SISSA

Scuola Internazionale Superiore di Studi Avanzati

Neuroscience Area – PhD course in

Functional and Structural Genomics

FOXG1 modulates LINE1 activity in the developing embryonic neocortex

Candidate:

Gabriele Liuzzi

Advisor:

Prof. Antonello Mallamaci

Academic Year 2021/2022

Abstract

Foxg1 is an ancient transcription factor mastering telencephalic development. *mLINE1* is a large retrotransposon family contributing to plasticity of neuronal genome. All *mLINE1's* share putative Foxg1 binding elements. Moreover, it is known that *mLINE1* transcription is upregulated in postnatal glutamatergic progenitors, whereas Foxg1 declines in embryonic progenitors of the same lineage. Based on that, we predicted that Foxg1 might limit *mLINE1* expression (and activity). Here we tested such prediction, with particular attention to its cellular and molecular features.

First, as expected, we found that *mLINE1*-mRNA encoded by all three retro-transposition-competent subfamilies (A, Gf and Tf) was increased in *Foxg1*-loss-of-function (-LOF) mouse pups. Next, to model articulation of *Foxg1*-dependent *LINE1* regulation, we developed a dedicated set of in vitro preparations, representing early-, mid- and late-phases of neuronogenesis. By profiling this set, we detected a progressive increase of *LINE1*-mRNA levels from neural stem cells (NSCs) up to neurons (Ns). Then, taking advantage of these preparations, we manipulated Foxg1 levels at different stages of the neuronogenic progression, according to specific neural cell type restrictions. We mapped changes in Foxg1 protein levels evoked of these manipulations, and we evaluated *LINE1*-mRNA levels originating from them. Integrated analysis of results pointed to selective *mLINE1* repression by Foxg1 in neuronal progenitors (NPs) and Ns.

To assess if Foxg1 modulation of *mLINE1*-mRNA takes place via direct gene trans-repression, we evaluated Foxg1 enrichment at *mLINE1* loci via ChIP. In wild type cultures, Foxg1 was enriched along the whole body of mLINE1's belonging to all three subfamilies. However, this enrichment was restricted to mid-neuronogenic cultures, and not detectable in early-neuronogenic ones, ruling out Foxg1 modulation of *mLINEs* in NSCs. Intriguingly, such enrichment was more pronounced upon artificial Foxg1 overexpression, suggesting that Foxg1 may physiologically tune *mLINE1* transcription. Furthermore, we found that Foxg1 overexpression in mid-neuronogenic cultures elicited a generalized decrease in transcription-activating marks (H3K4me3) and an increase of repressing ones (H3K9me3). These phenomena occurred at all *mLINE1* subfamilies loci, suggesting that Foxg1 largely impacts on *mLINE1* transcription via modifications of the corresponding epigenetic landscape.

Next, we moved to Foxg1 control (if any) of cellular *mLINE1*-DNA content. First, we found that this content was upregulated by about 25% in Tubb3^{+}, post-mitotic wild type neurons compared to their ancestors, in a retro-transcription-dependent way. Next, paradoxically, we also discovered that Foxg1 down-regulation elicited a co-linear reduction in *mLINE1*-DNA content, suggesting that Foxg1 is strictly needed to sustain physiological amplification of such DNA. This phenomenon was observed in vivo as well as in vitro. Puzzingly, Foxg1 overexpression led to variable outcomes, increased, unchanged or decreased *mLINE1*-DNA content, depending on different promoters driving it. Differential sensitivity of *mLINE1* transcription and retro-transcription to Foxg1 levels could account for these phenomena. Finally, mechanisms underlying Foxg1 impact on *mLINE1* DNA content were not yet investigated, however, we got evidence that direct Foxg1 protein/*mLINE1*-mRNA interaction could underlie it.

Contents

1. INTRODUCTION 1

1.1. Evolution of the mammalian brain.....1

1.2. Evolution of the six-layered mammalian neocortex.....3

1.3. Development and specification of neocortical patterning.....5

- 1.3.1. Spatio-temporal histogenesis of the murine cortex.....7
- 1.3.2. Neocortical precursors.....7
- 1.3.3. Genetic specification of neocortical progenitors.....8
- 1.3.4. Specification of projection neurons.....9

1.4. Foxg1: structure and function of the gene.....12

1.5. Foxg1 in neural precursor expansion 14

1.6. Foxg1 in the differentiation of pyramidal neurons.....15

1.7. Foxg1 as a transcriptional repressor.....17

1.8. FOXG1 and neurological disorders.....18

1.9. Transposable elements (TEs) in the mammalian genomes.....20

1.10. Structural and functional diversity of TEs.....21

1.11. Co-evolution between TEs and host genomes.....23

1.12. Evolution of *LINEs* from deuterostomes to mammals.....24

1.13. Evolutionary divergence of *LINE1* between mouse and human.....25

1.14. Structure of *LINE1*.....25

- 1.14.1. *LINE1* 5'UTR.....25
- 1.14.2. Murine *LINE1 subfamilies* 27
- 1.14.3. *LINE1* ORF1p.....28
- 1.14.4. *LINE1* ORF2p.....29
- 1.14.5. *LINE1* 3'UTR.....30

1.15. *LINE1* retro-transposition mechanism.....30

1.16. Regulation of *LINE1 retro-transcription*.....32

1.17. Genomic rearrangements generated during *LINE1* retro-transposition.....33

- 1.17.1. Intra-*LINE1* alterations.....33

1.18. Regulation of *LINE1*-expression.....34

1.18.1. Transcriptional regulation of *LINE1*-expression.....35

- 1.18.1.1. Transcription factors.....35

- 1.18.1.2. DNA methylation.....37
- 1.18.1.3. MeCP2 regulates neurogenesis and differentiation.....38
- 1.18.1.4. Histone modifications associated with retrotransposons.....39
- 1.18.1.5. piRNA, PIWI proteins and small RNAs derived from *LINE1*.....41

1.18.2. Post-transcriptional regulation of *LINE1* expression.....42

- 1.18.2.1. Endogenous host factors involved in post-transcriptional regulation of *LINE1*.....42
- 1.18.2.2. RNA interference (RNAi).....43
- 1.18.2.3. piRNA-mediated post-transcriptional silencing.....44

1.19. Somatic retro-transposition of *LINE1* in the Nervous System.....45

2. AIMS OF THE PROJECT 47

3. MATERIALS AND METHODS 48

3.1. Animal handling.....48

3.2. Lentiviral vectors.....48

3.3. HEK293T cell cultures.....50

3.4. *In vivo* dissection of *Foxg1*$^{+/-}$ and *Foxg1*$^{+/+}$ embryos.....50

3.5. Primary cortico-cerebral cell cultures.....51

3.6. Primary neuronal cell cultures.....52

3.7. Immunofluorescence: sample preparation.....52

3.8. Primary cortico-cerebral cell cultures immunofluorescence.....53

3.9. Image acquisition and analysis.....53

3.10. Evaluation of Foxg1 protein levels in distinctive cell types of engineered primary neural cultures.....53

3.11. RNA extraction.....54

3.12. DNAse treatment; Reverse Transcription and Real Time quantitative PCR.....54

3.13. RNA Immunoprecipitation (RIP) assay.....57

3.14. Chromatin Immunoprecipitation (ChIP) assay.....58

3.15. Fast genotyping of CD1-Wt male and female embryos.....59

3.16. Isolation of DNA from primary cortico-cerebral cell cultures.....59

3.17. Statistical analysis.....60

4. RESULTS

4.1. In vivo Foxg1 down-regulation of *mLINE1*-mRNA……..61

4.2. Modeling temporal progression of pallial *mLINE1*-mRNA in vitro……..62

4.3. Modeling Foxg1 regulation of *mLINE1*-mRNA in vitro, by constitutive Foxg1 manipulation……64

4.4. Modeling Foxg1 regulation of *mLINE1*-mRNA in vitro, by conditional Foxg1 overexpression……66

4.5. Mechanisms which underlie Foxg1 control of *mLINE1* transcription……71

4.6. Modeling temporal progression of pallial *mLINE1*-DNA copy numbers in vitro……75

4.7. Assessing Foxg1 impact on *mLINE1*-DNA copy numbers, both in vivo and in vitro……77

4.8. Hints about putative mechanisms underlying Foxg1 impact on *mLINE1*-DNA amplification……79

5. DISCUSSIONS AND CONCLUSIONS 81

5.1. Modelling *mLINE1* dynamics in primary pallial cultures……82

5.2. Foxg1 dampens *mLINE1*-mRNA expression in the developing rodent pallium……82

5.3. Foxg1 increases *mLINE1*-DNA copy number within the rodent developing pallium……84

5.4. Conclusions and significance……86

6. SUPPLEMENTARY MATERIALS 88

7. BIBLIOGRAPHY 91

Acknowledgments 113

1. Introduction

The neocortex constitutes the evolutionarily most recent region of the mammalian cerebral cortex. It is largely accepted that the profuse expansion of the cortical surface, and the defining of its intricate connections has enabled in higher vertebrates, including primarily the human species, unique cognitive abilities including abstract thinking, long-term planning, memory, verbal language, and elaborate tool-making skills (Amjad et al., 2010). This expansion involves growth in the number of neocortical neurons, which is highly regulated during development and primarily reflects the number of neurogenic divisions of distinct classes of progenitor cells. The extraordinary complexity of events leading to the formation of the neocortex is reflected in its embryonic development, which generates myriads of cells with divergent profiles in terms of morphology, function, and gene expression. This process by which new neurons are generated from immature cells is called neuronogenesis, and must be regulated with a precise spatiotemporal order, both to generate all neural cells and to assemble and integrate them properly into functional neural circuits. (Florio and Huttner, 2014). The integration of these neurons into exquisitely precise neural circuits underlies the foundation of organismal behavior and interaction not exclusively between the organism and its environment but also at a higher biological complexity, among individuals who are associated by the relevant ecological niche they belong to (Proekt et al., 2008). The cerebral cortex varies greatly in size and complexity among amniotes due to differences in the number and composition of neurons. These differences emerge during embryonic development as a result of changes in neural histogenesis, which is thought to recapitulate changes that occurred during evolution that culminated in the human neocortex (Proekt et al., 2008).

1.1. Evolution of the mammalian brain

The history of mammals begins at the primitive level with the appearance of a particular lineage of reptiles that had some of the distinctive and peculiar characters of mammals. (Kemp, 2005; Kielan-Jaworowska et al., 2005; Rowe, 2017).

The most ancestral primitive reptiles, now usually referred to as "stem amniotes", raised dichotomously starting from amphibians about 320 million years ago at the time the Carboniferous was about to end (Benton, 2015; Colbert et al., 2001) and after 15 mya (about 305 mya) split into two main groups, the sauropsid (or diapsid) and secondly the synapsid one (Figure 1). Modern reptiles and birds raised from the sauropsid clade, the synapsid clade led to the birth of the first mammals ~280 mya.

It has been proposed that the stem amniotes from which birds, reptiles and current mammals originated, had a cerebral cortex. A study model for the organization of this amniote cortex was proposed based on gene expression, cell migration patterns, neural connectivity, and

cytoarchitecture (Puelles et al., 2016). This cortex was found to be subdivided into a ventral part and three dorsal domains, namely a medial, a lateral, and a dorsal one. While the medial part in sauropsids is homologous to the formation of both the hippocampus and entorhinal cortex in mammals, the lateral cortex is homologous to piriform cortex, claustrum, and insular cortex. As the advent of reptiles established, the dorsal domain developed in the dorsal cortex, which is homologous to the "Wulst" or "hyperpallium" in birds (Reiner et al., 2004; Striedter, 2005) and eventually to the "neocortex" in mammals (Figure I1) (Puelles et al., 2016; Briscoe and Ragsdale, 2018). In mammals, the ventral portion originates the olfactory bulb along with the olfactory cortex. The six-layered neocortex, which in the mammalian class constitutes a key evolutionary hallmark, emerged biologically in the form of anlage from the dorsal cortex in reptilian ancestors, which is characterized by the presence of three layers. Of these, only one is characterized by the presence of cellular soma of both interneurons and pyramidal neurons (Bruce, 2010; Molnár, 2011; Aboitiz et al., 2002). Both reptiles and birds (sauropsids) are characterized by presence of a telencephalic macrostructure called Dorsal Ventricular Ridge (DVR) (Figure 1). At this level, the amount of visual, somato-sensory, and auditory sensory inputs is integrated, allowing this particularly circumscribed area to play a primary role in performing many of the functions unique to the mammalian neocortex. (Montiel et al., 2016; Striedter, 2005; Puelles et al., 2017; Medina, 2007). In birds, the cortex appears to be much more expanded than in reptiles but unlike the latter, it is organized into nuclei and not layers (Figure I1) (Dugas-Ford et al., 2012). Many references in the literature agreed in establishing that Wulst is homologous to the reptilian dorsal cortex but also to the typical mammalian neocortex. Despite this, its anatomy has smaller dimensions in the vastness of birds than the mammalian neocortex (Striedter, 2005).

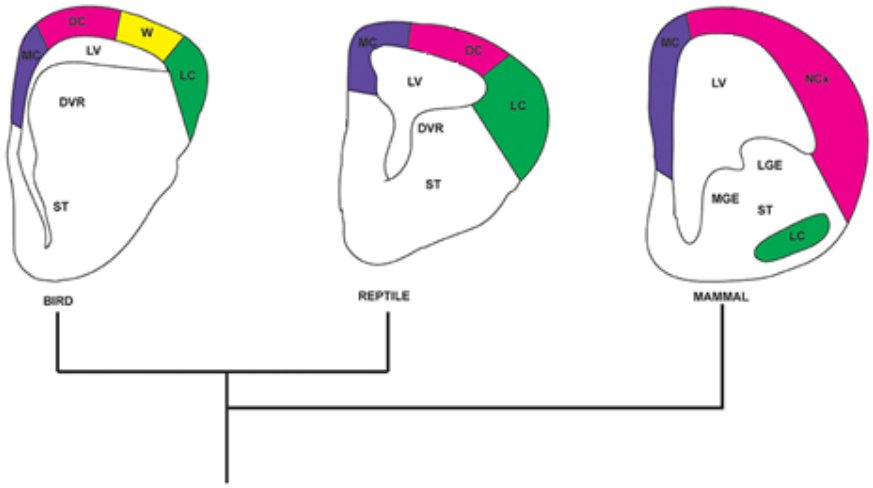


Figure I1: Schematics of color-coded evolutionary homology between regions of the forebrain evolved in reptiles, birds, and mammals, respectively. Both left and central images show sections of the forebrain of reptiles and birds where dorsal cortex (DC), medial cortex, lateral cortex, the dorsal ventricular region (DVR) and Wulst (W) are highlighted. The position of the striatum is also indicated (ST). An indicative coronal section of the mammalian forebrain is shown in the right section. The positional boundaries of the neocortex (NCx), medial cortex (MC), lateral cortex (LC), medial ganglionic eminence and lateral ganglionic eminence (MGE and LGE) are indicated. (Modified from Franchini, 2021).

Because of this reason, it has been proposed that a highly expanded DVR in birds (Figure I1) is just directly implicated in performing many of the cortical mammalian functions, being itself homologous to the mammalian claustrum and amygdala, whereas the dorsal cortex and Wulst are homologous to the mammalian neocortex (Franchini, 2021).

1.2. Evolution of the six-layered mammalian neocortex

During the early part of the Mesozoic era, the first mammals appeared. From this group, the egg-laying prototherians split very early around 200 mya, while the metatherians or marsupials separated more recently, about 150 mya before the phyletic separation that marks the advent of the eutherian or placental mammals. Regarding the characteristics of the six-layered neocortex, it has been established to date that all mammals belonging to the therian systematic subclass including marsupials and placentals, possess a six-layered neocortex. Indeed, marsupials have been shown to exhibit an organized subventricular zone (SVZ), established by the presence of basal progenitor cells as well as the expression of relevant gene sets resembling those found in eutherian mammals (Figure I2). This implies that the SVZ emerged prior to the phyletic divergence that marks the diversification between both prototherian and metatherian infraclasses (Cheung et al., 2010). The evolution and development of the six-layered neocortex required the selection of a mechanism to the formation of a highly organized and functionally ordered structure leading to an increased neural production along the course of embryonic neurogenesis (Figure I2). This process takes place at the very beginning with the emergence of a distinctive cell type that to date is considered a key player of laminar cortical development. In the mammalian embryonic cortex, the basal radial glia (bRG) constitutes the major progenitor cell type that originates in the ventricular zone (VZ) (Figure I2).

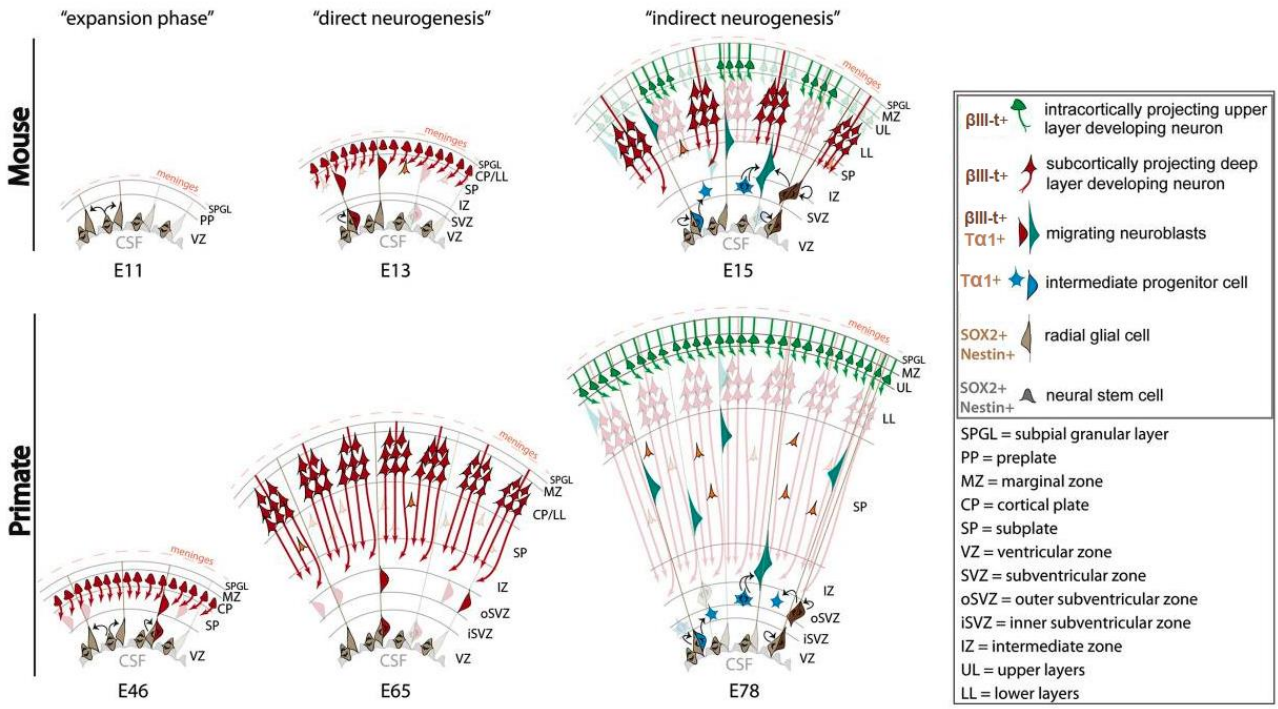


Figure I2: Model scheme showing the mechanisms by which neocortical neurogenesis is carried out comparing mouse and primates during development; and involves the prioritization of direct neurogenesis to indirect neurogenesis. Following an initial expansion phase characterized by symmetric divisions and expansion of the neural precursor pool, they subsequently originate a more committed neocortical stem cell type called radial glia (RG), which divide asymmetrically through "direct neurogenesis" drawing in the center. Following the progression of corticogenesis the RGs begin to massively originate outer radial glia and progenitor cells, (IPCs) and outer radial glia (oRGs) drawing to the right. This next stage of neurogenesis has been termed "indirect neurogenesis." (Modified from Deboer et al., 2013).

Within VZ, apical radial glia (aRG) produces daughter cells that can differentiate as post mitotic neuron through a mechanism known as direct neurogenesis, or remain proliferative in the form of progenitor. aRGs give rise to two types of basal progenitors: basal radial glia (bRG) and *classical* basal intermediate progenitors (IPs), which migrate to populate and define the outer and inner part of SVZ, respectively (Franchini, 2021). These basal progenitors are able in turn to divide and subsequently generate neurons (the whole process is referred to as "indirect neurogenesis") (Figure I2). Remarkably, "direct neurogenesis" produces neurons quickly, but also rapidly depletes the starting progenitor pool (Figure I2). This process is the mechanism responsible for the main production of pallial neuronal cells in reptiles and birds. Derived from the Diapside clade, these vertebrates do not possess an SVZ in the homologous region of the mammalian neocortex, where indirect neurogenesis occurs (see above). Based on this observation, it was hypothesized that in mammals, the evolution of the dual mechanism based on the co-presence of both direct and indirect neuronogenic mechanisms in the SVZ, could subtend the amplification in the number of neurons produced by increasing the rate and lengthening the period of neurogenesis (Fig. I2). Hence, this is supposed to represent the starting raw material for the expansion of the cerebral cortex in different mammalian lineages (Fig. I2) through a model involving the prioritization of direct neurogenesis over

indirect neurogenesis (Deboer et al., 2013). Many authors agree that the emergence of this mechanism may have been the key step leading to the advent and evolution of the six-layered neocortex (Roth and Dicke, 2013).

1.3. Development and specification of neocortical patterning

The nervous system begins its first phases of determination and development with the beginning of embryonic organogenesis. At the end of gastrulation, in fact, the embryo includes three leaflets, namely ectoderm, mesoderm and endoderm. Next, the ectoderm originates the neural plate, from which the nervous system is formed. First, this structure undergoes a folding along the central axis, such that it results in the formation of a concavity known as the neural groove, surrounded laterally by the neural folds. Subsequently, the union and fusion of the neural folds generates a new tubular structure known as the neural tube, which will lead to the formation of the three primary brain vesicles, the proencephalon (the forebrain), the mesencephalon (the midbrain), and the rhomboencephalon (hindbrain). The embryonic forebrain can be subdivided into telencephalon, diencephalon, hypothalamus, and retina. In turn, the telencephalon gives rise to basal nuclei, olfactory bulb, and cerebral cortex.

The telencephalon originates through the succession of three simultaneous phases. The first of these is represented by directional specification that occurs in two senses: both dorso-ventral and rostro-caudal. Second, fine regulation occurs that controls both differentiation and cellular mitotic rate. Third, as as a follow-up of these processes, lamination and arealization of the cerebral cortex ultimately result. These three steps depend on the presence of molecular factors that are synthesized and released by specific signaling centers, and also by morphogenic molecules. Among the major morphogens that are essential contributors in the development of the telencephalic cortex, there are Bone Morphogenetic Proteins (BMPs), Fibroblast Growth Factor 8 (FGF8), Sonic Hedgehog (SHH) family, Wing-less/Int (WNT) proteins and Reelin. These proteins diffuse according to concentration gradients leading to neocortex growth and regionalization. More in detail, the canonical Wnt signaling pathway plays a key role at the level of apical progenitors in the developing cortex. In 2002, Chenn and collegues found that transgenic mice expressing a stable form of beta-catenin in neural precursors developed macrocephaly as well as an increase in the surface of the cortical surface (Chenn and Walsh, 2002). While deletion of the *Fgf2* gene results in a decrease in the number of glutamatergic excitatory neurons, increased FGF signaling expands progenitor generation and leads to gyrification of the developing rostrolateral forebrain (Rash et al., 2013; Raballo et al., 2000; Korada et al., 2002). Bone morphogenetic proteins (BMP 2, 4, 5, 6, and 7) are secreted from the cortical rim and functionally interact with Wnts to induce dorsomedial patterning in the telencephalon and promote neuronal differentiation (Li et al., 1998; Bond et al., 2012). In the telencephalon, SHH is produced by Cajal-Retzius (CR) neurons in the marginal zone (MZ) and cortical plate interneurons (Komada et al., 2008; Huang et al., 2010). By means of transgenic mouse models,

Wang and collegues found an increase in the number of neuronal progenitors under constitutive Shh signaling, whilst loss of Shh signaling reduced their number and consequently, neocortical size (Wang et al., 2016). Through Reelin secretion, CR cells orchestrate radial migration of later born neurons from their region of origin to the cortical plate (CP) (Meyer, 2010).

Secreted ligands act on neurostem cells and progenitors and, depending on their concentration, they induce the expression of different transcription factors (TFs), which results in subsequent molecular parcelization of telencephalic and pallial territories. Along the ventral-dorsal as well as the antero-posterior axes they influence the differentiation of functional cortical regions. By employing a loss-of-function approach at this stage, several authors have identified that *Foxg1*, *CoupTF1*, *Emx2*, *Lhx2*, *Pax6*, Sp8, and *Lef1* show high transcriptional susceptibility to morphogens (Mallamaci, 2011; Ypsilanti and Rubenstein, 2016). Similarly, proteins encoded by the *Fezf2*, *Tbr1.2*, *Pax6*, *Ngn1/2, Emx1/2*, and *Satb2* genes selectively intervene in the activation of transcriptional programs targeted to glutamatergic neuronal lineage differentiation (Lai et al., 2013; Ypsilanti and Rubenstein, 2016). Notably, *Tbr2* regulates the transcription of a great many genes by going on to impact the proliferative and post-mitotic phase of committed neuronal progenitors (Hevner, 2019). Pax6 has also long been characterized for its intrinsic ability to regulate migration, neuronal maturation, and regional specification of the cortex at multiple different levels (Ypsilanti and Rubenstein, 2016).

The cerebral cortex can be divided into: neocortex, paleocortex, and archicortex. Mammalian neocortex comprises distinct cortical areas and develops diversified laminar patterns which are both retained during adulthood. The neocortex primarily handles afferent sensory information and it is organized in specialized regions for different functions. These areas include: primary somatosensory (S1), visual (V1), and auditory (A1) areas. Besides these, the neocortex includes other areas such as motor areas, secondary somatosensory areas, visual areas, and others that vary from lineage to lineage (Franchini., 2021). Despite different duration of neocorticogenesis in different orders of mammals, there is a clear basic similarity in the sequence of events that finely orchestrate it. The first stage of neocortical development is the partition of the cortical wall into an inner layer of proliferating cells, called the ventricular zone (VZ), and an outer layer including early born neurons, known as primordial plexiform layer (PPL) (Kriegstein and Alvarez-Buylla., 2009). Next, apical precursors located in VZ give raise to more basal descendants, accumulating in a new region above VZ, termed sub-ventricular zone (SVZ). Newly generated neurons, coming from VZ and SVZ, penetrate the PPL and split it into a superficial "marginal zone" (MZ) and a deeper "subplate" (SP). These neurons stop inbetween MZ and SP and form the so-called cortical plaque (CP), where they settle according to the inside-out rule (the earlier born the deeper). MZ and CP give rise to layer I and VI-II of mature cortex, respectively. SP partially disappears or is incorporated in layer VI. As the central phase of neuronogenesis is going to end, first astrogenesis and then pallial oligodendrocytogenesis are activated. In rodents, they will continue past birth (Borrell., 2019), so that, like other CNS structures, even neocortex is generated through three main, consecutive and partially overlapping, histogenetic waves (Kriegstein and Alvarez-Buylla., 2009; Angevine and Sidman., 1961; Rakic., 1974).

1.3.1. Spatio-temporal histogenesis of the murine cortex

In the mouse corticogenesis begins at embryonic day 9.5 (E9.5) with the initial proliferation of neuroepithelial cells that undergo symmetrical cell divisions to expand the multipotent stem compartment (Götz and Huttner., 2005; Smart., 1973). At this stage or slightly later, the telencephalic anlage consists of two dorsolateral convex expansions of the primary forebrain vesicle. Next, murine cortical neuronogenesis takes place between embryonic day 11 (E11) and E17 (Chen and McConnell, 1995). Initially, as the developing cortical epithelium thickens, the neuroepithelial cells elongate and differentiate into radial glial progenitors. A particular feature of these cells is the extension of an elongated cellular process that extends from the ventricular wall to the pial surface and plays a crucial role in guiding neurons to their final positions in the cortical plate, serving as a migratory scaffold (Rakic., 1972; Rakic., 2003). In addition to this, radial glia cells also function as neural stem cells that make major contributions to cortical neurogenesis by generating neurons according to a direct mode in the VZ, or indirectly through the production of proliferating intermediate progenitors (IPCs) (Anthony et al., 2004; Miyata et al., 2004; Noctor et al., 2001; Malatesta et al., 2000). In fact, intermediate progenitor cells (IP) (also known as basal progenitors because they divide without touching the apical surface) are proliferating progenitors that express the transcription factor Tbr2 and lead to the origin of most excitatory neurons that differentiate from the subventricular zone (SVZ). In rodents, they are the most important type of neuron-producing progenitor. They act primarily as transit-amplifying cells, undergoing limited proliferative divisions and more often dividing symmetrically to produce two neurons (Haubensak et al., 2004; Kowalczyk et al., 2009; Noctor et al., 2004). At E13-E14, the VZ grows in thickness due to intense mitotic activity by PVE cells, expanding the progenitor pool. From E14 to E16 the VZ shrinks and the SVZ expands. Newborn neurons continue to migrate to cortical plate, which thickens progressively, reaching its maximum width at the E18-E19.5 stage. At this terminal stage, all neuronal cells have exited the cell cycle and migrated to their final location within the developing cortex (Takahashi et al., 1996). At this final stage even the SVZ-derivative, white matter, including plenty of astroglial progenitors, has reached its maximal width.

1.3.2. Neocortical precursors

As mentioned in the preceding paragraph 1.3.1., within the early rodent pallium there are two main proliferative populations, namely the pseudostratified ventricular epithelium (PVE) and the secondary proliferative population (SPP). Primarily, the PVE is close to the ventricle and it is initially composed of neuroepithelial cells (NEC) that later acquire radial glial cell (RGC) nature. Both NECs and RGCs undergo interkinetic nuclear migration (IKNM), i.e. they synthesize DNA far from the

ventricle and give rise to mitosis close to it. As said, they constitute the stem cell pool from which the vast majority of mature neural types within the neocortex will be derived (Franchini, 2021). The SPP has a more heterogeneous structure. In the earliest stages, it includes neuronogenic progenitors (= intermediate precursor cells, IPCs, aka basal precursors, BPs). Later, it is largely populated by astrogenic progenitors. Both progenitor types do not display IKNM. Remarkably, it has also been shown that a substantial fraction of neuronal precursors located within the VZ and undergoing IKNM are not radial glial cells, but, conversely, committed neuronogenic progenitors. They are termed short neuronal precursors (SNPs) or "pin-like cells" (Hartfuss, 2001). IPCs and SNPs generate the vast majority of deep layer and upper layer projection neurons (Malatesta et al., 2003; Hartfuss, 2001). The biological relevance of the most important genes involved in the specification of neocortical progenitors will be discussed in more detail in the following section.

1.3.3. Genetic specification of neocortical progenitors

Following telencephalon induction, several genes are expressed at the level of the dorsolateral wall of the telencephalon that are essential to orchestrate the process of neocortical neurogenesis. These include *Emx2* (empty spiracles homolog 2), *Lhx2* (LIM homeobox 2), Foxg1 (Forkhead Box G1) and *Pax6* (paired box 6). Each of these genes is of crucial importance in determining the specification of progenitor cells that after terminal differentiation will give rise to projection neurons of the neocortex. Through the coordinated expression and activity of these four genes, while *Lhx2* and *Foxg1* act by repressing dorsal midline fates in maintaining the domain of neocortical progenitors, *Emx2* and *Pax6* repress ventral midline fates. Murine models loss of function for *Lhx2* display an enlargement of the medial choroid plexus (CPI) and cortical hem (CH) domains as well as the absence of progenitors that belong to the cortical domain and specifically possess a neocortical identity (Muzio et al., 2002). Similarly, the loss of the transcription factor Foxg1 prevents the activation of the neocortical program and leads to expansion of CH and archicortex (acx) (Muzio et al 2005). In particular, regarding Emx2 and Pax6, their expression lead to the formation of opposite concentration gradients, which overlap in the intermediate region of the dorsal telencephalon. These gradients are needed for proper development of cortical areas and are also required to establish the identity of dorsal progenitors (Mallamaci and Stoykova, 2006). In mice homozygous null for *Emx2* and *Pax6*, ventralization of cortical progenitors and consequently

absence of the neocortical domain (Ncx), archicortex (Acx), CH and CPI by embryonic day 14 (E14.5) has been observed and described (Muzio et al, 2002).

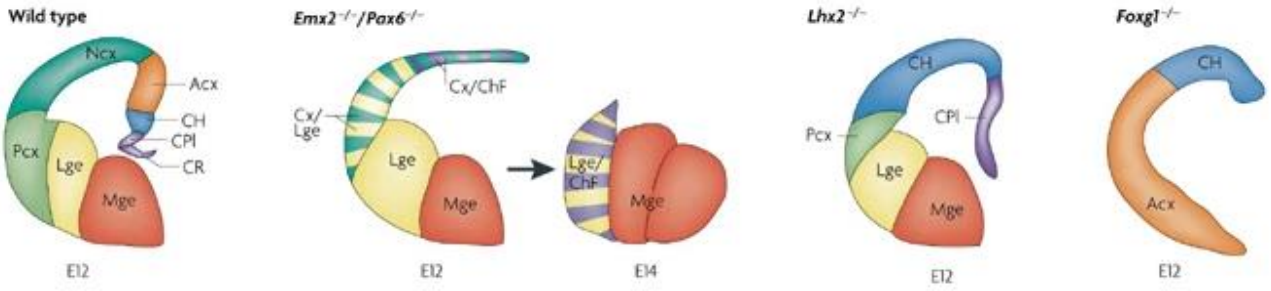


Figure I3: Crucial roles of Pax6, Lhx2, Emx2, and Foxg1 in determining the genetic specification of neocortical progenitors. Ncx, neocortex; Acx, archicortex; Cx, cortex,' CH, cortical border; CPI, choroid plexus; CR, choroid roof; chF, choroid field (choroid plexus and choroid roof); Lge, lateral ganglionic eminence; Mge, medial ganglionic eminence; Pcx, paleocortex. (Modified from Molyneaux et al., 2007).

1.3.4. Specification of projection neurons

Cortical neurons fall into two major categories: interneurons, establishing *local* connections, and projection neurons, extending their axons to *far*, intracortical, subcortical, or subcerebral targets. Projection neurons are glutamatergic cells, predominantly characterized by a pyramid-like morphology (Spruston, 2008). They are generated during neurogenesis from progenitors in the dorsolateral wall of the telencephalon and acquire their final shape via intermediate multipolar and bipolar morphologies, while leaving proliferative layers and migrating radially to the cortical plate (Anderson, 2002; Molyneaux et al., 2007). Conversely interneurons are GABAergic. They originate primarily from progenitors located in ventral telencephalon and reach the neocortical parenchyma thanks to a long tangential migration. Upon arrival, the interneurons move radially to enter the CP (LoTurco and Bai, 2006). They then segregate into their layer destinations depending on their molecular subtype, origin, and date of birth during early postnatal development (Ware et al., 1999; Rakic, 1972). In this way, different progenitor zones contribute to the variety of neuronal types in the neocortex.

Once generated at the level of the germinal region, projection neurons move to the cortical lamina, where they extend a primary axon generally facing the ventricular cortical wall region. The axonal growth cone may reach the thalamus, as in case of cortico-thalamic neurons belonging to layer VI. Alternatively, the cone may cross the cerebral peduncle, the pyramidal tract, the brainstem, and finally the spinal cord, that is the case of cortico-subthalamic projection neurons of layer Vb. Finally, in the case of axons projecting to ipsilateral or heterolateral cortical areas, the respective neurons will be confined to layers II-III and layer V consisting of projection neurons that establish internal cerebral cortex connections, i.e. cortico-cortical) (Schreyer and Jones, 1988). Beyond axonal

growth cones, secondary collateral cones originate from the primary axon after it has passed other targets (O'Leary and Terashima, 1988). All supernumerary and surplus connections are remodeled and eliminated. This causes neurons in the visual cortex to maintain their projections directed to dorsal mesencephalon, while subcerebral projection neurons in the sensorimotor cortex extend their axons solely to the caudal pons and spinal cord. Based on this pattern of early development, the possibility that certain sets of genes regulating early differentiation and specification are also shared among different types of subcerebral projection neurons cannot be ruled out (Inoue et al., 2004; Molyneaux et al., 2005). Recent advances and advances in understanding cortical development have highlighted the microtubule-associated protein Lszts as a master regulator of the cellular dynamics occurring at this level of neuronal specification. Indeed, it has been reported that Lszts acts at several levels among which first emerges its action on the bRGC and neurons derived from the asymmetric division of the apical radial glia (aRGC) so as to facilitate delamination and thus alter the organization of the apical junction. (Borrell, 2019). An additional TF was recently found to be strictly necessary for specification of the entire population of subcerebral projection neurons to occur. Fezf2 is expressed in all subcerebral projection neurons beginning at the onset of neurogenesis, and its persistence remains late into adulthood. In *Fezf2* $^{-/-}$ null mutants, subcerebral projection neurons are absent and consequently there are no connections between cortex and spinal cord (cortico-spinal) or between cortex and brainstem (cortico-pontine). According to the state of the art, at the end of the lamination phase *Fezf2* does not appear to be involved in the ability to affect the fate-choice of progenitors, leading them to generate significantly different fractions of glutamatergic neurons residing in layer V. Rather, Arlotta and colleagues support the hypothesis that Fezf2 may act to promote the next phase of neuronal specification (Arlotta et al., 2005). Nevertheless, through gain of function approach, *Fezf2* overexpression was found to be sufficient to cause the emergence of Foxg1 positive projection neurons in the deep layer (Chen et al., 2005). Ctip2 is a TF characterized by a C2H2 Zinc Finger domain reported previously to be expressed in specific regions of the CNS, including the neocortex. It has been identified as a crucial regulator of sub-cerebral axon extension. In homozygous null mouse models and loss of function of this protein, axons of sub-cerebral projection neurons show defects in axonal cone expansion and growth and in identifying the direction and direction of growth. This results in a reduced number of axons reaching the primary target, i.e., the brainstem and consequently presence of fasciculation. In addition to *Lszts*, *Fezf2* and *Ctip2*, another key TF that is very important and studied to play a role in the choice of the terminal target with which to establish axonal synapses in subcerebral projection neurons is *Otx1*. This protein is expressed in the deep layer progenitors in the VZ at the time when the upper layer neurons come to be generated. At this stage, it shows decreasing levels of expression in the VZ (Gray et al., 2004). *Otx1*$^{-/-}$ mice are found to be defective in the development of corticotectal projection neurons. Compared with *Fezf2* and *Ctip2*, *Otx1* might play a later role in the development of sub-cerebral projection neurons by controlling the branching and arborization density of axonal collateral growth cones (Arlotta et al., 2005).

A possible model for the generation of sub-cerebral projection neurons is described in Figure I3. Here, altogether, *Emx2*, *Foxg1*, *Lhx2* and *Pax6* dictate the neocortical histogenetic potential of

the progenitors commissioned to generate different glutamatergic neuronal classes (Figure I4). Next, at the stage when single subtypes of pyramidal neurons are generated, more specific TF-sets might specify their laminar identity. For example, *Fezf2* could specify the lineage of layer V sub-cerebral projection neurons, whereas *Otx1* and *Ctip2*, which are strongly associated with both sub-cerebral axon development and target recognition and axonal growth direction, could additionally specify further morphological features of sub-cerebral projection neurons (Figure I4) (Visel, 2004). To date, the search for novel effectors involved in layer-specific differentiation programs new work is still ongoing. Remarkably, however, the developmental step at which the commitment to activate layer-specific programs occurs is still controversial. In fact, although the existence of early precursors that were selectively able to generate projection neurons from deep layers VI-V or upper layers IV and II was initially identified, it was subsequently shown that *Fezf2* precursors within the early-stage VZ possess the differentiation potential to generate neurons belonging to both the upper and deep layers (Lein et al., 2007).

