cells; middle panel: photomicrograph showing the N-type-enriched SH-SY5Y population; right panel: photomicrograph showing the morphological changes of the N-type cells after 6 days of RA-induced differentiation.

We sought to unravel the gene expression changes occurring immediately after BDNF stimulation, considering that the genes involved in this initial phase play critical roles in long-

lasting neuronal changes during development by regulating many downstream target genes. In particular, we used microarray technology to survey long non-coding RNAs and coding RNAs involved in the BDNF function. With this scope, we preliminary validated our experimental model by assessing the BDNF high-affinity receptor *NTRK2* expression level. In fact, it is reported that RA induces the expression of the *NTRK2* in SH-SY5Y cells, making them responsive to BDNF (Kaplan et al., 1993; Encinas et al., 2000). The results of the qPCR analysis reported in the Figure 8, showed that the N-enriched population responded to RA stimulation increasing *NTRK2* expression with the highest level after 6 days of treatment.

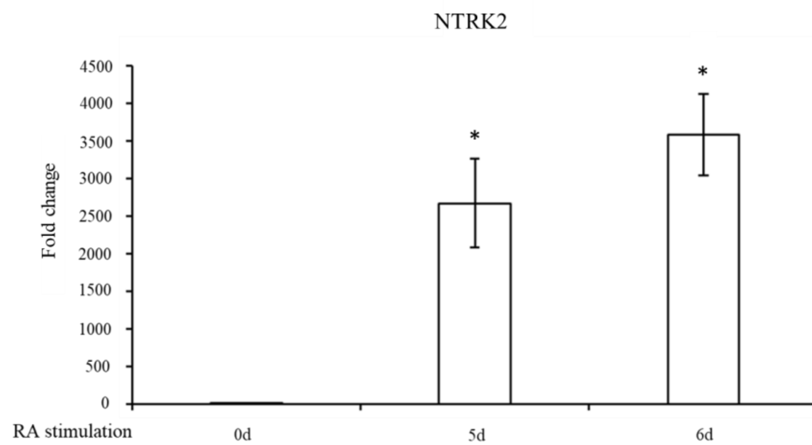

Fig 8. Analysis of *NTRK2* mRNA level after RA treatment by qPCR analysis. *NTRK2* expression level in N-enriched SH-SY5Y cells differentiated with RA for 5 and 6 days; time (0d) represented the undifferentiated cells. *NTRK2* gene expression level was normalized to the reference transcript (*GAPDH*) and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0d was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0d (*$p < 0.05$) is shown. (d): days of RA treatment.

Based on the qPCR results, we decided to perform the BDNF stimulation after 6 days of RA treatment. In order to identify the most appropriate time point for the immediate early response to BDNF treatment, we carried out a preliminary investigation of the expression pattern of three immediate early genes downstream of BDNF signalling *ARC* (Activity-Regulated Cytoskeleton-Associated Protein), *SIK1* (Salt-Inducible Kinase 1) and *NR4A1* (Nuclear Receptor Subfamily 4). The expression of all three analyzed genes was induced by BDNF, with the transcript level change showing a peak after 1h of the neurotrophin treatment (Figure 9).

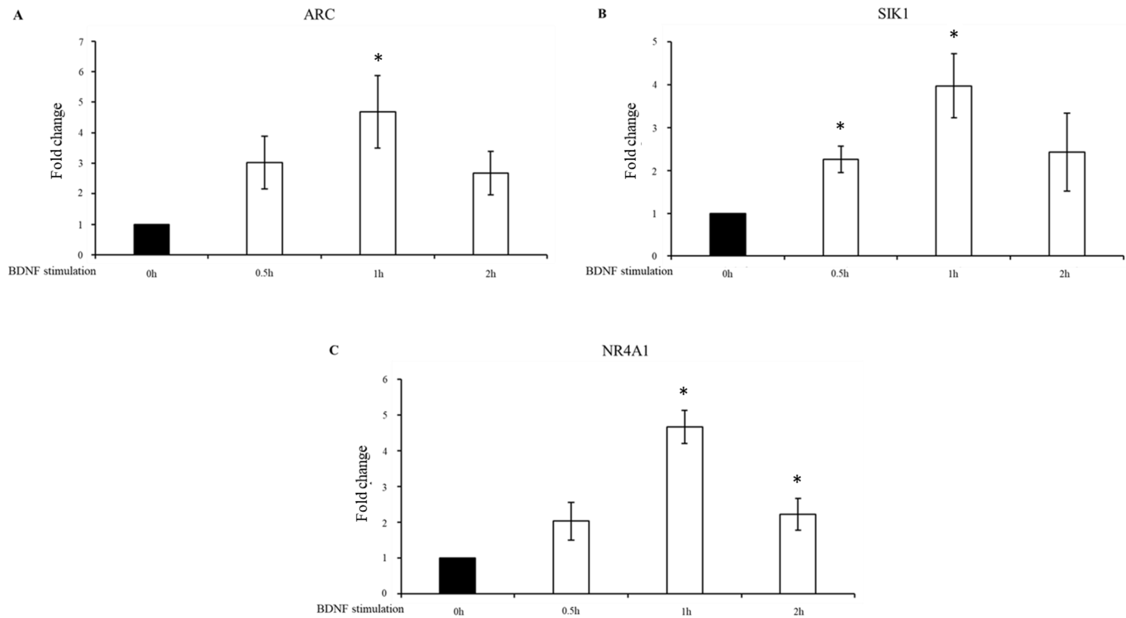

Fig 9. Expression analysis for 3 immediate early genes after BDNF stimulation. Expression level change in N-enriched SH-SY5Y cells after BDNF stimulation at the indicated time points for *ARC* **(A)**, *SIK1* **(B)**, and *NR4A1* **(C)**. Time (0h) represented cells after 6 days of RA-induced differentiation. Gene expression level was normalized to the reference gene (*GAPDH*) and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0h was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0h ($*p < 0.05$) is shown.

In light of these results, we carried out a microarray analysis on RNAs extracted after 1h of BDNF treatment. The array used for the analysis is characterized by probes designed on GENCODE v15 lncRNA annotation. This platform enables the analysis of gene expression changes for a total of 22,001 lncRNAs and 17,535 genes encoding proteins. I found that many lncRNAs and coding genes significantly (p-value < 0.05) changed their transcript level after 1h of BDNF treatment. A fold change of > 1.5 was found in 41 lncRNAs (25 up and 17 down regulated) and 40 mRNAs (31 up and 9 down regulated) (Table 5).

Transcript	number	differentially expressed	FC ≥ 1.5
lncRNA	22,001	155	41
mRNA	17,535	238	40

Table 5. Summary of microarray analysis.

A panel of both differentially expressed lncRNAs and mRNAs with fold change above 1.5 was validated by qPCR (Figure 10).

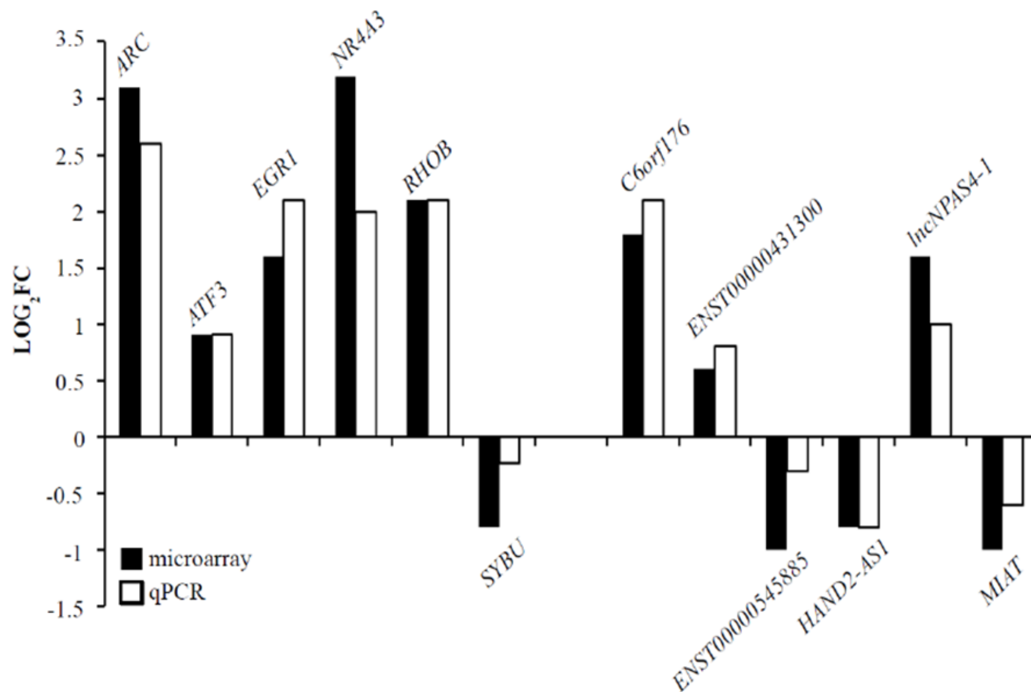

Fig 10. qPCR validation of some differentially expressed mRNA and lncRNAs (Aliperti and Donizetti, 2016).

The differentially expressed lncRNAs were classified in accordance with the definition reported in the Materials and Methods section. Most differentially expressed lncRNAs were included in the intergenic and antisense classes (Table 6).

Class	Number
Intergenic	19
Antisense	13
Overlapping	9
Bidirectional	1

Table 6. Summary of lncRNAs classification.

It is interesting to note that the vast majority of top regulated coding transcripts with fold changes above 2 are classical immediate early genes (IEGs); many of these genes are involved in different biological processes in response to various neural stimuli, such as *NR4A3*, *ARC*, *EGR1* and *DUSP5* (Table 7). Notably, two of the top regulated lncRNAs (*C6orf176* and *MIAT*) have also been shown to have an expression pattern that resembles the kinetics of immediate early response genes (Reitmair et al., 2012; Barry et al. 2014).

coding	regulation	FC	lncRNA	regulation	FC
NR4A3	up	9.2	C6orf176	up	3.8
ARC	up	8.3	lnc-NPAS4-1	up	3.1
RHOB	up	4.3	lnc-WDR1-1	up	2.8
FAM46A	up	3.6	IGFBP7-AS1	up	2.1
EGR1	up	3.0	lnc-ZSCAN10-4	up	2.1
DUSP5	up	3.0	MIAT-003	down	2.1
KLF10	up	2.3	MIAT-001	down	2.0
MAP3K14	up	2.0	lnc-RHOF-1	down	2.0
F3	up	2.0			

Table 7. Top regulated coding and long non-coding transcripts with fold change above2.

lncRNAs belong to a relatively recent class of RNA, and many important features regarding their origin and role remain largely unexplored. One of the most interesting debate is over the possibility that some lncRNAs encode for evolutionary new peptides/proteins. In this regard, we submitted the 41 differentially expressed lncRNAs to the analysis of CPAT software (http://lilab.research.bcm.edu/cpat/) (Wang et al., 2013) to assess their protein-coding potential (Table 8). We found that 38 lncRNAs have a coding probability value below the cut off (0.364), while the remaining 3 lncRNAs have a coding probability value above the cut off. Among these three transcripts, C6orf176 had the highest fold change in our microarray analysis.

Alternative gene name	RNA size	ORF size	Coding Probability	Coding Label	FC (Abs)
C6ORF176	1822	561	0.47	yes	3.40
RP11-867G23.10	637	123	0.006	no	3.11
RP11-448G15.3	3510	297	0.03	no	2.76
MIAT	9942	309	0.01	no	2.09
IGFBP7-AS1	1389	168	0.0001	no	2.09
RP11-473M20.16	748	120	0.003	no	2.07
MIAT	10215	300	0.02	no	2.03
AC084018.1	559	144	0.002	no	2.00
AC093673.5	648	78	0.03	no	1.95
AC084018.1	875	270	0.04	no	1.93
RP11-471J12.1	514	186	0.02	no	1.92
RP11-466F5.8	2149	159	0.03	no	1.87
MIAT	10142	300	0.02	no	1.84
HOXD-AS1	609	171	0.01	no	1.81
MIAT	10068	309	0.01	no	1.75
RP4-564F22.5	2004	363	0.01	no	1.74
MALAT1	480	120	0.007	no	1.73
NR_026991.1	2193	294	0.16	no	1.73
AC002310.7	1488	324	0.31	no	1.72
MALAT1	1519	141	0.002	no	1.71
AC084018.1	479	171	0.08	no	1.70

RP11-370A5.1	413	171	0.01	no	1.68
AC084018.1	1220	348	0.06	no	1.68
RAD51-AS1	718	105	0.004	no	1.67
SOX9-AS1	394	132	0.05	no	1.66
TTTY7	1367	180	0.01	no	1.65
AC084018.1	1521	246	0.06	no	1.65
AC046143.3	972	105	0.001	no	1.62
H1FX-AS1	917	252	0.05	no	1.61
CTC-444N24.8	1385	228	0.02	no	1.59
RAD51-AS1	419	171	0.006	no	1.58
RP11-182L21.2	3272	585	0.54	yes	1.58
RP11-535A19.2	444	93	0.04	no	1.56
ENST00000529707	1130	201	0.01	no	1.56
ENST00000589496	5702	783	0.91	yes	1.55
RP11-404P21.5	460	261	0.01	no	1.53
RP11-217B1.2	1109	204	0.01	no	1.52
LINC00324	2082	447	0.24	no	1.51
CTB-55O6.12	686	216	0.10	no	1.51
linc-RPL19-3	2365	255	0.01	no	1.50
RP11-473I1.10	8334	273	0.005	no	1.43

Table 8. Results of CPAT analysis (Aliperti and Donizetti, 2016).

Five lncRNAs with a fold change above 1.5 were hypothesized to play a role in physiological and pathological processes in neuronal cells and/or as a regulators of gene expression (Table 9).

Gene symbol	Regulation	Function
MALAT1 (Neat2)	downn	Control of the expression of genes involved in synapse function (Bernard et al., 2010). Down-regulation led to cell arrest in the G1/S or G2/M phase (Yang et al., 2013).
MIAT (Gomafu)	down	Down-regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing (Barry et al., 2014). Decreased in the medial prefrontal complex following fear conditioning and knockdown of promoted stress reactivity and anxiety-like behavior (Spadaro et al., 2015). Neurogenic commitment and neuronal survival, sustained overexpression of Miat promoted neuronal death (Aprea et al. 2013).
HAND2-AS1 (Dein)	down	Highly expressed in stage IVS neuroblastoma (Voth et al., 2007). Expression is neuroblastoma is coregulated together with HAND2 (Voth et al., 2009).
C6orf176 (LINC00473)	up	A possible regulatory function in response to cAMP signalling (Reitmair et al., 2012).
HOXD-AS1	up	It is induced by RA, could be regulated via PI3K/Akt pathway and controls genes involved in RA signaling, angiogenesis and inflammation (Yarmishyn et al., 2014).

Table 9. Differentially expressed lncRNAs that have been identified in literature (Aliperti and Donizetti, 2016).

These evidences supported the hypothesis of an involvement of lncRNAs in BDNF-mediated molecular effects. To assess the putative role of the differentially expressed lncRNAs, we examined the function of coding genes located near the lncRNAs in the genome. In fact, very often the expression of the lncRNAs affect *cis* regulation of nearby transcription. We first identified the neighbouring coding genes by using the GREAT tool (http://bejerano.stanford.edu/great/public/html/) (McLean et al., 2010), and then used this gene list together with the list of differentially expressed coding RNAs found in microarray survey analyse using the DAVID tool (http://david.abcc.ncifcrf.gov/) (Dennis et al., 2003). This tool outputs a wide range of information, including Gene Ontology (GO) and pathway enrichment, providing an overview of the functions and relationship of the differentially expressed genes. The most enriched biological processes are related to transcription regulation and RNA metabolic process, while there are also processes related to chromatin organization and function (Figure 11 A). The involvement of genomic loci of lncRNAs in transcription control and chromatin structure was also revealed by GO in terms of molecular function and cellular component (Figures 11 B and C). This analysis corroborated the hypothesis that lncRNAs and coding genes identified in the early phase of BDNF stimulation may act in concert to regulate the gene expression of downstream genes at a transcriptional level.

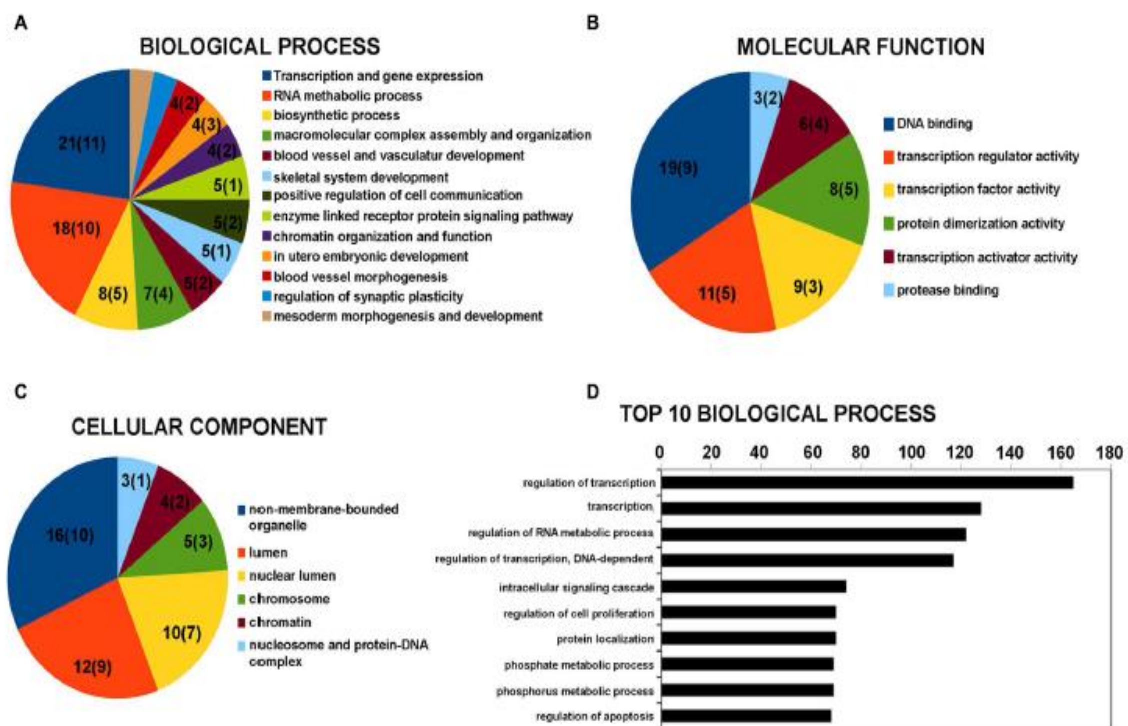

Fig 11. Gene Ontology (GO) enrichment analysis for lncRNAs by DAVID bioinformatics tool. (A) GO analysis of lncRNA-target genes + differentially expressed coding genes according to biological process. **(B)** GO analysis of lncRNA-target genes + differentially expressed coding genes according to molecular function. **(C)** GO analysis of lncRNA-target genes + differentially expressed coding genes according to cell component. **(D)** Top 10 biological processes for coding genes that are target of miRNA potentially regulated by the differentially expressed lncRNAs. Grouped GO terms are reported as following

with the ID of the single gene ontology term. transcription and gene expression: GO:0045449 + GO:0006350 + GO:0006355 + GO:0006357 + GO:0045893 + + GO:0045941 + GO:0010628 + GO:0045944; RNA methabolic process: GO:0051252 + GO:0051254 + GO:0045935 + GO:0051173 + GO:0010604; biosynthetic process: GO:0010557 + GO:0031328 + GO:0009891; macromolecular complex assembly and organization: GO:0065003 + GO:0043933 + GO:0034622 + GO:0034621; blood vessel and vasculatur development: GO:0001568 + GO:0001944; chromatin organization and function: GO:0006334 + GO:0031497 + GO:0065004 + GO:0034728 + GO:0006323 + GO:0006333; regulation of synaptic plasticity: GO:0048168 + GO:0048167; mesoderm morphogenesis and development: GO:0048332 + GO:0007498. The number on the pie chart indicates the number of differentially expressed coding genes associated to the GO term (or GO group), while the number in brackets indicates the number of genes located near the differentially expressed lncRNAs (Aliperti and Donizetti, 2016).

The list of differentially expressed lncRNAs was submitted to the lncRNome database (http://genome.igib.res.in/lncRNome) (Bhartiya et al., 2013), that contains information about 6,800 protein binding sites. These data were obtained from PAR-CLIP and CLIP-Seq analyses. The analysis returned a result for only a small percentage of the lncRNAs in question. In general, the presence of binding sites can be observed for a small group of proteins, such as PTB, PUM2, QKI, IGF2BP1, IGF2BP2, IGF2BP3, TNRC6A and TNRC6C (Table 10).

Sequence Name	Alternative gene name	Proteins (lncRNome)
ENST00000455853	LINC00473 (C6orf176)	PTB
ENST00000570843	RP11-473M20.16	PUM2, QKI, IGF2P1, IGF2BP2, IGF2BP3, TNRC6A, TNRC6C
ENST00000549329	HOXD-AS1	IGF2BP2
ENST00000544868	MALAT1	PTB, PUM2, QKI, IGF2BP1, IGF2BP2, IGF2BP3, TNRC6C
ENST00000508832	MALAT1	PTB, PUM2, IGF2BP1, IGF2BP2, IGF2BP3, TNRC6C
ENST00000511998	H1FX-AS1	IGF2BP1, IGF2BP2, IGF2BP3
ENST00000569473	RP11-217B1.2	PUM2, QKI, IGF2BP1, IGF2BP2, IGF2BP3, TNRC6A, TNRC6B, TNRC6C
ENST00000315707	LINC00324	PUM2, QKI, IGF2BP2, IGF2BP3, TNRC6A, TNRC6B, TNRC6C
ENST00000574616	RP11-473I1.10	PTB, IGF2P1, IGF2BP2, IGF2P3

Table 10. lncRNAs – proteins interaction.

Many pieces of evidence demonstrated that lncRNAs may also influence gene expression at a post-transcriptional level. The effect is achieved by interfering with miRNAs pathways as competitive endogenous RNAs (ceRNAs). In this process, thanks to complementary sequences, the lncRNA interact with miRNA that in turn is no longer available to bind its target mRNA, thus reducing the repression of this mRNA. lnCeDB (http://gyanxet-beta.com/lncedb/) provides a database of human lncRNAs (from the GENCODE version 19) that can potentially act as ceRNAs. Using this database, we found a list of miRNAs, of which seven are associated with neuropathologies as reported in the mir2disease database (Table 11). To acquire information on

biological processes affected by the putative lncRNA-miRNA interaction, we retrieved a list of experimentally validated mRNA targets of all previously-identified miRNAs and performed a functional analysis with the DAVID tool. In the top 10 biological processes, the most represented categories are related to the regulation of transcription and RNA metabolic processes (Figure 11 D).

miRNA symbol	neuropathology
hsa-miR-339-5p	Neurodegeneration
hsa-miR-433	Parkinson's disease
hsa-miR-133b	Parkinson's disease
hsa-miR-346	Schizophrenia
hsa-miR-328	Alzheimer's disease
hsa-miR-299-3p	Alzheimer's disease
hsa-miR-422a	Multiple sclerosis

Table 11. Putative miRNA targets on lncRNA and their involvement in neuropsychiatric diseases. (Aliperti and Donizetti, 2016).

PART II

EGR1 in neuronal differentiation

4.II.1 *EGR1*

In the above analysis, we provided an overview of a number of coding and non-coding genes involved in the BDNF function during neuronal differentiation. Among the coding genes that we found regulated by neurotrophin, *EGR1* is likely one of the most enigmatic concerning its role in neuronal differentiation. In fact, the function of EGR1 is particularly well-documented in the adult nervous system, where it acts as an activity-dependent transcription factor highly expressed in long-term potentiation (LTP) and long-term depression (LTD) and is often used as a marker of neuronal stimulation. Because of an apparent normal embryonic brain development in EGR1-KO mice and relatively few studies on its expression pattern in embryos (Watson and Milbrandt, 1990), the involvement of EGR1 in neuronal development during embryogenesis remains elusive. In addition, one limitation in the use of *in vivo* models such as EGR1-KO mice is questionable, mainly because it is difficult to exclude the fact that other transcription factors related to *EGR1* (that is *EGR2*, *EGR3*, and *EGR4*) may compensate for the loss of function of this gene. Recently, a fluorescent transgene approach provided evidence of a wide distribution of $Egr1^{+}$/$Sox2^{+}$ cells, suggesting a general role for EGR1 in cellular regulation during late embryonic/early postnatal brain development (Wells et al., 2011). Beyond this sporadic evidence, the hypothesis of a function for EGR1 in neurons development is mainly supported by the use of rat/mouse neuronal cancer cell lines as progenitor-type models, based on their ability to differentiate after specific stimulations. In this regard, the discovery of *EGR1* three decades ago was strictly related to its function in neuronal maturation (Milbrandt, 1987). In that seminal paper, Milbrandt (1987) described the identification of a cDNA clone for a transcript able to quickly respond to NGF stimulation in the rat pheochromocytoma PC12. NGF and other neurotrophic factors play an essential role in the regulation of the growth, survival, and differentiation of neurons. In this regard, the expression induction of EGR1 by different neurotrophic factors including BDNF, as reported in the present thesis in a human neuronal model, points out that this gene may play a key role in the transcriptional programs underlying survival and differentiation of neuronal cells. With the aim to explore the role of human EGR1 in neuronal maturation, we designed different molecular approaches to interfere with *EGR1* gene function. One of the main strategies was based on the CRISPR-Cas9 technology to knockout the *EGR1* gene in the human neuroblastoma cell line SH-SY5Y.

4.II.2 *EGR1* expression level during RA-dependent differentiation

We first sought to check if *EGR1* is involved in the early phases of differentiation during RA treatment. Different reports demonstrated that RA stimulation can induce an increase of *EGR1* expression level in different cell cultures (Edwards et al., 1991; Larsen et al., 1994; Lee and Kim, 2004). In this regard, we evaluated the RA induction of the *EGR1* gene in the N-enriched SH-SY5Y model used in this thesis. As shown in the Figure 12 A, we did not find a RA-dependent increase of *EGR1* transcript level in the early period after stimulation; a significant increase was observed after 24h and 48h of RA treatment with a corresponding increase in protein levels (Figures 12 A and B).

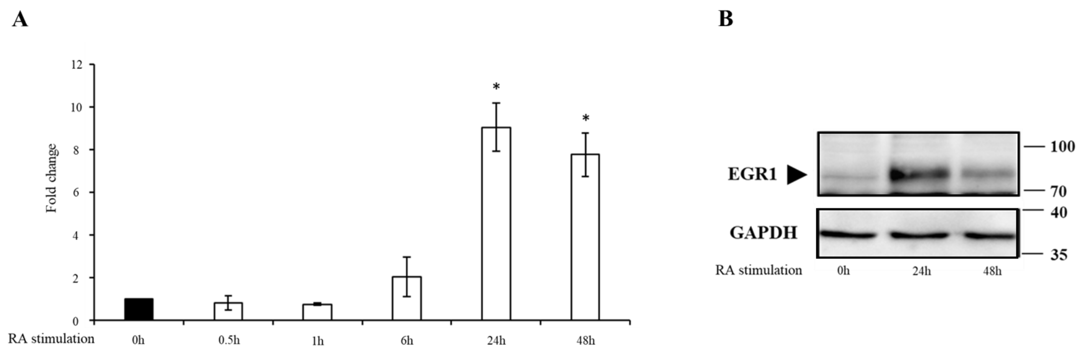

Fig 12. *EGR1* expression level during RA-treatment of N-enriched SH-SY5Y cells. (A) Transcript level changes examined by qPCR at the indicated time points. *EGR1* gene expression level was normalized to the reference transcript *GAPDH* and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0h was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0h (*$p < 0.05$) is shown. **(B)** EGR1 protein level analysed by Western Blot at the indicated time points. The level of GAPDH protein was used as a loading control.

4.II.3 EGR1 knockout cell line generation and validation

To functionally evaluate the role of *EGR1* in human neuronal differentiation, we used the N-enriched SH-SY5Y neuroblastoma cell line to generate a KO cell line for EGR1. The cells were transfected with the all-in-one plasmid pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene plasmid #62988), which includes both the coding region for the Cas9 and the coding sequence for the gRNA (Figure 13).

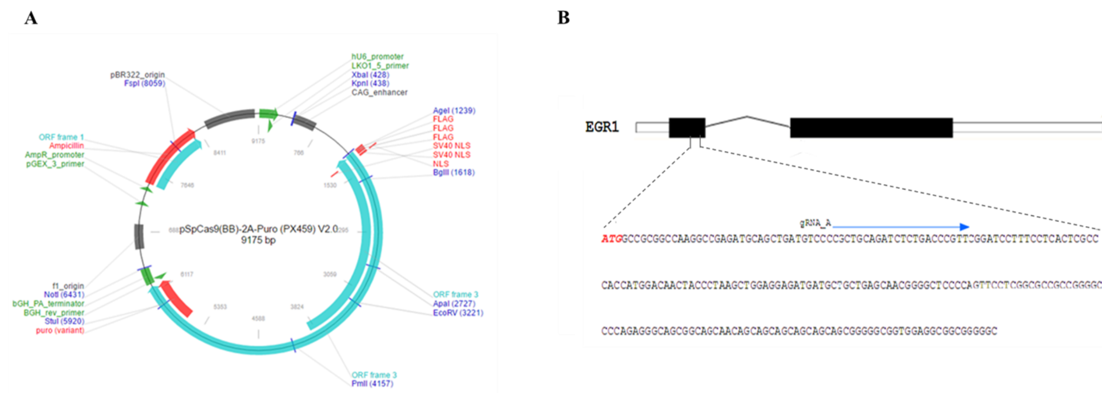

Fig 13. Basic elements of the CRISPR/Cas9 technology used to knockout the *EGR1* gene. (A) Graphical representation of the pSpCas9 (BB) -2A-Pure (PX459) V2.0 vector used for the CRISPR-Cas9 system (www.addgene.com). **(B)** Schematic representation of the *EGR1* transcript and the corresponding DNA sequence recognized by the gRNA. The rectangles represented the two exons of *EGR1* transcript with the black region representing the coding region. The ATG of the *EGR1* gene is highlighted in red, while the arrow indicated the gRNA target sequence.

After the puromycin selection of positive transfected clones, we isolated different single cell-derived populations. The target genomic region of two of the selected populations was amplified by PCR and subcloned in a vector to obtain the sequence of both the alleles separately and to confirm the KO generation. One of these two KO lines was used for further validation and analysis. As reported in Figure 13 C, this KO line carried two different genomic mutations: one single nucleotide insertion for one allele, and one single nucleotide deletion for the second allele (Figure 14).

Fig 14. Analysis of the *EGR1* genomic sequence in the KO cell line. Electropherograms of both WT and KO *EGR1* genomic sequence.

In both cases, the INDEL (INsertion/DELetion) determined the generation of a premature stop codon because of a frame-shift mutation. To further validate the EGR1-KO cell line, we employed western blotting analysis to provide evidence for the lack of functional protein production. The analysis showed that the protein level strongly increased in WT cells after 24h and 48h of RA

stimulation, while the corresponding protein was completely absent in KO cells (Figure 15). At that point, we wanted to further validate the EGR1-KO cell line and provide evidence that the loss of EGR1 protein production depended on the premature stop codon, generated by the genome editing strategy and not by misregulatation of transcript expression. The qPCR analysis of the EGR1-KO cell line showed that RA stimulation is still able to increase the transcript level after 24h and 48h of treatment, demonstrating that the *EGR1* transcript regulation and the upstream activating pathways were functional in the KO cell line (data not shown).