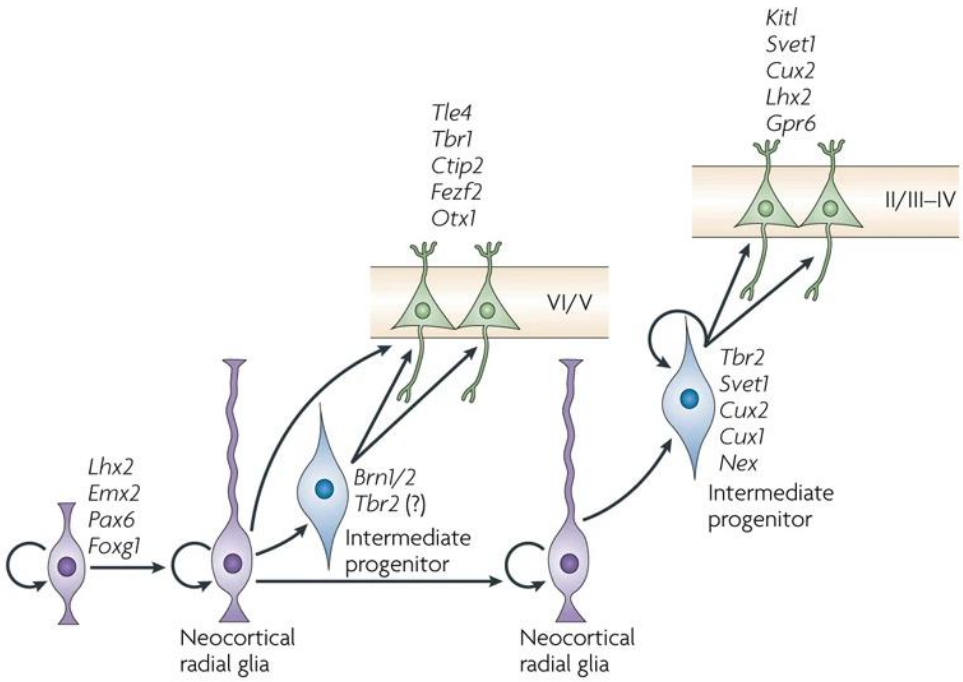


Figure I4: model proposed by Molineaux and collegues in 2007, which describes the generation of pyramidal neurons during the neocortical lamination process. As the course of development proceeds, the initial RGs progenitors laying at the level of the ventricular domain (VZ) generate IPs in the SVZ, which are able to originate both pyramidal neurons belonging to both deep and upper layers, depending on their genetic and laminar specification (Modified from Molineaux at al., 2007).

At the vertebrate level, *Foxg1* is one of the first genes to be expressed and restricted to the telencephalic field. Its persistence in the telencephalon extends from the early embryonic stage to the adult, and even single nucleotide mutations can cause severe brain impairment (Frisari et al., 2022; Wong et al., 2019). Over the past 3 decades, many studies have helped to characterize the various functions of Foxg1, ranging from control of telencephalic development to its pleiotropic

roles during cerebral cortex histogenesis (Figure I4). For these reasons, *Foxg1* is the main topic of this dissertation that focuses on its role in the development of the glutamatergic neuronal lineage (Florian et al., 2012; Cargnin et al., 2018).

1.4. *Foxg1*: structure and function of the gene

Foxg1, namely Forkhead Box G1, is a highly conserved winged-helix transcription factor acting as master regulator of telencephalic development. The *Foxg1* locus is located in mouse chr. 12q.B3, and human chr. 14q.13. In the mouse, Foxg1 encodes for two main mRNAs, an intronless 3.9kb long one, and a di-exonic, 2.9 kb long one. Both include a unique cds, encoding for a 481 aa long polypeptide. Sequence conservation among mammals is very high, both at polypeptide and mRNA level (in the latter case, in both cds and UTRs).

The Foxg1 polypeptide includes three beta strands, three terminal alpha helices and two intermediate loops (Figure I5) (Götz and Huttner, 2005). The the term "Forkhead domain", by which the protein domain binding to the major groove of DNA is named, is related to the morphological phenotype displayed by the brain of a Drosophila mutant, harboring a specific mutation within a gene encoding for a DNA-binding domain structurally similar to the Foxg1 one. Moreover, because of the three terminal α-helices flanked by two loops reminiscent of butterfly wings, it is possible to refer to this pattern of motifs in secondary structure of with the nominative of ''winged helix.'' This categorization on a conformational basis, is shared by Foxg1 with more than 100 different TFs, with functions also significantly different from each other. FOX TFs are direct DNA interactors. They share a DNA binding domain (DBD), which recognizes and binds the consensus matrix RYAAYA (R = A/G, Y = C/T) (Li et al., 2017; Brent et al., 2008), and adopts a winged helix conformation, with three helices, three filaments, and two less conserved winged rings (wing1 and wing2) (Chen et al., 2019). Upon direct interaction with DNA, the third helix (H3) fits into the main DNA groove in an orientation nearly perpendicular to the DNA duplex and contributes to most specific interactions. In addition, the N-terminal and the two wing loops of the DBD together operate divergent roles in DNA binding by interacting with phosphate groups. (Shuyan et al., 2020). In addition to the forkhead domain (FKHD), the Foxg1 protein harbors a 20aa-long Groucho (Gro)-binding domain (GBD), and a 10aa-long histone demethylase (KDM5B/JARID1B)-binding domain (JBD), the latter close to the C-terminal part (Figure I5). A striking feature of the N-terminal domain is its contribution to DNA-binding in addition to the classical FKHD. Thereby, Foxg1 seems to recognize and bind to canonical, FKHD recognition binding motifs, and to alternative, non-canonical DNA sequences (Akol et al., 2021).

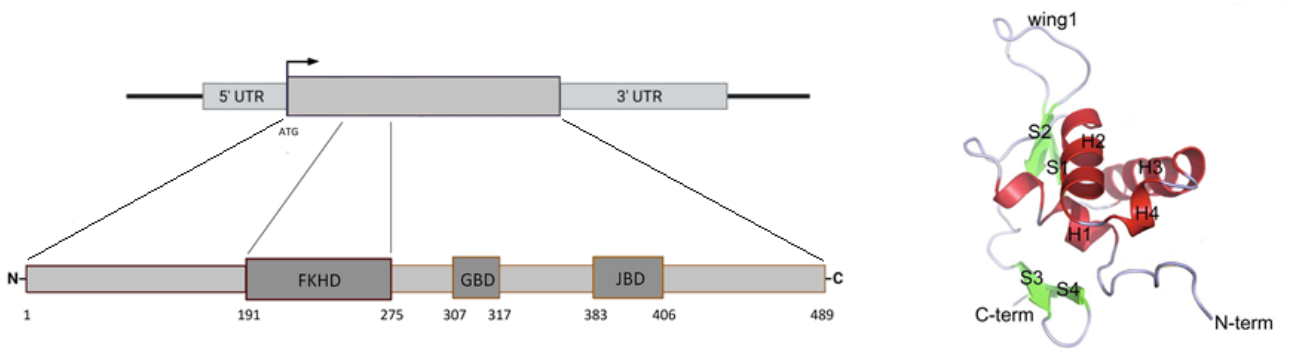


Figure I5: Schematic representation of Foxg1 gene and protein domains. (Left): representation of Foxg1 locus and of three functional domains (FHD, GBD, JBD). (Modified from Akol et al., 2022). (Right): example structure of the FHD domain, showing helical (H) sections and winged (W) sections. (Modified from Dai et al., 2020).

Both the telencephalic emergence at the rostral end of the neural tube and its subsequent Foxg1-mediated expansion have been suggested to be key determinants of higher-order information processing abilities peculiar to the vertebrate brain (Kumamoto and Hanashima, 2017). This transcriptional factor possesses a spectrum of multiple distinctive roles in the developing rostral telencephalon. Over the past three decades, numerous efforts have been made to advance *Foxg1* functional characterization with a particular focus on stem cell proliferation, dorsoventral cortex morphogenesis and patterning, cell cycle control, induction of lineage fate choices, regulation of neocortical cell differentiation, migration, and pathogenic implications of these pathways caused by mutations in the gene itself (Brancaccio et al., 2010). To date, among all the 44 defined forkhead family genes described to date (Hannenhalli and Kaestner, 2009), *Foxg1* holds relevance because it is the only TF that plays a key set of non-redundant roles in vertebrate brain development. Constitutively defective Foxg1 expression causes major impairment of telencephalic morphogenesis, including agenesy of subpallium and microcephaly (Xuan et al., 1995; Manuel et al., 2010). In the cortical domain, *Foxg1*$^{-/-}$ null genotype results in defective specification of neocortical and paleocortical anlagen and exaggerated production of Cajal-Retzius (CR) neurons. On the contrary, Foxg1 overexpression suppresses CR cell fate. (Hanashima et al., 2004; Muzio and Mallamaci, 2005).

1.5. Foxg1 in neural precursors expansion

Around day E8.0-8.5 of mouse development, *Foxg1* begins to be expressed in the anterior border of the neural plate. It is induced by rostral secretion of *Fgf8* after which *Foxg1* expression spreads on a gradient in the caudal direction according to a ventro-anterior to low dorso-rostral gradient. (Shimamura and Rubenstein, 1997).

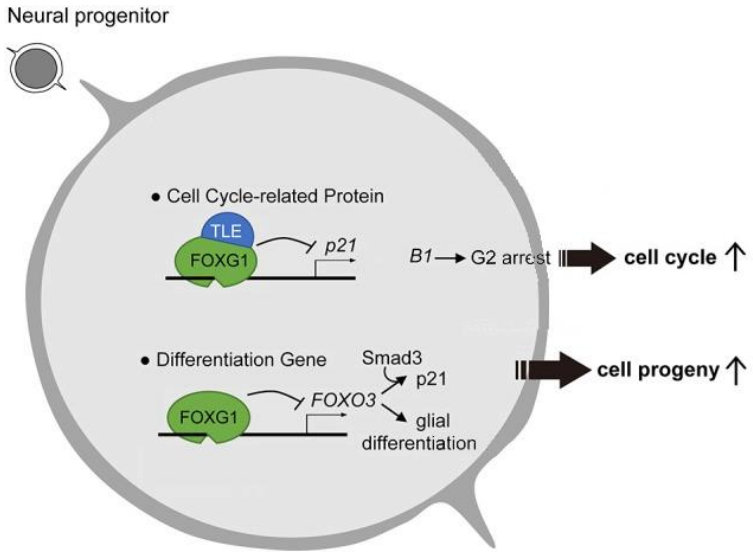


Figure I6: Roles of Forkhead Box G1 (Foxg1) in progenitor cell maintenance and suppression of differentiation. To prevent cell cycle exit, Foxg1 cooperates with Groucho/Transducin-like Split Enhancer (TLE) proteins to inhibit key cell cycle protein expression. Foxg1 also inhibits gene differentiation and DNA/histone modifiers to prevent neuronal differentiation. (Modified from Hou et al., 2020).

By the time of its transcriptional activation, *Foxg1* plays a key role in cell cycle control, regulating of the balance between proliferation and differentiation states of neural precursors. Indeed, it acts at an early stage to inhibit premature differentiation by means of cell cycle regulation (Xuan et al., 1995; Martynoga et al., 2005). *Foxg1* is expressed in the proliferating stem cells comparment of the neuroepithelium during neocortical development. (Brancaccio et al., 2010; Falcone et al., 2019; Xuan et al., 1995). Regarding the control of neuronal precursor kinetics, Foxg1 normal dosage is rigorously required to allow a correct progression of the cell cycle. An increase in Foxg1 expression levels generally promotes neural precursors self-renewal and inhibits their differentiation to post-mitotic neurons. Through the interaction with TLE proteins, Groucho (paragraph 1.4 and 1.7) is responsible for the inhibition cell cycle progression in neural precursors (Figure I6) (Falcone et al., 2019; Martynoga et al., 2005; Brancaccio et al., 2010). Regarding the significant brain volume loss in the hemizygous condition, it has been highlighted to date, that a lowering of Foxg1 levels is followed by a decrease in cells that express *Eomesodermin* (*Eomes/Tbr2*) and *ki67*. In such case there is also an increase in Cdk1a at early stages of corticogenesis (E13.5-E16.5 in mice) followed by an increase in cell cycle length (~E17. 5) toward the end of corticogenesis (Siegenthaler et al., 2008). Hou and collaborators proposed that since Cdk1a is also modulated by

the TGFß transduction pathway, Foxg1 might also cross-antagonize the TGFß transduction pathway (Hou et al., 2020). This finding was also corroborated by the fact that Fgf8 decreases its transcript levels in Foxg1 deficient embryos (Kumamoto and Hanashima, 2017). In both cases, the resulting phenotypic consequence would be a reduction in the population of Tbr2-positive progenitors (Hou et al., 2020).

On the other hand, Foxg1-mediated cell cycle regulation is also focused on its collateral roles in antagonizing key components of the cell cycle pathway such as FOXO/SMAD (Sma and Mad related protein) a complex that promotes the precursors cell-fate and differentiation towards cortical post-mitotic neurons (Vezzali et al., 2016; Seoane et al., 2004). This fine mechanism involving Foxg1 appears to be both DNA-binding-dependent and DNA-binding-independent (Hanashima et al., 2002). Indeed, it is achieved through both Foxg1 direct competition for the consensus sequence to DNA, and further by direct interaction with the FOXO/SMAD complex. This bimodal mechanism of action causes a lowering of Cdkn1a/p21 kinase inhibitor expression levels neural precursor cells to leave the cell-cycle (Figure I6). As a secondary consequence, this mechanism promotes an increase in the precursors compartment that keeps expressing *Foxg1* at consistent levels (Vezzali et al., 2016).

1.6. Foxg1 in the differentiation of pyramidal neurons

Migration of newborn glutamatergic neurons from periventricular proliferative layers to cortical plate is a complex process, including a number of stereotyped steps associated to specific changes in neuronal morphology. These are: primary somal translocation within proliferative layers; early, short bipolar migration along radial glia; multipolar tangential migration within the intermediate zone; secondary, long bipolar migration along radial glia; and final somal translocation, just prior to final neuronal settling within the cortical plate (Kon et al., 2017). Miyhoshi and Fishell observed that, during the articulation of these phases, *Foxg1* is subject to relevant changes in its expression dynamics (Miyhoshi and Fishell, 2012). They found that, at the onset of the multipolar phase, within the inner intermediate zone (E14.5 in mice), *Foxg1*-mRNA levels decrease significantly (Figure I7), and this is instrumental in transient derepression of two key genes inhibited by Foxg1, NeuroD1 and Unc5d. The authors showed that this phase of Foxg1 dowregulation is critical and indispensable for the transition to the late multipolar phase to occur. Indeed, if Foxg1 expression does not decline, then pyramidal progenitors cannot enter the late multipolar phase and remain transiently confined to the lower intermediate layer, delaying their entry into the cortical plate as well as the acquisition of laminar identity (Figure I7). (Miyhoshi and Fishell, 2012). Subsequently, the exit of the cells from the multipolar stage and their subsequent entry into the cortical plate is accompanied by an increase in *Foxg1* expression levels, which - from this stage onward - exhibit a rising trend. If such Foxg1 upregulation does not occur, neurons regress to the early multipolar phase and definitively fail to enter the cortical plate.

Therefore, the importance of the correct dosage of this gene is reflected through the timing and duration of its dynamic expression that allows partially committed progenitors to arrive to populate the cortical plate and insert themselves in the patterns of functional circuits proper to this neocortical layer. High levels of Foxg1 detectable in post-mitotic glutamatergic pyramidal neurons have been suggested to be crucial for their proper differentiation and function. In this regard, Foxg1 has recently been described to promote elongation and arborization of neocortical pyramidal dendrites, via dedicated molecular cascades including Hes1 and pCreb1 (Chiola et al., 2018).

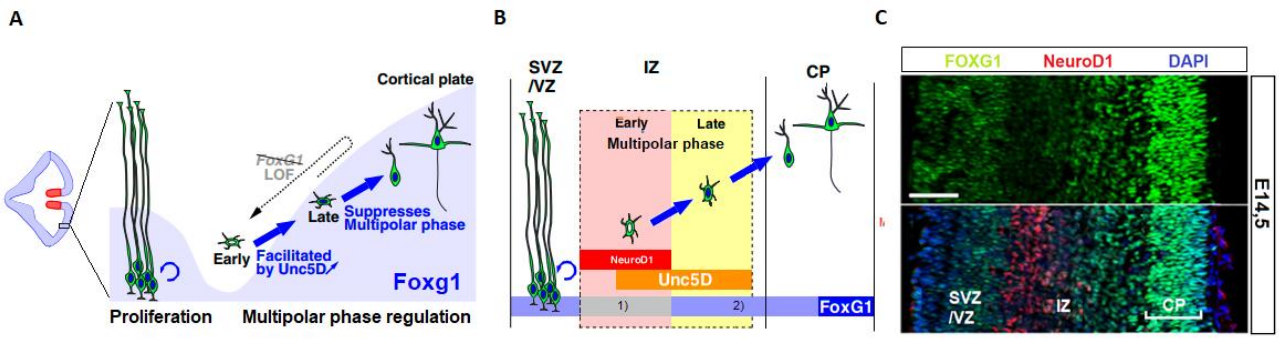


Figure I7: ***Foxg1*** **is regulated in a highly dynamic way during the transition of the multipolar phase and coordinates the development of pyramidal neurons**. (A) Graphical representation that summarizes Foxg1 expression dynamics along the neuronal glutamatergic lineage. Foxg1 time-dependent roles and functions are highlighted by dark blue arrows. Peculiarly at the level of the intermediate zone (IZ) Foxg1 encompasses a transient downregulation after which its expression is subsequently raised towards the end of the multipolar phase. Foxg1 levels depletion at this phase delays the stratification of neurons that reach and form the CP. (B) *Foxg1* temporal downregulation is necessary to allow *Unc5D* expression which increases the transition towards late multipolar phase (highlighted in yellow). (C) E14,5 cortical section comparing the expression of genes in precursors, progenitors (Neurod1 positive cells) and early pyramidal neurons as they transit from the VZ to the CP. *Foxg1* expression is downregulated in cells in the moment of NeuroD1 expression. (Modified from Myoshi and Fishell, 2012).

In addition to migration and morphological dynamics of pyramidal neurons, neocortical development also includes differential laminar specification of glutamatergic neurons belonging to different layers, including Cajal-Retzius pioneer neurons, deep layer neurons (DL) and upper layer neurons (UL). Even in such context *Foxg1* plays a cardinal role (Huh et al., 1999). As previously mentioned, *Foxg1* is induced by *Fgf8* expressed in the anterior neural ridge and subsequently expands caudally, therefore, the onset of Foxg1 expression represses several TF in an opposite rostro-caudal gradient. This allows Foxg1 and Lhx2 to bring about the cessation of CR cell production, thereby inducing progenitors to transition from CR cells to DL neuron identity in a spatiotemporally dependent manner (Hwang et al., 2009). At the same time, *Foxg1* promotes the onset of Ctip2 and *Fezf2* expression, both associated with deep layer (DL) determination, by downregulating *Tbr1*. Furthermore, it has recently been shown that high levels of *Foxg1* in the lower granule layer inhibit *Couptfl* transcription, resulting in further promotion of layer V differentiation programs, at the expense of layer IV ones (Hou et al., 2019). In this way, Foxg1 exerts a fine, multi-

level distributed regulation that orchestrates the proportions of progenitors that will undertake upper and lower layers differentiation programs.

To note, within differentiating neurons, Foxg1 is not exclusively involved in transcriptional control. It synergizes with the DDX5 helicase in promoting pre-miR-200 maturation, so eliciting a downregulation of cAMP-dependent protein kinase type II-beta regulatory subunit (PRKAR2B) and a derepression of postsynaptic function (Weise et al., 2019). Foxg1 is also detectable in the cytosol, where it has been recently shown to assist the translation of mRNA encoding for voltage-dependent membrane channels (Artimagnella et al., unpublished results). Last but not least, Foxg1 polypeptides (the full one or fragments of it) can be found in mitochondria, where they modulate membrane potential and organelle fission, ultimately impacting early neuronal differentiation (Pancrazi et al., 2015).

1.7. Foxg1 as a transcriptional repressor

Concerning mechanisms mediating Foxg1 impact on transcription, it has been shown to often act in association with TLE, which stands for Transducin/Groucho-like Enhancer of split, that contains a C-terminal WD40 repeat domain possessing affinity for Foxg1 and forms with Foxg1 a transcriptional repressor complex (Yao et al., 2001; Marçal et al., 2005). Interestingly, GRG6, namely Gro/TLE-Related Gene Product 6 family, is also provided with a WD40 domain and, by means of such domain, it prevents Foxg1 and TLE1 from physically interacting, so antagonizing the transcriptional inhibition imparted by Foxg1/TLE. Hence, GRG6 and TLE1 associate with Foxg1 and possess opposing roles during stem cell proliferation. Indeed, whilst TLE1 / Foxg1 overexpression is responsible to keep stem cells in a proliferative state, GRG6 / Foxg1 determine both proneural and differentiative functions (Marçal et al., 2005).

An analysis of Foxg1 genomic targets in neuronal NS cell lines highlighted a total of 6897 significant binding sites (Bulstrode et al., 2017). A subsequent motif-analysis revealed novel Foxg1 interacting genes that are essential in cell-fate and in the determination of the transcriptional programs of neural differentiating stem cells (Mateo et al., 2015). These sites were strongly enriched in genes with a histone-methyltransferase function. Consequently, *Foxg1* overexpression has been demonstrated to regulate key proteins that act as methylome-remodelers and facilitate the de-differentiation toward a more stem-prone cellular identity (Bulstrode et al., 2017).

Recent studies by Falcone and collegues, have shown that both in mice and humans, *Foxg1* overexpression in the neocortical neuronal precursor compartment induces a strong repression on astrogenic output (Figure I8). Specifically, Foxg1 has been observed to promote neuronal precursor

self-renewal and neurogenesis (Fig I6). Foxg1-mediated astrogenic inhibition activity can occur by multiple mechanisms.

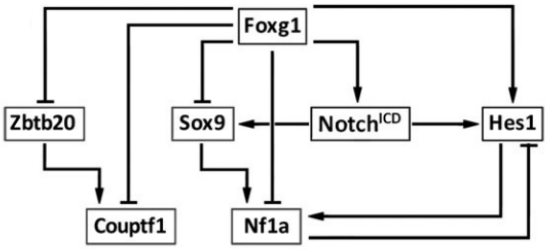


Figure I8: graphic map highlights Foxg1 transcriptional control of astrogenic genes. Foxg1 impact on early effectors endowing NSCs with astrogenic competence.

Foxg1 acts through repression in trans regulating TF genes that promote neuronal precursor to astrocyte differentiation (*Couptf1*, *Zbtb20*, *Nfia* and *Sox9*). Second, by direct interaction with chromatin, astroglial genes that implement the astroglial differentiation program (*Gfap*, *S100β*, *Aqp4*) are also subject to binding enrichment for Foxg1 at promoters. Through transcriptional profiling, the study highlighted the ability by Foxg1 to inhibit the trans-activating capacity of the pStat3-pSmad 1,5,8 complex known to act as a strongly pro-astrogenic trans-activator (Falcone et al., 2019; Lee et al., 2016). Moreover, *Foxg1* levels in neocortical precursors progressively decrease before the neuronogenic-gliogenic transition, suggesting an involvement of Foxg1 in the fine temporal synchronization of astrogenesis rates (Falcone et al., 2019).

1.8. *FOXG1* and neurological disorders

Foxg1 syndrome is a complex heterogeneous pathological entity, characterized by multiple neurodevelopmental alterations and neurocognitive aberrations. In particular, *FOXG1* turns out to be a highly dose-sensitive gene: its deficiency, as well as its excess in fact, severely affect normal brain development and function. Clinical phenotypes originate from production of a structurally abnormal Foxg1 protein and/or an altered dosage of it, resulting in an extremely complex spectrum of physical and neurological symptoms (Florian et al., 2011).

Rett syndrome (RTT) is a rare neurodevelopmental disorder, predominantly affecting females, resulting from mutations in the methyl-CpG binding protein 2 (*MeCP2*) gene located on the Xq28 locus in the majority of individuals affected by the disorder. The Rett syndrome, is characterized by profound cognitive impairment, poor communication skills, stereotyped hand movements and reduced growth from 6 to 18 months after a period of apparently normal development, including the acquisition of motor skills and language. In a small percentage of cases, estimated at around 10-15%, no *MeCP2* mutation can be found in developmental patients suffering from Rett syndrome. Precisely because of this reason, it has been suspected for some time that in these atypical forms of the disease, there is the involvement of other genes. In recent years, in this

regard, it has been discovered that some variants of Rett syndrome are caused by mutations in *FOXG1* (Fabio et al., 2014; Nomura, 2005). In humans, the first mutation related to *FOXG1* was described in a 7-year-old boy who presented the translocation T (2; 14) (P22; Q12), in which the rupture points of the DSB on the chromosome 14 was internal to the *FOXG1* CDS (Shoichet et al., 2005). To date, it has been reported that not only translocations, but also deletions at 14q12 locus, as well as the inactivating mutations that have the final effect of resulting in Foxg1 haploinsufficiency, were initially associated with the Rett syndrome. In this case, since both conditions lead to microcephaly, crisis, convulsive movement, sleep disorders, complete agenesis of the Corpus Callosum and poor cognitive skills, was initially chosen to make *FOXG1* clinical spectrum coincide with the same syndrome (Le Guen et al., 2011; Mencarelli et al., 2010). Although MeCP2 and *FOXG1* syndromes share phenotypic similarities in the clinical spectrum, the remarkable characteristics have distinguished the most common mutation phenotype Mecp2 congenital with the rarer *FOXG1* variant. Although the disease caused by *FOXG1* is less incident, there are strong differences between the two genes comparing ambulation, receptive language and sleep duration, with *FOXG1* subjects showing more serious disabilities (Chen et al., 2019). Therefore, the significant phenotypic differences between neurological disorder caused by *FOXG1* loss of function mutations and the Rett syndrome associated with *MeCP2* have aimed clinicians to designate Foxg1 syndrome as a distinct disorder. (Florian et al., 2012).

Conversely, even a small increase in Foxg1 dosage leads to neurological disorders. In this regard, West syndrome (WS) is a rare form of epilepsy characterized by infantile-onset flexor and extensor spasms, a high-amplitude EEG pattern with asynchronous spike activity called hypsarrhythmia associated with impaired psychomotor development (Striano et al., 2011). The disease can have several causes: in most cases, it depends on an anatomical/structural malformation of the brain, on genetic abnormalities such as aneuploidies trisomy 21 translocations or even gene mutations. In 10-20% of cases, the syndrome is cryptogenic, probably associated with abnormalities that have not yet been definitively identified. (Wong et al., 2019). In 2011, Brunetti-Pierri and colleagues, published a report of seven patients with duplications of the 14q12 region that included *FOXG1*. These patients had relatively severe neurodevelopmental phenotypes including intellectual disability, epilepsy, and severe language delay (Brunetti-Pierri et al., 2010). Previously, patients with West syndrome have been reported to possess duplications that included *FOXG1*. Along the same lines, results published by Bertossi and colleagues in 2013 show that patients carrying 14q12 duplications developed seizures, in the first months of life. Most of them developed infantile spasms measured by EEG. (Bertossi and al., 2013). The results of analyses of these patients insist that duplication of *FOXG1* 14q12, associated with a consequent increase in *FOXG1* dosage, may be instrumental in the origin of infantile spasms and neurocognitive impairment in patients with West syndrome (Tohyama et al., 2011; Amor et al., 2012). Recently, Tigani and colleagues reported that mice overexpressing Foxg1 in neocortical pyramidal cells exhibited electroencephalography (EEG) with a higher spike rate and were more prone to kainic acid (KA)-induced seizures. Consistently, primary cultures of neocortical neurons gain of function for FOXG1 were hyperactive and hyper-synchronized. This reflected an imbalanced expression of key genes

encoding for ion channels, gamma aminobutyric acid, and glutamate receptors, and was likely exacerbated by a pronounced depletion of interneurons. Based on this, the authors describe how small changes in *FOXG1* levels can result in a profound impact on pyramidal cell activity (Tigani et al., 2020).

1.9. Transposable Elements (TEs) in the mammalian genomes

For almost the entire 20th century, experimental biology has always considered genes as structurally stable entities defined according to a static linear model; ordered on chromosomes in the same way that beads are arranged adjacently and contiguously on a thread (Morgan TH, 1922). Gregor Mendel's contribution on inheritance in pea plants was rediscovered early in the last century, and since then, many researchers have begun to use maize and fruit flies to further classical genetics. During this period, Thomas Hunt Morgan together with his group, developed significant studies on the genetics of Drosophila melanogaster, while in 1913, Alfred Sturtevant produced the first genetic map of a chromosome (Sturtevant AH, 1913). In this same environment, Barbara McClintock made her first steps on the study of maize. By the end of 1940, she had spent many years studying the characteristics of genes when she discovered that some of them could be mobile. Her first methods of investigation in maize genetics consisted of a narrow spectrum of classical genetic techniques combined with microscopic analysis of the differences between maize chromosomes belonging to different colored caryopses. It was these studies that directed her to the discovery of a locus that interrupts the structural continuity of genetic material within some chromosomes and that has the ability to change its position within a chromosome itself. McClintock went on to discover other mobile elements that had such features, now known as transposons. She also discovered that depending on where they fit into a chromosome, these mobile elements could irreversibly alter the expression of other genes. The data she obtained on the first TEs were first published in the journal PNAS in 1950, "The origin and behavior of mutable loci in maize" (McClintock B, 1950). The veracity and true occurrence of these elements was initially considered by plant geneticists alone. The wide distribution of these mobile genetic elements, along with its implications, took decades to be widely validated and approved. Subsequently referred to as "jumping genes" they have been historically downgraded to the definition of "junk DNA" in mammalian genomes until particularly recent times (Orgel et al., 1980; Doolittle and Sapienza, 1980). The advent and modernization of molecular biology techniques, sequencing and molecular genetics have gradually led, since 1970, to the discovery of transposons also in microorganisms, starting from prokaryotes and viruses. To date it is known that at least 46% of human DNA, and 37% of mouse DNA are derived from TEs (Lander et al., 2001; Lindblad-Toh et al., 2005; Waterstone et al., 2002; McGrayne S.B, 1993). TEs are so well represented that they comprise most of the genetic material in some species. In particular, their extreme prevalence in all eukaryotes together with the increasing number of functional studies is revealing that TEs are biologically important components

in mammalian genomes. This has led to a marked reconsideration of the relevance of this unique discovery (Fedoroff N, 2012; Richardson et al., 2015).

1.10. Structural and Functional diversity of TEs

McClintock's findings also led to the elucidation of the ways in which TEs are able to bring about changes in the content of genetic information. First, it was possible to understand that TEs could play a relevant role in re-arranging the structure of genomes, while later a deeper knowledge allowed to realize that the positional information of the insertion site was decisive in producing a change in gene expression of sequences located in the molecular environment. For this reason, the presence and the enormous amount of TEs in eukaryotic genomes produces a strong impact on both their remodeling and evolution. To date, it is possible to divide eukaryotic TEs into two classes, based on the nature of the transposition intermediate, RNA (class 1) or DNA (class 2) (Figure I9). (Craig et al., 2002). Classical DNA transposons originally discovered by Barbara McClintock, move through a DNA Intermediate (McClintock, 1950; Fedoroff et al., 1983). The predominant class of TEs in most mammalian genomes, are mobilized via an RNA intermediate by a process called retro-transposition (Boeke et al., 1985). The transcript encoded by class 1 elements (mRNAs), constitutes the transposition intermediate (Figure I9). In contrast, in the case of DNA TEs, it is the transposon itself that relocates itself from one point to another in the genome by a "cut and paste" mechanism (Wessler, 2006).

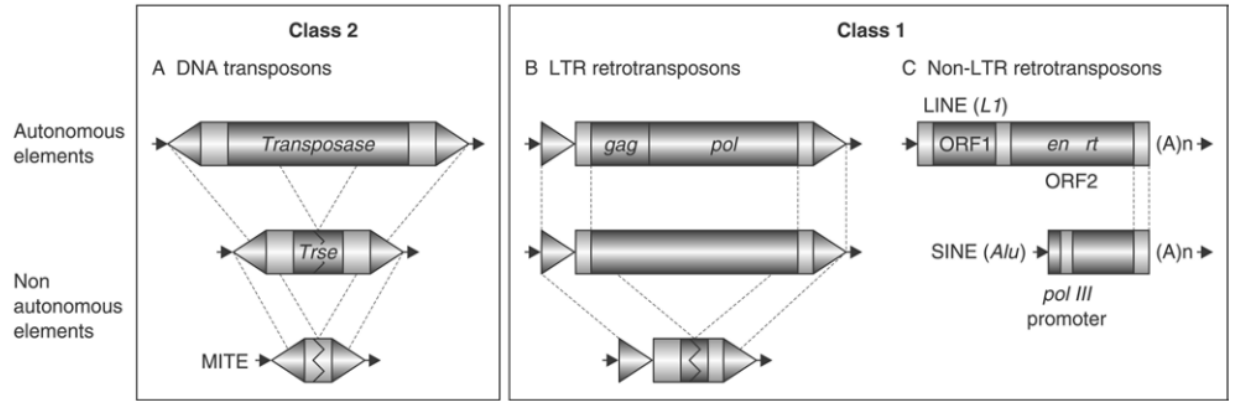


Figure I9: Characteristics and classification of eukaryotic transposable elements (TEs). TEs are separated in two classes. This depends on whether the molecular intermediate of the transpositional mechanism is formed by RNA: Class 1) or DNA: Class 2). (Modified from Wessler, 2006).

Each of the two classes consists of either autonomous or nonautonomous sequences. Non-autonomous transposons do not encode for key enzymes required for transposition, and their ability to transpose originate from the fact that they retain cis-active sequences implicated in transposition. As for autonomous transposons, they possess one or more ORFs that encode for the enzymatic apparatus required for transposase activity (Kuppers and Dalla-Favera, 2001). Following the integration of the vast majority of TEs, a small sequence (namely TSD, or Target Site Duplication) is found duplicated at the point where the transposon insertion into the new genomic region occurs. Class 2 transposons are characterized by a small-inverted repeat sequence in the terminal region, named "Terminal Inverted Transposition" (TIR) (covering a distance between 10 and 200 bp) as well as a sequence encoding an enzyme that catalyzes and enables the transposition mechanism. The enzyme itself, defined transposase binds in a selective way the consensus positioned at the terminals 5' and 3' respectively of its coding sequence as well as at the ends of the non-autonomous elements. At this level, the enzyme catalyzes the "ab-initio" detachment of the transposon from its original location and its reinsertion into a new chromosomal site (Makałowski et al., 2019). Based on the mechanism by which this process takes place, class 1 TEs are classified in LTR retro-transposons and non-LTR retro-transposons.

LTR retrotransposons possess two identical sequences that are positioned at the ends (LTR, standing for long terminar repeats) and can vary in size ranging from approximately 100 bp to several kilobases. If the retrotransposon is autonomous, it also contains two genes, called pol and gag. The former encodes for a multifunctional enzyme, which sequentially catalyzes protease activity; RNA-dependent DNA polymerase activity, ribonuclease H (RNase H) and integrase activity. Instead, gag encodes for a protein phylogenetically similar to the viral capsid. Transcription begins at the 5’ LTR and ends near the 3’ LTR originating a transcript that migrates into the cytoplasm (Lander et al., 2001). Since RNA plays a bimodal role at this level, the coding sequences contained in it are translated, while the double-stranded cDNA obtained as a result of catalysis by the pol protein is transported into the nucleus where it can then integrate into a new site in the genome. From the evolutionary perspective, the cumulative increase in genome size given by the summation of transposition events can be compensated through homologous recombination between identical LTRs that belong to whole elements (Skalka, 2014).

Non-LTR retrotransposons do not possess LTR elements (Figure I9). They can be classified in autonomous long interspersed elements (LINEs) and non-autonomous short-interspersed elements (SINEs). LINEs are structurally composed by two ORFs which originate a single bi-cistronic mRNA transcript. It is composed of two open reading frames: the first encoding for ORF1 (an RNA-binding protein with chaperone-like activity) and an ORF2 (endonuclease and reverse transcriptase activity). Both LINE and SINE mRNAs harbor a poly-adenylated tail at the 3' terminal (poly-A). Whereas each transcription event in a LINE begins by RNA pol II at the 5’ end of the element and ends downstream of its poly-A repeat sequence, in SINEs we find a small internal promoter near the 5’ end that is transcribed by RNA pol III. Unlike LINEs, SINEs constitute a highly diverse set of sequences that can measure from 90 bp to approximately 300 bp and are evolutionarily related to 7SL RNA as well as tRNA genes (Richardson et al., 2015).

1.11. Co-evolution between TEs and host genomes

Comparative study of genomic sequences between phylogenetically distinct groups shows that the TE occurrence changes significantly from species to species. In this respect, there are differences in cumulative distribution of different TEs-types present in the genome and in levels oftransposase activity. A particularly well-known case is represented by the yeast Saccharomyces cerevisiae, which possesses only LTR-type retrotransposons. On the other hand, it has been found that while in mammalian genomes class 2 DNA elements are almost negligible, non-LTR class 1 retrotransposons are predominant. On the other hand, if we evaluate the genomic activity of TEs, we can see that it varies from species to species. Existing mammalian genomes have few active TEs, and despite this, their activity is strongly characteristic across species (Capy et al., 2000; Seidl and Thomma, 2017). During biological evolution, TEs, together with host regulatory factors, have coexisted and coevolved together. As a result, a network of novel molecular interactions between endogenous host genome regulatory factors and novel heterologous naive transposon sequences has arisen over time (Aravin et al., 2007; McDonald JF, 1998). This allowed the host genome to partially counteract the negative effects of transposition over time, while moderately allowing it to replicate. This level of molecular control exerted by the host is maintained due to selective pressure to maintain and preserve genomic integrity for protective purposes. Among key mechanisms controlling TEs, there is the transcriptional dependence of several transposable elements on endogenous regulatory factors characteristic of the host, such as TFs, poly(A) and splicing effectors (Perepelitsa-Belancio and Deininger, 2003) as well as transposase-mediated autoregulation (Laski et al., 1986; Schlappi et al., 1994). Although DNA mobilization events are mainly deleterious to species fitness, rare germline-fixed events in the evolutionary history of each species have originated beneficial functions for their host. In this regard, several transposons have acquired the function of gene regulatory sequences that, in turn, control the expression of already originated and functional genes. (Sasaki et al., 2008). Single examples of selectively beneficial TE insertions with apparent regulatory functions have been described in non-mammalian species, especially in *Drosophila melanogaster* (Maside et al., 2002; Schlenke et al., 2004; Chung et al., 2007). With respect to mammals, pioneering studies initiated by Silva and colleagues in 2003 compared human-mouse orthologous sequences and suggested that a substantial fraction of ancestral repeats (which were inserted before Eutherian radiation) were subject to strong functional selective constraints over a long period of evolutionary time (Silva et al., 2003). Overall, it can be assumed that at the expense of obvious risks lined to transposition, such process has been converted into an evolutionary resource, allowing to improve fitness (exaptation) and sustain evolution of cellular physiology (Gould and Vrba, 1982).