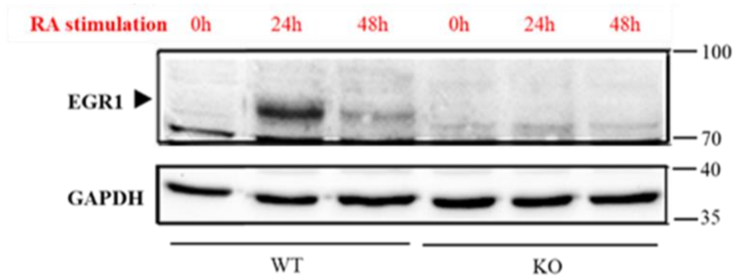

Fig 15. Validation of the EGR1 KO cell line by qPCR and western blotting analysis. EGR1 protein level analysed by Western Blot at the indicated time points in WT and KO cells. The level of GAPDH protein was used as a loading control.

4.II.4 KO cells are unable to properly differentiate under RA stimulation

To test whether the knockout of the EGR1 gene affected the RA-induced differentiation process, we followed the morphological changes after RA treatment by microscopy observation. EGR1-WT cells differentiated normally under 10μM RA stimulation, as evidenced by changes in the shape of soma and neurites elongation and branching (Figure 16). In contrast, EGR1-KO cells did not show any differentiation landmarks and underwent cell shrinkage and detachment (Figure 16).

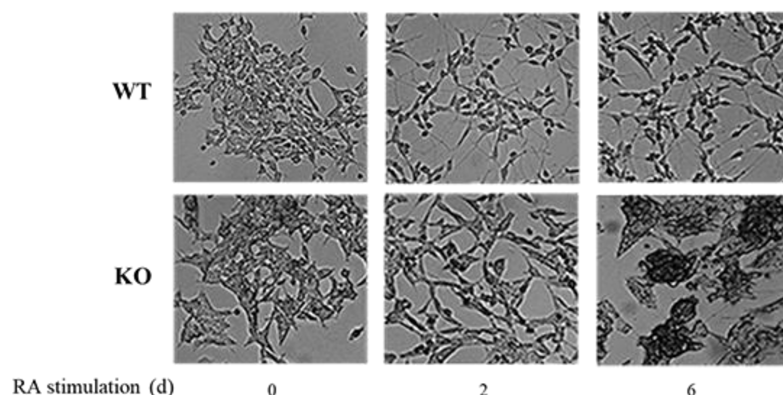

Fig 16. Effect of RA-induced differentiation on WT and KO cells. Time (0) represented starved cells before RA treatment at 10μM concentration. (d): days of RA treatment.

Interestingly, this effect was also confirmed with smaller concentrations (3 and 6 μM) of RA (Figure 17), demonstrating that EGR1 is required for a correct RA-induced differentiation.

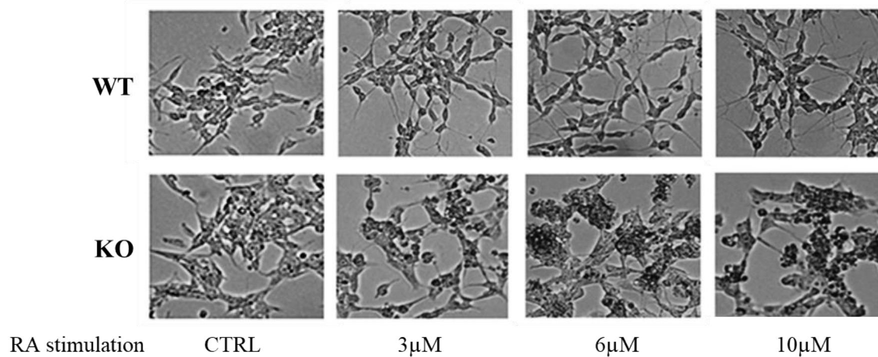

Fig 17. Effect of RA-induced differentiation on EGR1 WT and KO cells. Cells treated with RA at different concentrations for 6 days. CTRL represented cells grown in the presence of only DMSO.

4.II.5 Effects of EGR1-KO on the RA pathway

These results prompted us to investigate if the inability to differentiate under RA treatment depends on the misregulation of RA signalling. To make this determination, we analysed the expression levels of genes involved in the RA pathway including retinoic receptors (*RARα*, *RARβ*, and *RARγ*) and *CYP26a1*, a key component of RA metabolic inactivation. As shown in Figure 18, the expression level of all analysed genes increased after 2 days of morphogen treatment in EGR1-WT cells. Differently, *RARγ* showed a higher increase in EGR1-KO cells compared to WT cells and, worthy to note, *CYP26a1* showed an extremely lowered increased expression after 1 and 2 days of RA in KO than WT cells (Figure 18), corroborating the idea that KO cells are affected in the RA signalling pathway.

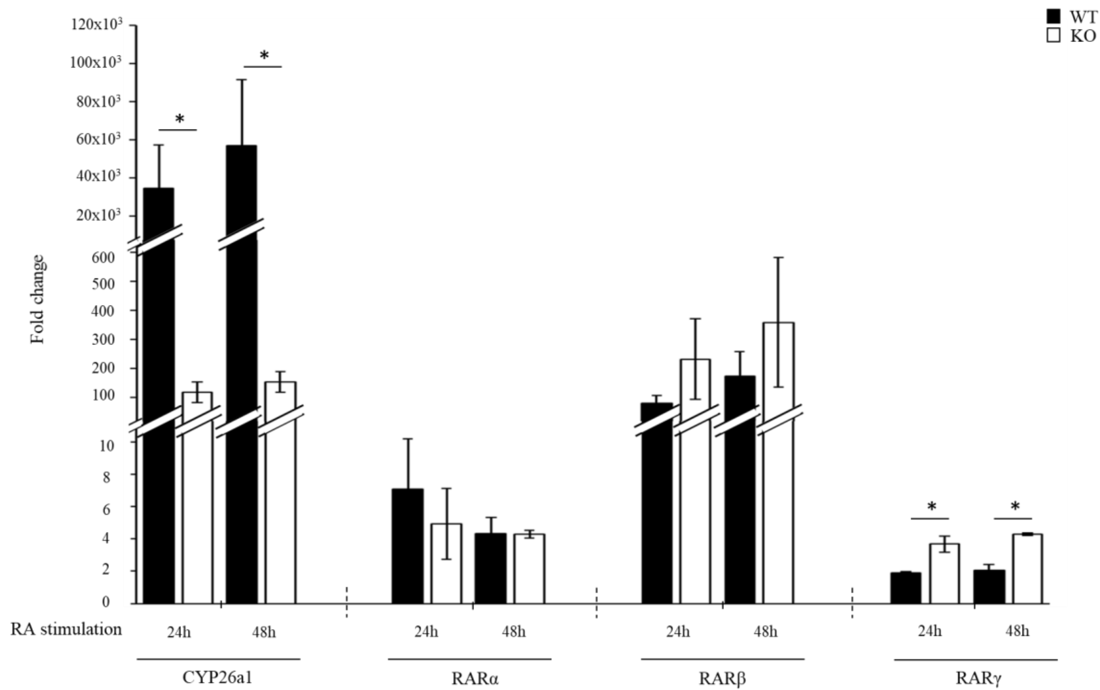

Fig 18. Expression pattern comparison of RA signalling genes in EGR1 WT and KO cells. Expression level change in EGR1 WT and KO cells after RA stimulation at the indicated time points for *CYP26a1*, *RARα*, *RARβ* and *RARγ*. Gene expression level was normalized to the reference transcript *GAPDH* and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0 was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference (*$p < 0.05$) is shown.

4.II.6 Dysregulation of BDNF system in EGR1-KO cells

During RA stimulation, SH-SY5Y normally increases the expression level of *NTRK2*, leading to the idea that RA may also affect the expression of *BDNF* and its receptors during differentiation. In this regard, we investigated mRNA expression level changes of *BDNF*, the high-affinity receptor *NTRK2* and the low-affinity *NGFR* (also known as p75). The transcript level of *BDNF* slightly increased after 2 days of RA treatment in both WT and KO cells, while a different pattern was observed for the receptors (Figure 19). The mRNA for the high-affinity *NTRK2* receptor progressively increased during RA treatment in WT cells, while a smaller increase was observed in KO after 1 and 2 days of RA treatment (Figure 19). Regarding the *NGFR* receptor, we observed no significant changes in the mRNA level after RA treatment in WT cells, and a large increase in KO cells at 1 and 2 days (Figure 19). These observations suggested that the effect of knockout of EGR1 in RA-induced differentiation may also involve BDNF signalling.

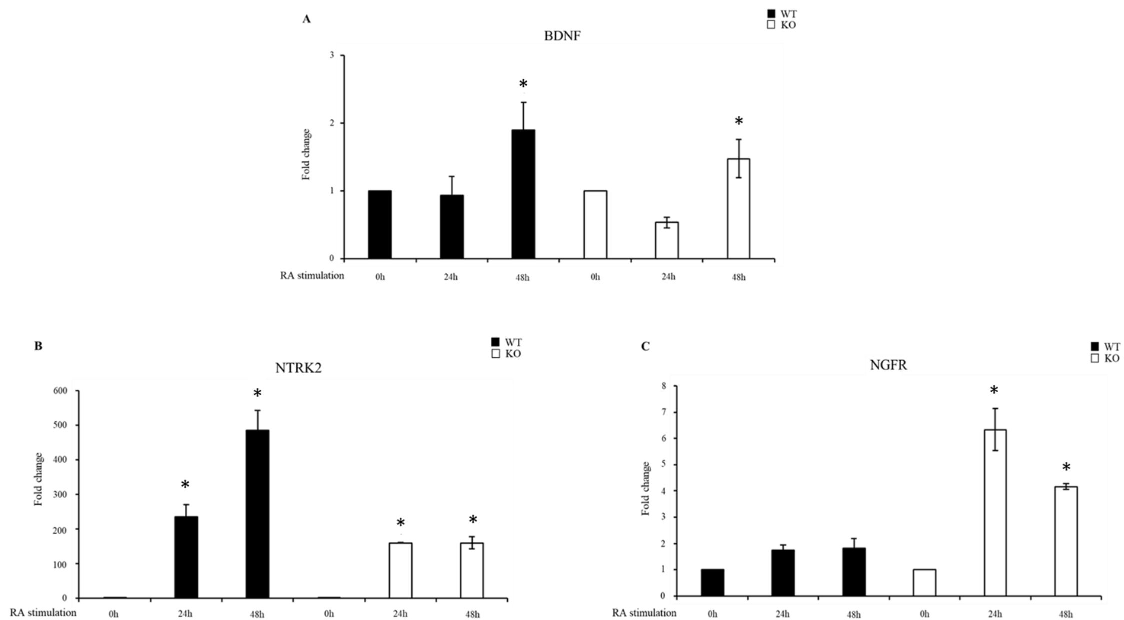

Fig 19. Expression pattern comparison of BDNF signalling genes in EGR1 WT and KO cells. Expression level change in EGR1 WT and KO cells after RA stimulation at the indicated time points for *BDNF* (**A**), *NTRK2* (**B**), and *NGFR* (**C**). Gene expression level was normalized to the reference transcript *GAPDH* and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0h was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0h

(*$p < 0.05$) is shown.

4.II.7 Effect of TPA stimulation on EGR1-KO differentiation

The effect of RA treatment on EGR1-KO highlighted that this transcription factor may be involved in the correct differentiation following RA stimulation. To further investigate the relevance of *EGR1* in neuronal differentiation, we wanted to broaden the analysis by using a different stimulating agent. As such, we employed TPA (12-O-tetradecanoyl-phorbol-13-acetate), another well-known differentiating agent that promotes cell proliferation arrest and induces differentiation of the human neuroblastoma SH-SY5Y cell line (Encinas et al., 2000). In this preliminary investigation, we compared the effect of TPA and RA using DMSO as a control. The microscopy analysis showed that cells stimulated with TPA or RA did not undergo a differentiation process and progressively led to cell death (Figure 20), demonstrating that EGR1-KO cells were unable to independently differentiate the stimulating agent used.

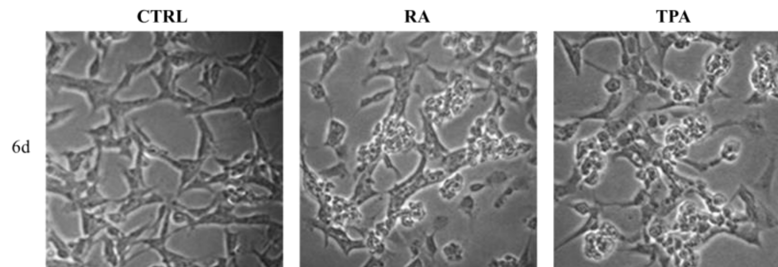

Fig 20. Effect of RA- and TPA-induced differentiation on KO cells. CTRL represented cells grown in the presence of only DMSO.

4.II.8 EGR1 knockout in iPSCs

The induced pluripotent stem cell (iPSC) technology has revolutionized many fields, in particular the study of the development and modelling of disease and cell therapy. Somatic cells can be reprogrammed in a pluripotent stem cell state, similar to the state present in very early embryogenesis through the transient expression of four transcription factors (Oct4, Sox2, Klf4 and c-Myc) (Takahashi and Yamanaka, 2006; Takahashi et al., 2007).

It is important to emphasize that such cells are diploid and karyotypically normal. They can self-renew for many cell divisions and can be differentiated into a wide range of different cell types. These characteristics lend themselves to the study of cell development and function in both normal and pathological states. In fact, since cells are reprogrammed at a very early stage of development, they can be used to monitor both developmental or differentiation defects, and the temporal sequence of events in the early stages of disease progression (Bassett, 2017).

The neuronal cells produced by iPSCs closely recapitulate the progression from early embryogenesis to late fetal periods *in vitro* and produce neuronal cells at various stages of maturity. These advantages make iPSCs a versatile tool for exploring early molecular and cellular phenotypes. In regard to the functional study of the transcription factor EGR1 in neuronal development in a more physiological system, it was interesting to try to extend the results obtained in the tumor cell line in the iPSCs. I spent 3 months in the laboratory of Dr. Michael Ziller at Max Planck Institute of Psychiatry (Munich) working with the HDF6 cell line. I transfected cells with the same vector used in SH-SY5Y cells. After the puromycin selection of positive transfected cells, I isolated different single clones and amplified the target genomic region by PCR. Unfortunately, none of the examined clones showed any mutation in the nucleotide sequence. It will therefore be necessary to make changes in the transfection protocol in order to increase editing efficiency.

4.II.9 Identification of a new alternative splicing isoform for *EGR1*

Among the strategies to investigate *EGR1* involvement in human neuronal differentiation, we planned to generate and use an expression vector for this transcription factor. This approach was designed to test the effects of EGR1 iper-expression in WT cells and to perform rescue experiments in KO cells. To clone the entire *EGR1* coding region, we designed a pair of primers spanning the start codon to the stop codon to use in PCR reactions. As a template, we used cDNA samples obtained by the retrotranscription of RNA extracted from N-type SH-SY5Y cells treated with RA for 6 days. As shown in the Figure 21, we surprisingly found two different amplicons: one of the expected size and an additional shorter one (Figure 21 A). We isolated and cloned both the PCR products to identify the corresponding sequences. As expected, the amplicon with the highest molecular weight corresponded to the entire coding sequence of *EGR1* reported in the genome databases; surprisingly, the shorter amplicon corresponded to the *EGR1* transcript characterized by a deletion of 414 nucleotides in length in the coding region (Figure 21 B). The bioinformatic translation of the alternative *EGR1* isoform showed that it retained the correct frame and encoded for a shorter protein compared to the canonical EGR1 (Figure 21 C). This alternative isoform lacks a protein region of 138 amino acids in length from the aa 141 to the aa 278.

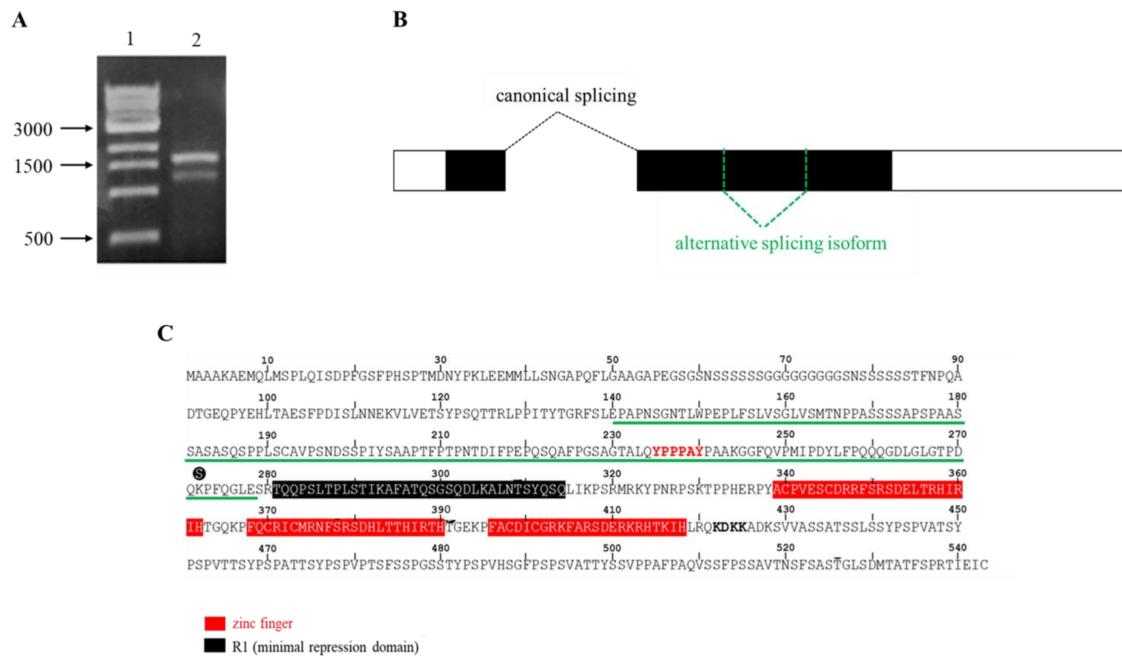

Fig 21. Identification of the *EGR1* splicing isoform. **(A)** Electrophoretic run of the PCR products related to *EGR1* CDS amplification. Lane 1: 1kb marker. Lane 2: PCR reaction for *EGR1* CDS. **(B)** Schematic representation of the two *EGR1* transcript isoforms. The rectangles represented the two exons of *EGR1* transcript with the black region representing the coding region. The alternative splicing was indicated by dashed green lines. **(C)** Amino acid sequence of the EGR1 protein. Underlined amino acid region represented the sequence missing in the putative alternative EGR1 isoform.

To provide a preliminary investigation of the relevance of this splicing isoform in neuronal differentiation, we analysed its expression pattern both downstream to BDNF and RA stimulation. To do so, we performed qPCR analysis by using specific primers designed to differentiate the two isoforms (see the Materials and Methods section for details). As shown in Figure 22, the transcript level of the isoform was unaffected during early RA treatment, while it increased after 24h and 48h of treatment in a manner similar to the canonical isoform (Figure 22 A). Regarding the expression pattern under BDNF stimulation, we performed the transcript level analysis at different time points. As shown in the Figure 22, the transcript level variation of the canonical *EGR1* isoform followed the typical kinetics of the immediate early genes, with a rapid increase until it reached a peak after 1h of stimulation followed by a rapid decrease (Figure 22 B.). The fold change of the splicing isoform at each point in time corresponded to that of the canonical isoform (Figure 22 B.). Both the expression pattern after RA and BDNF stimulation demonstrated that the alternative isoform may be involved in the *EGR1* function during neuronal differentiation.

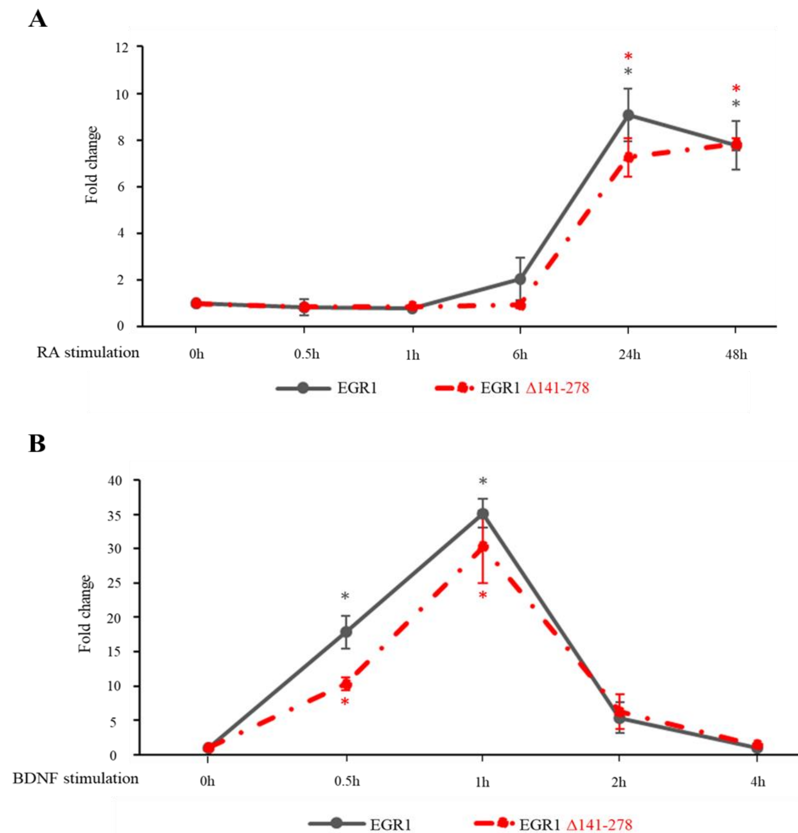

Fig 22. Expression pattern comparison of *EGR1* splicing isoforms under RA and BDNF stimulation. Expression level change in N-enriched SH-SY5Y cells after RA **(A)** and BDNF **(B)** stimulation at the indicated time points for the canonical *EGR1* (grey line) and the splicing isoform Δ141-278 (dashed red line). Time (0h) represented cells before RA treatment **(A)** or after 6 days of RA-induced differentiation **(B)**. Gene expression level was normalized to the reference gene (*GAPDH*) and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0h was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0h (*$p < 0.05$) is shown.

To gain preliminary insights into the molecular mechanism of action of the alternative EGR1 isoform, we analysed the cellular localization of this protein. To do so, we transfected expression vectors for both the canonical and alternative EGR1 isoform in HEK293T cells. Western blotting experiments on nuclear and cytoplasmic fraction showed that the EGR1 alternative isoform preferentially had nuclear localization (Figure 23 A), as also confirmed by immunofluorescence analysis (Figure 23 B). This result led us to hypothesize that the alternative EGR1 isoform may affect gene transcription.

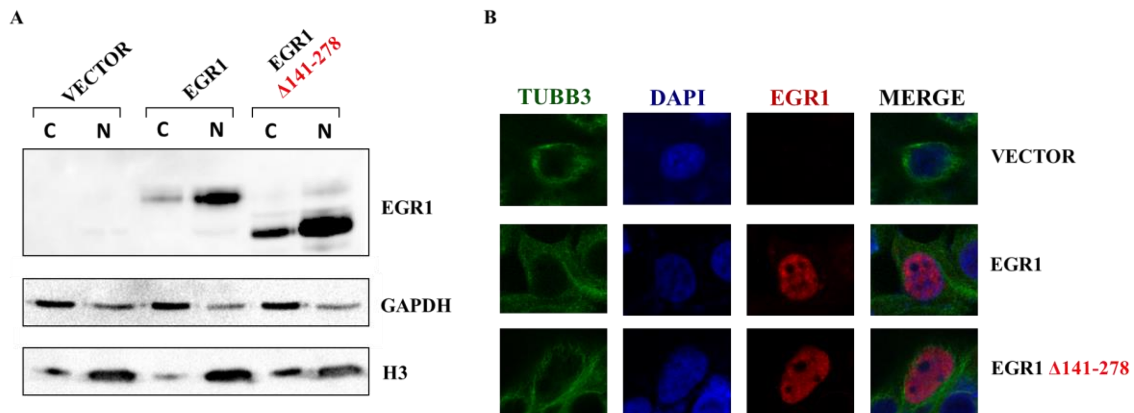

Fig 23. Cellular localization of EGR1 Δ141-278 splicing isoform. (A) Analysis of EGR1 localization by Western-blot. HEK293T cells transfected with both EGR1 constructs and empty plasmid were subjected to cellular fractionation. Distribution of canonical and Δ141-278 EGR1 isoforms was analysed in cytosolic (C) and nuclear (N) fractions. GADPH and histone H3 were respectively used as specific markers of the cytoplasmic and the nuclear fractions. **(B)** Localization of EGR1 in HEK293T cells transfected with empty plasmid, canonical and Δ141-278 EGR1 isoforms. Representative images of localization of EGR1 isoforms using anti-EGR1 antibody (red). The localization of nucleus was determined by immunofluorescence with 4 ,6-Diamidine-2-phenylindole dihydrochloride (DAPI). TUBB3 immunolocalization (green) was used as a marker of cytoplasmic distribution.

To evaluate this hypothesis, we analysed the expression level variation of different EGR1 target genes. We planned iper-expression experiments for both the canonical and alternative isoform and qPCR assays on putative target genes. After a preliminary survey to identify good candidate genes able to respond to EGR1 regulation, we selected two genes, *ARC* and *SIK1*. As shown in Figure 24 B, for both genes only the canonical EGR1 isoform was able to increase transcription, demonstrating that the alternative isoform lacks the ability to activate transcription. In addition, the co-transfection experiment demonstrated that the canonical EGR1-dependent increase of ARC transcript is lowed by the present of Δ141-278 (Figure 24 B).

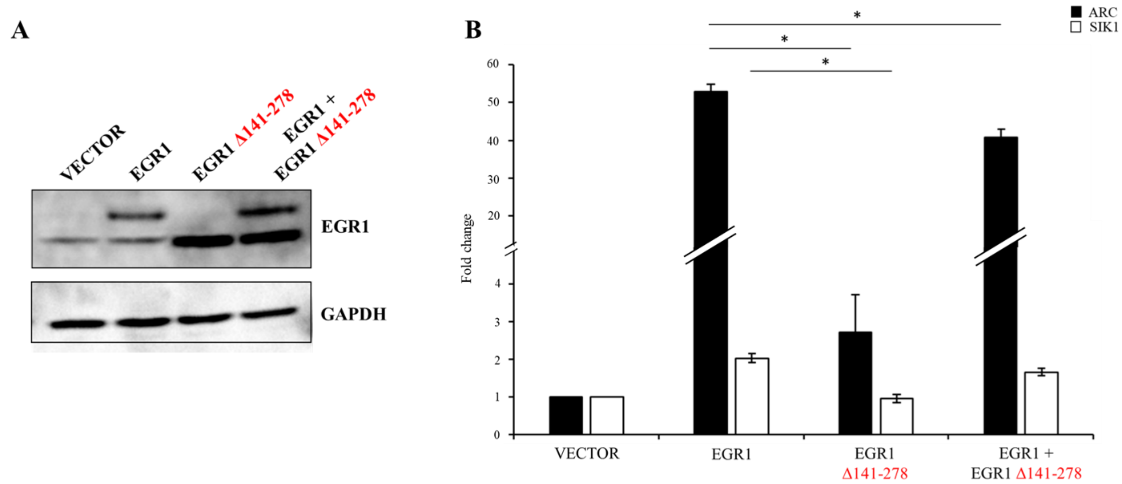

Fig 24. Expression pattern comparison in EGR1 and EGR1 Δ141-278 iper-expressing cells. (A) Analysis of the isoforms iper-expression in HEK293T cells transfected with both EGR1 constructs and empty plasmid by Western Blot. The level of GAPDH protein was used as a loading control. (**B**) Expression level change for *ARC* and *SIK1* transcripts in EGR1 and EGR1 Δ141-278 iper-expressing cells. Gene expression level was normalized to the reference transcript *GAPDH* and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample transfected with vector was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference in fold change (*$p < 0.05$) is shown.

PART III

Preliminary characterization of LINC00473 (C6orf176)

4.III.1 *LINC00473* overview

The *LINC00473* gene has 5 different alternative splicing isoforms (Figure 25).

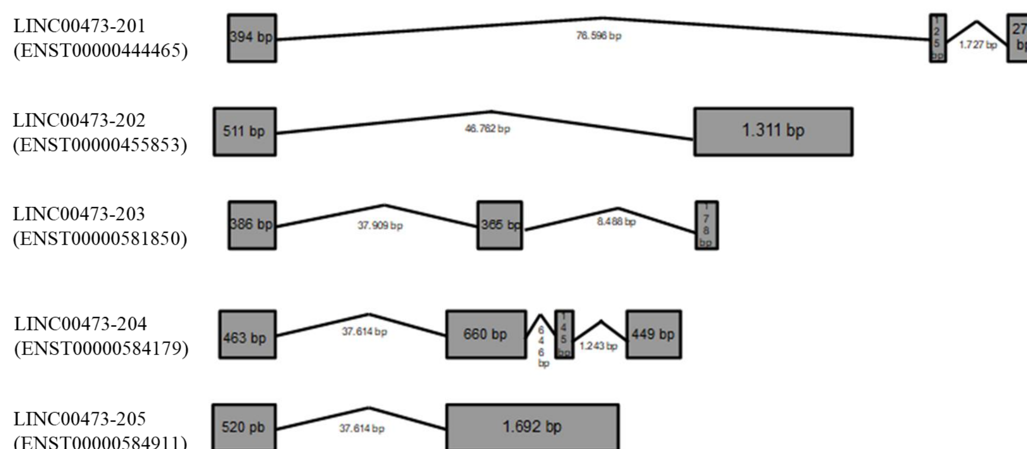

Fig 25. Schematic representation of the 5 different isoforms of the *LINC00473* gene.

The microarray platform used for the analysis discussed in Part I contained probes for all the *LINC00473* splicing isoforms. The microarray analysis results showed that only the *LINC00473-202* isoform increased the expression level after BDNF stimulation. Interestingly, the analysis of coding potential by the CPAT software argued the possibility that this isoform might encode for a relatively short protein (Table 12). The CPAT analysis was also extended to the other splicing isoforms of the *LINC00473* gene, demonstrating that only the *LINC00473-202* isoform had a significant coding potential probability (Table 12).

Sequence Name	RNA Size	ORF Size	Coding Probability	Coding Label
LINC00473-201	798	321	0.024	no
LINC00473-202	1822	561	0.478	yes
LINC00473-203	929	321	0.024	no
LINC00473-204	1717	321	0.023	no
LINC00473-205	2212	321	0.022	no

Table 12. Results of CPAT analysis for *LINC00473* isoforms.

This result led us to evaluate the evolutionary conservation of the putative encoded protein of the *LINC00472-202* isoform. We performed a TBLASTN search of the NCBI nucleotide database

using the putative amino acid sequence of the *LINC00473-202* isoform as bait. We found many sequences with high alignment scores; all of these sequences resulted as predicted by automated computational analysis from a genomic sequence and belonged to primate species (Figure 26). The amino acid alignment showed that the sequences can be subdivided in two groups: one that include a human sequence and is characterized by a putative protein length of 186 aa (185 for *Pan paniscus* and *Pan troglodytes*), and a second characterized by a putative longer protein containing an additional N-terminal region (Figure 26).