1.12. Evolution of *LINEs* from deuterostomes to mammals

Through the process of biological evolution, mobile elements are still under constant competition for the presence of host cell factors that may exert a secondary regulatory effect on their activity in the cellular context. Several authors agree that they result in short, sometimes replicationally independent, retroviral-derived genomes whose evolutionary success consists in maintaining the ability to spread across the genome and increase copy number by occupying new genomic loci (Sotero-Caio et al., 2017; Chalopin et al., 2015; Shao et al., 2019). LINE class retrotransposons are in genomes of animals and plants. Given the subject of this thesis, here we will mainly discuss these TEs in animal contexts (Ohshima, 2012; Ivancevic, 2016). Although the presence of LINEs has been documented in some genera of invertebrate metazoans, in which the mouth is derived from the embryonic blastopore (protostomes), tracking expansion and quantitative accumulation of new genomic copies can be found starting with deuterostomes, where the anterior digestive tract forms far from the blastopore. In the most primitive deuterostomes,such as tunicates and echinoderms, LINEs constitute a very small fraction of repeated elements (Kordis et al., 2006; Chalopin et al., 2015). Conversely, in known representatives of agnates and chondrichthyes, the relative genomic abundance of LINEs is no less than that of other DNA-TEs and corresponds to approximately 25-50% of all repeated elements (Metcalfe et al., 2012; Meyer et al., 2021). Based on molecular phylogeny, just in these two classes of chordates a relevant expansion of TE elements occurred, originating a pool of LINEs families (Smith et al., 2013; Venkatesh et al., 2014). Next, while in reptiles there was an increase in the number of emerging copies of the *CR1*, *BovB*, and *L2* families, in avian genomes, a sharp reduction in genome size took place with LINE/CR1 being the only remaining family (Kapusta and Suh, 2017; Boissinot and Sookdeo, 2016; Boissinot et al., 2019). In mammalian genomes, there was an accumulation of mutations and truncations of full-length copies (a problem that will be addressed in paragraph 1.16; 1.17 and 1.17.1) leading to a decrease in the number of intact elements (Smith et al., 1995). In this class, the most successful group of mobile elements is *LINE1*, documented as early as 160 mya, that became the most successful and active group, while many ancient repeats gradually disappeared from genomes as they evolved. To note, a significant fraction of *LINE1* transcripts that underwent incomplete retro-transcription and truncation underwent more moderate selective pressures that allowed the molecular evolution of CREs (Cis Regulatory Elements). Several current enhancers, silencers, and insulators originated from ancient "abortive residues" of ORF2p-operated catalysis (Mikkelsen et al., 2007). As for intact *LINE1s*, these are species-specific genomic elements. Metatheria (marsupials) and eutheria (placentals) resemble each other in terms of sequence and tendency of their genomes to evolve in common patterns. Conversely, monotremes do not contain any LINE1 sequences (Ivancevic et al., 2016; Warren et al., 2008). Primates separated from other ancestral mammals approximately 90-65 million years ago (mya) and are characterized by the distribution of *LINE1* PA and *LINE1* PB families (Glazko and Nei, 2003; Konkel et al., 2010). Within primates, the great anthropoid ape branch separated approximately 26 mya (Castellano and Munk, 2020).

Compared with human, the number of *LINE1* insertions with intact or partial length in primates is greater. However, concerning the absolute number of currently active *LINE1s*, it is higher in the

ORFs (Lee et al., 2007). In contrast, in humans, the active family is *LINE1* HS. Moreover, comparative analysis showed that the activity of human *LINE1* copies is significantly higher than that of chimpanzees (Lee et al., 2019; Lee et al., 2007).

1.13. Evolutionary divergence of *LINE1* between mouse and human

Humans and rodents segregated approximately 65-75 mya. Over the past 25 million years, *LINE1* retro-transposition has declined significantly in human lineage but not in mouse lineage (Gibbs et al., 2004; Waterston et al., 2002). At the moment, LINE1 account for approximately 18% of the mouse genome and 17% of the human genome (Waterston et al., 2002). As discussed in detail in subsequent paragraphs 1.14.1; 1.14.2; 1.17 and 1.17.1, most of *LINE1* elements are inactive (Voliva et al., 1983; Grimaldi et al., 1984) and therefore cannot be expressed. In humans, the number of full-length elements is in the order of 6,000, however it is estimated that only 146 are potentially active and capable of retro-transposition. In contrast, in mice, of approximately 7,000 full-length copies, 3,000 copies are found to be active and competent to retro-transpose (Rosser and An., 2012). Since the primate-roditor split, human-specific *LINE1* subfamilies contribute to only 8% of the human genome compared with over 16% of mouse-specific *LINE1s* (Waterston et al., 2002). Based on sequence-related phylogeny studies, Sookdeo and coworkers show that the general pattern of mouse evolution is similar to that of humans, suggesting that the nature of interactions between *LINE1* and its host may be similar in these two species. (Sookdeo et al., 2013). Despite slight differences at 5', mouse *LINE1* are structurally similar to human *LINE1* (Myers et al., 2002), and many findings on *LINE1* structure and function come from studies conducted in mice. Based on this evidence, the mouse has so far served as an indispensable model for studying *LINE1* expression over the past 2.5 decades, and much of what we know today about developmental *LINE1* expression patterns in vivo comes from mouse studies. Specifically, endogenous murine *LINE1* is expressed at different stages of mouse development in a tightly regulated manner (Rosser and An., 2012).

1.14. *LINE1* structure

1.14.1. *LINE1* 5' UTR

Although the sequences of the 5'UTR of murine *LINE1* differ moderately from that of human *LINE1* (Figure 1), conservation of RNA polymerase II responsivity allowed *LINE1* to expand its copy number genome-wide (Swergold., 1990; Becker and al., 1993). The 5' UTR region functions as a promoter for LINE1 and contains cis consensus for several TFs (Yang et al., 2003; Tchenio et al.,

2000). Notably, functional and structural studies have highlighted a binding matrix for YY1 at the most proximal end of the 5' UTR; a TF that is critical and necessary for transcription initiation.

Similarly to pol II-dependent transcripts, *LINE1* mRNA has a cap at the very 5' end, formed by 7-methyl guanosine which facilitates translation (Dmitriev et al., 2007; Athanikar et al., 2004). In addition to containing a promoter that exerts its activity in (5'-3') direction, the sequence also includes a second promoter also regulated by pol II that acts in the opposite direction (3'-5') (AS-promoter) (Richardson et al., 2015). Whether in human *LINE1,* this sequence named ORF0 (Figure I10) can be identified as a short region interposed between the promoter at the 5'UTR and the ORF1p coding sequence, in murine *LINE1*, the antisense promoter is located within ORF1 (Li et al., 2014; Zemojtel et al., 2007). The presence of a bidirectional promoter has as a direct consequence, to determine the also antisense synthesis and generation of chimeric transcripts formed by *LINE1* and genomic sequences flanking the 5' end (Nigumann et al., 2002). Nevertheless, according to some authors, antisense promoters would be involved in originating fusion proteins or regulating *LINE1* retro-transposition efficiency (Naufer et al., 2018; Freeman et al., 2019).

Although the sequences of the 5'UTR of murine *LINE1* differ moderately from that of human *LINE1* (Figure 1), conservation of RNA polymerase II responsivity allowed *LINE1* to expand its copy number genome-wide (Swergold., 1990; Becker and al., 1993). The 5' UTR region functions as a promoter for LINE1 and contains cis consensus for several TFs (Yang et al., 2003; Tchenio et al., 2000). Notably, functional and structural studies have highlighted a binding matrix for YY1 at the most proximal end of the 5' UTR; a TF that is critical and necessary for transcription initiation.

Similarly to pol II-dependent transcripts, *LINE1* mRNA has a cap at the very 5' end, formed by 7-methyl guanosine which facilitates translation (Dmitriev et al., 2007; Athanikar et al., 2004). In addition to containing a promoter that exerts its activity in (5'-3') direction, the sequence also includes a second promoter also regulated by pol II that acts in the opposite direction (3'-5') (AS-promoter) (Richardson et al., 2015). Whether in human *LINE1,* this sequence named ORF0 (Figure I10) can be identified as a short region interposed between the promoter at the 5'UTR and the ORF1p coding sequence, in murine *LINE1*, the antisense promoter is located within ORF1 (Li et al., 2014; Zemojtel et al., 2007). The presence of a bidirectional promoter has as a direct consequence, to determine the also antisense synthesis and generation of chimeric transcripts formed by *LINE1* and genomic sequences flanking the 5' end (Nigumann et al., 2002). Nevertheless, according to some authors, antisense promoters would be involved in originating fusion proteins or regulating *LINE1* retro-transposition efficiency (Naufer et al., 2018; Freeman et al., 2019).

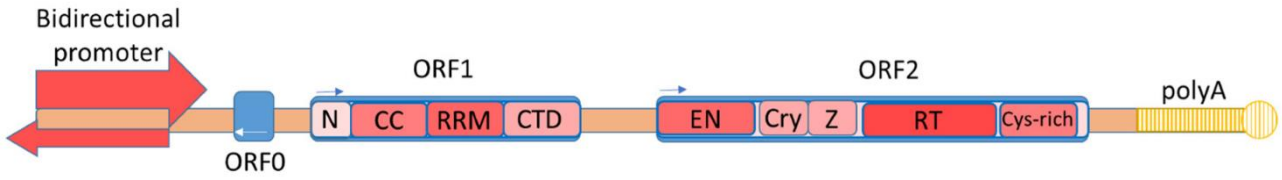


Figure I10: The structure of a full-length copy of *LINE1* retro-transposon. ORF1 consists of an N-terminal domain (N), a coiled-coil domain (CCD), an RNA recognition motive (RRM), and a C-terminal domain (CTD). ORF2 consists of endonuclease (EN), retro-transposase (RT), a cryptic domain (Cry), a Z-domain (Z), and a C-terminal domain with a cysteine-rich region (Cys-rich). (Protasova et al., 2021).

Although the function of these chimeric transcripts is not yet understood in a profound sense and requires further study, this unique feature to the *LINE1* 5'UTR occurs exclusively in elements that have conserved transposition competence and maintain the entire transposon sequence intact (Figure I10) (Macia et al., 2011; Luchetti et al., 2017). In this regard, Singer and colleagues, showed a particular condition related to fate committment during the differentiation progression of the glutamatergic neuronal lineage. In particular, by studying novel insertions, they clarified that the directional bivalence of the murine *LINE1* promoter can significantly affect the transcript levels of an adjacent gene, PSD93. This suggests that the transposase activity of *LINE1* might affect the mature clonal identities and neuronal output starting at the level of early neural precursors (Singer et al., 2010).

1.14.2. Murine *LINE1* subfamilies

Unlike human *LINE1* retrotransposons, the 5' UTR of mouse *LINE1* (*mLINE1*) contains promoter activity which is required for subfamily expression. Its own structure contains tandem repeats of ~200 bp in length called monomers. Based on the sequence and number of monomeric repeats in their respective 5' UTRs, it is possible to distinguish and classify five murine *LINE1* subfamilies (V, F, A, TF, and GF). As far as we know, V and F families are transposition-incompetent (D'Ambrosio et al, 1986), whereas A, Gf, and Tf subfamilies have to date retained the ability to retro-transpose, are still evolutionarily young, and are undergoing active expansion (Figure I11) (Rosser and An., 2012). A study by Zemojtel and coworkers reports that the subfamilies of mLINE1 that possess retro-transposase activity contain monomers of the same type (F-type) that distinguish the Gf and Tf families. In contrast, A-type monomers are characteristic and unique only to the A subfamily (Zemojtel et al., 2007). This means that the subcategorization of mLINE1 into subfamilies is based on both the nature and the repetition pattern of promoter monomers (Lander et al., 2001). The same Zemojtel et al. also reports that the GF subfamily has the longest average promoter size (~5.5 monomers), followed by the A families (~4.4 monomers) and last is TF (~4.1 monomers). As said in paragraph 1.13, ~3000 full-length elements are nowadays potentially capable of active retro-

transposition. Cell culture assays show that 900, 400 and 1800 of them are allotted to A, Gf and Tf subfamilies, respectively. (Schichman et al., 1993).

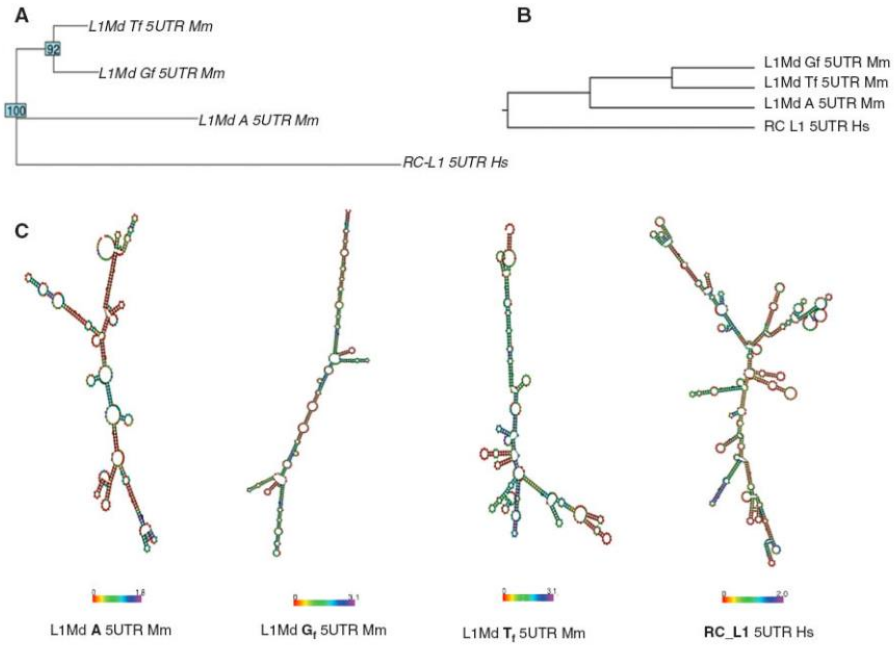


Figure I11: Phylogenetic cladistic comparison of sequences and secondary conformations between the 5'UTR of the three retro-transposition competent murine *mLINE1* subfamilies (A, Gf and Tf) compared to human. (A) Phylogenetic tree shows the evolutionary divergences between the different *mLINE1* subfamilies compared to human. Numbers on nodes indicate the percentage of the labeled node present in 93 replicates. (B) Multi-sequence alignment based on LocARNA. Tree height corresponds to highest or lowest scores that are indicative and proportional to lengths. (C) Secondary RNA folding patterns of *LINE1*. In this case, the color code corresponds to the amount of local entropy. Md A 5UTR Mm, *LINE1* subfamily of mouse A; L1Md Gf 5UTR Mm, *LINE1* subfamily of mouse Gf; L1Md Tf 5UTR Mm, *LINE1* subfamily of mouse Tf; RC_L1 5UTR Hs, human *LINE1* competent for retro-transposition (Modified from Bodak et al., 2014).

1.14.3. *LINE1* ORF1p

The protein encoded by Open Reading Frame 1 (ORF1p) is translated from *LINE1* polycistronic mRNA by the presence of the traditional cap (Dmitriev et al., 2007). Tt has a weight of ~40 kDa and is referred to as p40 in humans. Crystallographic and functional characterization of this protein has revealed that the N-terminus of ORF1p contains a coiled-coil domain, which facilitates trimerization of ORF1p molecules Martin et al., 2003; Callahan et al., 2012). The intermediate region of ORF1p contains an RNA recognition motif (RRM) that, with assistance from its C-terminal domain (CTD), is required for recognition of *LINE1*-mRNA (Khazina and Weichenrieder, 2009; Januszyk et al., 2007). In particular, missense mutations affecting highly conserved amino acid residues in both RRM and CTD abolish or negatively affect *LINE1* retro-transposition in cultured cells (Khazina and

Weichenrieder, 2009). Initial mappings regarding the cellular localization of this protein showed that in both species, ORF1p is found in the form of cytoplasmic ribonucleoprotein aggregates and physically interacts with single-stranded RNA without relying on consensus recognition (Kolosha and Martin, 2003). Therefore, ORF1p turns out to be required for the formation of the RNP-LINE1 mRNA complex (Kulpa and Moran, 2005). Interestingly, subsequent studies have provided insight into the fact that mRNA-associated RNPs complexes accumulate in cytoplasmic foci that appear electronically dense by electron microscopy and are closely associated with aggregates of stress-response proteins (Goodier et al., 2007). ORF1p in both mammals and nonmammals are characterized by mRNA chaperone activity (Evans et al., 2011). Although deletion of ORF1 does not abolish retro-transposition in cultured human cells (as it happens following ORF2p deletion) it has been hypothesized that ORF1p helps *LINE1* integration in vivo (Kajikawa et al., 2012).

1.14.4. *LINE1* ORF2p

The protein encoded by Open Reading Frame 2 (ORF2p) has a weight of ~150 kDa (Goodier et al., 2010; Taylor et al., 2013) and constitutes a multifunctional enzyme. In fact, it presents bimodal catalytic activity, thanks to the presence of two catalytic domains, both as endonuclease (EN) and reverse retro-transcriptase (RT), each of which is considered "*condicio sine qua non*" for the purpose of the fulfillment of each individual retro-transposition event. The endonuclease (EN) domain is located near the N-terminus of the protein and retains partial sequence homology with traditional apurinic and apiridinic endonucleases (APE) (Richardson et al., 2015). Recent studies highlight that EN operates a single endonucleolytic cut in a degenerated consensus sequence (5'-TTTT/AA-3'). Structural biochemistry studies highlight that this catalytic domain recognizes an inverted extra-helical adenine residue at the 3' end of the bond that is clipped to achieve the enzymatic cut (Weichenrieder et al., 2004). In contrast, the domain that catalyzes retro-transcription (RT) is located downstream of the EN domain in ORF2p and displays significant sequence homology when compared with RTs encoded by other non-LTR retro-transposons, LTR retro-transposons, and even retroviruses (Evgen'ev and Arkhipova., 2005). Importantly, *ORF2p shows preferential recognition in cis for its specific mRNA.* Although mutations in ORF1 and the EN domain may play a detrimental role on the entire retro-transposition mechanism, catalysis operated by the RT domain is maintained (Kulpa and Moran., 2006). Immunofluorescence analyses revealed that ORF2p co-localizes with both ORF1p and LINE1 mRNA in the cytoplasm (Taylor et al., 2013). If we look at the stoichiometric ratio and the ratio at which the interaction occurs in puncta, ORF1p appears to be particularly enriched within the complexes formed by RNPs and *LINE1* compared to ORF2p. However, the precise chain of events taking place from translation to retro-transposition is still unknown. Near the C-terminus, ORF2p possesses a domain strongly enriched in cysteine residues (C-domain). As evidence of the strong selective pressure acting on this sequence, it has been observed that the in vitro attempt to selectively replace cysteine residues with serines prevents the formation of the RNP complex and

strongly represses retro-transposition (Doucet et al., 2010). Although this observation has been verified in vitro, spontaneous mutations in the C domain do not result in a negative effect on retro-transcription. The exact function of this particular domain will have to be further investigated with the aim of elucidating its role in *LINE1* retro-transposition. (Piskareva et al., 2013). Finally, translation of human and mouse ORF2p may be initiated through different mechanisms. Although both murine and human ORF2 can be translated independently from AUG, in mouse (but not in humans), an Internal Ribosomal Entry Site (IRES) in the vicinity of the 3' region of ORF1, co-adjuvates the translation of ORF2. Based on this evolutionarily divergent trait, it is hypothesized that *LINE1s* have progressively adapted their sequence to the host cis-regulatory apparatus in order to employ translation mechanisms that evolved independently under selective pressures characteristic of different orders of host organisms (Alisch et al., 2006).

1.14.5. *LINE1* 3' UTR

The 3' UTR of *LINE1* measures ~206 bp in length and consists of a short-conserved sequence made up of multiple purine residues giving rise to a conformer termed G-quadruplex. The biology associated with this structural motif is not understood to date, however it is known that in vitro, although its removal can inhibit the catalytic activity of the RT domain of ORF2p, it does not impair the entire mechanism of retro-transposition (Richardson et al., 2015). Furthermore, the 3'UTR has a sequence of polyadenylated residues at the 3' terminal. At this level, polyadenylate-binding protein 1 (PABPC1) is critical for the formation of the *LINE1*-RNP complex as it attaches to the polyA tail of *LINE1* (Dai et al., 2012). Experimental verifications assessed that the signal from this poly(A) is often irrelevant and allows pol II to continue transcription in the 5'-3' direction until it separates from the template once it encounters polyadenylation sites of genes flanking the retrotransposon. The presence of a weak polyadenylation signal promotes the origination of chimeric transcripts not exclusively at 5' of *LINE1*, but also at 3'. A recent series of studies are raising new considerations regarding the 3'UTR of *LINE1* since both in human and mouse it appears to retain traces of promoter activity resulting in *LINE1* transcripts in different tissue types (Faulkner et al., 2009).

1.15. *LINE1* retro-transposition mechanism

The process of *LINE1* retro-transposition takes place through a "copy and paste" process named Target-Site Primed Reverse Transcription (TPRT; Figure I12) by which it amplifies itself and redistributes new copies throughout the genome. This is a mechanism originally discovered and characterized using the larval phenophase of the heterocercs *Bombyx mori* as a model study (Luan et al., 1993). Initially, the replicative cycle involves a new copy of the bicistronic RNA of the entire

LINE1 being transcribed in the presence of its strong promoter located in the 5'UTR. Next, *LINE1* mRNA is polyadenylated before migrating from the nucleus to the cytoplasm. It is at this level that ORF1p and ORF2p are translated and associate with *LINE1* mRNA forming the ribonucleoprotein complex (*LINE1*-RNP) previously reported (Figure I12) (see paragraphs 1.14.3 and 1.14.4) (Dai et al., 2012; Athanikar et al., 2004). To date, the detailed mechanism by which mRNA associated RNP complexes can return to the nucleoplasm is still not fully understood. What has been understood so far is that *LINE1* interacts with "endosomal selection complex with transport function" (ESCRT), and this interaction facilitates *LINE1* replication. Loss of the *LINE1*/ESCRT interaction does not impair RNP formation or enzymatic activity, but it may jeopardize RNP entrance in the nucleus, leading to reduced *LINE1* endonuclease activity in the nucleus and loss of retro-transposition. In addition to this, it is shown that the ORF1 protein also plays a role in this regard, being in fact able to interact with the karyopherins KPNA2 and KPNB1, as well as with other KPNA family proteins involved in the import of proteins through the pores of the nucleus (Horn et al., 2017; Freeman et al., 2019). In 2001, employing a retro-transposition assay in cultured human cells, Wei and coworkers showed that ORF1p and ORF2p possess selective specificity for interaction with the RNA encoding them. This mechanism may support "selfish retro-transposition", contributing to the fact that *LINE1s* remain transposition competent even despite the presence of a large abundance of non-functional *LINE1s* dispersed throughout the genome (Wei et al., 2001).

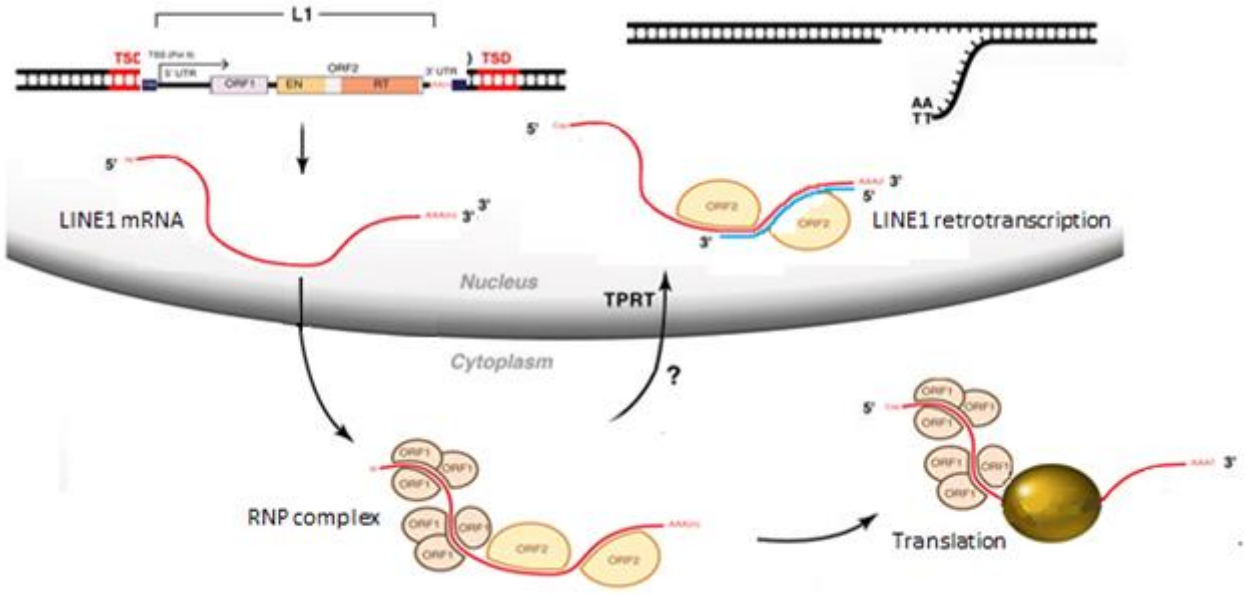


Figure I12: *LINE1* retro-trasposition mechanism. Following transcription, *LINE1* mRNA migrates into the cytoplasm. At this level associating with ribosomes, ORF1p and ORF2p are translated associate with single-stranded RNA forming the RNP complex. The *LINE1*-RNP complex migrates back into the nucleus. Here, the EN domain of ORF2p recognizes and cleaves the consensus sequence 3'-AA/TTTT-5' on the lower DNA strand. At the end of this step, the free 3' hydroxyl originated by EN is employed by the RT domain of ORF2p to use the single-stranded *LINE1*-mRNA as a template to produce the complementary *LINE1*-DNA. (Modified from Singer et al., 2010).

Following migration into the nucleus, the polyA region of the RNA of *LINE1* associates with the RT domain recognition site of ORF2p and in this way retro-transcription begins (Ichiyanagi and Okada, 2008). *LINE1* insertion is initiated by single-stranded DNA cleavage catalyzed by the EN domain of ORF2p. Upon recognition of the relevant consensus sequence, EN then cleaves the target DNA preferentially on the lower strand (the complementary strand). The cleaved 3' end is used as a primer to initiate RT-catalyzed reverse transcription. This process is called target-primed reverse transcription (TPRT) (Figure I12). The other strand of DNA could be cleaved by EN during or after synthesis of the first strand cDNA, followed by synthesis of the second strand cDNA. The second strand is generally cleaved downstream of the cleavage site of the first DNA strand. This distance causes target site duplications (TSDs), short direct repeats flanking both sides of *LINE1* (Cost et al., 2002; Feng et al., 1996).

1.16. Regulation of *LINE1* retro-transcription

It has been observed that a subset of host proteins involved in DNA repair and replication, including RUVBL proteins, participate in retro-transcription. Their absence causes a strong reduction in retro-transposition. In this regard, I refer the reader to the following work for further discussion (Flasch et al., 2019; Miyoshi et al., 2019; Gibson and Kraus., 2012). At the end of the retro-transcriptional step, PARP1/2 proteins interact with a replicative protein A complex called RPA (heterotrimeric protein A complex) that includes ORF2p, which binds to the cleaved site to prevent loss of the amino group from cytidine. This greatly facilitates and makes it possible to obtain terminal integration of a new copy of *LINE1*. In order to exert its biological activity, ORF2, further associates with several proteins. Among the most significant ones emerges the proliferating cell nuclear antigen factor (PCNA). This interaction is critical for retro-transposition (Miyoshi et al., 2019). It is relevant to report that both the decay factor UPF1 and the helicase MOV10 were found within the RNP-*LINE1* complex. Although knockdown of UPF1 levels by RNA interference increases the amount of mRNA and *LINE1* proteins, at the same time it results in reduced retro-transposition efficiency. In contrast, MOV10 showed marked inhibitory activity against retro-transcription. Finally, other studies place this finding in contrast to the ability of MOV10 to facilitate the interaction of the *LINE1*-RNP complex to UPF1 (Gregersen et al., 2014; Taylor et al., 2013; Taylor et al., 2018). Following integration of the reverse complementary strand of the new copy of *LINE1*, the second DNA strand on the parental hemihelix is clipped and the new complementary strand is immediately synthesized using host-encoded and expressed cell replication enzymes (Miyoshi et al., 2019). Although it was initially believed that the RNP-*LINE1* complex penetrates the nucleus during mitosis and integration of the new copy of the *LINE1* element into the genome occurs at S phase of the cell cycle, interestingly, in different tissues, retro-transposition of *LINE1* can occur at multiple stages of the cell cycle. In particular, in neuronal cell cultures, it has been shown that retro-transposition can occur even in cells that are no longer dividing. (Mita et al., 2018; Macia et al., 2017). Finally, retro-

transposition can also occur in special cases using an alternative endonuclease-independent mechanism (EN-independent ORF2p) in cells containing mutations in non-homologous end-joining (NHEJ) genes during DNA repair. This mechanism employs DNA breaks to initiate transcription (Morrish et al., 2002; Kopera et al., 2011). Typically, integration of a back-transcribed full-length copy is a particularly rare event; this occurs because new copies of *LINE1* are truncated at the 5'UTR (Coufal et al., 2011).

1.17. Genomic rearrangements generated during *LINE1* retro-transposition

As mentioned at the end of the previous paragraph 1.16, *LINE1*-mediated retro-transposition events are sometimes accompanied by rearrangements (e.g., 5' truncations and 5' truncations associated with inversion/deletion events) or genomic structural rearrangements. The characteristics of these phenomena suggest that host processes, such as DNA repair and/or DNA replication, may affect the structure of newly retro-transposed *LINE1s* (Szak et al. 2002).

1.17. 1. Intra-*LINE1* alterations

When considering each murine and human *LINE1* family within their respective genome reference sequences, only 15% of all intact *LINE1* insertions flanked by target site duplications (TSDs) represent full-length insertions, 85% are 5'-truncated, and 19% are both 5'-truncated and 5'-inverted (Lander et al., 2001; Szak et al. 2002; Venter et al., 2001). Intensive studies of *LINE1* replication kinetics have progressively directed researchers to divide the previously discussed mechanism of retro-transposition into two categories: conventional and abortive.

Conventional retro-transposition, illustrated in the more classical models and literature, describes what occurs when what retro-transposes is an intact *LINE1*, that is, a *LINE1* that includes within it every sequence from the 5'UTR to the 3'UTR. This type of retro-transposition is the only one capable of generating new retrotransposon-competent elements that may in turn be a source of propagation for new elements during the life course of a cell as well as in its progeny (Gilbert et al., 2005). These types of insertions have structural connotations that identify them, i.e. they occur at a consensus site for the EN domain of ORF2p, are flanked by target site duplications (TSDs), and finally terminate in 3' with a repeating site of A (poly-A) (see paragraph 1.14.5) (Beck et al., 2011). Abortive retro-transposition can result in 5'-truncated or inverted insertions at the 5'.

According to the state of the art, to date, all variations of *LINE1* retro-transposition (full-length, 5'-truncated, and 5'-inverted insertions) are considered TPRT-mediated, differing in attachment modalities on the 5' side of the TE. The 5' inversion is most likely to be achieved by a variant of TPRT called twin priming (Gilbert et al., 2005; Ostertag and Kazazian, 2001). In twin

priming, one end of the DNA can be annealed to the 3'-terminal of the *LINE1* RNA, while the other end is annealed to the same *LINE1* RNA internally. As a result of annealing, nucleotides derived from the target DNA or *LINE1* cDNA, termed micro-homologies (MHs), consist of small sequences of 10 nucleotides complementary to the junction between the 5'-TSD and the sequence of the inverted *LINE1*. These MHs are often observed at the 5' and 3' ends of 5'-reversed *LINE1* copies. After reverse transcription, *LINE1* RNA might be degraded, and the two cDNA strands are annealed to a region containing micro-homology (MH) domains at the reversal junction. This annealing is thought to underlie the formation of MHs at the inverted junctions.

As for 5' truncations, it is necessary to elucidate a priori the so-called "template jumping," proposed by the analysis of another family of non-LTR R2 retrotransposons, which is a very likely mechanism for the generation of full-length *LINE1* copies (Kurzynska-Kokorniak et al., 2007). In the template jumping model, the RT domain of ORF2p that synthesizes the first-strand cDNA moves (jumps) to the 3'-end of the target DNA to continue DNA polymerization. The integration of *LINE1* copies truncated at the 5' end has long been explained by the inability of the RT domain of *LINE1* to copy the entire sequence, either by dissociating prematurely from RNA or by competing with a cellular RNAase that digests RNA before the completion of reverse transcription (Ostertag and Kazazian, 2001). Subsequently, several mechanisms have been proposed to explain the origin of truncated 5' integrations. Template jumping could contribute to 5' truncated retro-transposition (Stage and Eickbush, 2009). Specifically, the double-TPRT model proposed that annealing the internal cDNA sequence with the 5' target site terminates reverse transcription and enhances second-strand cDNA synthesis (Martin et al., 2005). The authors describing this model point out that termination of retro-transcription prior to completion could generate an intermediate that assumes a characteristic "Y" conformation, which could be converted back to physiological conformation through cellular DNA repair mechanisms, including micro-homology-mediated end-joining (MMEJ) and non-homologous end-joining (NHEJ) (Morrish et al., 2002). Based on these observations, in 2005, Zingler and coworkers proposed that at least both mechanisms are responsible for attachment of the 5'-end of *LINE1* and initiation of second-strand synthesis (Zingler et al., 2005). However, with respect to both full-length and 5'-truncated *LINE1* insertions, neither the means by which the 3'-end of the cDNA is attached to chromosomal DNA nor the mechanisms that initiate second-strand synthesis are fully elucidated to date. (Ostertag and Kazazian., 2001).

1.18. Regulation of *LINE1* expression

The process of regulation of *LINE1* retro-transposon activity is a complex multifactorial orchestra, the coordinated activity of which is extremely important for the proper development of the organism. Factors controlling *LINE1* activity are involved in many biological processes, most notably cell division and neuronal function. Many cellular factors are intended to limit the activity and retro-transposition of *LINE1* in the genome. A much smaller number of factors positively

regulate *LINE1*. *LINE1* activity is required at certain developmental stages. The functional role of *LINE1* is intriguing and not yet fully elucidated. Recent studies have made an important step forward in understanding the molecular genetic mechanisms of retro-transposition and regulation of *LINE1* activity. However, many factors involved in the process of *LINE1* retro-transposition and regulation in diseases are still unknown. (Singer et al., 2010; Protasova et al., 2021). The process of ontogeny of the organism, involves changes in the regulation of *LINE1* activity. With particular reference to somatic cells, *LINE1* represents an agent potentially capable of perturbing the physiology of the organism. Based on this, the co-evolution of *LINE1* with the host cell proteome has favored the origination of a recent and increasingly specialized regulome, which induces almost complete transcriptional repression in cells (Aravin et al., 2007). Transcriptional and retro-transposase activity is suppressed early in embryogenesis according to several pathways: these limit the total amount of DNA accessible to regulation in trans either through the presence of regulatory complexes mediated by TFs, chromatin methylation (Muotri et al., 2005; Sanchez-Luque et al., 2019), and through histone modifications (Hunter et al., 2012; Muotri et al., 2005; Sanchez-Luque et al., 2019). The interaction between these three finely coordinated processes leads to the formation of transcriptionally inactive heterochromatin at this level (Sanchez-Luque et al., 2019).

1.18.1 Transcriptional regulation of *LINE1* expression:

1.18.1.1. Transcription factors

To date, there are several TFs that are able to interact with LINE1. Indeed, the 5'UTR of human LINE1 elements contains binding sites for YY1, RUNX3, SOX proteins, and p53 (Becker et al., 1993; Rosser et an., 2012; Tchenio, 2000). The consensus matrix of the transcription factor YY1 (Yin Yang-1) functions as a key element of the *LINE1* promoter to accurately direct the initiation of transcription and consequently, to maintain the integrity of the 5' end. In mice, it appears conserved in monomers of the F and T subfamilies, although it is mutated in the G subfamily (Goodier et al., 2001; Athanikar et al., 2004). Human *LINE1* possesses three binding arrays for the transcription factor Runt-Domain 3 (RUNX3). Of these, the third has regulatory significance on the *LINE1* antisense promoter. Although RUNX3 has important significance in the regulation of *LINE1*, it is considered unlikely to be involved in the origin of the neurogenic tropism that is characteristic of *LINE1* expression. The reason for this is that in addition to the central nervous system, RUNX3 is also expressed in many other tissue types. In mice, it is notable that in the subfamily that is most active during natural developmental progression (T-subfamily), a binding matrix for this TF is missing (Yang et al., 2003). Approximately 450bp downstream of the transcription initiation site of *LINE1*, p53, the onco-suppressor regulator of genome structural integrity also has two binding arrays that are strongly close together. Here, it was confirmed that overexpression of p53 results in increased

transcript levels of *LINE1* and this effect is achieved when it is the second cis matrix of p53 that is involved in binding. In this regard, reporter assays have also been validated by measuring the change in integrated copy number following retro-transcription (Harris et al., 2009). To date, there are mainly two TFs belonging to the SOX family (SRY-related HMG box containing) that possess a role of absolute relevance during neurogenesis. First, using luciferase as a reporter of Sox11 expression in HEK293T cells that physiologically do not express it, a tenfold increase in *LINE1* mRNA levels has been measured (Tchenio et al., 2000). In contrast to Sox11, Sox2 acts in the adult rat hippocampus as a strong transcriptional repressor of *LINE1* at the level of neural stem cells (NSCs). The mechanism is accomplished by recruitment of histone deacetylase 1 (HDAC1) to form a heterodimeric complex.

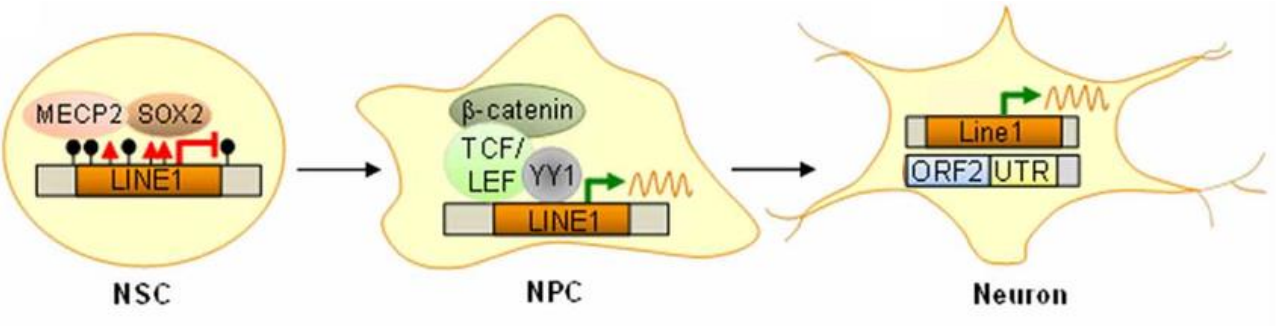


Figure I13: Time course diagram of the major factors involved in transcriptional repression of *LINE1* during glutamatergic lineage differentiation. In NSCs, *LINE1* is bound by MeCP2 together with SOX2. Subsequently, at the level of NPCs, SOX2 and MeCP2 separate from their binding arrays on the *LINE1* sequence and assist in the formation of transcriptionally active euchromatin following DNA methylation reduction on CpG islands. Simultaneously, TCF/LEF activates β-catenin and together with YY1, allows transcription of *LINE1* to continue in adult neurons. (Modified from Hoffmann et al., 2015).