Fig 26. Amino acid alignment of the LINC00473 putative encoded protein. *Homo sapiens* (A8K010), *Gorilla gorilla* (XP_018885655.1), *Pan paniscus* (XP_008971877.1), *Pan troglodytes* (XP_016812094.1), *Rhinopithecus bieti* (XP_017725042.1), *Piliocolobus tephrosceles* (XP_023079412.1), *Colobus angolensis palliates* (XP_011790480.1), *Papio Anubis* (XP_009204634.2), *Cercocebus atys* (XP_011933425.1), *Chlorocebus sabaeus* (XP_008005965.1), *Macaca fascicularis* (XP_015303993.1), *Macaca nemestrina* (XP_011749299.1). "*" indicates positions which have a single, fully conserved residue; ":" indicates conservation between groups of strongly similar properties; "." indicates conservation between groups of weakly similar properties.

4.III.2 *LINC00473* expression levels under RA and BDNF stimulation

The identification of the *LINC00473-202* expression under BDNF stimulation shows that this gene may play a role in the differentiation process. In this regard, as performed for the *EGR1* gene in the Part II, we extended the analysis of the expression level during RA-induced differentiation. The qPCR analysis showed that although the RA is able to induce a transcript level variation of this gene, this change was not significant, likely reflecting an overall very low expression level (Cycle threshold Ct ranging from 27 to 30).

In contrast, under BDNF stimulation the overall transcript level was higher compared to RA stimulation. In particular, we evaluated the *LINC00473-202* expression after different time points of BDNF stimulation. As shown in Figure 27, the transcript level increased starting from 1h and reached a peak after 2h of BDNF treatment and then decreased at 4h, with kinetics resembling those of the immediate early genes. Interestingly, the expression pattern comparison of the *EGR1* and *LINC00473* genes after BDNF stimulation showed that the expression of *EGR1* anticipated

that of *LINC00473* (Figure 27). This data led us to hypothesize that *EGR1* could be involved in the transcription of the *LINC00473* gene.

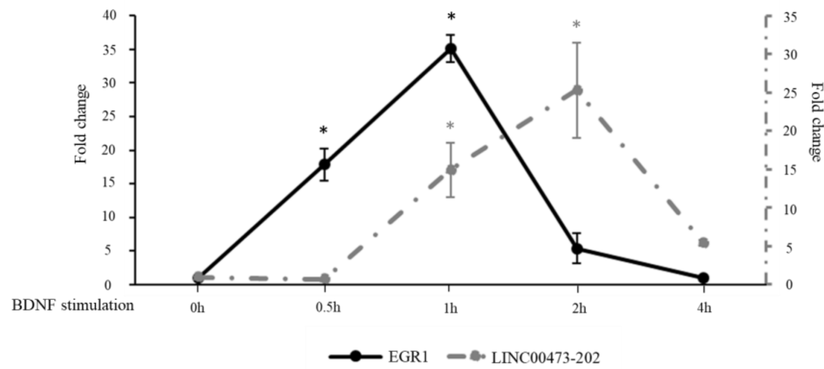

Fig 27. Expression pattern comparison of *EGR1* and *LINC00473* under BDNF stimulation. Expression level change in N-enriched SH-SY5Y cells after BDNF stimulation at the indicated time points for the canonical *EGR1* (black line) and *LINC00473* (dashed grey line). Time (0h) represented cells after 6 days of RA-induced differentiation. Gene expression level was normalized to the reference gene (*GAPDH*) and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0h was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0h (*$p < 0.05$) is shown.

4.III.3 *LINC00473* expression levels under FBS stimulation

In order to test whether *LINC00473-202*, as IEG, also varies in response to stimulation with FBS, we evaluated *LINC00473-202* expression after different time points of FBS stimulation. The transcript level increased starting from 1h and reached a peak after 2h of serum stimulation and then decreased at 4h, with kinetics resembling those of the immediate early genes (Figure 28). Furthermore, once again, the expression of *EGR1* anticipated that of *LINC00473* (Figure 28).

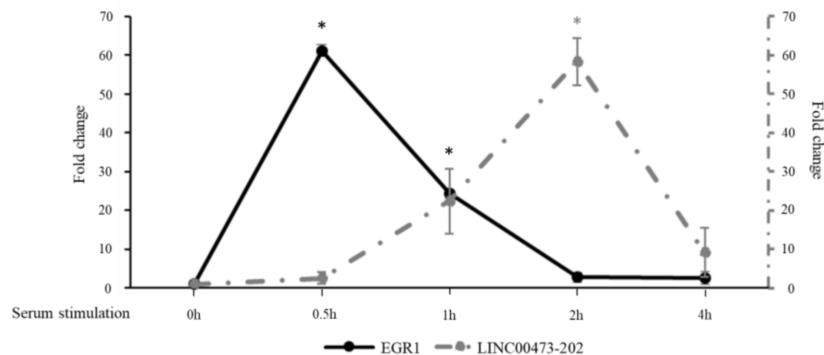

Fig 28. Expression pattern comparison of *EGR1* and *LINC00473* under FBS stimulation. Expression level change in HEK293T cells after serum stimulation at the indicated time points for the canonical *EGR1* (black line) and *LINC00473* (dashed grey line). Time (0h) represented cells before FBS treatment. Gene expression level was normalized to the reference gene (*GAPDH*) and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0h was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0h (*$p < 0.05$) is shown.

The putative regulation of *LINC00473* by EGR1 is corroborated by the analysis of the EGR1 binding site reported in the Ensembl database (https://www.ensembl.org/index.html) (Figure 29). In particular, Chip-seq results in H1ESC cells (human embryonic stem cells) demonstrated EGR1 binding at the level of the promoter region of the *LINC00473* gene (Figure 29).

Fig 29. Graphical representation of the binding site for EGR1 in the *LINC00473* promoter as reported in the genome browser Ensembl (https://www.ensembl.org/index.html).

4.III.4 LINC00473 KO

In order to obtain preliminary insights into the function of the *LINC00473* gene, we designed KO experiments by using CRISPR/Cas9 technology in HEK293T cells. The strategy was based on the use of an all-in-one vector AIO-PURO (Addgene plasmid #74630) with the Cas9 Nickase. This enzyme variant is able to generate a single-stranded break instead of a double-strand break, and when used with two adjacent gRNAs, it can lower the probability of off-target editing. In particular the gRNAs were design to remove a region inside the ORF of the putative encoding *LINC00473-202* isoform (Figure 30).

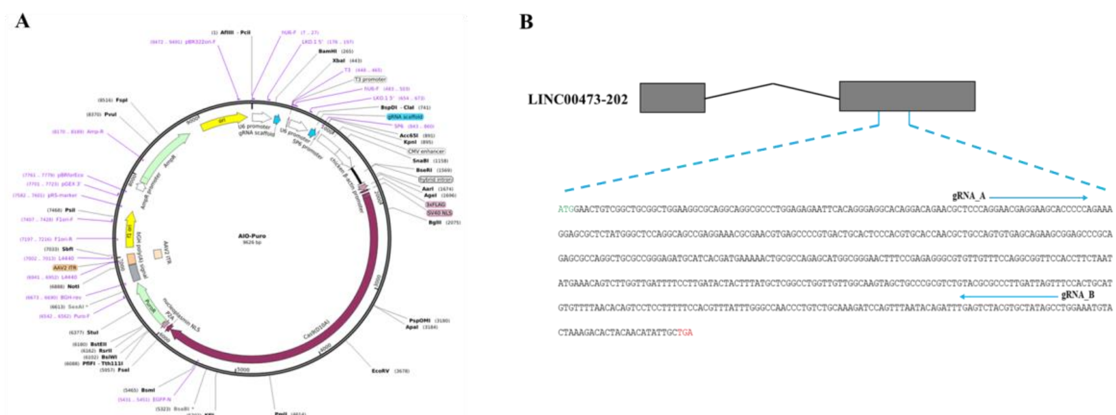

Fig 30. Basic elements of the CRISPR/Cas9 technology used to knockout the LINC00473 gene. (A) Graphical representation of the AIO-PURO vector used for the CRISPR-Cas9 system (www.addgene.com). **(B)** Schematic representation of the *LINC00473-202* transcript and the corresponding DNA sequence recognized by the gRNAs. The rectangles represented the two exons of *LINC00473-202*. The ATG of the *LINC00473* gene is highlighted in green; the STOP codon of the *LINC00473* gene is highlighted in red.

This approach allowed us to obtain deletion mutants as follows. Once gRNAs were cloned into the vector, we performed transfection experiments followed by antibiotic selection as reported in Materials and Methods section. By serial dilution of transfected cells, we obtained several single-cell derived populations to be genetically analysed for the mutation. We amplified the genomic region related to the putative mutation by using primers spanning the interested region. As shown in Figure 31 A, PCR products showed a generation of several mutant populations, two of which were homozygous for the deletion. One of these two population was also confirmed for the deletion mutation by sequencing and was used for the following analysis (Figure 31 B).

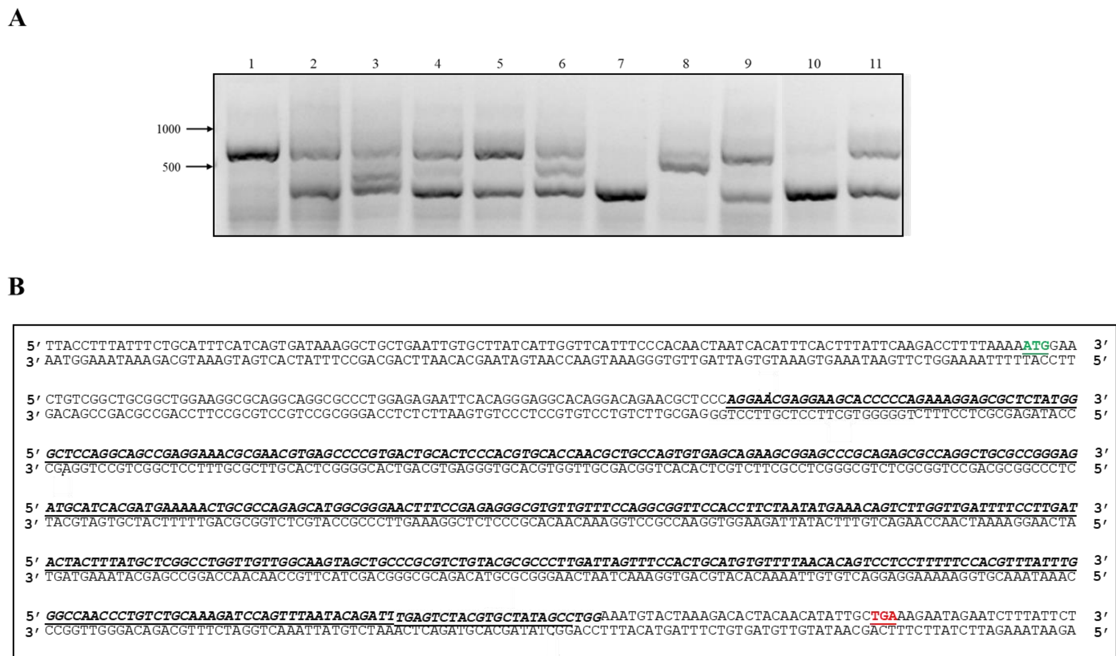

Fig 31. Analysis of the *LINC00473* genomic sequence in the KO cell line. (A) Electrophoretic run of the PCR products related to *LINC00473* amplification. Lane 1: WT. Lane 2: pool of KO populations. Lanes 3,4,5,6,8,9,11: heterozygous KO populations. Lanes 7,10: homozygous KO populations. **(B)** Genomic sequence of *LINC00473*. Underlined nucleotide region represented the sequence missing in the KO cell line. The ATG of the *LINC00473* gene is highlighted in green; the STOP codon of the *LINC00473* gene is highlighted in red.

As a preliminary investigation, we tested whether the knockout line for the *LINC00473* gene affected the regulation of other IEGs. As shown in Figure 32, the KO affected the level of the transcript variation of the *ARC* and *SIK1* gene but not of *EGR1*. These results suggested that the *LINC00473* gene could therefore be involved in gene expression regulation of some IEGs.

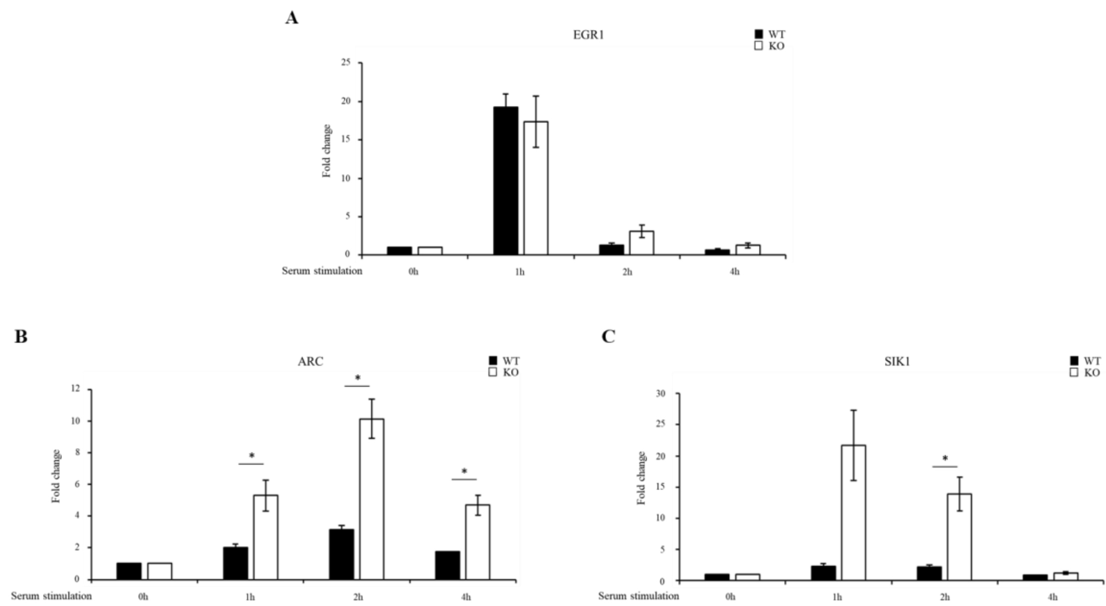

Fig 32. Expression analysis for 3 immediate early genes after FBS stimulation in LINC00473 WT and KO cells. Expression level change in LINC00473 WT and KO cells after serum stimulation at the indicated time points for the canonical *EGR1* (**A**), *ARC* (**B**), and *SIK1* (**C**). Time (0h) represented cells before FBS treatment. Gene expression level was normalized to the reference gene (*GAPDH*) and calculated by the $2^{-\Delta\Delta Ct}$ method. The sample at time 0h was used as a calibrator. The results from independent biological replicates are expressed as mean of fold change ± SEM. Statistical analysis of the qPCR data was carried out using a two-tailed t test. Significance of difference from time 0h (*$p < 0.05$) is shown.

5. DISCUSSION

Neuronal differentiation is characterized by diverse cellular events including cell cycle exit, the acquisition of specialized structures and functions, and survival of those neurons that have been able to establish correct and functional synapses during development. While many transcription factors involved in the regulation and coordination of these events have been characterized, many others remain to be explored. Very recently, long non-coding RNA emerged as another important element of gene expression regulation with an essential role in cellular differentiation. Long non-coding RNAs are a heterogeneous group of RNAs, many of which have emerged as regulators of genomic expression and stability. lncRNAs can influence gene expression through a wide variety of mechanisms, including transcriptional, post-transcriptional, and epigenetic regulation. The majority of lncRNAs are localized in the nucleus, where they can act as scaffolds for chromatin modifiers by interacting with chromatin-modifying complexes or as transcriptional co-regulators by binding to transcription factors (Rinn and Chang, 2012; Ulitsky and Bartel, 2013; Quan et al., 2017). Instead, cytoplasmic lncRNAs act as modulators on post-transcriptional regulation of genes through various mechanisms during RNA processing, such as mRNA editing, alternative splicing and others (Quan et al., 2017). Recently, lncRNAs have received widespread attention as a potentially new and crucial layer of gene expression regulation. lncRNAs of all kinds have been implicated in a range of developmental processes and diseases and remarkably, 40% of the annotated lncRNAs are specifically expressed in the brain (Derrien et al., 2012). Many transcriptome analyses have revealed that lncRNAs are differentially expressed over time in brain regions. They can therefore have different functions in neuronal development and activity (Wu et al., 2013). It has been shown that lncRNAs play indispensable roles in the development of CNS from neural early differentiation to late synaptogenesis (Briggs et al., 2015; Quan et al., 2017). An example is BDNF-AS, which represses the expression of BDNF, reducing neuronal outgrowth, differentiation, survival and proliferation both *in vitro* and *in vivo* (Modarresi et al., 2012). Knowledge of the mechanisms by which they act is still surprisingly in its infancy. The identification and characterization of coding and non-coding RNAs and their functional relationship is mandatory for unravelling the complex transcriptional programs underlying neuronal differentiation. In this regard, to preliminary address this item in a relatively easy model system of *in vitro* human neuronal differentiation, we employed SH-SY5Y, one of the most popular cell lines of human neuroblastoma in the neuroscience field (Pahlman et al., 1995; Kovalevich and Langford, 2013; Teppola et al., 2016). In fact, neuroblastoma cells can be induced to undergo neuronal differentiation by serum deprivation (Seeds et al., 1970), nerve growth factor (Jensen et al., 1988), or retinoic acid (RA) (Sidell et al., 1982). In the undifferentiated form, SH-SY5Y cells continuously proliferate, express immature neuronal markers, and are characterized morphologically by neuroblast-like, non-polarized cell bodies with few and truncated processes. Following treatment with differentiation-inducing agents, SH-SY5Y cells decrease in

proliferation rate, exit from the cell cycle, and become morphologically more similar to primary neurons with long processes (Kovalevich and Langford, 2013). RA-induced neuronal differentiation of SH-SY5Y cells is a well-established model for molecular investigation of neuronal differentiation. This treatment is associated with inhibition of proliferation and extension of neurites (Preis et al., 1988; Toselli et al., 1996; Encinas et al., 2000; Stio et al., 2001; Cheung et al., 2009; Kovalevick and Langford, 2013; Teppola et al., 2016). The sequential exposure of SH-SY5Y cells to RA and the neurotrophin BDNF increases the maturation process, highlighting an essential role for BDNF in neuronal differentiation (Encinas et al., 2000).

Brain-Derived Neurotrophic Factor (BDNF) has a key function in both neuronal development and the adult nervous system. The activity of BDNF affects many cellular processes in the nervous system including survival, growth, differentiation, synaptic plasticity of neurons, and long-term potentiation (LTP) (Casaccia-Bonnefil et al., 1999; Bibel and Barde, 2000; Xu et al., 2000; Huang and Reichardt, 2003; Bramham and Messaoudi, 2005; Gottmann et al., 2009; Minichiello, 2009; Park and Poo, 2013; Suliman et al., 2013; Mizui et al., 2014; Zagrebelsky and Korte, 2014; Leal et al., 2015). Knockout mice for BDNF usually die soon after birth and suffer developmental defects in the brain and sensory nervous system (Ernfors et al., 1995). Suppression of BDNF expression results in defective LTP and memory formation (Korte et al., 1995; Linnarsson et al., 1997; Ma et al., 1997; Mu et al., 1999). In contrast, treatment of hippocampal slices from BDNF knockout mice with recombinant BDNF completely reversed deficits in long-term potentiation and significantly improved deficits in basal synaptic transmission (Patterson et al., 1996). The fundamental role of BDNF for the central nervous system is also highlighted by its dysregulation in both neurological and psychiatric disorders disorders (Binder and Scharfman, 2004; Pruunsild et al., 2007; Nagahara and Tuszynski, 2011; Weissmiller and Wu, 2012). In addition to its major roles in supporting neuronal survival and function during development and in adulthood, BDNF exerts potent pro-survival and functional effects in models of neurological disease (Nagahara and Tuszynski, 2011). The effects of BDNF on synaptogenesis, synaptic plasticity, and cell survival have been characterized mainly at the post-transcriptional level (Leal et al., 2015), while little is known about the genes regulated by the neurotrophin. The necessity of BDNF for correct neuronal development is also evident for the differentiation paradigm of SH-SY5Y cells, where the addition of BDNF after 4-6 days of RA treatment ensures the production of a cell population with significantly more characteristics of mature neurons compared to RA treatment alone, with obvious morphological changes, including intricate and complex neurite structure (Encinas et al., 2000; Goldie et al., 2014). In the present thesis, we demonstrated that BDNF is able to involve several IEGs, belonging to both coding and non-coding genes, highlighting that the neurotrophin affects neuronal differentiation and survival by a plethora of mechanisms including gene transcription. In addition, our data are in agreement with a growing body of evidence in literature

on the role of lncRNAs in neuronal development (Lin et al., 2011; Roberts et al., 2014; D'haene et al., 2016; Quan et al., 2017).

Regarding the biological function of the BDNF-regulated lncRNAs, we provided evidence that they can affect the regulation of genes involved in gene transcription, highlighting the putative role of lncRNAs in orchestrating the immediate response to BDNF. The analysis of lncRNA-protein interaction in this thesis shows that these lncRNAs can influence different molecular processes including splicing. Genome-wide transcriptome analysis has revealed an exceptionally high level of alternative splicing in the mammalian brain (Yeo et al., 2004; Su et al., 2018). Alternative splicing is essential for cell differentiation, morphogenesis, the formation of complex neural networks and the regulation of synaptogenesis and plasticity (Norris and Calarco, 2012; Zheng and Black, 2013). Splicing regulation involves some specific neuron splicing factors and their interaction with ubiquitous factors (Raj and Blencowe, 2015; Vuong et al., 2016; Su et al., 2018). For example, among the proteins identified in our analysis, there were PTB proteins that are important for the stem cell-to neuron transition (Boutz et al., 2007; Vuong et al., 2016; Su et al., 2018). In particular, PTBP1 suppresses the splicing of a subset of neural targets to inhibit neuronal differentiation, while PTBP2 expression is elevated in differentiating neuronal cells and activates certain neural targets that promote differentiation (Boutz et al., 2007; Su et al., 2018). Other proteins implicated in the lncRNA-protein interaction coordinate the export, trafficking, and precise localization and translation of RNA in cells during embryonic development and in adults, suggesting their different physiological duties. Among them, we identified IGF2BP proteins, which belong to a large family of RBPs that exhibit different expression patterns during development (Cao et al., 2018). In particular, IGF2BP2 regulates the differentiation of neural precursor cells into mature neurons or glial cells. Its levels are very high in embryonic development and then decrease as development progresses (Cao et al., 2018). Another example is PUM2 (Pumilio RNA-Binding Family Member 2) that acts as a translational repressor during embryonic development and cell differentiation. In immature neurons, loss of Pum2 led to enhanced dendritic outgrowth and arborization. In mature neurons, Pum2 down-regulation resulted in a significant reduction in dendritic spines and an increase in elongated dendritic filopodia (Vessey et al., 2010).

The results obtained in this thesis work show that BDNF-regulated lncRNAs might also influence gene expression regulation through their interaction with miRNAs, a well-known class of ncRNAs highly expressed in the brain, in a spatially and temporally controlled manner (Fineberg et al., 2009). In particular, lncRNAs can affect gene expression by interfering with miRNAs pathways and acting as competing endogenous RNAs. Many miRNAs have been shown to be neuronal specific and involved in the control of neuronal differentiation, excitability and functions. Interestingly, aberrant expression and dysfunction of miRNAs have been associated

with the development of neurodegenerative diseases, such as Alzheimer's and Parkinson's. Overall, the analysis reported in Part I of this thesis gave an overview of the complex interplay between coding and non-coding RNA involved in the molecular mechanisms underlying gene expression regulation in the initial phase of BDNF-induced molecular cascade that affects neuronal differentiation.

Among the genes that we identified as induced under BDNF stimulation, EGR1 is one of the most enigmatic. It is highly expressed in the brain where it acts as a pleiotropic mediator whose encoding gene is a convergence point for many signaling cascades (Raf/MEK/ERK, Rho/ROCK/LIMK, FHL, cAMP/PKA and PKC). Upon induction, it contributes to several cell behaviours including proliferation, apoptosis, and differentiation in a cell type- and stimulus-specific manner (Thiel and Cibelli 2002; Pagel and Deindl, 2011). The role of EGR1 is particularly well-documented in the adult nervous system, where it plays important roles in learning and memory through regulation of genes that contribute to synaptic plasticity and long-term potentiation and is often used as a marker of neuronal stimulation. Moreover, it is considered to be involved in pathological states such as addiction, anxiety, and neuropsychiatric disorders (Knapska and Kaczmarek 2004; Penke et al., 2013; Veyrac 2014; Duclot 2017). The important roles of EGR1 in neuron function (supported by studies on KO mice) include impaired long-term memory in both spatial and non-spatial learning tasks (Wei et al., 2000; Jones et al., 2001), altered behavioural responses to persistent inflammatory pain, and impaired contextual extinction learning and normal fear acquisition relative to wild-type controls (Han, 2014). The involvement of EGR1 in neuronal development during embryogenesis has not yet been unravelled. In fact, EGR1-KO mice showed apparently normal embryonic brain development and relatively few studies on its expression pattern in embryos have been produced (Watson and Milbrandt, 1990). The hypothesis of an EGR1 function in neuron development is mainly supported by different studies employing neuronal cancer cell lines as progenitor-type models, based on their ability to differentiate after specific stimulations. In this regard, the discovery of EGR1 three decades ago was strictly related to its function in neuronal maturation. In that seminal paper, Milbrandt (1987) described the identification of a cDNA clone for a transcript able to quickly respond to NGF stimulation in the rat pheochromocytoma PC12 and established different important characteristics of EGR1: the nature of the encoded protein as a zinc finger transcription factor, typical features of immediate early genes such as the rapid and transient induction of the expression and its independence from de novo protein synthesis, and the involvement in neuronal differentiation as a neurotrophin-responsive gene. Neurotrophic factors play an essential role in the regulation of the growth, survival, and differentiation of neurons and astrocytes in the central and peripheral nervous system both during embryonic development and in adulthood (Oliveira et al., 2012; Vilar and Mira, 2016; Kashyap et., 2018). In this regard, the expression induction of EGR1 by different neurotrophic factors including NGF, GDNF (Milbrandt, 1987; Hu and Russek 2008;

Ahmadiantehrani and Ron 2013), and BDNF, as shown in the present thesis (Aliperti and Donizetti, 2016), point out that this gene might be a key element in the transcriptional programs underlying survival and differentiation of neuronal cells. The fundamental role of EGR1 in neuronal differentiation was corroborated by the results of the present thesis, where we took advantage of the generation of a KO cell line for EGR1 in the SH-SY5Y cell line. We showed that the EGR1-KO line was unable to differentiate under RA stimulation and in fact underwent cell death. The analysis of the gene expression levels of key elements of the RA pathway, including retinoic receptors (RARα, RARβ, and RARγ) and CYP26a1, an enzyme involved in RA metabolic inactivation, demonstrated that the KO for EGR1 affected the correct expression of these components during RA stimulation. In particular, it was interesting to note that CYP26a1 showed a very smaller increased expression after 2 days of RA in KO than WT cells. The expression levels of CYP26A1 are crucial in regulating intracellular levels of RA and in determining cell fate (Osanai and Petkovich, 2005). Therefore, it can be hypothesized that the low levels of CYP26A1 in KO cells cause an insufficient degradation of retinoic acid, with subsequent accumulation up to cytotoxic levels that induce cell death. In addition, we demonstrated the up-regulation of of RARγ in KO cells after 48 h of RA stimulation. The effect of this dysregulation is in line with data showing that RARγ appeared to control a differentiation-apoptosis switch in neuroblastoma cells (Ferrari et al.,1998). Considering that the BDNF system is induced by RA stimulation and is involved in differentiation and survival of neuronal cells, we also focused our attention on this system in RA-stimulated EGR1-KO cells. We observed that while the mRNA for BDNF increased in both WT and KO cells, the mRNA for NTRK2 increased less in KO cells than WT and, on the contrary, the mRNA level for NGFR was higher in KO than WT cells. NGFR signaling mediates different biologic effects compared to the NTRK receptors and is able to trigger programmed cell death (Casaccia-Bonnefil et al., 1996; Frade et al., 1996; Roux et al., 1999; Dechant and Barde, 2002). This led us to speculate that the unbalance of the two receptor levels may contribute to the RA-induced cell death in EGR1-KO cells.

Interestingly, in the present work, we identified an alternative splicing isoform for EGR1 never before described in literature. This isoform may be relevant for the EGR1 function in neuronal differentiation considering that its expression is regulated by RA and BDNF stimulation. It is worthy to note that the splicing event that produces this isoform belongs to a recently described splicing category that generates so-called "exitrons" (Marquez, 2015). Exitrons are defined as introns within protein-coding exons that, when retained, maintain the protein-coding potential of the transcript and contribute to transcriptome diversity in specific situations (Marquez, 2015; Park et al., 2018). The analysis of the putative protein encoded by the alternative splicing isoform showed that it maintains the nuclear localization signal and DNA binding domain. In fact, as shown by western blotting and immunolocalization analysis, the splicing isoform preferentially localized into the nucleus as the canonical EGR1 isoform. In addition, EGR1 lacked a region

belonging to N-terminal activation domain of EGR1 and corresponded to one of the two regions identified and described by Gashler and colleagues (1993) as important for its activity. This detail may explain the effect of the iper-expression of the EGR1 isoform that is unable to trigger the transcription of the two EGR1 target genes analysed in the present work (ARC and SIK1). This allows us to hypothesize that the alternative EGR1 isoform, when produced, may act as a negative competitor to the canonical EGR1 isoform as corroborated by the co-expression experiments. Although we obtained promising results, further investigations are needed to better characterize the expression pattern and the role of the alternative isoform in gene expression regulation.

Concerning the lncRNAs regulated by BDNF, we focused our attention on *LINC00473* and, in particular, on the splicing isoform *LINC00473-202*, which was the most differentially expressed under neurotrophin stimulation. In the vast majority of studies, *LINC00473* has been shown to act as an oncogene in tumours where it is upregulated and likely affects cell proliferation, colony formation, cellular invasion, and epithelial-mesenchymal transition (EMT) (Chen et al., 2016; Shi et al.,2017; Chen et al., 2018; Zhu et al., 2018). In the present thesis, we showed different expression kinetics; in particular, the expression pattern under BDNF stimulation reported in the present thesis aligns with one of the first published papers that described *LINC00473* as a primate-specific IEG able to respond to the treatment with EP2 and EP4 agonists (Reitmair et al., 2011). In addition, we provided preliminary observations that the expression of the *LINC00473* gene can be regulated by the transcription factor EGR1, providing an interesting functional relationship between these two BDNF-regulated genes. Reitmair and colleagues (2011) proposed *LINC00473* as a possible regulator of c-AMP-mediated gene expression since small interfering RNA-mediated *LINC00473* gene silencing experiments affected the expression of several cAMP-responsive genes. Interestingly, a recent study demonstrated the expression of *LINC00473* and other primate-specific lncRNAs in human iPSC-derived neuronal cells after activity (Pruunsild et al., 2017); in this paper, the authors argued that lineage-specific gain of lncRNAs might affect the synaptic activity-regulated gene program. lncRNAs can be translated and play a role in the evolution of new proteins. Our results demonstrated that the *LINC00473-202* isoform has significant coding potential and that the putative protein is highly conserved in primate species. The actual production of a protein by the translation of this RNA remains to be demonstrated in the cell line and experimental paradigms used in this thesis. Recently, C6orf176 (LINC00473) was identify by mass spectrometry as one of the proteins differentially expressed after the treatment of HT-29 cells with the antitumoral peptide CIGB-552, providing evidence for the existence of this protein (Rodríguez-Ulloa et al., 2015). Our preliminary results on the effects of the deletion of the CDS region of *LINC00473-202* on the transcript level changes of some IEGs after serum stimulation corroborated the hypothesis of a protein production and suggested that it can be involved in the regulation of gene expression regulation.