At this level, Sox2 binds a bipartite sequence motif that also contains a LEF binding site. (Singer et al., 2010). During the transition of hippocampal neuro-stem cells to neuronal progenitors (NPCs), *Sox2* expression is subject to downregulation. This causes a concomitant decrease in *Sox2* levels to result in a reduction in its trans contribution to transcriptional regulation, and consequently, an increase in *LINE1* transcription. The reporter assay by luciferase expression that allowed us to understand this indicates that not only Sox2, but also other factors are functionally associated with repression of human *LINE1* transcription (Muotri et al., 2005). Importantly, in humans, the two binding matrices for Sox2 are positionally coincident with the TCF/LEF (T-cell/lymphoid enhancer factor) matrix that afferent to canonical Wnt signaling. Specifically, TCF/LEF functions to adjuvate the expression of precise gene targets when acting in association with β-catenin. WNT3A activates β-catenin, which then accumulates in the nucleus where it forms an activating complex with TCF/LEF (Figure I13). The TCF/LEF complex can now bind to the LEF binding site that has been left vacated by Sox2. This condition leads to transcriptional activation of *LINE1* elements (Figure I13) (Singer et al., 2010).

Although the first characterization of these regulatory mechanisms was conducted in rat, high selective pressure on these regulatory mechanisms favored convergent evolution on these traits of the 5'UTR of *mLINE1*. As a result, the arrays of Sox2 and TCF/LEF are also conserved in both rat and human (Kuwabara et al., 2009). Thus, it appears that in the first two cellular compartments underlying the developmental progression of the glutamatergic lineage, neural stem cells and committed progenitors, Sox2 and TCF/LEF play reverse regulatory roles. Some authors reported that it is precisely through this alternating mechanism of inhibition followed by transcriptional activation in the transition from neuro-stem cells to neural progenitors, that the neuronal specificity of somatic *LINE1* retro-transposition is defined (Kuwabara et al., 2009; Muotri et al., 2005). TAR DNA-binding protein 43 (TDP43) is a highly conserved nuclear RNA/DNA-binding protein involved in the regulation of RNA processing and has been attributed a role of relevance in neurodegenerative diseases. Liu and colleagues have recently shown that in loss-of-function models for TDP43 significant changes occur that are involved in cytoplasmic RNA processing, epigenetic shaping of histones, and DNA damage. Specifically, through reduced levels of TDP43 there is transcriptional activation of chromatin around *LINE1* and a subsequent increase in *LINE1* single-stranded DNA content. (Liu et al., 2019).

1.18.1.2. DNA methylation

The biological significance of *LINE1* DNA methylation is associated with transcriptional control. Although it manifests from the stem precursor stage of both somatic cells and gametes, it persists through differential patterns as a cellular identity factor throughout the life of an organism (Bestor and Bourc'his, 2004). DNA methylation acts by directly preventing the binding of TFs to their target sites and/or altering chromatin structure through histone modification and nucleosome occupancy within gene promoter regions. *LINE1* 5'UTR promoters from humans, and mice (both type A and type F) are readily recognized as CpG islands (Loeb et al., 1986; Padgett et al., 1988). Indeed, the 5'UTR of *LINE1* possesses within it an RNA polymerase II-dependent promoter that directs *LINE1* mRNA transcription and is regulated by DNA methylation of a nearby CpG island (Swergold, 1990; Hata and Sakaki., 1997). This causes transcriptional repression of these elements to be maintained in different cell types. (Richardson et al., 2015). Under normal conditions, there are three categories of key "de novo" methyltransferases "writers" that operate catalysis by the addition of a methyl group at the carbon-5 level of cytosine (5mC) residues, namely Dnmt3a, Dnmt3b, and Dnmt3c. Nevertheless, at the level of embryonic germ cells, methylation of retrotransposons requires DNA methyltransferase 3-like (Dnmt3L) (Ooi et al., 2009), however not able per se to catalyze methylation. In this regard, it has been seen that at this specific developmental stage, mice deficient in Dnmt3L (similarly to mice defective for specific histone methylase) undergo meiosis arrest and transcription of an altered number of *LINE1* copies (Jönsson

et al., 2019). In contrast, Dnmt3a is prominently expressed in NSCs and NPs from the ventricular and subventricular areas of the developing cerebral cortex (E10.5-E17.5) (Moyon et al., 2016), whereas it is mostly absent in astrocytes (Feng et al., 2005). Dnmt3b can only be detected in the SVZ early (E10.5-E13.5), not later during development (E15.5) (Feng et al., 2005; Moyon et al., 2016). Dnmt3c is not expressed in the brain (Barau et al., 2016). In this regard, conditional deletions of Dnmt3a in the developing nervous system (Nes1-Cre dependent) caused shortened lifespan, attributed to massive postnatal motor neuron death (Nguyen et al., 2007). As for mouse embryos lacking Dnmt3b, they too have multiple developmental abnormalities, including rostral neural tube defects, and do not reach birth (Okano et al., 1999). Thus, normal neural development depends on the presence of both 3a and 3b methyl-transferases ab initio. The product of enzymatic catalysis, 5mC, is enriched at CpG dinucleotides and is a marker generally associated with transcriptional repression that has long been considered a stable and highly heritable marker (Branco et al., 2011). In this respect, CpG sites are recognized by members of the Methyl-Binding Domain (MBD) protein family, which in turn recruit histone deacetylases and generate a transcriptionally inactive chromatin structure. In particular, as a member of the MBD protein family, methyl-CpG binding protein 2 (MeCP2) can bind to the 5'UTR of methylated *LINE1* and represses its transcription in reporter assays (Yu, 2001; Muotri et al., 2010).

1.18.1.3. MeCP2 regulates neurogenesis and differentiation

Methyl CpG binding protein 2 (MeCP2) is a member of a complex family of proteins, the Methyl-CpG-Binding Domain (MBD) family of proteins. MeCP2 was first described as a classical transcriptional repressor. MeCP2 is a nuclear protein characterized by several functional domains that regulate the binding of MeCP2 to methylated DNA and subsequent transcriptional repression (Guy et al., 2011). The C-terminal portion contains two DNA binding regions (CTD α and β). The latter regions are involved in interaction with proteins that regulate chromatin structure resulting in hetero-chromatization of the genomic region of interest (Guy et al., 2011; Ghosh et al., 2010). Expression of *MeCP2* during early embryonic development suggests a role for the protein in neurogenesis and differentiation. Indeed, some studies indicate that *MeCP2* is required for these processes. In mice, at the level of cerebral cortex, the level of transcript is moderated at E10.5 (Pelka et al., 2005; Mnatzakanian et al., 2004) and is followed by a transient phase of markedly lowered mRNA levels at E14.5 (Diex-Roux et al., 2011). Finally, it increases during later stages of development. The gradual increase in *MeCP2* expression occurs with the completion of neurogenesis and differentiation and coincides with the specification of stages of neuronal maturation, such as dendritic growth and branching and morphogenesis of the dendritic tree (Shahbazian et al., 2002; Kishi and Macklis, 2005). *MeCP2* levels remain high throughout adulthood, which underscores its important function in the mature brain.

MeCP2 has been extensively studied in the context of Rett syndrome (RTT), a severe neurodevelopmental disorder caused by mutations in the X-chromosome-locus linked to *MeCP2*. RTT occurs in approximately in 1 out of 10,000 female births (see paragraphs 1.8). This disorder is characterized by developmental regression, including cognitive and motor development. While on the one hand, the 2009 study by Tsujimura and colleagues shows that ectopic expression of *MeCP2* in NPCs promotes differentiation toward a neuronal fate by suppressing astrocytic fate determination (Tsujimura et al., 2009). In agreement with this same work, neural stem cells derived from RTT patients differentiated preferentially into astrocytes. This phenotype was associated with altered expression of the astrocytic marker glial fibrillary acidic protein (GFAP). The promoter of this gene contains a CpG island that upon chromatin immunoprecipitation was found to be depleted of *MeCP2* (Fan et al., 2005). Consequently, reduced expression of neural markers was also observed during differentiation of mesenchymal stem cells obtained from an RTT patient. Alterations in the differentiation of *MeCP2*-deficient embryonic neurons have also been identified in mice. (Cobolli Gigli et al., 2018; Squillaro et al., 2012). The mouse pattern of delayed *MeCP2* expression was thought to underlie the apparently normal early development of RTT patients. Analysis of both *MeCP2* knockouts mice and mice with a truncating mutation in the *MeCP2* further supported this view. The mutant mice appeared normal until 5 to 6 weeks after birth, after which phenotypes associated with neuronal dysfunction began to appear (Shahbazian et al., 2002). In a landmark paper published in the journal Nature in 2010, Muotri and co-workers, starting with primary neuro-stem cells obtained from tissue and human induced pluripotent stem cells (hIPSc) revealed that patients with Rett syndrome (RTT) who possess *MeCP2* mutations are more prone to undergo an increased rate of *LINE1* transposition. (Muotri et al., 2010). It is plausible that DNA methylation is not the only mechanism in place for *LINE1* repression because, otherwise, the organism would have become very vulnerable during these two time periods due to uncontrolled *LINE1* expression and activity. Possible candidates include repressive histone modifications that are normally associated with transcriptional gene silencing. An alternative, but not necessarily mutually exclusive, possibility is that relaxed control allows for a physiological level of expression tolerated by the host. In some cases (i.e., during gametogenesis and neurogenesis), this level of expression may also be beneficial to the host (Rosser and An., 2012).

1.18.1.4. Histone modifications associated with retrotransposons

Nucleosomal core histones (histones H2A, H2B, H3, and H4) possess outer tails enriched in lysine (K) residues. It is precisely these N-terminal ends that are the targets of covalent post-translational modifications. These include methylation, acetylation and phosphorylation. The acquisition of these changes leads to the origin of a "histone code" that plays a regulatory role by establishing a selectivity in the recruitment of TFs and polymerases, which in turn determine the establishment of supramolecular complexes that can repress or assist the transcription also at the

spatial level, acting on the three-dimensional conformation of chromatin. These histone modifications are deposited and removed thanks to the presence of enzymes belonging to the class of transferases. While methyl and acetyltransferases operate by deposition of methyl and acetyl groups, respectively, in an opposite manner, deacetylases and demethylases catalyze the removal of these groups. In physiological systems, chromatin is represented by an ordered alternation of euchromatic and heterochromatic regions. These regions may be stable over time (constitutive) or they may be subject to dynamic remodeling, with alternation and presence of mixed and undefined areas especially during development and differentiation (facultative). (Saksouk et al., 2015; Trojer and Reinberg, 2007). Euchromatin constitutes that portion of a cell's genome that is necessary for encoding a proteome that meets the physiology of each cell type (Erdel and Rippe, 2018). Therefore, it is transcriptionally active, consisting of key genes for cellular metabolism, directly accessed by TFs, RNA polymerase, and associated with "activating" histone modifications including methylation in H3K4, in H3K36 and acetylations in H4K16 and H3K27. For example, the deposition of trimethylation of H3 lysine 4 (H3K4me3) strongly correlates with the activation of the euchromatic regulatory regions of promoters (Cruz et al., 2018). In an opposite manner, genomic regions consisting of heterochromatin are transcriptionally repressed, conformationally inaccessible to TFs, and associated with repressive histone modifications. Among these, those that particularly stand out are methylation in H3K27 and H3K9. Trimethylation of H3 lysine 27 (H3K27me3) is a marker of gene repression and exert a repressive function. Deposition of this histone modification is accomplished by the Polycomb 2 repression complex and several methyltransferases. A recent study showed that Polycomb Repressive Complex 2 (PRC2) plays crucial roles in regulating neurogenesis from the beginning phases of VZ germinal zone establishment. This protein complex acts as an epigenetic regulator of the promoter status of accessibility to trans-activators. Selective deletion of the PRC2 methyltransferase subunit, results in the reduction of H3K27me3 and profound transcriptional dysregulation, including a set of TFs directly involved in neuronal cell-type specification and differentiation (Guo et al., 2021). Recently, it has been described that targeting of these histone-modifying complexes within mammalian genomes can be mediated through several DNA-binding proteins, including the Foxg1 interactor, JARID2, and the transcription factor YY1 that confers transcriptional specificity to *LINE1*, (discussed previously in paragraphs 1.14.1 and 1.18.1.1). H3K9me3-tagged heterochromatin plays a pivotal role in neuro-stem cells at the onset of neurogenesis during lineage committment, and also a role in maintaining cell type identity (Nicetto et al., 2019). Based on chromatin immunoprecipitation (ChIP) experiments followed by sequencing, recent work has examined the deposition of methylation levels in H3K9 in relation to the expression of TEs in the hippocampus (Hunter et al., 2012). The data showed that following stressogenic stimuli, H3K9me3 levels undergo an increase that is selectively more intense at loci harboring TEs within them. At the transcriptional level, the authors show that what they discovered correlates with the downregulation and repression of many TEs.

1.18.1.5. piRNAs, PIWI proteins and small RNAs derived from *LINE1*

Another type of evolutionarily relevant factors that represent an effective defense tool possessed by the host to cope with undesired retro-transposition are the PIWI proteins, which represent a subclade of the Argonaute family of small RNA-binding proteins, interacting with 26-31 nucleotide-long RNAs called piRNAs. At this purpose, I suggest the reader to elaborate on this topic through the following articles (Aravin et al., 2008; Siomi et al., 2011). PIWI proteins and piRNAs have been identified in various somatic tissues, including neural cells, although it should be noted that their expression levels are significantly lower in these tissues than in the germline (Rojas-Rios and Simonelig, 2018). In rodents, PIWIL1 (MIWI) is expressed in multiple brain tissues, including the hippocampus and cortex, and controls dendritic spine development and morphogenesis (Lee et al., 2011; Zhao et al., 2015). In the brain tissues of mice, researchers have identified approximately 30,000 neuronal piRNAs (Ghosheh et al., 2016). These somatic piRNAs are shorter in length and tissue-specific, with a greater presence of unique piRNAs in the hippocampus than in the germline. Hippocampal tissues showed the highest expression of piRNAs compared to the cerebral cortex placed second, and following decreasingly, by the kidneys and liver (Perera et al., 2019). Although in many cases, the full molecular mechanisms underlying the activity and regulation of the neuronal piRNA pathway remain largely unknown, these findings suggest that the role of neuronal piRNAs may be associated with neurogenesis, learning, and memory (Kuramochi-Miyagawa et al., 2008). Primary piRNAs correspond to retro-transposons sense-strand transcripts and associate predominantly with MILIs in the cytoplasm. Secondary piRNAs are primarily antisense stranded with respect to retrotransposon transcripts and associate with MIWI2 in the nucleus. Together, MILI and MIWI2 act to amplify piRNAs through an alternating mechanism referred to as "Ping-Pong". Loss-of-function studies in mice suggest that piRNAs regulate de novo methylation at retrotransposons (Kuramochi-Miyagawa et al., 2008). In this regard, two research groups led by Aravin and Carmell, respectively, performing separate studies in germ lines highlighted that de novo methylation to retrotransposon sequences happens when piRNAs acts in association with MIWI2 (Aravin et al., 2008; Carmell et al., 2007). Independently, *LINE1* mRNAs were found to be up-regulated in cells from Mili and Miwi2 knockout mice and at the same time, this change in expression correlated with significant demethylation of sequences localized to the 5'UTR of LINE1 (Kuramochi-Miyagawa et al., 2008). During 2012, Saxena and colleagues found that global piRNA levels are higher in tissues of the mouse cerebellum possessing complete *MeCP2* depletion. As previously discussed, (see paragraph 1.18.1.3), neuronal *LINE1* activity increases as *MeCP2* levels decrease (Muotri et al., 2010). Therefore, patients who harbor inactivating mutations at *MeCP2* have an increased susceptibility to transposition, which could plausibly be the result of piRNA dysregulation. This suggests that PIWI/piRNA-mediated inhibition of transposase activity might be a mechanism strongly selected by biological evolution that contributes to differential transposition rates in the

regions involved. Indeed, the accumulation of disruptive transposases over the course of an organism's existence may also contribute to neural decline and the pathogenesis of various neurological diseases (Erwin et al., 2014).

1.18.2. Post-transcriptional regulation of *LINE1* expression

Expression control of *LINE1* elements can be also achieved post-transcriptionally. Specifically, control can occur in the cytoplasm through repression of translation of both ORF1p and ORF2p, through preventing the RNP-*LINE1* complex from reaching the nucleus (Herrmann et al., 2018; Guo et al., 2014), or - during the integration of new retro-transcribed copies - by modulating enzymatic apparatuses for DNA repair. Following transcriptional regulation, I will now turn to analyze regulatory mechanisms acting downstream of RNA transcription. Candidates for post-transcriptional regulation of *LINE1* expression include endogenous host factors that are found to interact directly or indirectly with the retrotransposon sequence, RNA interference (RNAi), piRNAs that also at this level are found to interfere with mechanisms downstream of transcription, and potential translational and post-translational mechanisms (Servant et al., 2017).

1.18.2.1. Endogenous host factors involved in post-transcriptional regulation of *LINE1*

There are several endogenous factors present in the host that are shown to result in a post-transcriptional regulatory effect of *LINE1*. Among these emerge TREX1 and some hydrolases involved in the deamination of cytidines. In the first case, TREX1 turns out to be an exonuclease expressed in both human and mouse whose substrate consists of single-stranded DNA hemi-helices (ssDNA) from 3' to 5' among which there is an abundance of retro-transcripts not yet integrated by TPRT. In 2008, Stetson and colleagues found that overexpression of TREX1 in cell cultures reduced *LINE1* retro-transposition to approximately 1/5 of its initial value. In the same study, TREX1-deficient mice originated 6-fold more than control ssDNA of *LINE1*. (Stetson et al., 2008). Similar results were confirmed in 2017 by Thomas and colleagues. They observed abundant extrachromosomal DNA in TREX1-deficient neural cells, of which *LINE1* sequences were a major fraction. TREX1-deficient neurons also exhibited increased apoptosis (Thomas et al., 2017). In addition to TREX1, the previously mentioned hydrolases, namely cystidine deaminase, causes hydrolytic cleavage of the -C-N= bond of cytidine with formation of uridine and ammonium ion. Thus, conversion of cytidine to uridine can lead to mutations. In this regard, both AID (activation-induced deaminase) and APOBEC1/3 (apolipoprotein B polypeptide-like catalytic 3 mRNA modification enzymes) are the cytidine deaminases mostly involved in both *mLINE1* DNA and mRNA sequence mutations. Looking

at mammals evolution, it is possible to highlight a clear correlation between the increase in the number of copies associated with isoforms of the APOBEC3 gene and at the same time, a drastic reduction in overall retro-transposition activity that culminates in primates with humans (Chiu and Greene, 2008). Indeed, several studies show that in cell cultures, only human APOBEC3 variants, but not the unique murine gene copy are able to limit *LINE1* retro-transposition (Bogerd et al., 2006; Esnault et al., 2005). Similarly, gain-of-function experiments for AID and APOBEC1 from several mammal species including humans, inhibit *LINE1* retro-transposition in human cells (MacDuff et al., 2009). These results show that regulation of APOBEC1/3 and TREX1 play functional roles in neuronal cells, where they may have evolved as innate defense mechanisms involved in neuronal protection in the CNS from *LINE1* retro-transposition (Ariumi, 2016).

1.18.2.2. RNA Interference (RNAi)

RNA interference is an evolutionarily very ancient and functional post-transcriptional regulatory mechanism that relies on small, 19-23nt long RNA effectors (Ghildiyal and Zamore, 2009). Discovered about three decades ago, small interfering RNAs (siRNAs) and microRNAs (miRNAs) are non-coding RNA sequences with important roles in gene regulation. Both target mRNAs (and lncRNAs) according to Watson-Crick-like rules, leading to target molecule destabilization and/or reduction of its translational output. siRNAs prevailingly affix single targets, miRNAs usually possess multiple targets. To get biologically functional and active effectors, precursors of both these small RNA species are enzymatically processed by RNase III Dicer. The guide strand becomes part of the RNA-mediated inhibition complex that recognizes complementary molecular targets and possibly fragments them (Yang and Kazazian, 2006). As described previously in paragraph 1.14.1, both the 5'UTR of human *LINE1* and the ORF1 of murine *LINE1* contain an antisense promoter. At this level, researchers verified that bidirectional transcription of the 5'UTR of *LINE1* retains the ability to generate double-stranded RNA specific to *LINE1* 5'UTR. In 2006, Yang and co-workers identified both right sense and antisense stranded RNAs complementary to the *LINE1* 5'UTR by Northern blot in HEK293T, a human mesenchymal-derived cell line (Rosser and An, 2012). Using a reporter assay, Yang and colleagues verify on the same cell line that the presence of the antisense promoter in 5'UTR leads to the destruction of *LINE1* mRNA by decreasing its cytoplasmic levels (Yang and Kazazian, 2006). In 2014, Li and colleagues also documented in mice that antisense transcription plays a substantial contribution in limiting *LINE1* retro-transposition (Li et al., 2014). These data suggest to date that the RNA self-interference that is conferred by *LINE1* through its antisense promoter could result from an evolutionary mechanism of co-option that has acquired over time the function of modulating the abundance of its own transcripts. It is important to investigate whether active siRNAs are derived solely from the *LINE1* 5'UTR or also from more terminal regions located toward the 3'UTR (Soifer, 2006). Remarkably, within murine embryonic stem cells, Dicer

knock-out results in hyper-accumulation of *LINE1* transcripts, that are translated and lead to increasd *LINE1* retro-transposition (Bodak et al., 2016).

1.18.2.3. piRNA-mediated post-transcriptional silencing

PIWI-piRNA complexes silence their target genes at both the transcriptional level and at the post-transcriptional level (Di Giacomo et al., 2013; Toth et al., 2016; Weick and Miska, 2014). At transcriptional level, such silencing often occurs through both de novo DNA methylation as well as the recruitment of repressive chromatin modifications (see paragraph 1.18.1.3 and 1.18.1.4). At post-transcriptional level, it prevalently takes place through the cleavage of targeted mRNA transcripts by PIWI's endonuclease activity. Piwi proteins are proposed to be piRNA-guided endonucleases that initiate secondary piRNA biogenesis; however, the actual contribution of their endonuclease activities to transposon silencing was until recently unknown. In this respect, De Fazio and collaborators were the first who found that Mili (a member of Piwi proteins) is selectively required for *LINE1* silencing (De Fazio et al., 2011). Later, in 2019, Kim and colleagues revealed that PIWI proteins SMEDWI-2 and SMEDWI-3 are involved in a mutual co-participation to degrade active transposons (Kim et al., 2019). At the germline level, piRNA pathway components localize to two distinct types of cytoplasmic granules in fetal gonocytes that possess mRNA-degrading components within them. In *Mael*-deficient gonocytes, the latter fail to localize to cytoplasmic granules and de novo DNA methylation of *LINE1* 5'UTR sequences is delayed (Rosser and An, 2012). Further support for piRNA-mediated posttranscriptional silencing of *LINE1* comes from an experiment in which the expression of a single *LINE1* antisense-specific piRNA is down-regulated (Aravin et al., 2009). In contrast to global piRNA knockout studies, down-regulation of this specific *LINE1* piRNA leads to *LINE1* de-repression in the absence of any obvious effect on spermatogenesis and male fertility (Xu et al., 2008). Numerous recent studies have revealed that PIWI and piRNAs are not restricted to gonadal cells, but are active also elsewhere, with special emphasis on neuronal cells. It is now known that in mouse neurons, the piRNA pathways perform multiple functions and act during the development of the nervous system. Here their presence is required for inhibition of axon regeneration (Zhang et al., 2015; Kim et al., 2018), modulation of contextual fear memory (Leighton et al., 2019) and regulation of the response to inflammatory damage caused to Schwann cells (Sohn et al., 2019). Despite this, due to the increasingly proven functional redundancy that the role of piRNAs possesses in addition to the regulation of the expression of TEs, new studies are now aimed at clarifying and elucidating in an ever more in-depth way the role through which they perform regulatory mechanisms associated particularly to *LINE1* in the development of the central nervous system will be strongly required.

1.19. Somatic retro-transposition of *LINE1* in the nervous system

The first evidence demonstrating in mammals the ability of *LINE1* to retro-transpose in the CNS comes from the study conducted by Fred H. Gage and his collaborators, through which it was established that the engineered construct of human *LINE1* is able to undergo retro-transposition both in adult rat NPC hippocampal cell cultures and in vivo, in the mouse CNS (Muotri et al., 2005). Remarkably, this finding was next confirmed in neuronal cultures originating from in vitro differentiation of hESCs and human fetal brain cells. Following the discovery that *MeCP2* plays a relevant role in the transcriptional regulation of *LINE1* in neuro-stem cells, as it promotes its repression, the authors show that this determined TF is downregulated during differentiation both in comparison to progenitors and postmitotic adult neurons (Yu et al., 2001). This finding initially suggested that there might be a lower level of methyl enrichment on the *LINE1* promoter during differentiation (Coufal et al., 2009; Thomas and Muotri, 2012). Confirming this initial hypothesis, a study was published in 2010 by these authors highlighting how increased transcript levels of *LINE1* occur in the absence of MeCP2. The same group in 2010 also looked specifically at the phase transition from stem neuron precursors to neuronal progenitors, confirming at this specific level the correlation between the absence of MeCP2 and increased *LINE1* retro-transposition (Muotri et al., 2010; Erwin et al., 2014).

Subsequently, people began to address the problem of quantifying the change in retro-transposed copy number (CNV) in human. For this purpose, Taqman technology was chosen to be used (Nellåker and Karlsson., 2006). In order to characterize the relative abundance of *LINE1* retro-transposed genomic copy number, Coufal and colleagues, established and refined a protocol that allowed them to detect that specifically in the hippocampus, a highly neurogenic niche, there was an average of 80 more copies than in other regions of the central nervous system, as well as compared to tissues from distal organs such as the liver and heart (Coufal et al., 2009; Richardson et al., 2014). qPCR-based CNV analysis allows for a targeted assessment of the relative abundance of retro-transposable elements at the cell/tissue-specific level as well as at the temporal level in a developmental context, however it does not provide information about the chromosomal locations where the new integrations have taken place. Conversely, with new state-of-the-art sequencing techniques it is possible to accurately map insertional sites and integration (Richardson et al., 2014). In this regard, Faulkner and collaborators devised a high-throughput protocol named retrotransposon capture sequencing (RC-seq) with the aim of quantifying and simultaneously mapping *LINE1*, Alu, and SVA. This technique allowed to identify and characterize, at the level of hippocampus and caudate nucleus, elements of ancient as well as recent insertion (Faulkner et al., 2009). In 2012, Evrony and coworkers conducted a single-cell study of human neurons in which a positional characterization of *LINE1* was conducted on a total of 300 neurons of cortical origin and from the caudate nucleus of three subjects. This established that although the majority of neurons considered had no somatic insertions, a somatic insertion rate of less than 0.6 events/cell occurred in the cerebral cortex and caudate nucleus (Evrony et al., 2012). Now, such a low insertion rate of

LINE1 reported by Evrony and colleagues in this study, raised perplexity when compared with Coufal's qPCR assay in 2009 in which the de novo insertion frequency was about 80 per hippocampal neuron. Different reasons may account for such difference. It is possible that the interindividual biological variability lead to a different frequency of new *LINE1* insertions, possibly due to speicific epigenetic and environmental factors. Next, different CNS structures may be characterized by different retro-transposition rates (Coufal et al., 2009; Baillie et al., 2011). Moreover, a single-cell study might even be less informative and even misleading if performed on an erroneous cell sampling (Erwin et al., 2014). Finally, retro-transposition rates are related to cellular metabolism, resulting to be lower in association to a sedentary lifestyle (Muotri et al., 2009, Thomas et al., 2012). In summary, the study of somatic *LINE1* retro-transposition in the central nervous system with new technologies has enabled the finding of *LINE1* activity in the healthy brain, but also in individual affected by long-lasting neurological disorders (Suarez et al., 2018).

Finally, a key issue, whether *LINE1* retro-transposition requires cell proliferation or can also occurr in postmitotic neurons, remained for a long time unanswered. Only in 2017, Macia et al. (Macia et al., 2017) provided the first rigorous demonstration that an engineered human *LINE1* element encoded by an adenoviral vector is able to retro-transpose in $NeuN^+$ post-mitotic elements.

2. Aims of the project

Primary aim of this work was investigating Foxg1 modulation of *mLINE1* expression in the developing murine neocortex, with particular attention to its cellular and molecular articulation.

Second aim of such work was assaying Foxg1 impact on dynamics of *mLINE1*-DNA content in the murine embryonic pallium.

These topics were addressed in vivo as well as in primary neocortical cultures, engineered by lentiviral somatic transgenesis, and profiled by ad hoc, cellular and molecular assays.

3. Materials and Methods

3.1. Animal handling

Animal handling and subsequent procedures were in accordance with European and Italian laws [European Parliament and Council Directive of 22 September 2010 (2010/63/EU); Italian Government Decree of 4 March 2014, n° 26]. Experimental protocols were approved by SISSA OpBA (Institutional SISSA Committee for Animal Care). *Foxg1*$^{+/-}$ (Hebert and McConnell 2000); and wild type CD1 males were mated to wild type CD1 females (purchased from Envigo Laboratories, Italy) and maintained at the SISSA mouse facility. E11.5 embryos were staged by timed breeding and vaginal plug inspection. Rosa26$^{(pCAG\text{-}flSTOP\text{-}Cas9\text{-}2P2\text{-}Egfp)}$ E11.5 embryos were distinguished from their wild type littermates by inspection under fluorescence microscope. *Foxg1*$^{+/-}$ E11.5 embryos were distinguished from their wild type littermates by PCR genotyping as previously described (Muzio and Mallamaci 2005). Pregnant females were sacrificed by cervical dislocation. Embryonic neocortices were harvested and dissected in sterile ice-cold 1x DPBS (Gibco) supplemented with 0.6% D-glucose.

3.2. Lentiviral vectors

Third generation self-inactivating (SIN) lentiviral vectors (LVs) were generated as previously described (Follenzi and Naldini, 2002) with some adjustments. In brief, HEK293T cells were co-transfected with the transfer vector plasmid plus three auxiliary plasmids (pMD2 VSV. G; pMDLg/pRRE; pRSV-REV), in the presence of LipoD293TM (SigmaGen). The conditioned medium was collected after 24 and 48 hours, filtered and ultracentrifuged at 50000 RCF on a fixed angle rotor (JA25.50 Beckmann Coulter) for 150 min at 4°C. Lentiviral pellets were then resuspended in PBS 1X without BSA (Gibco). LVs were titrated by Real Time quantitative PCR after infection of HEK293T cells, as previously reported (Sastry et al., 2002). One end point fluorescence titrated LV was included in each PCR titration session and PCR-titers were adjusted to fluorescence-equivalent titers throughout the study. Where necessary, specific lentiviral plasmids were constructed with basic cloning techniques. DNA manipulations (extraction, purification and ligation), bacterial cultures and transformations were performed according to standard methods. Restriction and modification enzymes were obtained from New England Biolabs and Promega; DNA fragments were purified from agarose gel by QIAquick Gel Extraction Kit (Qiagen); plasmid preparations were performed through DNA Plasmid Purification Kit (Qiagen). Plasmids were grown in E. Coli, Xl1-blue or ElectroMAX™ Stbl4™ Competent Cells (Invitrogen).

LVs used for this study were referred to throughout the thesis according to the standard nomenclature: LV: pX-GOI, where pX is the promoter and GOI is the gene of interest.

They were:

- LV_pPgk1-rtTA2S-M2 (Spigoni et al., 2010).
- LV_pNes-rtTA-M2, aka pNes/hsp68-rtTA2S-M2 (Brancaccio et al., 2010).
- LV_pTα1-rtTA2S-M2 (Brancaccio et al., 2010).
- LV_pSyn-rtTA2S-M2 (Tigani et al., 2020).
- LV_TREt-PLAP (Falcone et a1., 2019).
- LV_TREt Foxg1 (Raciti et al. 2013).
- LV_pU6 shFoxg1 (Sigma SHCLND-NM 008241, TRCN0000081746).
- LV_pU6 shCtrl (Chiola et al., 2019).
- L V_pTα1-mCherry (Brancaccio et al., 2010);
- LV_U6p-αLACZ.sgRNA [obtained by cloning the TGCGA ATACG CCCAC GCGAT guide (Platt et al., 2014) into BsmBI-cut, lentiCRISPR_v2ΔEB; lentiCRISPR_v2ΔEB was previously obtained starting from lentiCRISPR_v2 (Addgene #52961), deleting the EF1a-prom-Cas9 fragment by EcoRI/BamHI digestion, filling-in and religating].
- LV_U6p-α Foxg1. S2.sgRNA sgRNA [obtained by cloning the GCCCG TCGGG CCGGA CGAGA guide (Mall et al., 2017) into BsmBI-cut, lentiCRISPR_v2ΔEB]
- LV_U6p-αFoxg1. S3. sgRNA [obtained by cloning the CCCCG ACGCC TGGGT GATGC guide (Mall et al., 2017) into BsmBI-cut, lentiCRISPR_v2ΔEB]
- LV_TREt-has Foxg1 W308X [built by transferring the AgeI-XhoI fragment from pUC57 “h-F1-W308X” (Gene Universal) into the AgeI-SalI digested LV_TREt-IRES2-EGFP (Falcone et al., 2015)].

Performances of LV transgenes were monitored by qRT-PCR. Results are summarized in **Table S1**.

3.3. HEK293T cell cultures

HEK293T cells were used for lentiviral production, quantitative titration (Brancaccio et al. 2010). HEK293T cells were cultured in DMEM glutaMAX – 10x FBS – 1x Pen-Strep on 6-multiwell and plated at $1.2x10^6$ cells/well. Cells were kept in culture for 3 days. Plasmid transfection was performed using LipoD293 (SigmaGen laboratories, SL100668) at DIV1 according to manufacturer's instructions.

3.4. In vivo dissection of *Foxg1*$^{+/-}$ and *Foxg1*$^{+/+}$ embryos

Fast genotyping procedure:

In experiments using mice heterozygous for *Foxg1* expression, neocortical primordium was dissected from *Foxg1*$^{+/-}$ P0 and *Foxg1*$^{+/+}$ P0 pups. Each dissection from a single embryo was performed individually to prevent any sample contamination of biological replicates (each individual embryo). For initial genotyping, a skin fragment from each corresponding embryo was collected and used for rapid genotype screening by qualitative PCR (Muzio and Mallamaci 2005). Extraction of genomic DNA as well as the preparation of the reaction mix were performed in accordance with the protocol provided by KAPA HotStart Mouse Genotyping Kit Roche (KK7351) employed without variations.

P0 cortices dissociation:

Of the two initial neocortices present in each cerebral hemisphere derived from a single embryo, the first one was used for mRNA profiling and quantitative assessment of *LINE1*-mRNA levels. The second one, was used to extract genomic DNA dedicated to copy number quantification in the two genotypes (*Foxg1*$^{+/-}$ and *Foxg1*$^{+/+}$). Neocortical tissue from newborn P0 pups was isolated cut into small pieces for 5 minutes (min) in the smallest volume of ice 1X PBS - 0.6%. Each dissociated neocortical primordium isolated from each embryo was carefully maintained isolated in a 24-well mw (Falcon) in the presence of 500uL of 1X PBS - 0.6% D-glucose on ice.

Preparing samples for RNA profiling:

Fragmented hemi-neocortices derived from *Foxg1*$^{+/-}$ and *Foxg1*$^{+/+}$ mice dedicated to profiling m*LINE1*-mRNA levels were dissolved in TRIzol using hand-held mechanical homogenizers and stored at -80 degrees Celsius before continuing with total RNA extraction; DNase treatment, retro-transcription and RT-qPCR (paragraph 3.11; 3.12).

Preparing samples for DNA profiling:

Fragmented hemi-neocortices derived from *Foxg1*$^{+/-}$ and *Foxg1*$^{+/+}$ mice dedicated to profiling m*LINE1*-DNA levels were enzymatically and mechanically dissociated to single cells by employing 2% trypsin followed by incubation at 37^{O}C for 5 minutes and gentle pipetting. For each biological replicate, number of dissociated cells were quantified in a Burker-chamber and neural precursors were then standardized to the number of $1,0*10^6$ cells each. stored at -80 degrees Celsius before continuing with DNA extraction (paragraph 3.16).

3.5. Primary cortico-cerebral cell cultures

Depending on the assay, cultures were set as follows:
Neocortical primordium was dissected depending on the type of experiment from E11.5 or E12,5 mouse embryos (depending on the experiment) and mechanically dissociated to single cells by gentle pipetting. The number of dissociated cells was quantified in a Burker chamber and neural precursors were then plated in 24-multiwell plates (Falcon). When required, they were infected with a mix of LTVs with molteplicity of infection (number of viruses per cell, m.o.i.) equal to 8. As previously described (Brancaccio et al., 2010), this m.o.i. is sufficient to infect the almost totality of neural cells in these conditions. TetSU trangenes modulabili per induzione, were activated starting from the beginning at DIV0 by 2µg/mL doxicicline (Sigma #D9891-10G) and maintained at the same levels by means of hemi-integration of di doxicicline every 2 days. The dissection/infection day was referred to as "day in vitro 0" (DIV0). As for Foxg1+/- heterozygous mice, cell cultures were not infected. Precursor cells were cultured under pro-proliferative (1:1 DMEM-F12, 1X Glutamax (Gibco), 1X N2 supplement (Invitrogen), 1mg/ml BSA, 0.6% Glucose, 2 µg/ml mouse heparin (Stemcell technologies), 1X Pen/strept (Gibco), 10pg/ml Fungizone (Gibco), 20 ng/ml bFGF (invitrogen), 20 ng/mL EGF (Invitrogen)]. Primary cortico-cerebral differentiated cells were cultured adopting three different types of protocols developed with the aim of obtaining a predominant enrichment of neuronal precursor cells (NSCs); neuronal progenitors (NPs) and ultimately neurons (Ns) respectively.

Protocol I (Prot-I) early-neurogenic cultures: cells were plated in pro-proliferative medium (see before) at a density of 300 cells / uL in both protocols involving 1 or 3 days in vitro (DIV1; DIV3). No FBS was added to the cultures. When required, they were infected with a mix of LTVs with molteplicity of infection (number of viruses per cell, m.o.i.) equal to 8. In the case cells were not infected, harvesting of the samples for both processing and molecular/optical analysis took place between the 21st and 22nd hours following cell plating.
Protocol II (Prot-II) mid-neuronogenic cultures: cells were plated in pro-proliferative medium (see before) at a density of 300 cells / uL in protocols involving 3 days in vitro (DIV3)., Cells were differentiated by adding FBS directly to the growth medium in which cells were plated to obtain a final concentration of 5%. The addition of FBS occurred between the 21st and 22nd hours following cell plating. When required, they were infected with a mix of LTVs with molteplicity of infection (number of viruses per cell, m.o.i.) equal to 8. Only in the case of Foxg1 KO experiments using Cas9-expressing mice, the m.o.i of infection used was 12 for each virus driving sgRNAs whether as for the ctrl- it was 24.
Protocol III (Prot-III) late-neuronogenic cultures: cells were plated in pro-proliferative medium (see before) at a density of 100 cells/µL in protocols involving 9 days in vitro (DIV9). Cells were differentiated by adding FBS directly to the growth medium in which cells were plated to obtain a final concentration of 5%. The addition of FBS occurred between the 21st and 22nd hours following

cell plating. When required, they were infected with a mix of LTVs with molteplicity of infection (number of viruses per cell, m.o.i.) equal to 8. Only in the case of Foxg1 KO experiments using Cas9-expressing mice, the m.o.i of infection used was 13 for each virus driving sgRNAs whether as for the ctrl- it was 26. From DIV4 to DIV9, replacement of half of the medium contained in each growth well with new medium free of growth factors (bFGF and EGF) and enriched with FBS at a final 5% concentration was performed daily. In the experiment shown in figure 11, graph 5, Lamivudine (L1295-10MG, Sigma Aldrich) was cronically added at a final concentration of 10 µM to the final medium of each well.

3.6. Primary neuronal cell cultures

Cortical tissue from E16.5 mice was chopped to small pieces for 5 minutes (min), in the smallest volume of ice-cold 1X PBS - 0,6% D-glucose - 5mg/ml DNaseI (Roche) solution. After enzymatic digestion in 2.5X trypsin (Gibco) - 2mg/ml DNaseI for 5 min, and its inhibition with DMEM-glutaMAX (Gibco) – 10X FBS (Euroclone) - IX Pen-Strep, cells were spinned down and transferred to differentiative medium [Neurobasal-A, IX Glutamax (Gibco), 1X B27 sup- plement (Invitrogen), 25qM L-glutamate (Sigma), 25pM §-Mercaptoethanol (Gibco), 2% FBS, 1X Pen/Strept (Gibco), l0pg/ml fungizone (Gibco)]. Cells were counted and plated as follows: in case of RNA profiling, and western blot experiments, cortico-cerebral cells were plated onto 0.1mg/ml poly-L-Lysine (Sigma) pre-treated 12-multiwell plates (Falcon) at 8x10” cells/well in 0.6-0.8 m1 differentiative medium. In general, when indicated, lentiviral infection was done at DIV1. TetON regulated transgenes were activated by 2qg/ml doxycyclin (Clontech, 63131l) administration. 10qM Cytosine §-D-arabinofuranoside (AraC; Sigma) was acutely added to the medium at DIV1. Cells were kept in culture 8 days.

3.7. Immunofluorescence: sample preparation

On the respective culture day scheduled for analysis, the cultures were dissociated using 0.3% trypsin for 4 minutes at RT. Beginning with the third minute of enzymatic treatment, each well of 24- Multiwell was individually gently pipetted with p200 a total of 10 times. Protease activity was stopped by serum inhibition using a 1:1 dilution of the cell lysis suspension in DMEM-glutaMAX - 10x FBS - 1x Pen-Strep medium. The number of dissociated cells from each biological replicate were independently quantified in a Burker chamber and then plated separately in 24-multiwell plates (Falcon) at the density of 200 cells/µL on previously poly-lysine 12mm coverslips. Once plated, 24x-multiwells were placed in the incubator at 37^{o}C for 1 hour. At the end, after medium removal, the cells contained in each well were fixed in 4% PFA at RT for 15 minutes and subsequently washed 3 times in PBS1x.

3.8. Primary cortico-cerebral cell cultures immunofluorescence

Neural cell cultures were fixed by ice-cold 4% PFA for 15-20 min and washed 3 times in lX PBS. Samples were subsequently treated with blocking mix (IX PBS; 10% FBS; l mg/ml BSA; 0.1 to Triton X-100) for at least 1 hour at room temperature (RT). After that, incubation with primary antibodies was performed in blocking mix, overnight at 4°C. The day after, samples were washed 3 times in IX PBS - 0.1to Triton X-100 for 5 minutes and then incubated with secondary antibodies in blocking mix, for 2 hours at RT. Samples were finally washed 3 times in IX PBS - 0.1to Triton X-100 for 5 minutes, and subsequently counterstained with DAPI (4', 6'-diamidino-2- phenylindole) and mounted in Vectashield Mounting Medium (Vector). The following primary antibodies were used: α-Sox2, rabbit polyclonal, (clone 2Q178, ab6142, Abcam); α-Tubb3, mouse monoclonal, (clone Tuj1, Covance, MMS-435P, 1:1000); α-Foxg1 antibody (rabbit polyclonal, ab18259 Abcam, 1:500); α-RFP (Red Fluorescent Protein), rat monoclonal (ABIN334653; Antibodies online, 1:500); α-EGFP (Enhanced Green Fluorescent Protein), chicken policlonal (GTX13970, Tebubio, 1:1000). Secondary antibodies were conjugates of Alexa Fluor 488, Alexa Fluor 594 and Alexa 647 (Invitrogen, 1:500).

3.9. Image acquisition and analysis

Immuno-stained cultured cells were photographed on a Nikon Eclipse TI microscope, equipped with a 20X objective through the Hamamatsu 1394 ORCA-285 camera (Figures. 2), and through the Nikon C1 confocal system (Fig. 6). Photos were collected as 1344x1024 pixel images, or as 3qm Z-stacks (step of 0.3pm) of l024x1024 pixel images, respectively. Images have been analyzed using Volocity 5.5.1 and Microsoft Excel softwares. Final results have been achieved by processing data using Microsoft Excel and Adobe Photoshop CS6 softwares. Details about analysis are provided in the legend of each figure.

3.10. Evaluation of Foxg1 protein levels in distinctive cell types of engineered primary neural cultures

Cultures were acutely transduced by LV_pTα1-mCherry, firing in NPS and Ns, at moi=8. Later, NSCs, NPs and Ns were recognized by their mCherry-Tubb3-, mCherry+Tubb3- and mCherry-±Tubb3+ profiles, respectively. Moreover, for each cell, nuclear Foxg1 protein content was quantified by Volocity 5.5.1 analysis of aFoxg1-IF signal.

Data referring to an equal number of cells from multiple biological replicates of control and Foxg1-manipulated samples (5 and 6, respectively) were collected, cumulatively ranked and employed to establish boundaries between contiguous (aFoxg1-IF signal) deciles. Here "biological replicates" are aliquots of neural cells originating from the same starting pool, each independently transduced and cultured. Next, distinctive cell types (NSCs, NP, and Ns) of different genotypes

(control or mis-expressing Foxg1), falling within different decile bins were quantified, normalized against total cells of the same type and genotype, and finally plotted against decile number. Cumulatively, >17,000 neural cells were scored for this analysis.

3.11. RNA extraction

Total RNA was extracted from both Primary cortico-cerebral differentiating cell culturesAnd primary neural cultured cells (Figures 2,3,4,10 and Figures 5, 14) using TRIzol Reagent (Thermofisher) according to the manufacturer's instructions. RNA was precipitated using isopropanol and Gly- coBlue (Ambion) overnight at -80°C. After two washes with 75% ethanol, the RNA was resuspended in 20µl sterile nuclease-free deionized water. Agarose ge1 electrophoresis and spectro-photometric measurements (DS-11, DeNovix) were employed to estimate its concentration, quality and purity.

3.12. DNase treatment, Reverse Transcription and Real-Time quantitative PCR

RNA preparations from total RNA samples, and RIP-derived samples were treated by TURBO•M DNase (2U/pl) (Ambion) for 1h at 37°C, following manufacturer's instructions. cDNA was produced by reverse transcription (RT) by Superscript III TM (Invitrogen) according to manufacturer's instructions, in the presence of random hexamers. Then, the RT reaction was diluted 1:5 (in case of RIP, and total RNA samples), and 1µl of the resulting cDNA was used as substrate of any subsequent quantitative PCR (qPCR) reaction. Limited to intron-less amplicons and for RIP-derived samples, negative control PCRs were run on RT(-) RNA preparations. qPCR reactions were performed by the SsoAdvanced SYBR Green SupermixT"platform (Biorad), according to manufacturer's instructions. For each transcript under examination and each sample, cDNA was qPCR-analyzed in technical triplicate, and results averaged. In case of total RNA, mRNA levels were normalized against the geometrical average which was obtained from the RT-PCR quantification of the following amplicons: Rpl10a: Gapdh; Cltc and Rn18S (see primer sequences reported below in the same paragraph). In case of both ChIP and RIP samples, IP samples were normalized against Inputs. Data analysis was performed using Microsoft Excel. Details about analysis are provided in the legend of each figure.

The following oligonucleotides have been employed in this study:

Rpl10a/F,R (Zhou et al., 2010).
5’ CAG CAG CAC TGT GAT GAA GCC AAG G 3’
5’ GGG ATCT GCT TAA TCA GAG ACT CAG AGG 3’

mGapdh/F,R
5’ ATC TTC TTG TGC AGT GCC AGC CTC GTC 3’
5’ GAA CAT GTA GAC CAT GTA GTT GAG GTC AAT GAA GG 3’

mmuCltc/F,R
5' CTT GCT CAG CGT TTG GAA AAA CAC GAA CTC AT 3'
5' GCC AAT TCT GTG TCT TTA GAT TCA GAG GCA TAC 3'

mmuRn18s/F,R
5' AAT TGA CGG AAG GGC ACC ACC AGG AGT 3'
5' GCC ATG CAC CAC CAC CCA CGG AAT C 3'

Foxg1(cds)/F,R
5’ GAC AAG AAG AAC GGC AAG TAC GAG AAG C 3’
5’ GAA CTC ATA GAT GCC ATT GAG CGT CAG G 3’

Foxg1(Mid)/F,R
5’ GAC AAG AAG AAC GGC AAG TAC GAG AAG C 3’
5’ GAA CTC ATA GAT GCC ATT GAG CGT CAG G 3’

Couptf1_Foxg1-BS.a/F,R
5’ CCA TCT GAA TGC AAA ACG TAA GCA TG 3'
5’ ATG GAC CTC GCT CTA CTA ATT TGT TGG 3'

Gfap_Foxg1-BS.b/F,R
5’ GCA GGA ACT CCT GAG GCT GGC AGC ATT 3'
5 GTG TAA ACA AAA TTA CAG GAG CTT GTG TTG AAC TTG C 3'

Nfia_Foxg1-BS.h1/F,R
5’ CAG AAC TTG GTG GAT GGT GAG TTG GCA G 3'
5’ TGT GAG TTT GAT GAA TGT GAG GTC TGG GAC 3'

S100b_Foxg1-BS.x2/F,R
5’ TTG TGA GCC ACC CAA TGT GGT TCT GG 3'
5’ CCA CAA ATG TGT TGG AAA GAT ATA ACT CAA AG 3'

Sox9_Foxg1-BS.h1/F,R
5’ CCG TGA TTG GCC CGA GGT ATC TAA CGT G 3'
5’ AGT TGT CGC TCC CAC AGA AGT TTC CAG G 3'

A5utr.t3L(oM)/F,R
5' GAA TAC TCT GCC CAC TGA AAC TAA GGA GA 3'
5'GTT AAG AGC TTG TTC TTC AGG TGA CTC TGT 3'

Gf5utr.t3L(oM)/F,R
5' GGA GGT CCA AAC ACC AGA TAA CTG TAC ACC 3'
5' AGC AGG CGT GGG AGA CAA GCT CTC TT 3'

Tf5utr.t1L(M)/F,R
5' GAG AGT CTG TAC CAC CTG GGA ACT GC 3'
5' CCG AAG AAG AAG GCC CAA AAC AGG AC 3'

*Orf2/F,R **
5' AGC TTT YAT YCC AGG GAT GCA GGG AT 3’
5' TGG GTG TTG GAT MTT GTC AAA TGC TT 3’

*** Y. Cytosine / Thymine (pyrimidine)*
M. Adenine / Cytosine

AM.mL1-pan3utr/F,R
5' GGT GGA ACA ACA TTA TGA ACT AAC CAG TAC C 3'
5' CCA ATG GGC CTC TCT TTC CAG TGA TG 3'

Cdkl5/F,R
5' GCA CTA CTT GAA AAT AAA GGA CGA AGG ATG GC 3’
5' CAG GCT GCG AGA TGA ACG TGA ATA TGA 3’

MeCP2/F,R
5' GTC CCC TGC CCA CAC TAA GTA TGA CAT 3’
5' GGG TCA CAT TGG GTA GCT GAA GGA GAT 3’

Uty/F,R
5' TGC TAT ACT GAT CTC ACA AGG AAC AAA GAG CA 3’
5' GGA GCA CCG ATT CTG GGA ACT TAA CAG 3’

3.13. RNA Immunoprecipitation (RIP) assay

Primary neuronal cells were set up as described above (see section 3.6 “Primary neuronal cell cultures”) and detailed in Fig 15. For each RIP reaction, 10µl of protein A/G Dynabeads (Thermofisher, 492024) were coupled with l0pg of anti-protein of interest (POI; anti-Foxg1 ChIP-grade, rabbit polyclonal, Abcam, ab18259; anti\-GFP, rabbit polyclonal, Abcam, ab290), or 10µg of rabbit IgG (Millipore, 12370) as control, according to manufacturer's protocols. Pre- clearing beads were prepared omitting antibody coupling. At DIV8, cells were washed once with ice-cold lX PBS. For each well (12-multiwe11 plate) was added 75µ1 ice-cold lysis buffer for 10 min on ice. Afterwards, cells were scraped and lysed by vigorously pipetting up and down without the creation of bubbles. The lysate derived from 10 wells (about 8x106 cells; corre- sponding to pair anti-POI/IgG biological samples), was twice “pipetted up and down and kept 10 min on ice”, then it was centrifuged at 2000g for 10 min at 4°C, and then centrifuged at 16000g for 10 min at 4°C. The resulting supernatant of each sample was incubated with pre- clearing beads (pre-equilibrated in lysis buffer) for 30 min at 4°C on roller-shaker. Then, preclearing beads were separated with a magnet, and supernatant was incubated with antibody-coupled beads (pre-equilibrated in lysis buffer), overnight at 4°C on roller-shaker. The l09r of supernatant (Input, IN-RIP) was stored on ice. The day after, beads were collected with a magnet and washed five times with 0.5ml of ice-cold high-salt buffer. [Lysis buffer: 25mM TRIS-HC1, 150mM KCl (Ambion), l0mM MgCl2 (Ambion), 1'to (vol/vol) NP-40 (Thermo Fisher Scientific), IX EDTA-free protease inhibitors (Roche), 0.5 mM DTT (Invitrogen), 10 ul/ml rRNasin (Promega), 10 µl/ml Superasin (Applied Biosystems). High-salt buffer: 25mM TRIS-HCl, 350mM KCl (Ambion), 10mM $MgCl_2$ (Ambion), I9r (vol/vol) NP-40 (Thermo Fisher Scientific), IX EDTA-free protease inhibitors (Roche), 0.5mM DTT (Invitrogen)]. For each sample, RNA immunoprecipitated (IP-RIP) and Input were extracted with Trizol" LS reagent according to manufacturer's instructions. A back extraction was used to improve yield. RNA was precip- itated using isopropanol and GlycoBlue over-night at -80°C. After two washes with 75% ethanol, the RNA was resuspended in l0ul sterile nuclease-free deionized water. Agarose gel electrophoresis and spectrophotometric measurements (NanoDrop ND-l000) were employed to estimate quantity, quality and purity of the resulting RNA.

3.14. Chromatin Immunoprecipitation (ChIP) assay

The Chromatin Immunoprecipitation followed by quantitative Polymerase Chain Reaction assays (ChIP-qPCRs) were performed on chromatin extracted from primary cortico-cerebral differentiating cell cultures engineered as shown in Figures 8,9. The starting material used for each immunoprecipitation was $3,0*10^5$ cells in the case of chromatin immunoprecipitations in which α-Foxg1 antibody was employed. In contrast, in chromatin immunoprecipitations in which antibodies against epigenetic factors (H3K4me3; H3K9me3; H3K27ac and MeCP2) was used, the number of

cells used for each immunoprecipitation was $1,0*10^5$. For each ChIP assay, chromatin was fixed by 1% formaldehyde for 10 min at RT. After cell lysis, fixed chromatin was sonicated by a Soniprep 150 apparatus according to the following settings:

- (on ice; 5s ON, 55 s OFF; oscillation amplitude 5 µm; 4 cycles), giving rise to ~1Kb fragments in the case of chromatin immunoprecipitations in which α-Foxg1 antibody was employed.
- (on ice; 5s ON, 55 s OFF; oscillation amplitude 5 µm; 5 cycles), giving rise to ~500 bp fragments in the case of chromatin immunoprecipitations in which antibodies for epigenetic factors were employed. (**Fig. M1**).

ChIP analysis was performed according to the MAGnifyTM Chromatin Immunoprecipitation System protocol (Invitrogen), with minor modifications. Agarose gel electrophoresis was employed to estimate quality of sonicated chromatin (see **Fig. M1**).

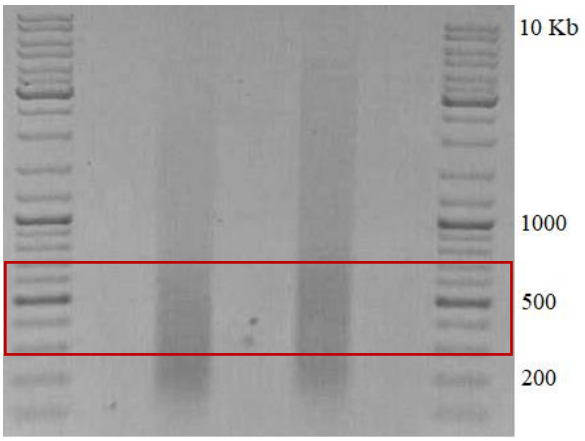


Figure M1: Example of assessment of chromatin fragmentation.
Numbers indicate the corresponding to molecular weights. Red box indicates the fragmented chromatin size samples.

Sonicated chromatin was immunoprecipitated for 2h at 4°C in a final volume of 100 µL, keeping the tubes in a rotating device using the following antibodies:

10 µg/reaction of an α-Foxg1 antibody (rabbit polyclonal, ab18259 Abcam).

3 µg/reaction of an α-histone H3K4me3 antibody (rabbit polyclonal, ab8580), Abcam.

3 µg/reaction of an α-histone H3K9me3 antibody (rabbit polyclonal, cat.n: 39161), Active Motif.

3 µg/reaction of an α-histone H3K27ac antibody (rabbit polyclonal, ab177178), Abcam.

3 µg/reaction of an α-MecP2 antibody (rat polyclonal IgG2a serotype, cat.n: 61291), Active Motif.

Next, immunoprecipitated DNA was purified according to the manufacturer's instructions. Last, 1/30 of each immunoprecipitated (IP) DNA sample was amplified by qPCR. For each sample, qPCRs were performed in technical triplicate. Averages were normalized against input chromatin and further normalized against IgG controls. IgG controls were provided by the MAGnifyTM Chromatin Immunoprecipitation System protocol (Invitrogen) except for α-MecP2 antibody: in this case we used Purified Rat IGG2A (029688) Thermo Fischer. Experiments were performed at least in biological triplicate. Results were evaluated by Student's t-test, via Excel software.

3.15. Fast-genotyping of CD1 wt male and female embryos

Cortical tissue from embryos at E11.5 was cut into small pieces for 5 minutes (min) in the smallest volume of ice 1X PBS - 0.6%. Each dissociated neocortical primordium isolated from each embryo was carefully maintained isolated in a 24-well mw (Falcon) in the presence of 500uL of 1X PBS - 0.6% D-glucose on ice. A skin fragment from each corresponding embryo was collected and used for rapid genotype screening by qualitative PCR using a pair of diagnostic primers directed against the *Uty* gene located on the Y chromosome (not showed results). Extraction of genomic DNA as well as the preparation of the reaction mix were performed in accordance with the protocol provided by KAPA HotStart Mouse Genotyping Kit Roche (KK7351) employed without variations. Following genotype determination, two unique pools including samples belonging to each genotype (male and female) were obtained. This was followed by dissociation and plating according to "Protocol II" discussed in section 2.5. "Primary cortico-cerebral differentiating cell cultures", and subsequently section 2.3. "Isolation of DNA from Primary cortico-cerebral differentiating cell cultures". Quantitative PCR assessment was performed against *Mecp2* and *Cdkl5* (results shown in **Fig. 11**), and the values obtained were normalized against euchromatic Gfap/Rpl10a autosomal locus.

3.16. Isolation of DNA from primary cortico-cerebral cell cultures

For each cultured well dissociated by the use of 0.3% trypsin, or in the case of cells dissociated from tissue (paragraph 3.4), cells belonging to each single biological replicate were counted individually in a burker chamber. Following each count, $1x10^6$ cells were used as starting material for genomic DNA extraction. Experiments were conducted pertaining to the "FlexiGene DNA Kit; Quiagen" protocol with the following modifications.

In the experiments indicated as "high" and "very High Pk",

1) Compared to the standard quantity indicated in the FlexiGene DNA Kit protocol, a 6x amount (high Pk) as well as 12x (very high PK) of reconstituted (1x = 0,1 mg/mL) protease (Proteinase K, recombinant, PCR Grade, Roche) was used for each sample.
2) The 10-minute incubation at 65 degrees Celsius was extended for a duration of 6 hours.
3) Following precipitation, 3 washes of the pellet in 70% ethanol were conducted.
4) Following resuspension of the purified pellet in aqueous solution, the relative concentration of nucleic acid was quantified by spectrophotometer (DS-11, DeNovix) and quantification by qPCR was conducted on samples diluted to 10ng/uL, of which 1uL was used for each amplification reaction.

In case of exctracted genomic DNA, *mLINE1*-DNA levels were normalized against the geometrical average obtained from the PCR quantification of the following amplicons from the same samples: Gfap; S100β and Nfia. (see primer sequences reported in paragraph 3.12).

3.17. Statistical analysis

When not otherwise stated, experiments were performed at least in biological triplicate. Statistical tests employed for result evaluation, and definitions of n (biological replicate) are provided in each Figure.

4. Results

4.1. In vivo *Foxg1* down-regulation of *mLINE1*-mRNA.

It was reported that a substantial fraction of Foxg1 protein is stably bound to chromatin, suggesting it might act as a structural component of it, implicated in long term gene repression (De Filippis et al., 2011). Next, pallial glutamatergic progenitors were described as neural precursors characterized by complementary trends of Foxg1 and *mLINE1*-mRNA expression (Kuwabara et al., 2009; Miyoshi and Fishell, 2012). Finally, motif enrichment analysis (MEA) by Jaspar software (Mathelier et al., 2015) revealed a high score, putative Foxg1 binding site (RTAAACAW) within *mLINE1-orf2* cds (Sookdeo et al., 2013). All this prompted us to hypothesize a major Foxg1 role in inhibition of *mLINE1* trascription. To probe this inference, we compared *mLINE1*-mRNA levels in the neocortex of P0 *Foxg1*$^{-/+}$ mice (Hébert and McConnell, 2000) with wild type controls. As expected, we found a substantial *mLINE1* upregulation in these mutants. It characterized the diagnostic amplicon "*mLINE1*.orf2", shared by all *mLINE1* sub-families (+19.7±1.6%, p<0.014, n=7,7), as well "*mLINE1*.5'UTR.A, .Gf, and .Tf" amplicons, peculiar to the three corresponding transposition-competent sub-families (+17.2±0.3%, p<0.008, n=7,7; +33.0±0.1%; p<10^{-4}, n=7,7; +15.8±0.5%, p<0.006, n=7,7; respectively) (**Fig. 1**).

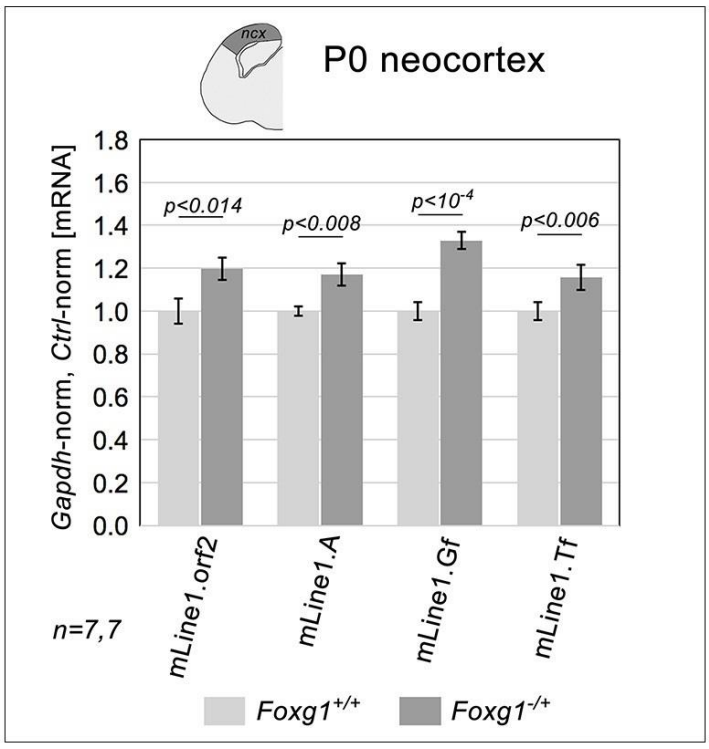
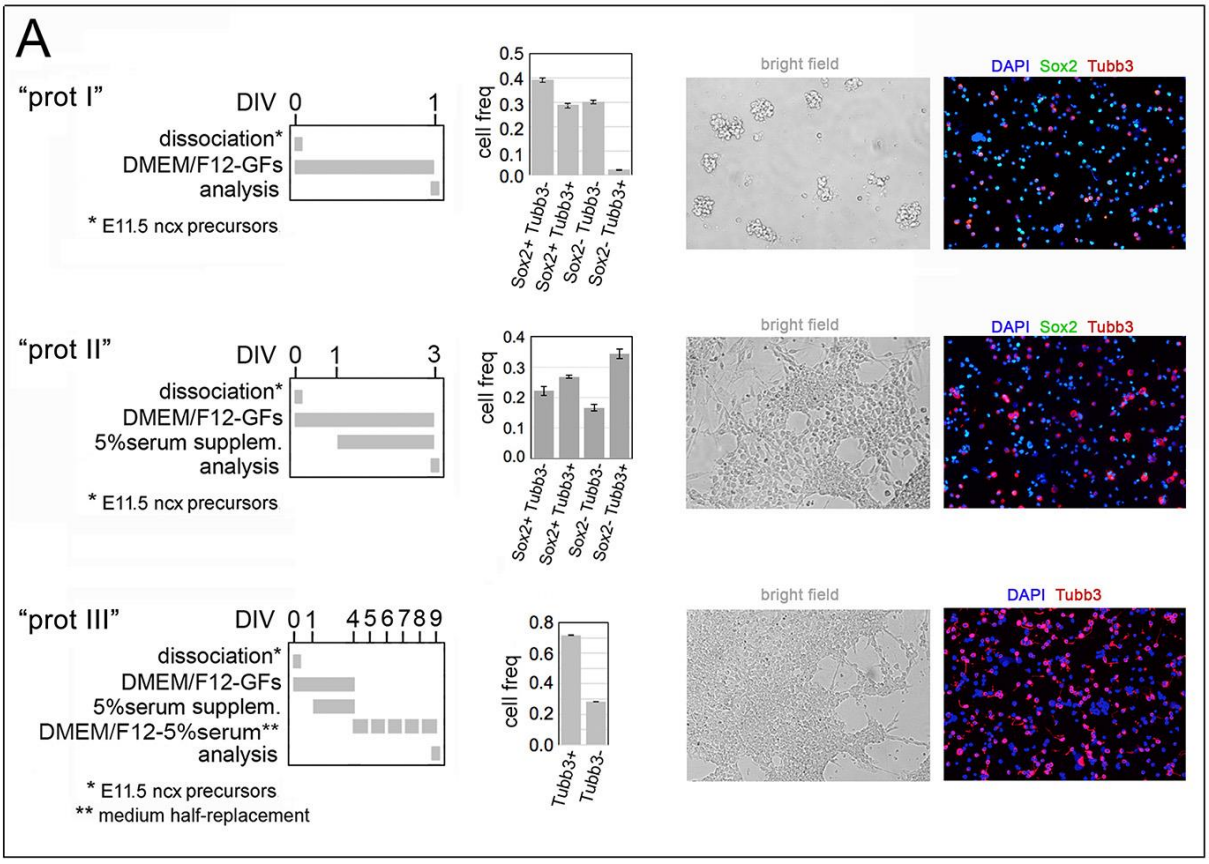


Figure 1. *mLINE1* transcripts levels in neocortex of *Foxg1*$^{-/+}$ mouse neonates and controls. RT-PCR quantitation of pan-*LINE1* diagnostic amplicon "orf2", and family-specific amplicons, "A", "Gf" and "Tf". Data double normalized against *Gapdh* and wild-type controls. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. neocortices taken from distinct pups.

4.2. Modeling temporal progression of pallial *mLINE1*-mRNA in vitro

To ease the modeling of Foxg1 control of *mLINE1* expression, we developed three protocols, "type-I", "-II" and "-III", aimed at generating primary neural cultures representative of early, mid and late phases of pallial neuronogenesis, respectively (**Fig. 2A**). Displaying progressively longer durations, such protocols differed for terminal exposure of neural cells to "pure pro-proliferative", "mixed pro-proliferative/pro-differentiative", and "pure pro-differentiative" media, respectively. Type I cultures gave substantial fractions of $Sox2^{+}Tubb3^{-}$ presumptive NSCs (39.3±0.9%; n=3) and $Sox2^{-}Tubb3^{-}$ neuronogenic progenitors (30.1±0.7%; n=3), with limited $Tubb3^{+}$ neuronal output (30.6±1.2%; n=3). Prevalence of these two precursors was lowered (22.2±1.4% and 16.6±1.1%, respectively; n=3) in Type II cultures, characterized by more frequent neurons (61.2±2.0%; n=3). As expected, neuronal prevalence further increased in Type III cultures (71.7±1.0%; n=3). Next, we profiled these cultures for *mLINE1*-mRNA expression levels, by qRT-PCR. An increasing progression ("type I" < "type II" ≤ "type III") emerged at the diagnostic amplicon "*mLINE1*.orf2", shared by all *k* sub-families (specifically, type I-normalized values were: 1.00±0.11, 19.85±2.59, and 21.65±3.27, respectively, with $p_{(I\text{-}vs\text{-}II)}<10^{-5}$, $p_{(II\text{-}vs\text{-}III)}<0.34$, n=6,6,6). As expected, similar progressions were also detectable at sub-family-specific amplicons, "*mLINE1*.5'UTR.A (Type I-normalized values: 1.00±0.39, 13.01±1.49, and 49.87±8.57, respectively, with $p_{(I\text{-}vs\text{-}II)}<10^{-5}$, $p_{(II\text{-}vs\text{-}III)}< 10^{-3}$, n=6,6,6), .Gf (Type I-normalized values: 1.00±0.10, 11.16±3.10, and 33.31±7.14, respectively, with $p_{(I\text{-}vs\text{-}II)}<10^{-2}$, $p_{(II\text{-}vs\text{-}III)}< 10^{-2}$, n=5,6,6), and .Tf" (Type I-normalized values: 1.00±0.05, 23.41±7.26, and 72.48±11.49, respectively, with $p_{(I\text{-}vs\text{-}II)}<10^{-2}$, $p_{(II\text{-}vs\text{-}III)}< 10^{-2}$, n=5,6,5) (**Fig. 2B**). To note, despite similarity among progressions, their magnitude largely differed among sub-families (with Tf>A>Gf). This points to differential, sub-family- specific transcriptional tuning of m*LINE1*-mRNAs. Moreover, *mLINE1*.orf2 signal reached its maximum in advance compared to 5'UTR, sub-family-specific amplicons. This further suggests that "head-less", retro-transposition-incompetent *mLINE1* relics (Hedgell et al., 1987; Deininger et al., 2017) may substantially contribute to the former.


A
"prot I"
DIV 0 1
dissociation*
DMEM/F12-GFs
analysis
* E11.5 ncx precursors
cell freq
0.5
0.4
0.3
0.2
0.1
0.0
Sox2+ Tubb3-
Sox2+ Tubb3+
Sox2- Tubb3-
Sox2- Tubb3+
bright field
DAPI Sox2 Tubb3
"prot II"
DIV 0 1 3
dissociation*
DMEM/F12-GFs
5%serum supplem.
analysis
* E11.5 ncx precursors
0.4
0.3
0.2
0.1
0.0
Sox2+ Tubb3-
Sox2+ Tubb3+
Sox2- Tubb3-
Sox2- Tubb3+
bright field
DAPI Sox2 Tubb3
"prot III"
DIV 0 1 4 5 6 7 8 9
dissociation*
DMEM/F12-GFs
5%serum supplem.
DMEM/F12-5%serum**
analysis
* E11.5 ncx precursors
** medium half-replacement
0.8
0.6
0.4
0.2
0.0
Tubb3+
Tubb3-
bright field
DAPI Tubb3

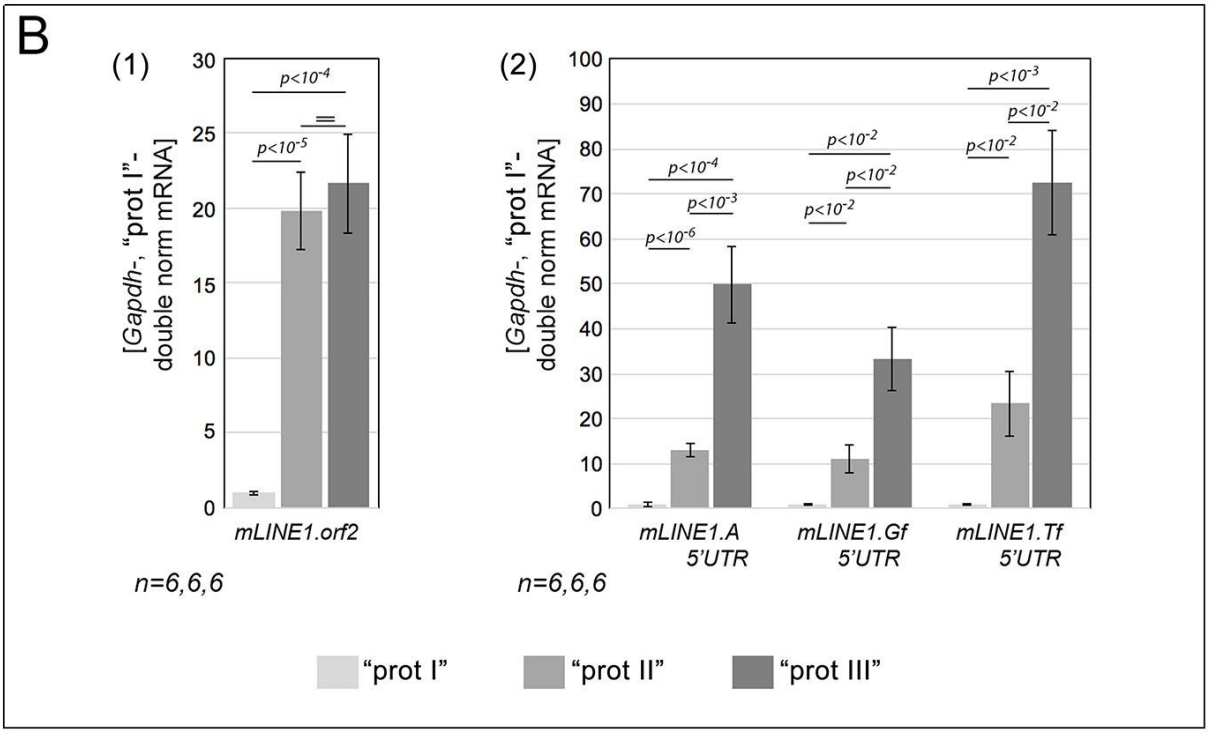

B
(1)
[Gapdh-, "prot I"-double norm mRNA]
30
25
20
15
10
5
0
p<10-4
p<10-5
mLINE1.orf2
n=6,6,6
(2)
[Gapdh-, "prot I"-double norm mRNA]
100
90
80
70
60
50
40
30
20
10
0
p<10-3
p<10-2
p<10-2
p<10-4
p<10-2
p<10-3
p<10-2
p<10-6
mLINE1.A 5'UTR
mLINE1.Gf 5'UTR
mLINE1.Tf 5'UTR
n=6,6,6
"prot I"
"prot II"
"prot III"

Figure 2. In vitro modeling of neuronogenic progression and corresponding *mLINE1*-mRNA expression. In (**A**) shown are three protocols aimed at achieving primary cultures, (I) enriched in neural stem cells (NSCs, $Sox2^{+}Tubb3^{-}$) and neuronogenic progenitors (NPs, $Sox2^{-}Tubb3^{-}$), (II) including comparable fractions of NSCs, NPs and neurons (Ns, $Tubb3^{+}$), and (III) highly enriched in Ns [cultures eventually exposed to GFs, GFs and serum, and serum, respectively]. In (**B**) RT-PCR quantitation of pan-*LINE1* diagnostic amplicon "orf2" (1), and family-specific amplicons, "A", "Gf" and "Tf" (2), in neural cultures set according to protocols I, II and II. Data double normalized against *Gapdh* and "protocol I" values. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. independently cultured cell aliquots, originating from pooled, wild-type E11.5 neocortical primordia.

Next, we profiled these cultures for *mLINE1*-mRNA expression levels, by qRT-PCR. An increasing progression ("type I" < "type II" ≤ "type III") emerged at the diagnostic amplicon "*mLINE1*.orf2", shared by all *mLINE1* sub-families (specifically, type I-normalized values were: 1.00±0.11, 19.85±2.59, and 21.65±3.27, respectively, with $p_{(\text{I-vs-II})}<10^{-5}$, $p_{(\text{II-vs-III})}<0.34$, n=6,6,6). As expected, similar progressions were also detectable at sub-family-specific amplicons, "mLINE1.5'UTR.A (Type I-normalized values: 1.00±0.39, 13.01±1.49, and 49.87±8.57, respectively, with $p_{(\text{I-vs-II})}<10^{-5}$, $p_{(\text{II-vs-III})}< 10^{-3}$, n=6,6,6), .Gf (Type I-normalized values: 1.00±0.10, 11.16±3.10, and 33.31±7.14, respectively, with $p_{(\text{I-vs-II})}<10^{-2}$, $p_{(\text{II-vs-III})}< 10^{-2}$, n=5,6,6), and .Tf" (Type I-normalized values: 1.00±0.05, 23.41±7.26, and 72.48±11.49, respectively, with $p_{(\text{I-vs-II})}<10^{-2}$, $p_{(\text{II-vs-III})}< 10^{-2}$, n=5,6,5) (**Fig. 2B**). To note, despite similarity among progressions, their magnitude largely differed among sub-families (with Tf>A>Gf). This points to differential, sub-family- specific transcriptional tuning of m*LINE1*-mRNAs. Moreover, *mLINE1*.orf2 signal reached its maximum in advance compared to 5'UTR, sub-family-specific amplicons. This further suggests that "head-less", retro-transposition-incompetent *mLINE1* relics (Hedgell et al., 1987; Deininger et al., 2017) may substantially contribute to the former.

4.3. Modeling Foxg1 regulation of *mLINE1*-mRNA in vitro, by constitutive *Foxg1* manipulation

To assess the suitability of our primary neural cultures to model Foxg1 control of *mLINE1* transcription in vitro, we set up "type II", mid-neuronogenic cultures (**Fig. 3A,B**, left), we constitutively manipulated their *Foxg1* expression levels downwards and upwards (**Table S1**), by CRISPR-Cas9 and Tet^{ON} technologies, respectively (**Fig. 3A**,**B**, left), and we finally evaluated *mLINE1*-mRNA levels. Results were fully consistent with those emerging from profiling of P0 $Foxg1^{-/+}$ pups. Specifically, *Foxg1*-LOF cultures displayed a systematic upregulation of *mLINE1*-mRNAs, +25.9±7.4% (p<0.004), +21.6±10.9% (p<0.054), and +18.7±6.7% (p< 0.017), as for sub-families A, Gf and Tf, respectively, with n=10,11 (**Fig. 3A**, graph (1)). Reverse, more robust changes were detectable upon generalized *Foxg1* upregulation. In this case, mLINE1.orf2 qRT-PCR signal was decreased by 41.2±7.7% (p<0,023, with n=6,6), sub-family-specific signals, A, Gf and Tf, were diminished even

more, by 61.6±5.0% ($p<10^{-4}$, with n=6,6), 67.5±4.6% ($p<10^{-4}$, with n=6,6), and 59.3±7.5% ($p<0,001$, with n=6,6), respectively (**Fig. 3B**, graph (1)).