In conclusion, this thesis work provided the identification of coding and lncRNAs involved in human neuronal development and gave insights on their functional relationship and the molecular mechanism of action.

6. REFERENCES

Adams K.W., et al. (2017). Role for Egr1 in the Transcriptional Program Associated with Neuronal Differentiation of PC12 Cells. PLoS One. 12(1):e0170076.

Ahmadiantehrani S. and Ron D. (2013). Dopamine D2 receptor activation leads to an up-regulation of glial cell line-derived neurotrophic factor via Gβγ-Erk1/2-dependent induction of Zif268. J Neurochem. 125(2):193-204.

Aliperti V. and Donizetti A. (2016). Long Non-coding RNA in Neurons: New Players in Early Response to BDNF Stimulation. Front Mol Neurosci.9:15.

Aprea J., et al. (2013). Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment. EMBO J. 32, 3145–3160.

Aydemir O., et al. (2005). The effect of chronic antidepressant treatment on serum brain-derived neurotrophic factor levels in depressed patients: a preliminary study. Prog Neuropsychopharmacol Biol Psychiatry. 29(2):261-5.

Bahrami S. and Drabløs F. (2016). Gene regulation in the immediate-early response process. Adv Biol Regul. 62:37-49.

Bale T.L., et al. (2010). Early life programming and neurodevelopmental disorders. Biol. Psychiatry, 68, 314–319.

Barry G., et al. (2014). The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing. Mol Psychiatry. 19, 486-494.

Bassett A.R. (2017). Editing the genome of hiPSC with CRISPR/Cas9: disease models. Mamm Genome. 28(7-8):348-364.

Bernard et al. (2010). A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. EMBO J. 29(18):3082-93.

Bhartiya D., et al. (2013). lncRNome: a comprehensive knowledgebase of human long noncoding RNAs. Database (Oxford). 2013:bat034.

Bibel M. and Barde Y. A. (2000). Neurotrophins: key regulators of cell fate and cell shape in the vertebrate nervous system. Genes Dev. 14, 2919-2937.

Binder D.H. and Scharfman H.E. (2004). Brain-derived neurotrophic factor. Growth Factors. 22(3):123-31.

Boutz P. L., et al. (2007). A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. Genes Dev. 21, 1636–1652.

Bramham C.R. and Messaoudi E. (2005). BDNF function in adult synaptic plasticity: the synaptic consolidation hypothesis. Progress in Neurobiology. 76, 99–125.

Briggs J. A., et al. (2015). Mechanisms of long non-coding RNAs in mammalian nervous system development, plasticity, disease, and evolution. Neuron 88, 861–877.

Budday S., et al. (2015). Physical biology of human brain development. Front Cell Neurosci. 8;9:257.

Cao J., et al. (2018). The Roles of Insulin-Like Growth Factor 2 mRNA-Binding Protein 2 in Cancer and Cancer Stem Cells. Stem Cells Int. 2018:4217259.

Casaccia-Bonnefil P., et al. (1999). Neurotrophins in cell survival/death decisions. Adv. Exp. Med. Biol. 468, 275-282.

Chen Z., et al. (2016). cAMP/CREB-regulated LINC00473 marks LKB1-inactivated lung cancer and mediates tumor growth. J Clin Invest. 126(6):2267-79.

Chen Z., et al. (2018). CRTC1-MAML2 fusion-induced lncRNA LINC00473 expression maintains the growth and survival of human mucoepidermoid carcinoma cells. Oncogene. 37(14):1885-1895.

Cheung Y.T., et al. (2009). Effects of all-trans-retinoic acid on human SH-SY5Y neuroblastoma as in vitro model in neurotoxicity research. Neurotoxicology. 30(1):127-35.

Chou C.H., et al. (2018). miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res. 46(D1):D296-D302.

Cheung I., et al. (2010). Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. Proc Natl Acad Sci U S A. 107(19):8824-9.

Clark B.S. and Blackshaw S. (2014). Long non-coding RNA-dependent transcriptional regulation in neuronal development and disease. Front Genet. 5:164.

Colantuoni C., et al. (2011). Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 478, 519–523.

Cooper J. A. (2008). A mechanism for inside-out lamination in the neocortex. Trends in Neurosciences, 31(3), 113–119.

Craige A.M., et al. (2006). How to build a central synapse: clues from Cell culture. Trends Neurosci. 29,8–20.

Cunha A.B., et al. (2006). Serum brain-derived neurotrophic factor is decreased in bipolar disorder during depressive and manic episodes. Neurosci Lett. 398(3):215-9.

D'haene E., et al. (2016). Identification of long non-coding RNAs involved in neuronal development and intellectual disability. Sci Rep. 6:28396.

Das S., et al. (2014). lnCeDB: database of human long noncoding RNA acting as competing endogenous RNA. PLoS One. 9(6):e98965.

Davis S., et al (2003). How necessary is the activation of the immediate early gene zif268 in synaptic plasticity and learning? Behav Brain Res. 142(1-2):17-30.

Dechant G. and Barde Y.A. (2002). The neurotrophin receptor p75(NTR): novel functions and implications for diseases of the nervous system. Nat Neurosci. 5(11):1131-6.

Dennis G., et al. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 4(5):P3.

Derrien T., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution and expression. Genome Res. 22, 1775–1789.

Duclot F. and Kabbaj M. (2017). The Role of Early Growth Response 1 (EGR1) in Brain Plasticity and Neuropsychiatric Disorders. Front Behav Neurosci. 11:35.

Duman R.S. and Voleti B. (2012). Signaling pathways underlying the pathophysiology and treatment of depression: novel mechanisms for rapid-acting agents. TrendsNeurosci. 35, 47–56.

Edwards S.A., et al. (1991). The transcription factor, Egr-1, is rapidly modulated in response to retinoic acid in P19 embryonal carcinoma cells. Dev Biol. 148(1):165-73.

Encinas M., et al. (2000). Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells. J Neurochem. 75, 991-1003.

Ernfors P., et al. (1995). Studies on the physiological role of brain-derived neurotrophic factor and neurotrophin-3 in knockout mice. Int J Dev Biol. 39(5):799-807.

Faria M.C., et al. (2014). Increased plasma levels of BDNF and inflammatory markers in Alzheimer's disease. J Psychiatr Res. 53:166-72.

Farris S., et al. (2014). Selective localization of Arc mRNA in dendrites involves activity- and translation-dependent mRNA degradation. J. Neurosci. 34, 4481–4493.

Ferrari N., et al. (1998). Retinoic Acid Receptor γ1 (RARγ1) Levels Control RARβ2 Expression in SK-N-BE2(c) Neuroblastoma Cells and Regulate a Differentiation-Apoptosis Switch. Mol Cell Biol. 18(11): 6482–6492.

Fineberg S.K., et al. (2009). MicroRNAs potentiate neural development. Neuron. 64(3):303-9.

Frade J.M. and Barde Y.A. (1999). Genetic evidence for cell death mediated by nerve-growth factor and the neurotrophin receptor p75 in the developing mouse retina and spinal cord. Development 126:683–90.

Friedman W. (2000). Neurotrophins induce death of hippocampal neurons via the p75 receptor. J Neurosci 20:6340–46.

Gashler A.L., et al. (1993). A novel repression module, an extensive activation domain, and a bipartite nuclear localization signal defined in the immediate-early transcription factor Egr-1. Mol Cell Biol. 13(8):4556-71.

Goldie B.J., et al. (2014). BDNF and the maturation of posttranscriptional regulatory networks in human SH-SY5Y neuroblast differentiation. Front Cell Neurosci. 8:325.

Gonul A.S., et al. (2005). Effect of treatment on serum brain-derived neurotrophic factor levels in depressed patients. Eur Arch Psychiatry Clin Neurosci. 255(6):381-6.

Gottmann, K., et al. (2009). BDNF signaling in the formation, maturation and plasticity of glutamatergic and GABAergic synapses. Experimental Brain Research. 199, 203–234.

Green D.R., et al. (2014). Cell biology. Metabolic control of cell death. Science 345, 1250256.

Hallbook F. (1999). Evolution of the vertebrate neurotrophin and Trk receptor gene families. Curr Opin Neurobiol 9:616–21.

Han S., et al. (2014). Impaired extinction of learned contextual fear memory in early growth response 1 knockout mice. Mol Cells. 37(1):24-30.

Holsinger R.M., et al. (2000). Quantitation of BDNF mRNA in human parietal cortex by competitive reverse transcription-polymerase chain reaction: decreased levels in Alzheimer's disease. Brain Res Mol Brain Res. 76(2):347-54.

Howells D.W., et al. (2000). Reduced BDNF mRNA expression in the Parkinson's disease substantia nigra. Exp Neurol. 166(1):127-35.

Hu Y. and Russek S.J. (2008). BDNF and the diseased nervous system: a delicate balance between adaptive and pathological processes of gene regulation. J Neurochem. 105(1):1-17.

Huang E.J. and Reichardt L.F. (2001). Neurotrophins: roles in neuronal development and function. Annu Rev Neurosci. 24:677-736.

Huang E. J., and Reichardt L. F. (2003). Trk receptors: roles in neuronal signal transduction. Annu. Rev. Biochem. 72, 609-642.

Huang Z. (2009). Molecular regulation of neuronal migration during neocortical development. Molecular and Cellular Neurosciences, 42(1), 11–22.

Ip B.K., et al. (2010). Investigating gradients of gene expression involved in early human cortical development. J. Anat. 217, 300–311.

Jaffe A.E., et al. (2015). Developmental regulation of human cortex transcription and its clinical relevance at single base resolution. Nat. Neurosci. 18, 154–161.

Jensen L. and Linn S. (1988). A reduced rate of bulky DNA adduct removal is coincident with differentiation of human neuroblastoma cells induced by nerve growth factor. Mol Cell Biol. 8(9):3964-8.

Ji Y., et al. (2005). Cyclic AMP controls BDNF-induced TrkB phosphorylation and dendritic spine formation in mature hippocampal neurons. Nat Neurosci 8: 164–172.

Jiang Q., et al. (2009). miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res. 37(Database issue):D98-104.

Jiang M., et al. (2013). Small-molecule TrkB receptor agonists improve motor function and extend survival in a mouse model of Huntington's disease. Hum Mol Genet. 22(12):2462-70.

Johnson M.B., et al. (2009). Functional and evolutionary insights into human brain development through global transcriptome analysis. Neuron 62, 494–509.

Jones M.W., et al. (2001). A requirement for the immediate early gene Zif268 in the expression of late LTP and long-term memories. Nat Neurosci. 4(3):289-96.

Kang H.J., et al. (2011). Spatio-temporal transcriptome of the human brain. Nature 478, 483–489.

Kaplan D.R., et al. (1993). Induction of TrkB by retinoic acid mediates biologic responsiveness to BDNF and differentiation of human neuroblastoma cells. Eukaryotic Signal Transduction Group. Neuron. 11(2):321-31.

Kaplan D.R. and Miller F.D. (2000). Neurotrophin signal transduction in the nervous system. Curr Opin Neurobiol 10:381–91.

Karege F., et al. (2002). Decreased serum brain-derived neurotrophic factor levels in major depressed patients. Psychiatry Res. 109(2):143-8.

Kashyap M.P., et al. (2018). Drug Targets in Neurotrophin Signaling in the Central and Peripheral Nervous System. Mol Neurobiol. 55(8):6939-6955.

Kempermann G., et al. (2004). Milestones of neuronal development in the adult hippocampus. Trends Neurosci. 27(8):447-52.

Kerr J.F., et al. (1972). Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Br. J. Cancer 26, 239–257.

Kim S., et al. (2018). Synapse development organized by neuronal activity-regulated immediate-early genes. Exp Mol Med. 50(4):11.

Knapska E. and Kaczmarek L. (2004). A gene for neuronal plasticity in the mammalian brain: Zif268/Egr-1/NGFI-A/Krox-24/TIS8/ZENK? Prog Neurobiol. 74(4):183-211.

Korb E. and Finkbeiner S. (2011). Arc in synaptic plasticity: from gene to behavior. Trends Neurosci. 34, 591–598.

Korte M., et al. (1995). Hippocampal long-term potentiation is impaired in mice lacking brain-derived neurotrophic factor. Proc Natl Acad Sci U S A. 92(19):8856-60.

Kovalevich J. and Langford D. (2013). Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. Methods Mol Biol. 1078:9-21.

Lambert N., et al. (2011). Genes expressed in specific areas of the human fetal cerebral cortex display distinct patterns of evolution. PLoS ONE 6, e17753.

Larsen F.G., et al. (1994). Retinoic acid induces expression of early growth response gene-1 (Egr-1) in human skin in vivo and in cultured skin fibroblasts. J Invest Dermatol. 102(5):730-3.

Leal G., et al. (2015). Regulation of hippocampal synaptic plasticity by BDNF. Brain Res. 1621, 82-101.

Lee J.H. and Kim K.T. (2004). Induction of cyclin-dependent kinase 5 and its activator p35 through the extracellular-signal-regulated kinase and protein kinase A pathways during retinoic-acid mediated neuronal differentiation in human neuroblastoma SK-N-BE(2)C cells. J Neurochem. 91(3):634-47.

Levi-Montalcini R. (1964). The nerve growth factor. Annals of the New York Academy of Sciences, 118, 149–170.

Lin M., et al. (2011). RNA-Seq of Human Neurons Derived from iPS Cells Reveals Candidate Long Non-Coding RNAs Involved in Neurogenesis and Neuropsychiatric Disorders. PLoS One. 6(9):e23356.

Linnarsson S., et al. (1997). Learning deficit in BDNF mutant mice. Eur J Neurosci. 9(12):2581-7.

Lister R., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. Science 341, 1237905.

Lu B., et al. (2005). The yin and yang of neurotrophin action. Nat Rev Neurosci. 6:603–614.

Ma Y.L., et al. (1997). Brain-derived neurotrophic factor antisense oligonucleotide impairs memory retention and inhibits long-term potentiation in rats. Neuroscience. 82(4):957-67.

Marin O. and Rubenstein J. L. (2001). A long, remarkable journey: tangential migration in the telencephalon. Nature Reviews. Neuroscience, 2(11), 780–790.

Marquez Y., et al. (2015). Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity. Genome Res. 25(7):995-1007.

Matsumoto T., et al. (2008). Biosynthesis and processing of endogenous BDNF: CNS neurons store and secrete BDNF, not pro-BDNF. Nat Neurosci. 11:131–133.

Mattick J.S. and Rinn J.L. (2015). Discovery and annotation of long noncoding RNAs. Nat Struct Mol Biol. 22(1):5-7.

Maze I., et al. (2014). Analytical tools and current challenges in the modern era of neuroepigenomics. Nat. Neurosci. 17, 1476–1490.

Mazin P., et al. (2013). Widespread splicing changes in human brain development and aging. Mol. Syst. Biol. 9, 633.

McLean C.Y., et al. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 28(5):495-501.

Melo C.V. et al. (2013). BDNF regulates the expression and distribution of vesicular glutamate transporters in cultured hippocampal neurons. PLoS One. 8(1):e53793.

Michalski B. and Fahnestock M. (2003). Pro-brain-derived neurotrophic factor is decreased in parietal cortex in Alzheimer's disease. Brain Res Mol Brain Res. 111(1-2):148-54.