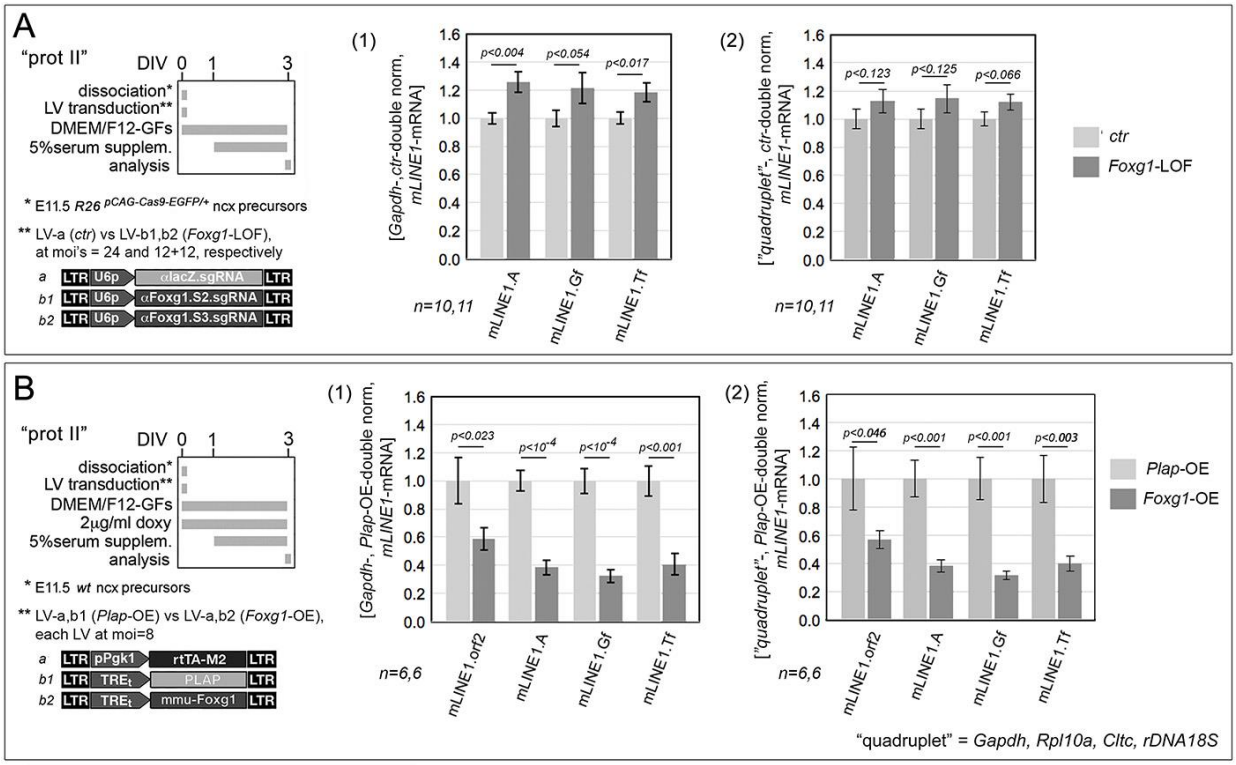


Figure 3. Impact of constitutive *Foxg1* manipulation on *mLINE1* expression in mid-neuronogenic pallial cultures. In (**A**) and (**B**), shown are outcomes of *Foxg1* loss-of-function (*Foxg1*-LOF) and *Foxg1* overexpression (*Foxg1*-OE), respectively. To left, protocols and lentiviral vectors employed. Transgenes driven by U6p and Pgk1p, two constitutively firing promoters, responding to RNA-pol III and II, respectively. To right, results. RT-PCR quantitation of pan-*LINE1* diagnostic amplicon "orf2", and family-specific amplicons, "A", "Gf" and "Tf", in neural cultures set according to protocol II. Data double normalized against *Gapdh* and control values (1), or against gene quadruplet (*Gapdh*, *Rpl10a*, *Cltc* and *rDNA 18S*) and control values (2). Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. independently cultured and engineered cell aliquots, originating from pooled, *R26*$^{pCAG\text{-}Cas9\text{-}EGFP/+}$ (A) and *wild-type* (B), E11.5 neocortical primordia.

To secure these results (obtained upon usual *Gapdh* mRNA normalization), we alternatively normalized *mLINE1* qRT-PCR values against a specific "gene quadruplet". This included three RNA-pol II-transcribed genes (*Gapdh*, *Rpl10a* and *Cltc*), characterized by comparable expression profiles in apical precursors (APs), basal progenitors (BPs), early neurons (eNS) and late neurons (lNs) (Telley et al., 2016) (**Table S2A**), as well as by poor sensitivity to *Foxg1* manipulation (Artimagnella and Mallamaci, 2020 i.e doi:10.5281/zenodo.3739467) (**Table S2B**).

The "quadruplet" also included RNA-pol I-transcribed *rDNA-45S*, from which the large majority of cell RNA complement is generated. Interestingly, *mLINE1*-mRNA normalization against "quadruplet" did not change the emerging outcome of *Foxg1*-OE manipulation, which reproduced the prominent systematic decrease of *mLINE1* signals observed upon *Gapdh* normalization (**Fig. 3B**, graph (2)). As

for *Foxg1*-LOF cultures, their "quadruplet"-normalized *mLINE1* signals again showed a consistent upwards trend (which - however - did not reach formal statistical significance; **Fig. 3A**, graph (2)).
Altogether these data point to a negative impact of Foxg1 on *mLINE1*-mRNA expression. Moreover, mirror phenotypes displayed by *Foxg1*-LOF and -OE cultures further suggest that Foxg1 *does contribute* to fine physiological tuning of these levels.

4.4. Modeling Foxg1 regulation of *mLINE1*-mRNA in vitro, by conditional *Foxg1* overexpression

To get insights into temporal and intra-clonal articulation of Foxg1 repression of mLINE1 expression, we run additional *Foxg1*-OE assays. For these assays we employed different, early- mid- and late-neuronogenic, cultures, and we restricted *Foxg1* transgene expression by means of cell-type specific promoters, pNes, pTa1 and pSyn, active in NSCs, NPS and Ns, and Ns, respectively (Brancaccio et al., 2010; Tigani et al., 2020).
In case of early-neuronogenic cultures, Foxg1 overexression driven by either pNes or the ubiquitous pPgk1 gave rise to a substantial and generalized down-regulation of mLINE-mRNAs. Normalized against "quadruplet", mLINE1.orf2, .A, .Gf, and .Tf signals were decresed by 29.14±4.06% ($p<10^{-4}$, n=9,9), 27.28±3.82% ($p<10^{-3}$, n=9,9), 23.22±5.02% ($p<10^{-2}$, n=9,9), and 22.26±5.54% ($p<10^{-2}$, n=9,9), in case of pNes-manipulated cultures, by 26.82±4.53% ($p<10^{-2}$, n=6,6), 31.01±6.72% ($p<10^{-2}$, n=6,6), 38.45±6.27% ($p<10^{-2}$, n=6,6), and 32.21±5.01% ($p<10^{-2}$, n=6,6), in case of pPgk1-manipulated cultures, respectively.
Next, mid-neuronogenic Foxg1-OE cultures gave different results depending on the promoter driving the Foxg1 transgene. "Quadruplet"-normalized, mLINE1.orf2, .A, .Gf, and .Tf signals were decresed by 6.90±7.50% ($p<0.238$, n=6,8), 18.13±4.63% ($p<10^{-2}$, n=6,8), 23.36±3.53% ($p<10^{-2}$, n=6,8), and 16.11±4.31% ($p<0.05$, n=6,8), in Foxg1-OE^{pTa1} cultures. They were prevalently unaffected in Foxg1-OE^{pSyn} cultures, except *mLINE1*.orf2 and .Gf signals, reduced by 7.53±2.25% ($p<0.06$, n=6,6), and 10.74±2.08% ($p<0.02$, n=6,6), respectively.

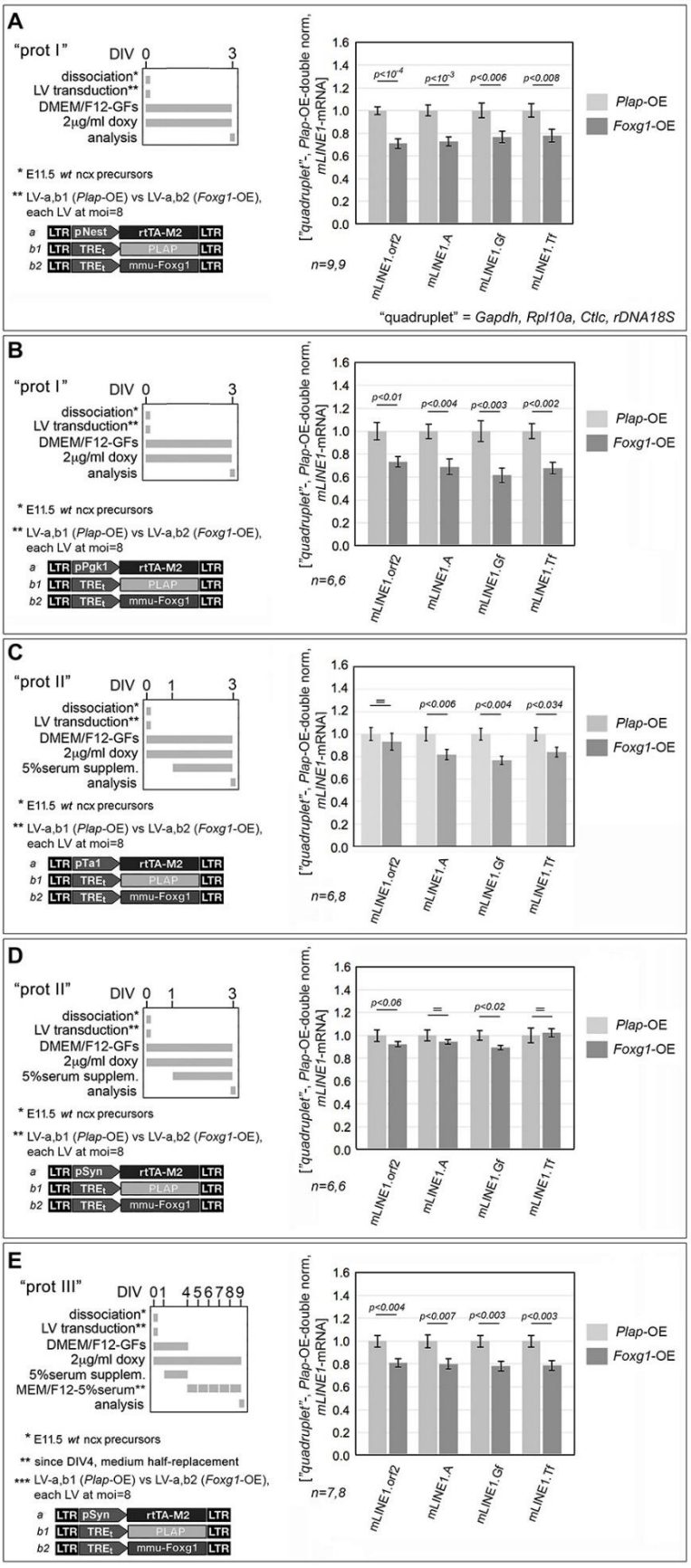


Figure 4. Impact of conditional *Foxg1* overexpression on *mLINE1*-mRNA levels, in progressively more advanced neuronogenic pallial cultures. In (**A**,**B**), (**C,D**) and (**E**), shown are outcomes of *Foxg1* overexpression (*Foxg1*-OE), in early-, mid- and late-neuronogenic cultures, set according to type I, II and III protocols, respectively. To left, protocols and lentiviral vectors employed. Transgenes driven by pNes, pTα1 and pSyn promoters, active in NSCs, NPs/Ns and Ns, respectively, and ubiquitously firing pPgk1 promoter. To right, results. RT-PCR quantitation of pan-*LINE1* diagnostic amplicon "orf2", and family-specific amplicons, "A", "Gf" and "Tf", in neural cultures set according to protocol II. Data double normalized against gene quadruplet (*Gapdh*, *Rpl10a*, *Cltc* and *rDNA 18S*) and control values. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. independently cultured and engineered cell aliquots, originating from pooled *wild-type*, E11.5 neocortical primordia.

Last, late-neuronogenic cultures over-expressing *Foxg1* under the control of pSyn displayed a robust *mLINE1*-mRNA down-regulation, with "quadruplet"-normalized, mLINE1.orf2, .A, .Gf, and .Tf signals decresed by 19.17±3.77% ($p<10^{-2}$, n=7,8), 19.89±4.29% ($p<10^{-2}$, n=7,8), 21.84±4.29% ($p<10^{-2}$, n=7,8), and 21.29±4.02% ($p<10^{-2}$, n=7,8), respectively.

Albeit collectively pointing to a generalized *Foxg1*-dependent down-regulation of *mLINE1* expression throughout the entire neuronogenic process, these data deserve however a more careful interpretation. Specifically, we have to distinguish between the case of *late* neuronogenic cultures manipulated by a *pSyn-driven Foxg1* transgene (**Fig. 4E**) and *other cultures* (**Fig. 4A-D**). In fact, in the former case the promoter was activated in cells occupying a terminal position along the neuronogenic sequence and the culture was allowed to age for a sufficiently long time, so that a robust peak of Foxg1 protein was reasonably achieved within the same cell type where the promoter fires. Conversely, in the latter cases, the promoters were active in transient precursor types, within relatively short-lived cultures. In this way, actual upregulation of Foxg1 protein could have taken place in a cell type different from the one where the promoter was active, or there could have been not enough time to get an appreciable protein upregulation. In the light of these considerations, the interpretation of data obtained in Foxg1-OE^{pSyn}, late neuronogenic cultures is straightforward, pointing to a consistent *neuronal* inhibition of *mLINE1* elements belonging to all sub-families by Foxg1. Remarkably, this inference was corroborated by the results of supplemental *Foxg1* manipulations, upwards and downwards, constitutively performed in pure neuronal cultures fully depleted of glial cells by araC supplementation (**Fig. 5**).

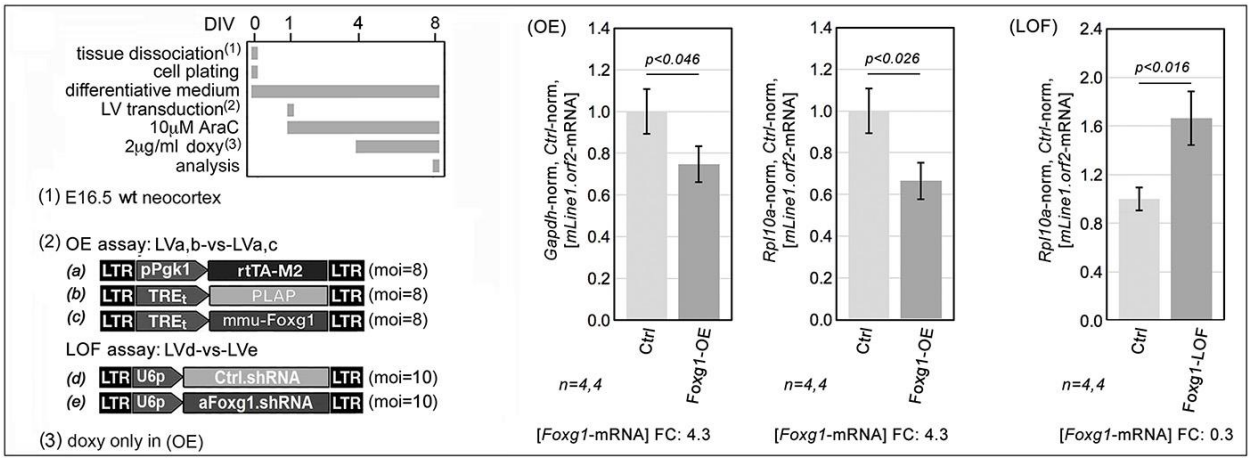


Figure 5. Impact of constitutive *Foxg1* overexpression on *mLINE1*-mRNA levels, in primary *neuronal* cell cultures. To left, protocols and lentiviral vectors employed, to right, results. Analysis run in pure neuronal cultures, obtaind by araC supplementation. Foxg1 constitutively overexpressed (OE) by a pPgk1-driven transgene, or down-regulated (LOF) by RNAi. *mLINE1*-mRNA data double normalized against *Rpl10a* (or *Gapdh*) and control samples. Scalebars, sem's. Statistical evaluation of results by t-test (one-tail, unpaired). *n* = number of biological replicates, i.e. independently cultured and engineered cell aliquots originating from a common neural precursor pool. At the bottom of each graph, shown are approximate fold changes in *Foxg1*-mRNA levels elicited by OE or LOF manipulation.

Here, Foxg1 overexpression reduced the *mLINE1*.orf2 qRT-PCR signal, by 25.21±8.66% (p<0.05, n=4,4) upon *Gapdh*-normalization (**Fig. 5**, graph (1)), by 33.56±8.71% (p<0.03, n=4,4) upon *Rpl10a*-normalization (**Fig. 5**, graph (2)), whereas *Foxg1*-knockdown increased such signal, by 66.44±21.89% (p<0.02, n=4,4) (**Fig. 5**, graph (3) upon *Rpl10a*-normalization. Beyond their consistency with Foxg1-OE^{pSyn}results, these last data suggest that Foxg1 protein does participate in *physiological tuning* of neuronal *mLINE1* transcription. To get insights into mechanisms underlying the outcome of conditional *Foxg1* overexpression in early- and mid-neuronogenic cultures, we scored Foxg1 protein expression within NSCs, NPs and Ns compartments of differently engineered cultures, we quantified their sizes, and we finally compared results of this analysis with *mLINE1*-mRNA dynamics peculiar to corresponding cultures. We found that, within early-neuronogenic cultures, pNes-driven *Foxg1* elicited a prominent increase of NPs (39.16±1.14% vs 11.93±2.58%, $p<10^{-5}$, n=6,5) at expenses of NSCs (57.81±1.53% vs 84.58±2.81%, $p<10^{-5}$, n=6,5), while keeping Ns to a minimum (**Fig. 6A** (1)). Moreover, it gave rise to a differently distributed Foxg1 protein modulation in NPs and NSCs: about 12% of NPs moved form Foxg1-deciles 7-10 to decile 1, about 35% of NSCs from deciles 2-5 to decile 1 (**Fig. 6A** (2)). As such, these data did not allow us to disambiguate NSC and NP contributions to early *mLINE1* dynamics shown in **Fig. 4A** (**Fig. 7**). Next, we found that, within mid-neuronogenic cultures, both pTα1- and pPgk1-driven *Foxg1* transgenes impacted only marginally on sizes of the main population compartments, both eliciting a moderate shrinkage of the neuronal one (from 66.09±1.88 to 59.96±1.82, with p<0.03 and n=5,6, as well as 62.00±0.89 to 53.64±1.59, with p<0.01 and n=3,6, respectively) (**Fig. 6B** (1) and **6C** (1)). More intriguingly, they led to differently shaped mis-distributions of Foxg1 protein. Specifically, upon pTα1-driven transgene delivery, Foxg1 protein upregulation was *less pronounced* in NPs (9% of cells moving from deciles 4-3 to deciles 2-1) compared to Ns (12% of cells moving from Foxg1-deciles 10-6 to decile 1) (**Fig. 6B** (2)). Conversely, following pPgk1-driven transgene delivery, Foxg1 protein upregulation was *more pronounced* in NPs (about 16% cells moving from Foxg1-deciles 5-2 to decile 1) than in Ns (about 12% cells moving from deciles 7-2 to decile 1 (**Fig. 6C** (2)).

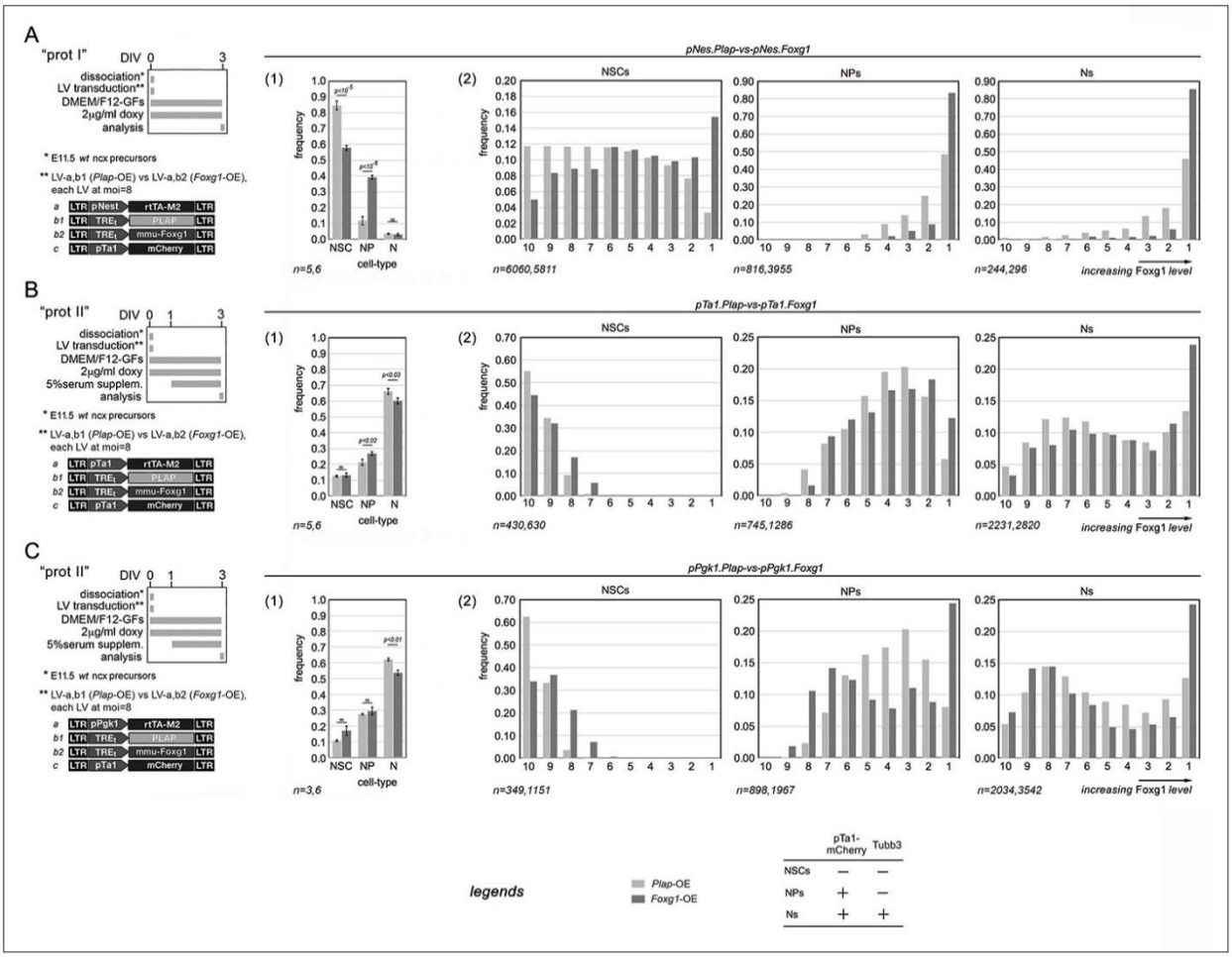


Figure 6. Quantification of Foxg1 protein levels in distinctive neural precursor types, upon *Foxg1* over-expression driven by lentiviral vectors and cell-type specific promoters. Throughout the figure, to left, protocols and lentiviruses employed, to right, results. Assays run in early- (**A**), mid- (**B,C**) and late- (**D**) neuronogenic cultures, set according to type I, II and III protocols, respectively. Cultures over-expressing *Foxg1* or a control (*Plap*). Foxg1-encoding transgenes driven by pNes (A), pTa1 (B) and pSyn (D) promoters, active in NSCs, NPs/Ns and Ns, respectively, and ubiquitously firing, pPgk1 promoter (C). NSCs, NPs and Ns recognized on the basis of their ^{pTa1}mCherry^{-}/Tubb3^{-}, ^{pTa1}mCherry^{+}/Tubb3^{-}, ^{pTa1}mCherry$^{\pm}$/Tubb3^{+} profiles, respectively. Foxg1 cell content evaluated by quantitative immuno-fluorescence. In graphs (1), shown are prevalences of NSCs, NPs and Ns in neural cultures set according to the different protocols. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. independently cultured and engineered cell aliquots, originating from pooled, *wild-type* E11.5 neocortical primordia. In graphs (2), shown are frequencies of NSCs, NPs and Ns falling within distinct Foxg1 expression deciles, i.e. equi-numerous bins characterized by decreasing (from 1 to 10) Foxg1 expression levels. n is the number of cells profiled, evenly pooled from the biological replicates referred to as for graph (1) series.

Together with stronger *mLINE1* inhibition elicited by pPgk1-driven *Foxg1* compared to pTa1-driven transgene (**Fig. 3B** and **4C**), this scenario suggests that Foxg1 down-regulation of *mLINE1*-mRNA detectable in mid-neuronogenic cultures preferentially takes place in NPs (**Fig. 7**).

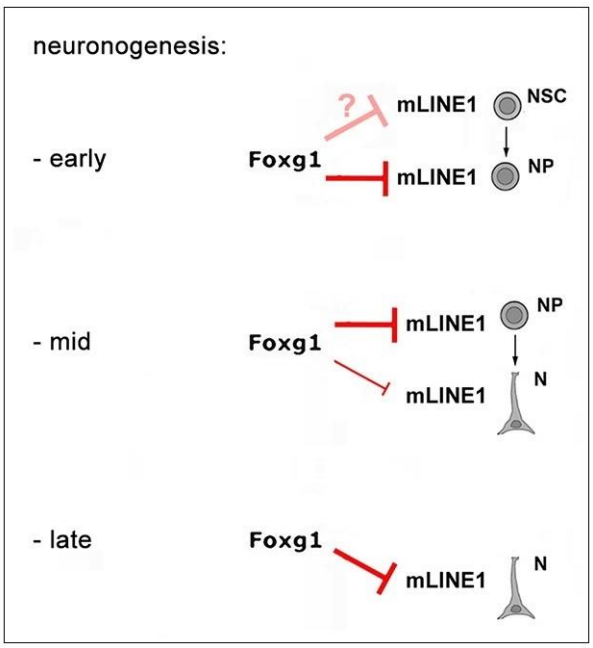


Figure 7. Synopsis of Foxg1 inhibition of *mLINE1* expression within the pallial neuronogenic lineage. Tentative, temporal articulation of Foxg1 control of *mLINE1* expression, based on integrated critical evaluation of Fig. 3-6 data.

4.5. Mechanisms which underlie Foxg1 control of *mLINE1* transcription (Fig. 8,9,10)

Foxg1 is known to often act as a transcriptional repressor (Seoane et al 2004 ; Falcone et al 2019). We wondered if *mLINE1*-mRNA down-regulation exerted by it could originate from direct transcriptional trans-repression. To this aim we set early neuronogenic cultures, wild type as well as *Foxg1*-OE, and we evaluated Foxg1-enrichment at their *mLINE1* loci, by chromatin immuno-precipitation (ChIP) -qPCR. We found that, compared to IgG controls, wild type cultures displayed an increasing Foxg1-enrichment trend at sub-family-specific "*mLINE1*.5'UTR.A, .Gf, and .Tf" amplicons ($p_{5'UTR.A}$<0.14, $p_{5'UTR.Gf}$<0.16, and $p_{5'UTR.Tf}$<0.10, with n=4,4), but not at shared "*mLINE1*.orf2" and "*mLINE1*.3'UTR" amplicons. This trend was detectable and statistically significant at all diagnostic amplicons in *Foxg1*-OE cultures ($p_{5'UTR.A}$<0.014, $p_{5'UTR.Gf}$<0.01, $p_{5'UTR.Tf}$<0.002, p_{orf2}<0.02, $p_{3'UTR}$<0.01, with n=4,4) (**Fig. 8A**).

Similar assays were run on chromatin of mid-neuronogenic cultures. In this case an increasing Foxg1-enrichment trend over IgG controls was detectable at *all* diagnostic amplicons, regardless of "Foxg1 genotype" of cultures, and such enrichment often reached formal statistical significance ("wild-type cultures": $p_{5'UTR.A}$<0.02, $p_{5'UTR.Gf}$<0.09, $p_{5'UTR.Tf}$<0.04, p_{orf2}<0.02, $p_{3'UTR}$<0.02, with *n*=5,5 or - case Gf - *n*=5,4; *Foxg1*-OE cultures: $p_{5'UTR.A}$<0.04, $p_{5'UTR.Gf}$<0.04, $p_{5'UTR.Tf}$<0.06, p_{orf2}<0.06, $p_{3'UTR}$<0.10, with n=4,5 or - case Gf - *n*=4,4) (**Fig. 8B**).

This scenario points to direct *mLINE1* trans-repression by Foxg1. It can be related to prominent *mLINE1* down-regulation occurring in NPs compartment (**Fig. 7**) and large sizing of such compartment, which specifically characterizes *Foxg1*-OE early-neuronogenic cultures (**Fig. 6A** (1)) and *all* their mid-neuronogenic counterparts (**Fig. 6B** (1)).

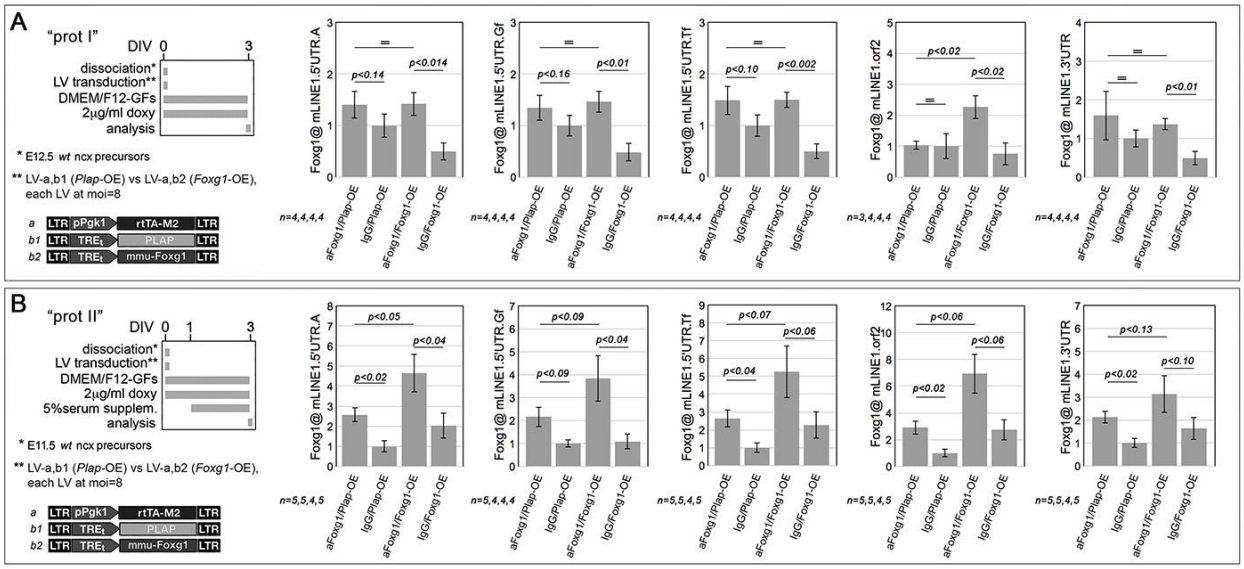


Figure 8. Chromatin immunoprecipitation (ChIP) profiling of Foxg1 protein enrichment at *mLINE1* elements. To left, protocols and lentiviral vectors employed, to right, results. Analysis run in early- (**A**) and mid- (**B**) neuronogenic cultures, set according to type I and II protocols, respectively. Cultures constitutively overexpressing *Foxg1* (or a *Plap* control), driven by the pPgk1 promoter. Chromatin immunoprecipitation performed by anti-Foxg1 antibody and control IgG. PCR quantitation of pan-*LINE1* diagnostic amplicons "*mLINE1*.orf2" and "*mLINE1*.3'UTR", and family-specific "*mLINE1*.5'UTR" amplicons, ".A", ".Gf" and ".Tf". Results normalized against input chromatin. Scalebars, sem's. Statistical evaluation of results by t-test (one-tail, unpaired). *n* = number of biological replicates, i.e. independently cultured and engineered cell aliquots originating from a common neural precursor pool.

It has been shown that dynamics of a number of epigenetic marks can be instrumental to tight control of *mLINE1* transcription (Guler et al., 2017; Bulut-Karslioglu et al., 2014; Rangasamy, 2013; Day et al., 2010; Kim et al., 2014; Protasova et al., 2021, Muotri et al., 2010, He et al., 2019). Moreover, inspection of the public Biogrid database (https://thebiogrid.org) revealed that Foxg1 physically interacts with a number of effectors modulating configurational and conformational state of chromatin, including Histone deacetylase 2 (HDAC2), Lysine-specific demethylase 5B (KDM5B), Lysine-specific demethylase 1A (KDM1A), Chromodomain helicase DNA binding protein 4 (CHD4), SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1 and 5 (SMARCA1 and 5). Therefore, Foxg1 might exert its impact on *mLINE1* transcription just by modulating the epigenetic state of *mLINE1* chromatin. To address this issue, we scored chromatin extracted from from mid-neuronogenic cultures, wild and *Foxg1*-OE, for its enrichment at *mLINE1* loci for a number of key epigenetic markers: H3K4me3, H3K9me3, H3K27ac, MeCP2 (**Fig. 9A,B**).

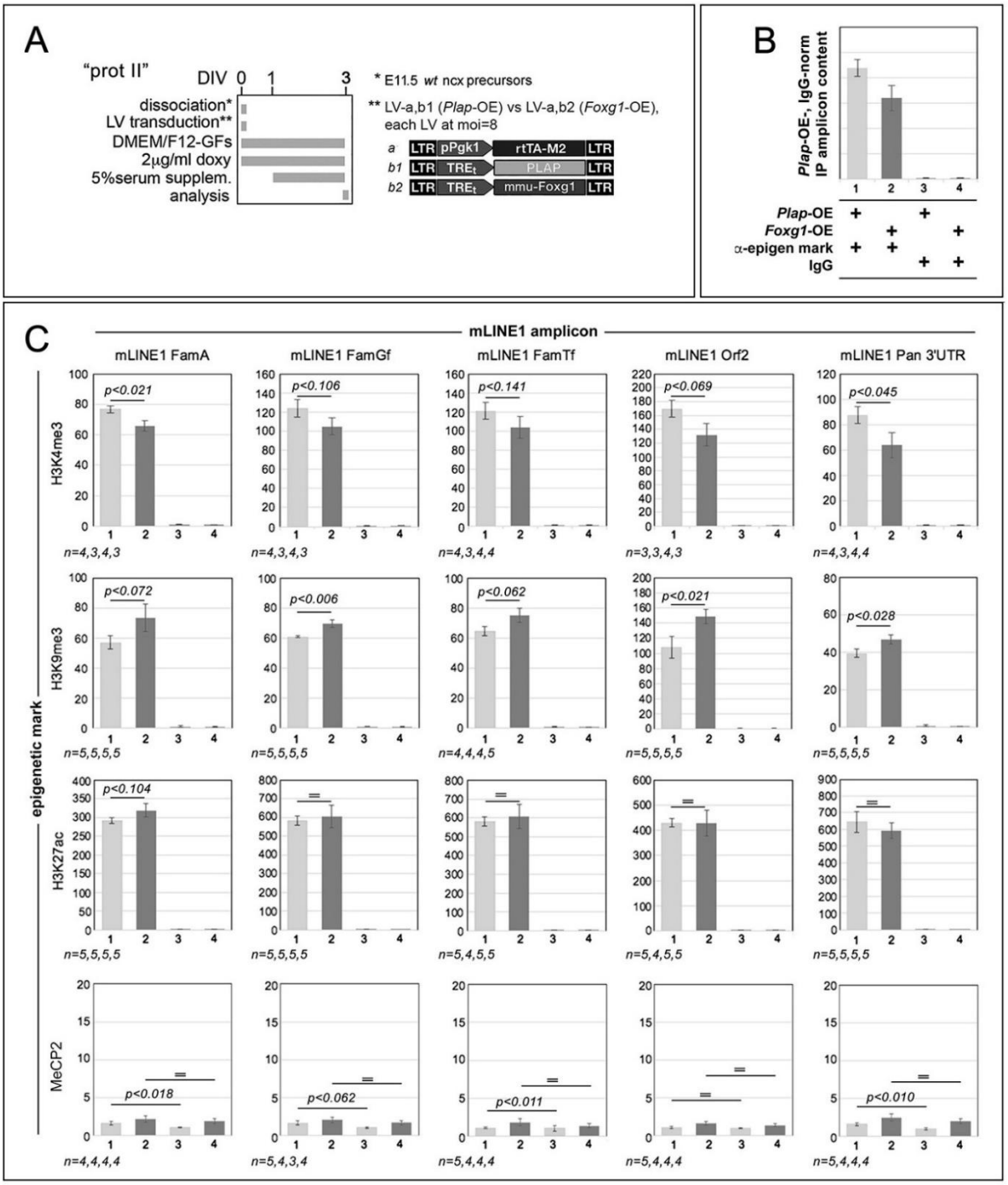


Figure 9. Chromatin immunoprecipitation (ChIP) profiling of H3K4me3, H3K9me3, H3K27ac, and MeCP2 enrichment at *mLINE1* elements. In (**A**), protocols and lentiviral vectors employed, in (**B**) an illustrative prototype of result histograms, in (**C**) results. Analyses run in mid-neuronogenic cultures, set according to a type II protocol. Cultures constitutively overexpressing *Foxg1* (or a *Plap* control), driven by the pPgk1 promoter. Chromatin immunoprecipitation performed by anti- H3K4me3, -H3K9me3, -H3K27ac, and -MeCP2 antibodies and their isotypic IgG controls. PCR quantitation of pan-*LINE1* diagnostic amplicons "*mLINE1*.orf2" and "*mLINE1*.3'UTR", and family-specific "mLINE1.5'UTR amplicons", ".A", ".Gf" and ".Tf". Results normalized against input chromatin. Scalebars, sem's. Statistical evaluation of results by t-test (one-tail, unpaired). *n* = number of biological replicates, i.e. independently cultured and engineered cell aliquots originating from a common neural precursor pool.

We found that H3K4me3, H3K9me3, H3K27ac were remarkably enriched at all diagnostic amplicons analyzed, *mLINE1*.5'UTR.A, .Gf, .Tf, .orf2 and.3'UTR, regardless of culture genotype (normalized against technical IgG controls, enrichments fell within the approximate 50-600 range). Conversely, MeCP2 enrichment was barely appreciable and, compared to IgG controls, it reached formal statistical significance only in a subset of cases, limited to *Plap*-OE control cultures ($p_{5'UTR.A}$<0.02, $p_{5'UTR.Tf}$<0.02, $p_{3'UTR}$<0.01, with *n*=4,4, n=5,4 and n=5,4, respectively). Next, H3K4me3 enrichment was systematically reduced in *Foxg1*-OE compared to *Plap*-OE cultures ($p_{5'UTR.A}$<0.021, $p_{5'UTR.Gf}$<0.106, $p_{5'UTR.Tf}$<0.141, p_{orf2}<0.069, $p_{3'UTR}$<0.045, with *n*=4,3 and - in case of orf2 - n=3,3). Conversely, an opposite dynamic occurred in case of H3K9me3 ($p_{5'UTR.A}$<0.072, $p_{5'UTR.Gf}$<0.006, $p_{5'UTR.Tf}$<0.062, p_{orf2}<0.021, $p_{3'UTR}$<0.028, with *n*=5,5 and - in case of Tf - n=4,4). Finally, H3K27ac did not show any genotype-specific trend (**Fig. 9C**).

Altogether, these results are consistent with the hypothesis that Foxg1 modulation of *mLINE1* transcription is mediated by pervasive changes in the epigenetic state of these elements, namely a decrease of transcription-promoting H3K4me3 marks and an increase of heterochomatic H3K9me3 ones. Low *mLINE1* enrichment for MeCP2 might reflect relatively low expression of this protein at mid-neuronogenic stages (Diez-Roux et al., 2011).

Finally, to shed further light on mechanisms mediating Foxg1 impact on *mLINE1* transcription, we took advantage of the neuropathogenic *FOXG1*^{W308X} allele, encoding for a prematurely truncated protein, missing binding domains for the Groucho/Tle co-repressor and the *KDM5B*-encoded JARID1B H3K4me2/3-demethylase. Delivered to mid-neuronogenic cultures as a Tet^{ON}-driven transgene, *FOXG1*^{W308X} reduced *mLINE1*-mRNA levels by 19.23±3.30% (*p*<0.006; *n*=5,5), 14.46±4.32% (*p*<0.013; *n*=5,5), and 8.40±2.59% (*p*<0.061; *n*=5,5), as evaluated at diagnostic amplicons "orf2", "5'UTR.A" and "5'UTR.Tf", respectively (**Fig. 10**). This down-regulation was far less pronounced compared to the one elicited by wild-type Foxg1, suggesting that the latter could largely stem from Foxg1-dependent recruitment of JARID1B and/or Groucho/Tle to *mLINE1* chromatin (**Fig. 3B**).

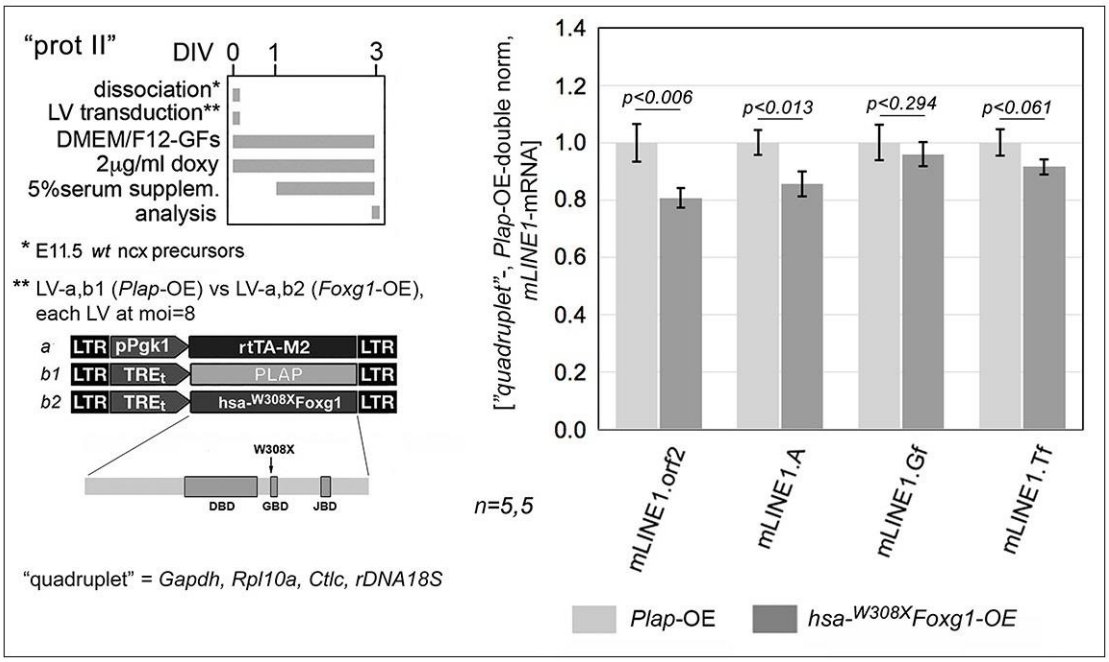


Figure 10. Modulation of *mLINE1*-mRNA levels in murine mid-neuronogenic pallial cultures overexpressing the *hsa-*^{W308X}*Foxg1* mutant allele. To left, protocol and lentiviral vectors employed. Transgenes driven by constitutively firing pPgk1 promoters. The human (hsa) mutant allele in order encodes for a prematurely truncated protein, including the DNA-binding domain (DBD), but not the Groucho- and Jarid-binding domains (GBD and JBD, respectively). RT-PCR quantitation of pan-*LINE1* diagnostic amplicon "*mLINE1*.orf2", and family-specific amplicons, "*mLINE1*.A", "*mLINE1*.Gf" and " *mLINE1*.Tf", in neural cultures set according to a type II protocol. Data double normalized against gene quadruplet (*Gapdh*, *Rpl10a*, *Cltc* and *rDNA 18S*) and control values. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. independently cultured and engineered cell aliquots, originating from pooled, *wild-type* E11.5 neocortical primordia.

4.6. Modeling temporal progression of pallial *mLINE1* DNA copy-numbers in vitro (Fig. 11)

We wondered if, in addition to its inhibition of *mLINE1* transcription, Foxg1 might further impact on retro-transcription of these elements. To get preliminary information about natural dynamics of *mLINE1* DNA within the developing embryonic pallium, we scored early-, mid- and late-neuronogenic cultures for their cumulative *mLINE1* copy number, by SYBR Green qPCR (**Fig. 11A**). We found that this number did not change between early- and mid-neuronogenic cultures, and, compared to the former, it was increased in late-neuronogenic cultures by 35.08±3.65% (p<0.001, n=5.7) (**Fig. 11B**, (1)). Actually, these results were obtained on DNA prepared by an enhanced sample digestion procedure ("high PK"), aimed at extracting DNA with comparable efficacy regardless of the compaction state of chromatin. To assess actual fulfillment of this requirement, an ad hoc test was run a posteriori. Specifically, X-chromosomal, *lyonizable*, *Mecp2* and *Cdkl5* sequences were quantified in DNA extracted from female and male tissues according to "high PK" protocol, normalizing them against the euchromatic *Gfap* autosomal locus.

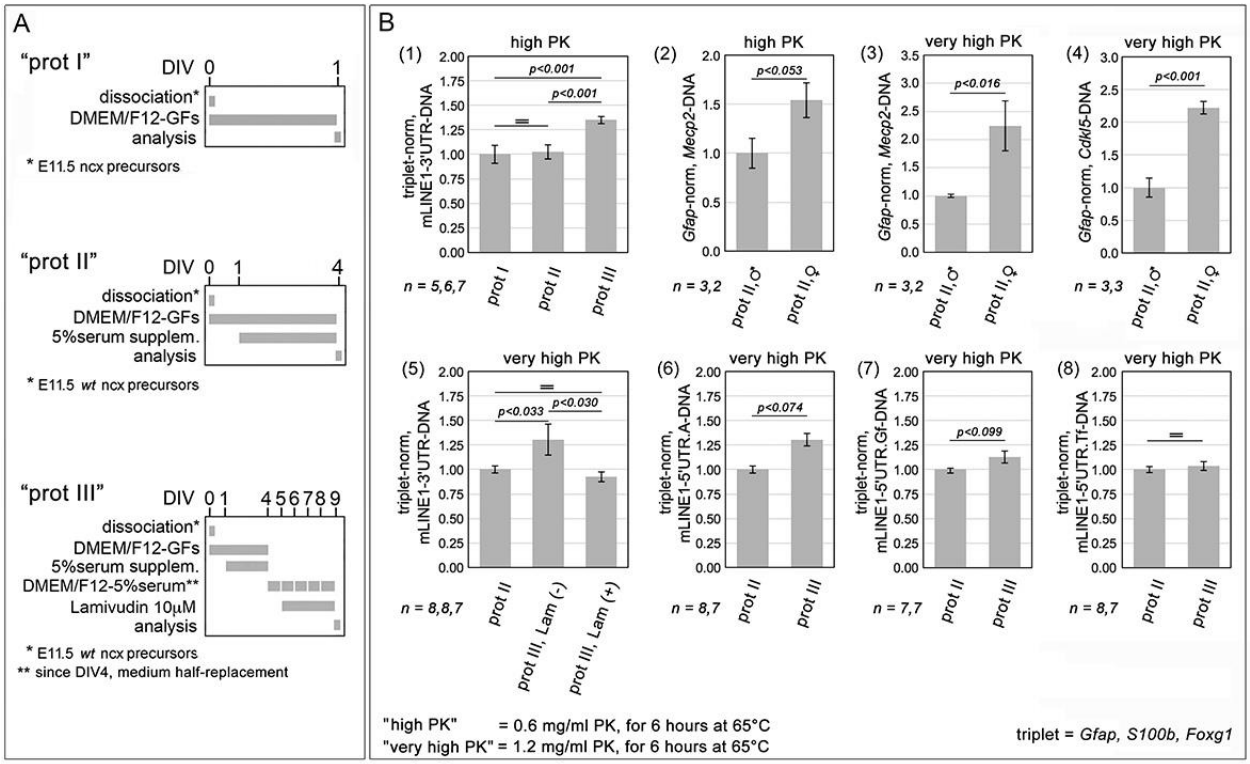


Figure 11. *mLINE1*-DNA copy number variations (CNV) in early- mid- and late-neuronogenic murine pallial cultures. (A) protocols, **(B)** results. Early- mid- and late-neuronogenic cultures set by means of type I, II and III protocols, respectively. DNA extraction performed by a strong ("high PK") or a very strong procedure ("very high PK"). DNA CNVs assessed by quantitative PCR, followed by normalization against endogenous *Gfap*, *S100b* and *Foxg1* (gene "triplet"). In graph (1), shown is the total mLINE1 copy number variation detectable in late- compared to mid-neuronogenic cultures. In graph (5), shown is the suppression of such variation elicited by the pan-RT inhibitor lamivudine. In graphs (6-8) shown are *mLINE1* family-specific DNA copy number variations, as detected by "*mLINE1*.A", "*mLINE1*.Gf" and "*mLINE1*.Tf" diagnostic amplicons. Finally, in graphs (2-4) provided is a comparative assessment of "high PK" and "very high PK" procedures for their capability to extract with similar efficacy genomic DNA, regardless of chromatin accessibility. Such assesment was done by quantifying *Mecp2* and *Cdkl5* sequences, amplified from female and male X chromosome regions, and normalizing them against the euchromatic *Gfap* autosomal locus. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. independently cultured and engineered cell aliquots, originating from pooled, *wild-type* E11.5 neocortical primordia.

As an internal technical control, a similar evaluation was performed upon replacing the "high PK" protocol with a further enhanced version of it, referred to as "very high PK". Interestingly, the normalized female-to-male *Mecp2* signal ratio, equalling 1.54±0.18 in case of "high PK"-extracted DNA ($p_{♂♀}$<0.053, *n*=3,2) (**Fig. 11B**(2)), arose to 2.25±0.45 ($p_{♂♀}$<0.016, *n*=3,2) in case of "very high PK" samples (**Fig. 11B**(3)), close to the theoretically expected value of 2. A similar 2.22±0.10 ratio was also obtained for *Cdkl5* ($p_{♂♀}$<0.001, *n*=3,3) (**Fig. 11B**(4)), suggesting that the "very high PK" protocol may be preferable to the other one and prompting us to adopt it in subsequent assays. To secure **Fig. 11B**(1) results, we re-compared mid- and late-neuronogenic cultures for their *mLINE1*

copy number, upon extracting DNA by the "very high PK" protocol. As a control, to verify that the arousal of *mLINE1* content detected in late cultures originated from somatic retro-transcription of these elements, we included in this last assay late-neuronogenic cultures pre-treated by chronic lamivudine (**Fig. 11A**). "Very high PK" samples replicated the outcome of "high PK" ones, with *mLINE1* copy number increased in late-neuronogenic cultures by 1.31±0.16-folds compared to their mid-neuronogenic counterparts (p<0.04, n=8,8). Remarkably this increase was fully suppressed by lamivudine (**Fig. 11B**(5)). As somatic retro-transcription has been reported to generate new mLINE1 copies missing the 5'UTR, assays referred to in **Fig. 11B**(1,5) relied on the "3'UTR" diagnostic amplicon. In this respect, as a control, we finally compared mid- and late neuronogenic cultures for their mLINE1 content by means of sub-family-specific 5'UTR amplicons. As expected "5'UTR.Gf" and "5'UTR.Tf" signals did not display any change (**Fig. 11B**(7,8)). An increasing trend was conversely detected at the "5'UTR.Gf" amplicon **(Fig. 11B**(6)), suggesting that RT failure might differentially affect LINEs belonging to different families.

4.7. Assessing Foxg1 impact on *mLINE1*-DNA copy numbers, both in vivo and in vitro (Fig. 12, 13, 14).

We have shown that *mLINE1* copy number increases during neocortical neuronogenesis progression, and SybrGreen technology is robustly adequate to detect this phenomenon. To further investigate the role (if any) of *Foxg1* in this process, we compared *mLINE1* copy numbers in neocortices of *Foxg1*$^{-/+}$ neonates and their littermate wild type controls. Normalized against *Gfap* and *Nfia,* such number turned out to be decreased in *Foxg1*-LOF samples by 7.50±0.85% (p<10^{-3}, n=6,8), compared to controls (**Fig. 12**).

Intriguingly, this variation roughly corresponds to 1/3 of the increment in *mLINE1* copies we found associated to neuronogenic progression in wild-type pallial cultures (**Fig. 11B**(1,5)). Moreover, it occurred in mutants characterized by *Foxg1*-mRNA levels reduced by only 33.59±6.17% (p<0.02, n=8,7) compared to wild-type controls (**Table S1**). As such, this copy number decrease may represent a real biological phenomenon and it deserves further attention.

To confirm this prediction, we repeated *mLINE1* copy number evaluation in primary, late-neuronogenic cultures manipulated by CRISPR-Cas9 technology, which allowed us to achieve a more pronounced, -65.37±2.68% (p<10^{-7}, n=8,8) *Foxg1* down-regulation, (**Table S1**).

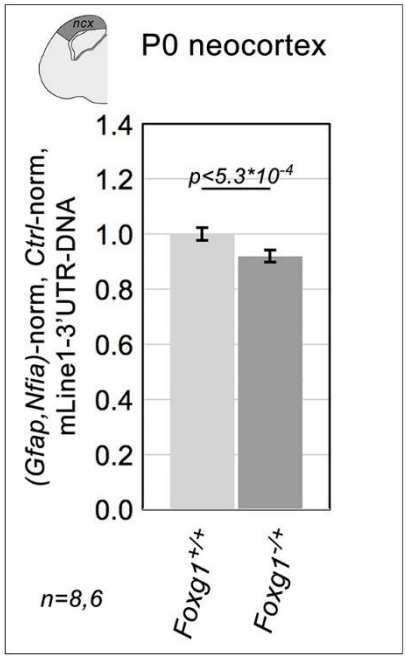


Figure 12. *mLINE1* DNA copy numbers in neocortex of *Foxg1*$^{-/+}$ mouse neonates and controls. PCR quantitation of the pan-*LINE1* diagnostic amplicon "*mLine1*.3'UTR". Data double normalized against the *Gfap/Nfia* gene doublet, and wild-type controls. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. neocortices taken from distinct pups.

Remarkably, in this case, *mLINE1* copy number was reduced by 15.23±2.19% (upon normalization against *Gfap* and *Nfia*, with $p<10^{-3}$, n=8,8) (**Fig. 13B**(2)), corresponding to almost 1/2 of the "physiological" increment referred to above (**Fig. 11B**(1,5)). Altogether, these data point to Foxg1 as a pivotal player promoting *mLINE1* DNA amplification. To note, such retro-transcription-promoting activity emerged despite the concomitant, Foxg1-induced down-regulation of *mLINE1*-mRNA, namely the *starting template* from which new *mLINE1* DNA is generated.

To corroborate these findings and further achieve a more comprehensive picture of Foxg1 role in tuning mLINE1 amplification, we overexpressed it in late-neuronogenic cultures, under the control of three promoters active in neurons, pTa1, pPgk1 and pSyn, and we evaluated the impact of these manipulations on *mLINE1* DNA copy number. We found that such number was increased by pTa1-driven *Foxg1* (37.11±16.35%, with p<0.03, and n=9,9), not affected by pPgk1-driven *Foxg1*, and down-regulated by pSyn-driven *Foxg1* (-20.18±6.51%, with p<0.03, and n=6,9) (**Fig. 13B**(3-5)). These results are puzzling. It is possible that, still consistent with retro-transcription-promoting Foxg1 activity emerging from LOF assays and reflecting different strength of the three promoters employed, such scenario can result from differential sensitivity of *mLINE1* transcription and retro-transcription to actual Foxg1 levels.

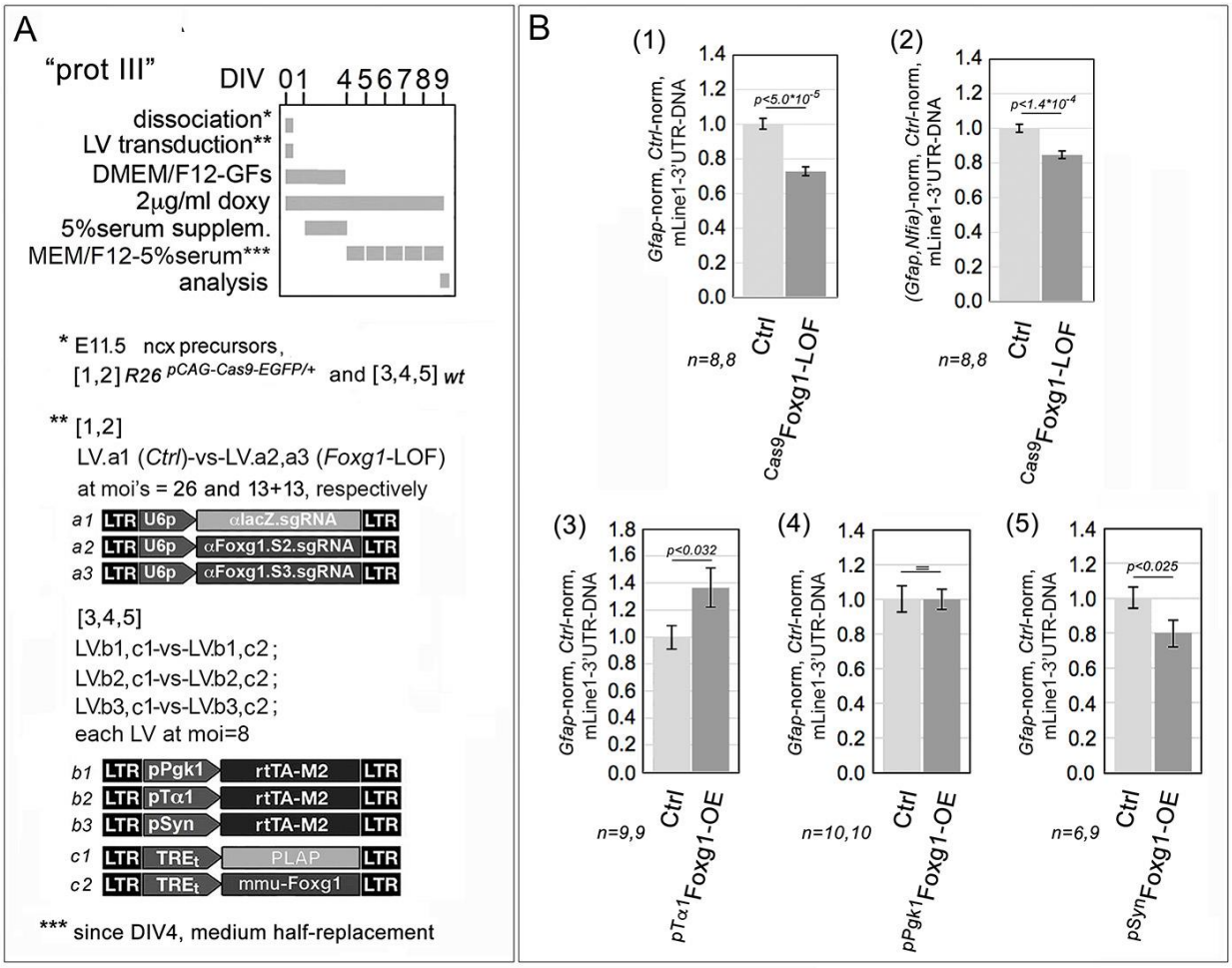


Figure 13. Impact of *Foxg1*-OE and *Foxg1*-knock-down (KD) on *mLINE1* DNA copy numbers in late-neuronogenic pallial cultures. In (**A**), protocols (with lentiviruses employed), in (**B**) results. CRISPR-Cas9 machinery driven by constitutively active U6p (sgRNAs) and R26/pCAG (Cas9), Foxg1 transgene led by constitutive Pgk1p as well as by NP/N-restricted pTa1 and N-restricted pSyn. Neural cultures set according to a "type III" protocol. Assesment of total m*LINE1* copies by quantitation of the "*mLINE*.3'UTR" amplicon. Data double-normalized, against *Gfap* (or the *Gfap/Nfia* gene doublet) and control values. Error-bars representing sem's. Statistical significance of results evaluated by t-test (one-tailed, unpaired). *n* is the number of biological replicates, i.e. independently cultured and engineered cell aliquots, originating from pooled, *R26*$^{pCAG\text{-}Cas9\text{-}EGFP/+}$ (1,2) and *wild-type* (3-5), E11.5 neocortical primordia.

4.8. Hints about putative mechanisms underlying Foxg1 impact on *mLINE1* DNA amplification

We wondered: how could Foxg1 impact on *mLINE1* retro-transcription rates? We considered two possible scenarios: (1) it acts indirectly, as a "professional transcription factor" modulating transcription of key effectors mastering retro-transposition; (2) it is straightly implicated in retro-transcription, through physical interaction with retro-transcription factors and/or *mLINE1*-mRNA.

As for (1), we inspected a database of genes mis-regulated upon Foxg1 overepression in neocortical neuronal cultures (Artimagnella and Mallamaci, 2020; i.e doi:

10.5281/zenodo.2849656). We found that mRNA encoding for Apobec1, a well-known inhibitor of *mLINE1* retro-transposition, is halved in *Foxg1*-OE samples (**Table S3A**). In this way, Foxg1 might promote retro-transposition, mitigating such inhibition.

As for (2), at first, we interrogated the public Biogrid database for Foxg1 interactors implicated in retro-transposition control. We found as many as three established antagonizers of *mLINE1* retro-transposition, namely Mov10, Ddx39a and Trp53, among known Foxg1 interactors (**Table S3B**). Foxg1 might counteract their activity, by sequestering them or preventing them from interacting with *mLINE1*-mRNA. Next, predicting a possible Foxg1/*mLINE1*-mRNA interaction, we run a proof-of-principle RNA-immunoprecipitation (RIP) assay, where we quantified sub-family-specific *mLINE1*-5'UTR.A amplicon in anti-Foxg1-immunoprecipitated RNA. Consistently with our prediction, we found an enrichment of such amplicon, which exceeded IgG controls by more than three-folds ($p<10^{-3}$, n=3,5) (**Fig. 14**).

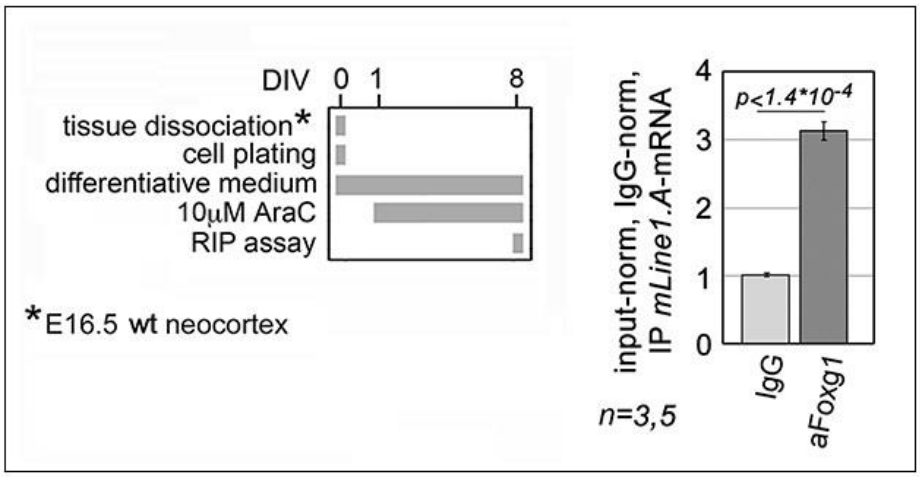


Figure 14. Foxg1-protein/*mLINE1*-mRNA interaction in neocortical neurons, as evaluated by RNA immunoprecipitation-quantitative PCR (qRIP-PCR) assay. To left, protocol, to right, results. Diagnostic amplicon, mLINE1.5'UTR.A. Results double normalized, against input-RNA and IgG-IP samples. Througout figure, n is the number of biological replicates, i.e. independently cultured and engineered preparations, originating from a common neural cell pool. Statistical evaluation of results performed by t-test, one-tailed and unpaired. Errors bars indicate s.e.m.

5. Discussions and Conclusions

Foxg1 masters telencephalic development, exerting a highly pleiotropic control over it. *mLINE1* is a prominent retrotransposon family providing a key contribution to plasticity of neuronal genomic DNA. Remarkably, *mLINE1's* share putative Foxg1 binding elements. Moreover, *mLINE1* transcription was described to be activated in committed progenitors of the postnatal glutamatergic lineage, whereas Foxg1 was reported to be transiently switched off in progenitors of the corresponding embryonic lineage. Based on that, we predicted that Foxg1 might limit *mLINE1* activity. In this study, we tested such prediction by paying special attention to its cellular and molecular articulation. Results can be summarized as follows.

As expected, we found that mLINE1-mRNA encoded by all three retro-transposition-competent subfamilies (A, Gf and Tf) was increased in *Foxg1*-LOF mouse neonates (**Fig. 1**). To model temporal and lineage articulation of *Foxg1*-dependent *LINE1* regulation, we developed an integrated culture-set, representing early-, mid- and late-phases of neuronogenesis in vitro. This set showed a progressive increase of *LINE1* expression from NSCs, through NPs up to neurons (**Fig. 2**). Next, we manipulated Foxg1 levels, downwards and upwards, at different stages of the neuronogenic progression and according to distinctive neural cell type restrictions (**Fig. 3-5**). We mapped the primary molecular outcome of this manipulation to NSCs, NPs and Ns (**Fig. 6**). We finally scored global *LINE1*-mRNA levels resulting from these manipulations (**Fig. 3-5**). Integrated analysis of results provided us robust evidence that Foxg1 represses *LINE1* expression, selectively in NPs and Ns (**Fig. 7**).

We expected that Foxg1 modulation of *mLINE1*-mRNA takes place via direct gene trans-repression. To test this prediction, we investigated Foxg1 recruitment to *mLINE1* chromatin via ChIP. We found a pervasive Foxg1 enrichment at all levels (5', mid and 3') of mLINE1's belonging to all three subfamilies. This enrichment was peculiar to mid-neuronogenic preparations and absent in wild-type early-neuronogenic ones, pointing to poor Foxg1 modulation of *mLINEs* in NSCs. Moreover, where detectable, such enrichment was Foxg1-dose-dependent, suggesting that Foxg1 may physiologically tune *mLINE1* transcription (**Fig. 8**). Intriguingly a neuropathogenic, prematurely truncated FOXG1 protein, unable to interact with Groucho and KDM5B, displayed only an attenuated mLINE1 inhibitory effect (**Fig. 10**), suggesting that, in addition to straight regulation of transcription, Foxg1 may impact LINE1 expression by shaping the epigenetic landscape at *mLINE1* loci. Consistently, Foxg1 overexpression in mid-neuronogenic cultures elicited a robust and pervasive decrease in transcription-promoting marks (H3K4me3) as well as an increase of repressing ones (H3K9me3), at all mLINE1 subfamilies loci (**Fig. 9**).

Next, we moved to Foxg1 control (if any) of *mLINE*-DNA copy number. First we found that *mLINE1*-DNA content was increased by about 25% in Tubb3^{+} post-mitotic neurons, in a retro-transcription-dependent way (**Fig. 11**). Next, unexpectedly, we discovered that Foxg1 down-regulation, in vivo as well as in vitro, elicited a remarkable, co-linear reduction in mLINE1-DNA content, suggesting that Foxg1 is strictly needed to achieve physiological amplification of such DNA

(**Fig. 12,13**). Puzzingly, Foxg1 overexpression led to variable outcomes, depending on different promoters driving it. This points to differential sensitivity of *mLINE1* transcription and retro-transcription to Foxg1 levels (**Fig. 13,15**). So far, mechanisms underlying Foxg1 impact on *mLINE1* retro-transcription were not investigated. However, we got preliminary evidence that such impact can take place via direct Foxg1 protein/*mLINE1*-mRNA interaction (**Fig. 14**).

5.1. Modelling *mLINE1* dynamics in primary pallial cultures

Upon the initial discovery of Foxg1 impact on *mLINE1* expression in the neonatal brain, we worked at developing tools adequate to model the complexity of the underlying phenomena in vitro. Characterized by an early decline of NSCs (from about 40% to 20%) and a progressive increase of neurons (from about 30% up to >70%), "early-", "mid-" and "late-neuronogenic" culturing protocols employed in our study (**Fig. 2A**) provided us a useful in vitro recapitulation of pallial neuronogenesis, suitable for dissecting Foxg1 impact on *mLINE1* dynamics. Inspection of such dynamics in wild type cultures (**Fig. 2B**) prompted us to two main considerations. At first, orf2 signal arousal in "mid-neuronogenic cultures" compared to delayed 5'UTR signal increase in late-neuronogenic preparations suggests that 5'-truncated *mLINE1* elements (Voliva et al, 1983) can substantially contribute to the difference between to the former and the latter. Embedded within other transcription units (Bodak et al. 2014), a large subset of these elements may reflect the expression trend of such units, upregulated during early neuronogenic progression. Next, mRNA expression gains undergone by individual *mLINE1* subfamilies from early to late neuronogenic stages (about 50, 30 and 70 in case of A, Gf and Tf subfamilies) are colinear with numbers of transpositionally-competent units belonging to each family (900, 400, and 1,800, respectively (Schichman et al., 1993), however not proportional to them. This suggests that sub-family-specific regulatory mechanisms are in place.

5.2. Foxg1 dampens *mLINE1*-mRNA expression in the developing rodent pallium

Comprehensive analysis of complex transcriptional outcomes of *Foxg1* manipulation in primary pallial cultures (**Fig. 3-5** and **10**) shows a pervasive inhibitory effect elicited by of this factor on *mLINE1*-mRNA expression.

To note, *mLINE1*-mRNA upregulation detectable in neonatal *Foxg1*-LOF brains peaked in Gf subfamily (**Fig. 1**), namely the very same displaying the narrowest increase in late-neuronogenic wild type cultures (**Fig. 2B**(2)). This suggests a possible involvement of Foxg1 in late selective silencing of Gf elements. However, Foxg1-engineered cultures did not clearly replicate sub-family dynamics peculiar to *Foxg1*-LOF brains (**Fig. 3**) (possibly because of their intrinsic developmental variability). That suggests caution on this inference.

Concerning Foxg1 impact on *mLINE1* expression, we were interested in identifying distinct neural types (NSCs, NPs, Ns) where such phenomenon occurs. To this aim, we modulated Foxg1 levels in cultures representing more and more advanced phases of neuronogenesis, therefore differentially enriched in these types (**Fig. 3-5**). In case of OE manipulations, we also delivered them by distinctive, neural cell type-specific promoters (pPgk1, pNes, pTa1, pSyn) (**Fig. 4**). Next, we investigated relationships between amplitudes of *mLINE1* mis-regulation achieved and culture engineering protocols employed. In this way, we straightly obtained robust evidence of *Foxg1*-driven mLINE1 down-regulation specifically taking place in neurons (**Fig. 7C**), namely the "terminal" point of the NSC→NP→N progression. Unfortunately, this approach was not sufficient to take reliable conclusions about NSCs and NPs, i.e. the "transient" states of such progression. In fact, we reasoned that, when employing promoters which fire in NSCs and NPs, the upregulation of Foxg1 protein can occur in a cell type *different* from the one where the promoter fires. Moreover, *Foxg1* manipulation can alter the relative prevalence of these types. To address intricacies originating from these issues, we additionally evaluated *actual* Foxg1 protein levels within NSC, NP and N compartments of differently engineered cultures, and we quantified sizes of such compartments (**Fig. 6**). Finally, we compared results of this analysis with global *mLINE1*-mRNA dynamics peculiar to the corresponding cultures. This integrated approach allowed us to get safe evidence of *Foxg1*-driven *mLINE1* downregulation in NPs (**Fig. 7B**). As for NSCs, we were not able to take definitive conclusions (**Fig. 7A**).

Moreover, we were interested in molecular mechanisms mediating the negative impact exerted by Foxg1 on *mLINE1*-mRNA levels. In this respect, Foxg1 recruitment to mLINE1 transcription units (at 5'UTR, orf2 and 3'UTR) points to a direct transcriptional inhibition exerted by the former on the latter (**Fig. 8**). Intriguingly, such recruitment was pervasively detectable in mid-neuronogenic wild type cultures (**Fig. 8B**), including about 90% of NPs+Ns and only 10% of NSCs (**Fig. 6A**(1)), but not in early-neuronogenic wild type ones (**Fig. 8A**), characterized by a specular distribution of NPs+Ns and NSCs (**Fig. 6B**(1),**C**(1)). This suggests that Foxg1 regulation of *mLINE1* takes place in NPs and Ns, but not in NSCs. [In this respect, the appearence of an aFoxg1-ChIP signal in early-neuronogenic *Foxg1*-OE cultures (**Fig. 8A**) likely reflected exaggerated Foxg1 levels available in such cultures and/or *Foxg1*-driven enlargment of their NP compartment (reaching here about 1/3 of neural population) (**Fig. 6A**(1))]. In this way, by further integrating results of *mLINE1*-mRNA expression analysis, quantitative profiling of Foxg1-dependent NSC, NP and N compartments' dynamics, and ChIP analysis of Foxg1 recruitment at *mLINE1* loci, we can infer that Foxg1 inhibition of *mLINE1* transcription occurs in NPs and Ns, but not in NSCs.

Next, we noticed that, within *Foxg1*-OE mid-neuronogenic cultures, differences among aFoxg1 and IgG ChIP signals showed a consistent upwards trend at all diagnostic amplicons tested, compared to their *Plap*-OE controls (**Fig. 8B**). This suggests that Foxg1 concentration dictates recruitment rates of such protein at *mLINE1* loci and, consequently, mLINE1 transcription. This is fully consistent with bi-directional responses displayed by *mLINE1*-mRNA levels to *Foxg1*-LOF and -OE manipulations (**Fig. 3**), further enforcing physiological *Foxg1* role in tuning of *mLINE1* transcription.

According to Jaspar 2016, the consensus FOXG1-binding matrix (RTAAACAW, MA0613.1) largely overlaps the corresponding FOXA1 one (WRWRTAAAYA, MA0047.2). Remarkably, FOXA1 is a key TF sitting on *LINE1*-5'UTR and promoting *LINE1* transcription in senescent cells (De Cecco et al., 2019). It is tempting to speculate that, during embryonic life, in the absence of FoxA1 expression, Foxg1 might take advantage of such matrices to exert its trans-repressing role. Beyond their transient impact on transcription, transcription factors may elicit a more long-lasting effect on gene expression by mediating the recruitment of specific epigenetic effectors to chromatine. These effectors include enzymes modulating DNA methylation (Hervouet et al., 2010; Pacaud et al., 2014), histone methylation (Arnold et al., 2013), histone acetylation (Mokrani et al., 2006; Weill et al., 2003), chromatine conformation (Muratcioglu et al., 2015; Wang et al., 1995). This could apply also to Foxg1 control of *mLINE1* transcription. Actually, stable *mLINE1* repression requires both 5methyl-cytosine (5mC) (Muotri et al 2010) and H3K9me3 (Bulut-Karslioglu et al., 2014). Inspired by Biogrid Foxg1 interaction data (as detailed in Results), we profiled Foxg1-OE cultures and controls for H3K4me3, H3K9me3, H3K27ac, and MeCP2 enrichment at *mLINE1* loci. Consistently with Foxg1 *mLINE1*-mRNA downregulation, we found a pervasive depletion in H3K4me3 and a widepsread enrichment in H3K9me3, both more prominent at diagnostic amplicon orf2 (**Fig. 9**). High H3K27ac levels were detected in both controls and *Foxg1*-OE samples (**Fig. 9**), pointing to a bivalent state of chromatine, amenable to both silencing and transcription (He et al., 2019). Unexpectedly, MeCP2 enrichment was generally very low (**Fig. 9**), likely because of transient decline of this protein at mid-neuronogenic stages (Diex-Roux et al., 2011). Intriguingly, *mLINE1*-mRNA downregulation driven by the neuropathoathogenic *FOXG1*^{W308X} allele (missing both GBD and JBD) was milder than *mLINE1* repression elicited by the wild type allele, (expecially in case of Gf family) (**Fig. 3B** and **10**), suggesting that pleiotropic mechanisms are needed to achieve full Foxg1 control of *mLINE1* transcription.

5.3. Foxg1 increases *mLINE1*-DNA copy number within the rodent developing pallium.

We have found that, while moving from mid- to late-neuronogenic cultures, the *mLINE1* DNA content is augmented by >25% and this phenomenon is suppressed by lamivudine (**Fig. 11B** (1,5)). That suggests that a net *mLINE1* DNA amplification takes place in neurons and it requires retro-transcription. Remarkably, this phenomenon is not restricted to the 3'UTR diagnostic amplicon, but, in case of subfamily A, it extends to 5'UTR (**Fig. 11B** (6,7,8)), further suggesting that TEs belonging to this subfamily may be more efficiently retro-transcribed. Next, we found that heterozygosity for a *Foxg1-null* allele (resulting into a -35% variation of *Foxg1*-mRNA) decreases neonatal *mLINE1*-DNA content by about 8% (**Fig. 12**), a more drastic *Foxg1* down-regulation by CRISPR-Cas9 (leading to a 70% decline of *Foxg1*-mRNA) diminishes *mLINE1*-DNA by about 17% (**Fig. 13** (1,2)). Such linear relationship between *Foxg1*-mRNA and *mLINE1*-DNA as well as the fact that 17% almost equals 2/3 of the increment undergone by *mLINE1*-DNA in wild-type pallial neurons, both suggest that Foxg1 is necessary for net neuronal amplification of *mLINE1*-DNA. Puzzlingly, when we overepressed Foxg1, we observed a complex effect on *mLINE1*-DNA content, that was increased, unaffected or

decreased, depending on the culturing schedule and the promoter driving the Foxg1 transgene (**Fig. 13B** (3,4,5)). Such scenario could originate from complex Foxg1 impact on the two tandemly arranged processes needed for *mLINE1* amplification, namely transcription and retro-trascription (**Fig. 15A-C**).