Milbrandt J. (1987). A nerve growth factor-induced gene encodes a possible transcriptional regulatory factor. Science. 238(4828):797-9.

Millan M.J. (2013). An epigenetic framework for neurodevelopmental disorders: From pathogenesis to potential therapy. Neuropharmacology, 68, 2–82.

Miller J.A., et al. (2014). Transcriptional landscape of the prenatal human brain. Nature 508, 199–206.

Minatohara K., et al. (2016). Role of Immediate-Early Genes in Synaptic Plasticity and Neuronal Ensembles Underlying the Memory Trace. Front Mol Neurosci. 8:78.

Ming G.L. and Song H. (2011). Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron. 70(4):687-702.

Minichiello L. (2009). TrkB signalling pathways in LTP and learning. Nat Rev Neurosci. 10,850-860.

Mizui T., et al. (2014). The Biological Actions and Mechanisms of Brain-Derived Neurotrophic Factor in Healthy and Disordered Brains. Neuroscience & Medicine. 5, 183-195.

Modarresi F., et al. (2012). Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. Nat. Biotechnol. 30, 453–459.

Morris K.V. and Mattick J.S. (2014). The rise of regulatory RNA. Nat. Rev.Genet. 15, 423–437.

Mowla S.J., et al. (1999). Differential sorting of nerve growth factor and brain-derived neurotrophic factor in hippocampal neurons. J Neurosci. 19:2069–2080.

Mowla S.J., et al. (2001). Biosynthesis and post-translational processing of the precursor to brain-derived neurotrophic factor. J Biol Chem. 276:12660–12666.

Mu J.S., et al. (1999). Deprivation of endogenous brain-derived neurotrophic factor results in impairment of spatial learning and memory in adult rats. Brain Res. 835(2):259-65.

Nadarajah B. and Parnavelas J. G. (2002). Modes of neuronal migration in the developing cerebral cortex. Nature Reviews. Neuroscience, 3(6), 423–432.

Nagahara A.H. and Tuszynski M.H. (2011). Potential therapeutic uses of BDNF in neurological and psychiatric disorders. Nat Rev Drug Discov. 10(3):209-19.

Noctor S. C., et al. (2001). Neurons derived from radial glial cells establish radial units in neocortex. Nature, 409(6821), 714–720.

Noctor S. C., et al. (2002). Dividing precursor cells of the embryonic cortical ventricular zone have morphological and molecular characteristics of radial glia. The Journal of Neuroscience, 22(8), 3161–3173.

Nord A.S., et al. (2015). Genomic perspectives of transcriptional regulation in forebrain development. Neuron 85, 27–47.

Norris A. D. and Calarco J. A. (2012). Emerging roles of alternative pre-mRNA splicing regulation in neuronal development and function. Front. Neurosci. 6:122.

Numata S., et al. (2012). DNA methylation signatures in development and aging of the human prefrontal cortex. Am. J. Hum. Genet. 90, 260–272.

O'Carroll D. and Schaefer A. (2013). General principals of miRNA biogenesis and regulation in the brain. Neuropsychopharmacology 38, 39–54.

Oliveira S.L., et al. (2013). Functions of neurotrophins and growth factors in neurogenesis and brain repair. Cytometry A. 83(1):76-89.

Oppenheim R. W. (1989). The neurotrophic theory and naturally occurring motoneuron death. Trends in Neurosciences, 12(7).

Osanai M. and Petkovich M. (2005). Expression of the retinoic acid-metabolizing enzyme CYP26A1 limits programmed cell death. Mol Pharmacol. 67(5):1808-17.

Pagel J.I. and Deindl E. (2011). Early growth response 1--a transcription factor in the crossfire of signal transduction cascades. Indian J Biochem Biophys. 48(4):226-35.

Pahlman S., et al. (1995). Differentiation and survival influences of growth factors in human neuroblastoma. Eur J Cancer. 31A(4):453-8.

Pakkenberg B. and Gundersen H. J. (1997). Neocortical neuron number in humans: effect of sex and age. The Journal of Comparative Neurology, 384(2), 312–320.

Parikshak N.N., et al. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell 155, 1008–1021.

Park H., and Poo M.M. (2013). Neurotrophin regulation of neural circuit development and function. Nat Rev Neurosci. 14, 7-23.

Park E., et al. (2018). The Expanding Landscape of Alternative Splicing Variation in Human Populations. Am J Hum Genet. 102(1):11-26.

Parnavelas J. G., et al. (2002). The origin and migration of cortical neurons. Progress in Brain Research, 136, 73–80.

Patterson S.L., et al. (1996). Recombinant BDNF rescues deficits in basal synaptic transmission and hippocampal LTP in BDNF knockout mice. Neuron. 16(6):1137-45.

Pawson T. and Nash P. (2000). Protein-protein interactions define specificity in signal transduction. Genes Dev 14:1027–47.

Peng S., et al. (2005). Precursor form of brain-derived neurotrophic factor and mature brain-derived neurotrophic factor are decreased in the pre-clinical stages of Alzheimer's disease. J Neurochem. 93(6):1412-21.

Penke Z., et al. (2013). Zif268/Egr1 gain of function facilitates hippocampal synaptic plasticity and long-term spatial recognition memory. Philos Trans R Soc Lond B Biol Sci. 369(1633):20130159.

Pérez-Cadahìa B., et al. (2011). Role of MSK1 in the malignant phenotype of Ras-transformed mouse fibroblasts. J Biol Chem. 286(1):42-9.

Pletikos M., et al. (2014). Temporal specification and bilaterality of human neocortical topographic gene expression. Neuron 81, 321–332.

Preis P.N., et al. (1988). Neuronal cell differentiation of human neuroblastoma cells by retinoic acid plus herbimycin A. Cancer Res. 48(22):6530-4.

Pruunsild P., et al. (2007). Dissecting the human BDNF locus: bidirectional transcription, complex splicing, and multiple promoters. Genomics. 90(3):397-406.

Pruunsild P., et al. (2017). Networks of Cultured iPSC-Derived Neurons Reveal the Human Synaptic Activity-Regulated Adaptive Gene Program. Cell Rep. 18(1):122-135.

Quan Z., et al. (2017). Regulatory Roles of Long Non-Coding RNAs in the Central Nervous System and Associated Neurodegenerative Diseases. Front Cell Neurosci. 11:175.

Raj B. and Blencowe B. J. (2015). Alternative splicing in the mammalian nervous system: recent insights into mechanisms and functional roles. Neuron 87, 14–27.

Rakic P. (1972). Mode of cell migration to the superficial layers of fetal monkey neocortex. The Journal of Comparative Neurology, 145(1), 61–83.

Ran F.A., et al. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 8(11):2281-2308.

Reichardt L.F. and Fariñas I. (1997). Neurotrophic factors and their receptors. Roles in neuronal development and function. Molecular and Cellular Approaches to Neural Development. New York: Oxford Univ. Press p. 220-63.

Reitmair A., et al. (2012). C6orf176: a novel possible regulator of cAMP-mediated gene expression. Physiol Genomics. 44(2):152-61.

Rinn J. L. and Chang H. Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. 81, 145–166.

Roberts T.C., et al. (2014). The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease. Philos Trans R Soc Lond B Biol Sci. 369(1652).

Rodríguez-Ulloa A., et al. (2015). Proteomic Study to Survey the CIGB-552 Antitumor Effect. Biomed Res Int. 2015:124082.

Ronan J.L., et al. (2013). From neural development to cognition: Unexpected roles for chromatin. Nat. Rev. Genet. 14, 347–359.

Roux P.P., et al. (1999). p75 neurotrophin receptor expression is induced in apoptotic neurons after seizure. J Neurosci. 19(16):6887-96.

Sala C. et al. (2003). Inhibition of dendritic spine morphogenesis and synaptic transmission by activity-inducible protein Homer1a. J. Neurosci. 23, 6327–6337.

Seeds N.W., et al. (1970). Regulation of axon formation by clonal lines of a neural tumor. Proc Natl Acad Sci U S A. 66(1):160-7.

Shi C., et al. (2017). The long noncoding RNA LINC00473, a target of microRNA 34a, promotes tumorigenesis by inhibiting ILF2 degradation in cervical cancer. Am J Cancer Res. 7(11):2157-2168.

Shibata M., et al. (2015). From trans to cis: transcriptional regulatory networks in neocortical development. Trends Genet. 31,77–87.

Sidell N. (1982). Retinoic acid-induced growth inhibition and morphologic differentiation of human neuroblastoma cells in vitro. J Natl Cancer Inst. 68(4):589-96.

Silbereis J.C., et al. (2016). The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. Neuron. 89(2):248-68.

Somel M., et al. (2009). Transcriptional neoteny in the human brain. Proc. Natl. Acad. Sci. USA 106, 5743–5748.

Somel M., et al. (2011). MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. PLoS Biol. 9, e1001214.

Spadaro P.A., et al. (2015). Long non coding RNA-directed epigenetic regulation of gene expression is associated with anxiety-like behaviour in mice. Biol.Psychiatry 78, 848–859.

St Laurent G. and Wahlestedt C. (2007). Noncoding RNAs: couplers of analog and digital information in nervous system function? Trends Neurosci. 30(12):612-21.

Stiles J. and Jernigan T.L. (2010). The basics of brain development. Neuropsychol Rev. 20(4):327-48.

Stio M., et al. (2001). Synergistic anti-proliferative effects of vitamin D derivatives and 9-cis retinoic acid in SH-SY5Y human neuroblastoma cells. J Steroid Biochem Mol Biol. 77(4-5):213-22.

Su C.H., et al. (2018). Alternative Splicing in Neurogenesis and Brain Development. Front Mol Biosci. 5:12.

Suliman S., et al. (2013). Brain-Derived Neurotrophic Factor (BDNF) protein levels in anxiety disorders: systematic review and meta-regression analysis. Front Integr Neurosci. 7:55.

Sun C., et al. (2018) Long non-coding RNA MIAT in development and disease: a new player in an old game. J Biomed Sci. 25(1):23.

Takahashi K. and Yamanaka S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126:663–676.

Takahashi K., et al. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131:861–872.

Teppola H., et al. (2016). Morphological Differentiation Towards Neuronal Phenotype of SH-SY5Y Neuroblastoma Cells by Estradiol, Retinoic Acid and Cholesterol. Neurochem Res. 41(4):731-47.

Tessarz P. and Kouzarides T. (2014). Histone core modifications regulating nucleosome structure and dynamics. Nat. Rev. Mol. Cell Biol. 15, 703–708.

Thiel G. and Cibelli G. (2002). Regulation of life and death by the zinc finger transcription factor Egr-1. J Cell Physiol. 193(3):287-92.

Tushir J.S. and Akbarian S. (2014). Chromatin-bound RNA and the neurobiology of psychiatric disease. Neuroscience. 264:131-41.

Toselli M., et al. (1996). Functional changes in sodium conductances in the human neuroblastoma cell line SH-SY5Y during in vitro differentiation. J Neurophysiol. 76(6):3920-7.

Ulitsky I. and Bartel D. P. (2013). lincRNAs: genomics, evolution, and mechanisms. Cell 154, 26–46.

Valiente M. and Marin O. (2010). Neuronal migration mechanisms in development and disease. Current Opinion in Neurobiology, 20(1), 68–78.

van de Vondervoort I.I., et al. (2013). Long non-coding RNAs in neurodevelopmental disorders. Front Mol Neurosci. 6:53.

van der Kolk N.M., et al. (2015). BDNF polymorphism associates with decline in set shifting in Parkinson's disease. Neurobiol Aging. 36(3):1605.e1-6.

Velazquez F.N., et al. (2015). c-Fos importance for brain development. Aging (Albany NY). 7(12):1028-9.

Vessey J.P., et al. (2010). Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function. Proc Natl Acad Sci U S A. 107(7):3222-7.

Vilar M. and Mira H. (2016). Regulation of Neurogenesis by Neurotrophins during Adulthood: Expected and Unexpected Roles. Front Neurosci. 10:26.

Vissers L.E., et al. (2016). Genetic studies in intellectual disability and related disorders. Nat. Rev. Genet. 17, 9–18.

Veyrac A., et al. (2014). The transcription factor Zif268/Egr1, brain plasticity, and memory. Prog Mol Biol Transl Sci. 122:89-129.

Volders P.J. (2015). An update on LNCipedia: a database for annotated human lncRNA sequences. Nucleic Acids Res. 43(8):4363-4.

Voth H., et al. (2007). Identification of DEIN, a novel gene with high expression levels in stage IVS neuroblastoma. Mol Cancer Res. 5(12):1276-84.

Voth H., et al. (2009). Co-regulated expression of HAND2 and DEIN by a bidirectional promoter with asymmetrical activity in neuroblastoma. BMCMol.Biol. 10:28.

Vuong C. K., et al. (2016). The neurogenetics of alternative splicing. Nat. Rev. 17, 265–281.

Wang L., et al. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res. 41(6):e74.

Watson M.A. and Milbrandt J. (1990). Expression of the nerve growth factor-regulated NGFI-A and NGFI-B genes in the developing rat. Development. 110(1):173-83.

Wei F., et al. (2000). Role of EGR1 in hippocampal synaptic enhancement induced by tetanic stimulation and amputation. J Cell Biol. 149(7):1325-34.

Weissman T., et al. (2003). Neurogenic Radial Glial Cells in Reptile, Rodent and Human: from Mitosis to Migration. Cerebral Cortex, 13(6), 550–559.

Weissmiller A.M. and Wu C. (2012). Current advances in using neurotrophic factors to treat neurodegenerative disorders. Transl Neurodegener. 1(1):14.

Wells T., et al. (2011). Transcription Mapping of Embryonic Rat Brain Reveals EGR-1 Induction in SOX2 Neural Progenitor Cells. Front Mol Neurosci. 4:6.

Willsey Silbereis A.J., et al. (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell 155, 997–1007.

Wilusz J.E., et al. (2009). Long noncoding RNAs: functional surprises from the RNA world. Genes Dev. 23(13):1494-504.

Wu P., et al. (2013). Roles of long noncoding RNAs in brain development, functional diversification and neurodegenerative diseases. Brain Res. Bull. 97:69-80.

Xiao B. et al. (1998). Homer regulates the association of group 1 metabotropic glutamate receptors with multivalent complexes of homer-related, synaptic proteins. Neuron 21, 707–716.

Xu B., et al. (2000). The role of brain-derived neurotrophic factor receptors in the mature hippocampus: modulation of long-term potentiation through a presynaptic mechanism involving trkB. Journal of Neuroscience. 20, 6888–6897.

Yamaguchi Y. and Miura M. (2015). Programmed cell death in neurodevelopment. Dev Cell. 32(4):478-90.

Yang F., et al. (2013). MALAT-1 interacts with hnRNP C in cell cycle regulation. FEBS Lett. 587, 3175–3181.

Yarmishyn A.A., et al. (2014). HOXD-AS1 is a novel lncRNA encoded in HOXD cluster and a marker of neuroblastoma progression revealed via integrative analysis of noncoding transcriptome. BMC Genomics. 15 Suppl 9:S7.

Yeo G., et al. (2004). Variation in alternative splicing across human tissues. Genome Biol. 5:R74.

Yeo W. and Gautier J. (2004). Early neural cell death: dying to become neurons. Dev Biol. 274(2):233-44.

Zagrebelsky M. and Korte M. (2014). Form follows function: BDNF and its involvement in sculpting the function and structure of synapses. Neuropharmacology. 76, 628-638.

Zheng S. and Black D. L. (2013). Alternative pre-mRNA splicing in neurons: growing up and extending its reach. Trends Genet. 29, 442–448.

Zhu S., et al. (2018). LINC00473 antagonizes the tumour suppressor miR-195 to mediate the pathogenesis of Wilms tumour via IKKα. Cell Prolif. 51(1).

Zuccato C., et al. (2008). Systematic assessment of BDNF and its receptor levels in human cortices affected by Huntington's disease. Brain Pathol. 18(2):225-38.