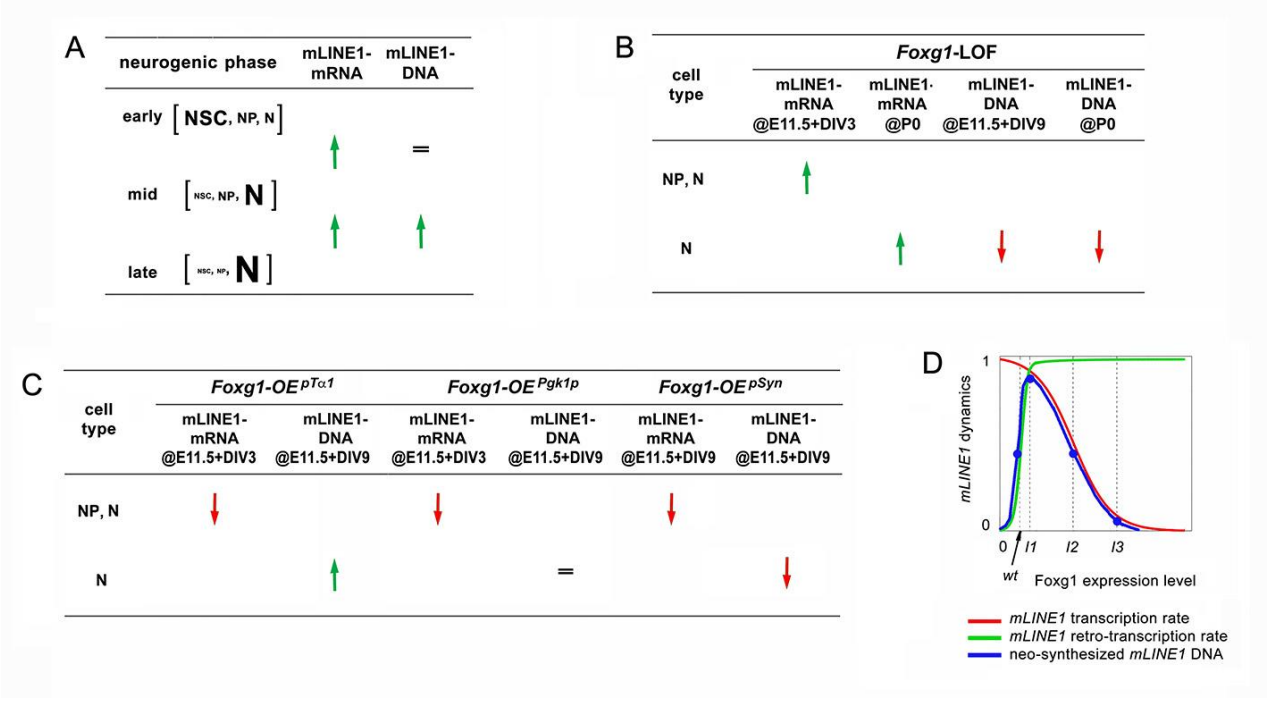


Figure 15. Foxg1 promotion of *mLINE1* retro-transcription in neuronal progenitors and neurons, inferred on the basis of primary Fig. 1-4,11-13 data. (A) Normal progression of *mLINE1*-mRNA and -DNA copy number in differentiating pallial cultures, according to Fig. 2 and 11 data. **(B)** Comparative impact of *Foxg1* down-regulation on *mLINE1*-mRNA and -DNA copy number, as emerging from Fig. 1, 3, 12 and 13 data. **(C)** Comparative impact of *Foxg1* up-regulation (constitutive or restricted to specific neuronogenic cell types) on *mLINE1*-mRNA and -DNA copy number, based on Fig. 3, 4 and 13 results. **(D)** Graphical representation of a tentative model accounting for not-monotonic impact of *Foxg1* expression levels on final DNA output of *mLINE1* retrotranscription. In red and green, hypothetical functional relationships taking place between Foxg1 levels and cumulative, transcriptase and retro-transcriptase activities available for *mLINE1*-DNA amplification. In blu, the hypothetical *mLINE1*-DNA output. This was assumed to be linearly related to the product of the two aforementioned enzymatic activities.

Specifically, we propose that even modest fluctuations of Foxg1 levels around the baseline elicit prominent and similarly oriented effects on retro-transcription, whereas a far more pronounced upregulation of it over the baseline is necessary to get maximal inhibition of transcription. In this way, the use of distinct promoters may lead to different *mLINE1*-DNA outputs, straightly depending on promoters' strength (**Fig. 15D**). To note, our evaluation of *mLINE*-DNA content was performed in bulk, i.e, not distinguishing among TEs integrated in chromosomes and free TE molecules, double or single-stranded, floating in nucleus or cytoplasm. Hence, changes in *mLINE1*-DNA signal we detected in aging wild type cultures as well as in *Foxg1*-engineered mice and cultures may represent different underlying phenomena, of distinct biological value and neuropathological relevance (Thomas et al., 2017; Gasior et al., 2006; Suarez et al., 2017). We plan

to address these issues by integrated employment of retrotransposon-capture/genomic DNA sequencing (Sanchez-Luque et al., 2016; Moore and Haber, 1996), fractioning of nuclear and cytosolic DNA (Tilgner et al., 2012; Mishra et al., 2011), and immuno/enzymatic selection of ss- and ds-DNA molecules (Welmann et al., 2005; Rekvig, 2022).

Finally, concerning molecular mechanisms underlying Foxg1 impact on m-*LINE1*-DNA amplification, they could rely on enhanced retro-transcription and/or increased stability of neo-generated DNA copies. We presently ignore such mechanisms. However, by inspection of public data, we found that Foxg1 trans-represses *Apobec3*, and chelates Mov10, Ddx39a and Trp53, suggesting it might prevent the negative impact that these four effectors exert on *mLINE1* retro-transcription (see Table S3 and references therein). Moreover, we got preliminarly evidence that Foxg1 binds to *mLINE1*-5'UTR in wild type pallial neurons. This points to a likely direct regulatory effect of it on *mLINE1*-mRNA metabolism. Starting from these hints, such issue will be the subject of a dedicated follow-up study.

5.4. Conclusions and significance

We have shown that Foxg1 robustly inhibits pallial *mLINE1*-mRNA expression, by binding *mLINE1* chromatin and exerting a deep impact on its epigenetic state. In this way, Foxg1 adds to a specific TF-set controlling *mLINE1* transcription in CNS. In this context, it recalls Sox2. However, Sox2 represses *LINE1* transcription in NSCs, whether Foxg1 is mainly active in neuronal progenitors and neurons. Moreover, Sox2 is expressed in the apical compartment of the whole neuraxis, Foxg1 is mainly confined to the telencephalon. In this respect, we presently ignore both *mLINE1* dynamics outside telencephalon and *LINE1*-repressing mechanisms (if any), possibly vicariating Foxg1 in more caudal CNS.

Unexpectedly, we also discovered that Foxg1 is needed to achieve the somatic upregulation of *mLINE1*-DNA copy number which normally occurs in developing pallial tissue. We presently ignore the underlying mechanisms. Physical interaction between Foxg1 protein and *mLINE1*-mRNA might be instrumental to such phenomenon.

Beyond its physiological articulation in the developing rodent embryo, the relationship between *FOXG1* and *LINEs* might be implicated in etiopathogenesis of the human FOXG1 syndrome. Deficient FOXG1 activity linked to hemizygosity or heterozygosity for LOF-alleles might lead to *LINE1*-mRNA upregulation, supernumerary or GOF *FOXG1* alleles might result in exaggerated *LINE1*-DNA neo-synthesis. Both scenarios are of potential neuropathogenic interest. However, our study was mainly run-on mouse, not human Foxg1. Moreover, the number of transcriptionally autonomous/retro-transpositionally competent human *LINE1* elements is about only 1/20 of its murine counterpart. For these reasons, it is still presently hard to predict relevance of these phenomena to human neurological health.

Finally, the impressive similarity between FOXG1- and FOXA1-binding DNA matrices raises the possibility that, within the neonatal/juvenile brain, a moderate availability of the former may

competitively antagonize the activity of the latter, hence dampening *LINE1* expression, limiting inflammation, and postponing senescence.

6. Supplementary Material

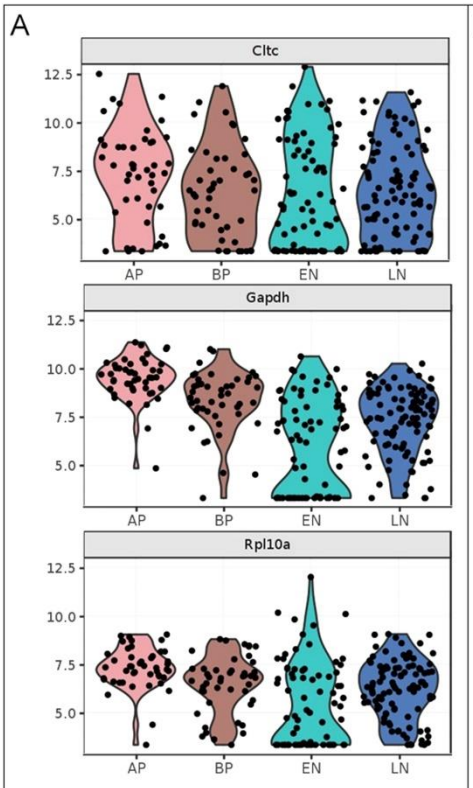
Table S1

ref		*Foxg1-mRNA manipulation*	
Fig. #	*panel, graph*	*ctr-norm Foxg1-mRNA levels*	*normalizer(s)*
1	---	0,64	*Gapdh*
3	A, graph 1	0.22	*Gapdh*
	A, graph 2	0.64	*gene quadruplet**
	B, graph 1	2.83	*Gapdh*
	B, graph 2	2.84	*gene quadruplet**
4	A	2.74	*gene quadruplet**
	B	3,63	*gene quadruplet**
	C	4.02	*gene quadruplet**
	D	NA	*gene quadruplet**
	E	NA	*gene quadruplet**
6	graph 1	4.84	*Gapdh*
	graph 1.1	4.27	*Rpl10a*
	graph 2	0.28	*Rpl10a*
10	---	14.08	*gene quadruplet**
12	---	0.66	*Gapdh*
13	graphs 1,2	0.35	*Gapdh*
	graph 3	NA	NA
	graph 4	3.05	*Gapdh*
	graph 5	NA	NA

**(Gapdh; Rpl10a; Cltc and rDNA 18S)*

Table S1. Foxg1-mRNA dynamics upon OE- and LOF-manipulations of this study. Here, reported are control-normalized *Foxg1*-mRNA levels in key experiments of this study, each with reference Figures and panels, and normalizer gene(s) employed for their evaluation.

A



http://genebrowser.unige.ch/science2016
doi:10.1126/science.aad8361

B

Locus	log2FC	logCPM	P-value	FDR	Gene Name
ENSMUSG00000047126	0,01	9,73	0,94	0,96	Cltc

Locus	log2FC	logCPM	P-value	FDR	Gene Name
ENSMUSG00000057666	-0,16	9,44	0,15	0,24	Gapdh

Locus	log2FC	logCPM	P-value	FDR	Gene Name
ENSMUSG00000037805	0,07	6,50	0,68	0,77	Rpl10a

doi:10.5281/zenodo.3739467

Table S2. "Quadruplet" RNA-pol II-transcribed normalizers: fluctuations of their mRNA levels throughout normal neuronogenesis as well as upon *Foxg1* manipulation in differentiating neurons. (A) Violin plots representing single cell mRNA levels, in pallial apical precursors (AP), basal precursors (BP), early neurons (EN) and late neurons (LN). Downloaded from http://genebrowser.unige.ch/science 2016 (Telley et al., 2016). **(B)** Modulation of average mRNA levels in differentiating neuronal cultures overexpressing *Foxg1* (Artimagnella and Mallamaci, 2020 i.e doi:10.5281/zenodo.3739467).

A

Foxg1-responsive gene (ref 1)	*impact on retro-transcription/transposition*	*mRNA dynamics upon Foxg1 overexpression in ncx neurons*				
		Locus	*log2FC*	*logCPM*	*P-value*	*FDR*
Apobec1	limits Line1 retrotransposition (ref 2,3,4)	ENSMUSG00000040613	-1.01	2.02	0.00	0.00

B

Foxg1 interactor (ref 5)	*impact on retro-transcription/transposition*	*Foxg1-protein interaction details (Biogrid #)*
MOV10	limits Line1 retrotransposition (ref 6)	1507467
DDX39A	limits Line1 retrotransposition (ref 6)	1507432
TRP53	limits Line1 retrotransposition (ref 7,8)	1507470

refs

1 doi: 10.5281/zenodo.2849656.
2 doi: 10.1093/nar/gkr124
3 doi: 10.1146/annurev.immunol.26.021607.090350
4 doi: 10.1073/pnas.0603313103
5 https://thebiogrid.org/
6 doi: 10.1371/journal.pgen.1002941
7 doi: 10.1101/gad.343186.120
8 doi: 10.1038/onc.2009.246

Table S3. Putative mediators of Foxg1 impact on *mLINE1* retro-transcription. Shown are selected genes impacting on *mLINE1* retro-transcription/transposition rates, (**A**) whose mRNA levels are sensitive to *Foxg1* overexpression, or (**B**) whose protein products physically interact with Foxg1 protein.

References

Aboitiz, F., Montiel, J. and López, J., 2002. Critical steps in the early evolution of the isocortex: Insights from developmental biology. Brazilian Journal of Medical and Biological Research, 35(12), pp.1455-1472.

Akol, I., Gather, F. and Vogel, T., 2022. Paving Therapeutic Avenues for FOXG1 Syndrome: Untangling Genotypes and Phenotypes from a Molecular Perspective.

Alisch, R., Garcia-Perez, J., Muotri, A., Gage, F. and Moran, J., 2006. Unconventional translation of mammalian LINE-1 retrotransposons. Genes & Development, 20(2), pp.210-224.

Amor, D., Burgess, T., Tan, T. and Pertile, M., 2012. Questionable pathogenicity of FOXG1 duplication. European Journal of Human Genetics, 20(6), pp.595-596.

Anderson, S., 2002. Distinct Origins of Neocortical Projection Neurons and Interneurons In Vivo. Cerebral Cortex, 12(7), pp.702-709.

ANGEVINE, J. and SIDMAN, R., 1961. Autoradiographic Study of Cell Migration during Histogenesis of Cerebral Cortex in the Mouse. Nature, 192(4804), pp.766-768.

Anthony, T., Klein, C., Fishell, G. and Heintz, N., 2004. Radial Glia Serve as Neuronal Progenitors in All Regions of the Central Nervous System. Neuron, 41(6), pp.881-890.

Aravin, A., Hannon, G. and Brennecke, J., 2007. The Piwi-piRNA Pathway Provides an Adaptive Defense in the Transposon Arms Race. Science, 318(5851), pp.761-764.

Aravin, A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K., Bestor, T. and Hannon, G., 2008. A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice. Molecular Cell, 31(6), pp.785-799.

Aravin, A., van der Heijden, G., Castañeda, J., Vagin, V., Hannon, G. and Bortvin, A., 2009. Cytoplasmic Compartmentalization of the Fetal piRNA Pathway in Mice. PLoS Genetics, 5(12), p.e1000764.

Archibald, J., 2006. Zofia Kielan-Jaworowska, Richard L. Cifelli, and Zhe-Xi Luo, Mammals from the Age of Dinosaurs Origins, Evolution, and Structure Columbia University Press, New York, 2004, 630 pp., $195 (hard cover). ISBN 0-231-11918-6. Journal of Mammalian Evolution, 13(2), pp.147-149.

Ariumi, Y., 2016. Guardian of the Human Genome: Host Defense Mechanisms against LINE-1 Retrotransposition. Frontiers in Chemistry, 4.

Arlotta, P., Molyneaux, B., Chen, J., Inoue, J., Kominami, R. and Macklis, J., 2005. Neuronal Subtype-Specific Genes that Control Corticospinal Motor Neuron Development In Vivo. Neuron, 45(2), pp.207-221.

Arnold, P., Schöler, A., Pachkov, M., Balwierz, P., Jørgensen, H., Stadler, M., van Nimwegen, E. and Schübeler, D., 2012. Modeling of epigenome dynamics identifies transcription factors that mediate Polycomb targeting. Genome Research, 23(1), pp.60-73.

Artimagnella, O. and Mallamaci, A., 2020. COMPARATIVE TRANSCRIPTOMIC PROFILING OF MURINE, Foxg1-GOF E16.5+DIV8 NEOCORTICAL NEURONS AND CONTROLS. *Laboratory of Cerebral Cortex Development, SISSA, Trieste, Italy*.

Athanikar, J., 2004. A YY1-binding site is required for accurate human LINE-1 transcription initiation. Nucleic Acids Research, 32(13), pp.3846-3855.

Baillie, J., Barnett, M., Upton, K., Gerhardt, D., Richmond, T., De Sapio, F., Brennan, P., Rizzu, P., Smith, S., Fell, M., Talbot, R., Gustincich, S., Freeman, T., Mattick, J., Hume, D., Heutink, P., Carninci, P., Jeddeloh, J. and Faulkner, G., 2011. Somatic retrotransposition alters the genetic landscape of the human brain. Nature, 479(7374), pp.534-537.

Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F. and Bourc'his, D., 2016. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science, 354(6314), pp.909-912.

Beck, C., Garcia-Perez, J., Badge, R. and Moran, J., 2011. LINE-1 Elements in Structural Variation and Disease. Annual Review of Genomics and Human Genetics, 12(1), pp.187-215.

Becker, K., Swergold, G., Ozato, K. and Thayer, R., 1993. Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element. Human Molecular Genetics, 2(10), pp.1697-1702.

Belancio, V., Deininger, P. and Roy-Engel, A., 2009. LINE dancing in the human genome: transposable elements and disease. Genome Medicine, 1(10), p.97.

Benton, M., 2015. Palaeodiversity and formation counts: redundancy or bias?. Palaeontology, 58(6), pp.1003-1029.

Benton, M., Donoghue, P., Vinther, J., Asher, R., Friedman, M. and Near, T., 2015. Constraints on the timescale of animal evolutionary history. Palaeontologia Electronica,.

Bertossi, C., Cassina, M., De Palma, L., Vecchi, M., Rossato, S., Toldo, I., Donà, M., Murgia, A., Boniver, C. and Sartori, S., 2014. 14q12 duplication including FOXG1: Is there a common age-dependent epileptic phenotype?. Brain and Development, 36(5), pp.402-407.

Bertsch McGrayne, S. and Vorburger Mielczarek, E., 1993. Nobel Prize Women in Science: Their Lives, Struggles, and Momentous Discoveries. American Journal of Physics, 61(12), pp.1155-1155.

Bodak, M., Cirera-Salinas, D., Yu, J., Ngondo, R. and Ciaudo, C., 2017. Dicer, a new regulator of pluripotency exit and LINE-1 elements in mouse embryonic stem cells. FEBS Open Bio, 7(2), pp.204-220.

Bodak, M., Yu, J. and Ciaudo, C., 2014. Regulation of LINE-1 in mammals. Biomolecular Concepts, 5(5), pp.409-428.

Boeke, J., Garfinkel, D., Styles, C. and Fink, G., 1985. Ty elements transpose through an RNA intermediate. Cell, 40(3), pp.491-500.

Bogerd, H., Wiegand, H., Hulme, A., Garcia-Perez, J., O'Shea, K., Moran, J. and Cullen, B., 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. Proceedings of the National Academy of Sciences, 103(23), pp.8780-8785.

Boissinot, S. and Sookdeo, A., 2016. The Evolution of Line-1 in Vertebrates. Genome Biology and Evolution, p.evw247.

Boissinot, S., Bourgeois, Y., Manthey, J. and Ruggiero, R., 2019. The Mobilome of Reptiles: Evolution, Structure, and Function. Cytogenetic and Genome Research, 157(1-2), pp.21-33.

Bond, A., Bhalala, O. and Kessler, J., 2012. The dynamic role of bone morphogenetic proteins in neural stem cell fate and maturation. Developmental Neurobiology, 72(7), pp.1068-1084.

Borrell, V., 2019. Recent advances in understanding neocortical development. F1000Research, 8, p.1791.

Bourc'his, D. and Bestor, T., 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature, 431(7004), pp.96-99.

Brancaccio, M., Pivetta, C., Granzotto, M., Filippis, C. and Mallamaci, A., 2010. Emx2 and Foxg1 Inhibit Gliogenesis and Promote Neuronogenesis. Stem Cells, 28(7), pp.1206-1218.

Branco, M., Ficz, G. and Reik, W., 2011. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. Nature Reviews Genetics, 13(1), pp.7-13.

Brent, M., Anand, R. and Marmorstein, R., 2008. Structural Basis for DNA Recognition by FoxO1 and Its Regulation by Posttranslational Modification. Structure, 16(9), pp.1407-1416.

Brunetti-Pierri, N., Paciorkowski, A., Ciccone, R., Mina, E., Bonaglia, M., Borgatti, R., Schaaf, C., Sutton, V., Xia, Z., Jelluma, N., Ruivenkamp, C., Bertrand, M., de Ravel, T., Jayakar, P., Belli, S., Rocchetti, K., Pantaleoni, C., D'Arrigo, S., Hughes, J., Cheung, S., Zuffardi, O. and Stankiewicz, P., 2010. Duplications of FOXG1 in 14q12 are associated with developmental epilepsy, mental retardation, and severe speech impairment. European Journal of Human Genetics, 19(1), pp.102-107.

Bulstrode, H., Johnstone, E., Marques-Torrejon, M., Ferguson, K., Bressan, R., Blin, C., Grant, V., Gogolok, S., Gangoso, E., Gagrica, S., Ender, C., Fotaki, V., Sproul, D., Bertone, P. and Pollard, S., 2017. Elevated FOXG1 and SOX2 in glioblastoma enforces neural stem cell identity through transcriptional control of cell cycle and epigenetic regulators. Genes & Development, 31(8), pp.757-773.

Bulut-Karslioglu, A., De La Rosa-Velázquez, I., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galán, C., Winter, G., Engist, B., Gerle, B., O'Sullivan, R., Martens, J., Walter, J., Manke, T., Lachner, M. and Jenuwein, T., 2014. Suv39h-Dependent H3K9me3 Marks Intact Retrotransposons and Silences LINE Elements in Mouse Embryonic Stem Cells. Molecular Cell, 55(2), pp.277-290.

Bulut-Karslioglu, A., De La Rosa-Velázquez, I., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galán, C., Winter, G., Engist, B., Gerle, B., O'Sullivan, R., Martens, J., Walter, J., Manke, T., Lachner, M. and Jenuwein, T., 2014. Suv39h-Dependent H3K9me3 Marks Intact Retrotransposons and Silences LINE Elements in Mouse Embryonic Stem Cells. Molecular Cell, 55(2), pp.277-290.

Bulut-Karslioglu, A., De La Rosa-Velázquez, I., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galán, C., Winter, G., Engist, B., Gerle, B., O'Sullivan, R., Martens, J., Walter, J., Manke, T., Lachner, M. and Jenuwein, T., 2014. Suv39h-Dependent H3K9me3 Marks Intact Retrotransposons and Silences LINE Elements in Mouse Embryonic Stem Cells. Molecular Cell, 55(2), pp.277-290.

Callahan, K., Hickman, A., Jones, C., Ghirlando, R. and Furano, A., 2011. Polymerization and nucleic acid-binding properties of human L1 ORF1 protein. Nucleic Acids Research, 40(2), pp.813-827.

Capy, P., Gasperi, G., Biémont, C. and Bazin, C., 2000. Stress and transposable elements: co-evolution or useful parasites?. Heredity, 85(2), pp.101-106.

Cargnin, F., Kwon, J., Katzman, S., Chen, B., Lee, J. and Lee, S., 2018. FOXG1 Orchestrates Neocortical Organization and Cortico-Cortical Connections. Neuron, 100(5), pp.1083-1096.e5.

Carmell, M., Girard, A., van de Kant, H., Bourc'his, D., Bestor, T., de Rooij, D. and Hannon, G., 2007. MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. Developmental Cell, 12(4), pp.503-514.

Castellano, D. and Munch, K., 2020. Population Genomics in the Great Apes. Methods in Molecular Biology, pp.453-463.

Chalopin, D., Naville, M., Plard, F., Galiana, D. and Volff, J., 2015. Comparative Analysis of Transposable Elements Highlights Mobilome Diversity and Evolution in Vertebrates. Genome Biology and Evolution, 7(2), pp.567-580.

Chen, D., Wang, C., Li, M., She, X., Yuan, Y., Chen, H., Zhang, W. and Zhao, C., 2019. Loss of Foxg1 Impairs the Development of Cortical SST-Interneurons Leading to Abnormal Emotional and Social Behaviors. Cerebral Cortex, 29(8), pp.3666-3682.

Chen, D., Wang, C., Li, M., She, X., Yuan, Y., Chen, H., Zhang, W. and Zhao, C., 2019. Loss of Foxg1 Impairs the Development of Cortical SST-Interneurons Leading to Abnormal Emotional and Social Behaviors. Cerebral Cortex, 29(8), pp.3666-3682.

Chen, Y., 2005. Reelin Modulates NMDA Receptor Activity in Cortical Neurons. Journal of Neuroscience, 25(36), pp.8209-8216.

Chenn, A. and McConnell, S., 1995. Cleavage orientation and the asymmetric inheritance of notchl immunoreactivity in mammalian neurogenesis. Cell, 82(4), pp.631-641.

Chenn, A. and Walsh, C., 2002. Regulation of Cerebral Cortical Size by Control of Cell Cycle Exit in Neural Precursors. Science, 297(5580), pp.365-369.

Cheung, A., Kondo, S., Abdel-Mannan, O., Chodroff, R., Sirey, T., Bluy, L., Webber, N., DeProto, J., Karlen, S., Krubitzer, L., Stolp, H., Saunders, N. and Molnár, Z., 2009. The Subventricular Zone Is the Developmental Milestone of a 6-Layered Neocortex: Comparisons in Metatherian and Eutherian Mammals. Cerebral Cortex, 20(5), pp.1071-1081.

Chiola, S., Do, M., Centrone, L. and Mallamaci, A., 2018. Foxg1 Overexpression in Neocortical Pyramids Stimulates Dendrite Elongation Via Hes1 and pCreb1 Upregulation. Cerebral Cortex, 29(3), pp.1006-1019.

Chiu, Y. and Greene, W., 2008. The APOBEC3 Cytidine Deaminases: An Innate Defensive Network Opposing Exogenous Retroviruses and Endogenous Retroelements. Annual Review of Immunology, 26(1), pp.317-353.

Chung, H., Bogwitz, M., McCart, C., Andrianopoulos, A., ffrench-Constant, R., Batterham, P. and Daborn, P., 2007. Cis-Regulatory Elements in the Accord Retrotransposon Result in Tissue-Specific Expression of the Drosophila melanogaster Insecticide Resistance Gene Cyp6g1. Genetics, 175(3), pp.1071-1077.

Cobolli Gigli, C., Scaramuzza, L., De Simone, M., Rossi, R., Pozzi, D., Pagani, M., Landsberger, N. and Bedogni, F., 2018. Lack of Methyl-CpG Binding Protein 2 (MeCP2) Affects Cell Fate Refinement During Embryonic Cortical Development. Cerebral Cortex, 28(5), pp.1846-1856.

Cost, G., 2002. Human L1 element target-primed reverse transcription in vitro. The EMBO Journal, 21(21), pp.5899-5910.

Coufal, N., Garcia-Perez, J., Peng, G., Marchetto, M., Muotri, A., Mu, Y., Carson, C., Macia, A., Moran, J. and Gage, F., 2011. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proceedings of the National Academy of Sciences, 108(51), pp.20382-20387.

Coufal, N., Garcia-Perez, J., Peng, G., Yeo, G., Mu, Y., Lovci, M., Morell, M., O'Shea, K., Moran, J. and Gage, F., 2009. L1 retrotransposition in human neural progenitor cells. Nature, 460(7259), pp.1127-1131.

Craig, N., Craigie, R., Gellert, M. and Lambowitz, A., 2002. Mobile DNA II, (Washington, D.C.: ASM Press), pp.1–1204.

Craig, N., Craigie, R., Gellert, M. and Lambowitz, A., 2022. Mobile DNA II. ASM Press, +1204 p + 49 pl(55581), pp.209–0.

Crow, Y. and Rehwinkel, J., 2009. Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. Human Molecular Genetics, 18(R2), pp.R130-R136.

Cruz, C., Della Rosa, M., Krueger, C., Gao, Q., Horkai, D., King, M., Field, L. and Houseley, J., 2018. Tri-methylation of histone H3 lysine 4 facilitates gene expression in ageing cells. eLife, 7.

Dai, L., Taylor, M., O'Donnell, K. and Boeke, J., 2012. Poly(A) Binding Protein C1 Is Essential for Efficient L1 Retrotransposition and Affects L1 RNP Formation. Molecular and Cellular Biology, 32(21), pp.4323-4336.

Dai, S., Qu, L., Li, J. and Chen, Y., 2021. Toward a mechanistic understanding of DNA binding by forkhead transcription factors and its perturbation by pathogenic mutations. Nucleic Acids Research, 49(18), pp.10235-10249.

D'Ambrosio, E., Waitzkin, S., Witney, F., Salemme, A. and Furano, A., 1986. Structure of the highly repeated, long interspersed DNA family (LINE or L1Rn) of the rat. Molecular and Cellular Biology, 6(2), pp.411-424.

Day, D., Luquette, L., Park, P. and Kharchenko, P., 2010. Estimating enrichment of repetitive elements from high-throughput sequence data. Genome Biology, 11(6), p.R69.

De Cecco, M., Ito, T., Petrashen, A., Elias, A., Skvir, N., Criscione, S., Caligiana, A., Brocculi, G., Adney, E., Boeke, J., Le, O., Beauséjour, C., Ambati, J., Ambati, K., Simon, M., Seluanov, A., Gorbunova, V., Slagboom, P., Helfand, S., Neretti, N. and Sedivy, J., 2019. L1 drives IFN in senescent cells and promotes age-associated inflammation. Nature, 566(7742), pp.73-78.

De Fazio, S., Bartonicek, N., Di Giacomo, M., Abreu-Goodger, C., Sankar, A., Funaya, C., Antony, C., Moreira, P., Enright, A. and O'Carroll, D., 2011. The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature, 480(7376), pp.259-263.

De Filippis, R., Pancrazi, L., Bjørgo, K., Rosseto, A., Kleefstra, T., Grillo, E., Panighini, A., Cardarelli, F., Meloni, I., Ariani, F., Mencarelli, M., Hayek, J., Renieri, A., Costa, M. and Mari, F., 2011. Expanding the phenotype associated with FOXG1 mutations and in vivo FoxG1 chromatin-binding dynamics. Clinical Genetics, 82(4), pp.395-403.

Deininger, P., Morales, M., White, T., Baddoo, M., Hedges, D., Servant, G., Srivastav, S., Smither, M., Concha, M., DeHaro, D., Flemington, E. and Belancio, V., 2016. A comprehensive approach to expression of L1 loci. Nucleic Acids Research, 45(5), pp.e31-e31.

Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A. and O'Carroll, D., 2013. Multiple Epigenetic Mechanisms and the piRNA Pathway Enforce LINE1 Silencing during Adult Spermatogenesis. Molecular Cell, 50(4), pp.601-608.

Diez-Roux, G., Banfi, S., Sultan, M., Geffers, L., Anand, S., Rozado, D., Magen, A., Canidio, E., Pagani, M., Peluso, I., Lin-Marq, N., Koch, M., Bilio, M., Cantiello, I., Verde, R., De Masi, C., Bianchi, S., Cicchini, J., Perroud, E., Mehmeti, S., Dagand, E., Schrinner, S., Nürnberger, A., Schmidt, K., Metz, K., Zwingmann, C., Brieske, N., Springer, C., Hernandez, A., Herzog, S., Grabbe, F., Sieverding, C., Fischer, B., Schrader, K., Brockmeyer, M., Dettmer, S., Helbig, C., Alunni, V., Battaini, M., Mura, C., Henrichsen, C., Garcia-Lopez, R., Echevarria, D., Puelles, E., Garcia-Calero, E., Kruse, S., Uhr, M., Kauck, C., Feng, G., Milyaev, N., Ong, C., Kumar, L., Lam, M., Semple, C., Gyenesei, A., Mundlos, S., Radelof, U., Lehrach, H., Sarmientos,

P., Reymond, A., Davidson, D., Dollé, P., Antonarakis, S., Yaspo, M., Martinez, S., Baldock, R., Eichele, G. and Ballabio, A., 2011. A High-Resolution Anatomical Atlas of the Transcriptome in the Mouse Embryo. PLoS Biology, 9(1), p.e1000582.

Dmitriev, S., Andreev, D., Terenin, I., Olovnikov, I., Prassolov, V., Merrick, W. and Shatsky, I., 2007. Efficient Translation Initiation Directed by the 900-Nucleotide-Long and GC-Rich 5′ Untranslated Region of the Human Retrotransposon LINE-1 mRNA Is Strictly Cap Dependent Rather than Internal Ribosome Entry Site Mediated. Molecular and Cellular Biology, 27(13), pp.4685-4697.

Doolittle, W. and Sapienza, C., 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature, 284(5757), pp.601-603.

Doucet, A., Hulme, A., Sahinovic, E., Kulpa, D., Moldovan, J., Kopera, H., Athanikar, J., Hasnaoui, M., Bucheton, A., Moran, J. and Gilbert, N., 2010. Characterization of LINE-1 Ribonucleoprotein Particles. PLoS Genetics, 6(10), p.e1001150.

Doyle, G. and Berrettini, W., 2017. Failure to Replicate an Association of a LINE-1 Element in ERI1 Exoribonuclease Family Member 3 (ERI3) with Schizophrenia. Neuropsychopharmacology, 42(13), pp.2471-2471.

Dugas-Ford, J., Rowell, J. and Ragsdale, C., 2012. Cell-type homologies and the origins of the neocortex. Proceedings of the National Academy of Sciences, 109(42), pp.16974-16979.

Edgell, M., Hardies, S., Loeb, D., Shehee, W., Padgett, R., Burton, F., Comer, M., Casavant, N., Funk, F. and Hutchinson 3rd, C., 1987. The L1 family in mice. Prog Clin Biol Res, 251, pp.107-29.

Edmondson, J. and Hatten, M., 1987. Glial-guided granule neuron migration in vitro: a high-resolution time- lapse video microscopic study. The Journal of Neuroscience, 7(6), pp.1928-1934.

Edmondson, J. and Hatten, M., 1987. Glial-guided granule neuron migration in vitro: a high-resolution time- lapse video microscopic study. The Journal of Neuroscience, 7(6), pp.1928-1934.

Erdel, F. and Rippe, K., 2018. Formation of Chromatin Subcompartments by Phase Separation. Biophysical Journal, 114(10), pp.2262-2270.

Erwin, J., Marchetto, M. and Gage, F., 2014. Mobile DNA elements in the generation of diversity and complexity in the brain. Nature Reviews Neuroscience, 15(8), pp.497-506.

Esnault, C., Heidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A., Heidmann, T. and Schwartz, O., 2005. APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature, 433(7024), pp.430-433.

Evans, J., Peddigari, S., Chaurasiya, K., Williams, M. and Martin, S., 2011. Paired mutations abolish and restore the balanced annealing and melting activities of ORF1p that are required for LINE-1 retrotransposition. Nucleic Acids Research, 39(13), pp.5611-5621.

Evgen'ev, M. and Arkhipova, I., 2005. <i>Penelope</i>-like elements – a new class of retroelements: distribution, function and possible evolutionary significance. Cytogenetic and Genome Research, 110(1-4), pp.510-521.

Evrony, G., Cai, X., Lee, E., Hills, L., Elhosary, P., Lehmann, H., Parker, J., Atabay, K., Gilmore, E., Poduri, A., Park, P. and Walsh, C., 2012. Single-Neuron Sequencing Analysis of L1 Retrotransposition and Somatic Mutation in the Human Brain. Cell, 151(3), pp.483-496.

Fabio, R., Colombo, B., Russo, S., Cogliati, F., Masciadri, M., Foglia, S., Antonietti, A. and Tavian, D., 2014. Recent insights into genotype–phenotype relationships in patients with Rett syndrome using a fine grain scale. Research in Developmental Disabilities, 35(11), pp.2976-2986.

Falcone, C., Santo, M., Liuzzi, G., Cannizzaro, N., Grudina, C., Valencic, E., Peruzzotti-Jametti, L., Pluchino, S. and Mallamaci, A., 2019. Foxg1 Antagonizes Neocortical Stem Cell Progression to Astrogenesis. Cerebral Cortex, 29(12), pp.4903-4918.

Fan, G., Martinowich, K., Chin, M., He, F., Fouse, S., Hutnick, L., Hattori, D., Ge, W., Shen, Y., Wu, H., ten Hoeve, J., Shuai, K. and Sun, Y., 2005. DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling. Development, 132(15), pp.3345-3356.

Fanning, T. and Singer, M., 1987. LINE-1: A mammalian transposable element. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression, 910(3), pp.203-212.

Faulkner, G., Kimura, Y., Daub, C., Wani, S., Plessy, C., Irvine, K., Schroder, K., Cloonan, N., Steptoe, A., Lassmann, T., Waki, K., Hornig, N., Arakawa, T., Takahashi, H., Kawai, J., Forrest, A., Suzuki, H., Hayashizaki, Y., Hume, D., Orlando, V., Grimmond, S. and Carninci, P., 2009. The regulated retrotransposon transcriptome of mammalian cells. Nature Genetics, 41(5), pp.563-571.

Fedoroff, N., 2012. Transposable Elements, Epigenetics, and Genome Evolution. Science, 338(6108), pp.758-767.

Fedoroff, N., Wessler, S. and Shure, M., 1983. Isolation of the transposable maize controlling elements Ac and Ds. Cell, 35(1), pp.235-242.

Feng, J., Chang, H., Li, E. and Fan, G., 2005. Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system. Journal of Neuroscience Research, 79(6), pp.734-746.

Feng, Q., Moran, J., Kazazian, H. and Boeke, J., 1996. Human L1 Retrotransposon Encodes a Conserved Endonuclease Required for Retrotransposition. Cell, 87(5), pp.905-916.

Flasch, D., Macia, Á., Sánchez, L., Ljungman, M., Heras, S., García-Pérez, J., Wilson, T. and Moran, J., 2019. Genome-wide de novo L1 Retrotransposition Connects Endonuclease Activity with Replication. Cell, 177(4), pp.837-851.e28.

Floreani, L., Ansaloni, F., Mangoni, D., Agostoni, E., Sanges, R., Persichetti, F. and Gustincich, S., 2022. Analysis of LINE1 Retrotransposons in Huntington's Disease. Frontiers in Cellular Neuroscience, 15.

Florian, C., Bahi-Buisson, N. and Bienvenu, T., 2011. FOXG1-Related Disorders: From Clinical Description to Molecular Genetics. Molecular Syndromology,.

Florio, M. and Huttner, W., 2014. Neural progenitors, neurogenesis and the evolution of the neocortex. Development, 141(11), pp.2182-2194.

Franchini, L., 2022. Genetic Mechanisms Underlying Cortical Evolution in Mammals.

Freeman, B., Sokolowski, M., Roy-Engel, A., Smither, M. and Belancio, V., 2019. Identification of charged amino acids required for nuclear localization of human L1 ORF1 protein. Mobile DNA, 10(1).

Frisari, S., Santo, M., Hosseini, A., Manzati, M., Giugliano, M. and Mallamaci, A., 2022. Multidimensional Functional Profiling of Human Neuropathogenic FOXG1 Alleles in Primary Cultures of Murine Pallial Precursors. International Journal of Molecular Sciences, 23(3), p.1343.

Gasior, S., Wakeman, T., Xu, B. and Deininger, P., 2006. The Human LINE-1 Retrotransposon Creates DNA Double-strand Breaks. Journal of Molecular Biology, 357(5), pp.1383-1393.

Gazin, C., 1970. Evolution of the Vertebrates. A History of the Backboned Animals through Time. Edwin H. Colbert. Second edition. Wiley, New York, 1969. xviii + 542 pp., illus. $12.95. Science, 168(3933), pp.815-815.

Ghildiyal, M. and Zamore, P., 2009. Small silencing RNAs: an expanding universe. Nature Reviews Genetics, 10(2), pp.94-108.

Ghosh, R., Horowitz-Scherer, R., Nikitina, T., Shlyakhtenko, L. and Woodcock, C., 2010. MeCP2 Binds Cooperatively to Its Substrate and Competes with Histone H1 for Chromatin Binding Sites. Molecular and Cellular Biology, 30(19), pp.4656-4670.

Ghosheh, Y., Seridi, L., Ryu, T., Takahashi, H., Orlando, V., Carninci, P. and Ravasi, T., 2016. Characterization of piRNAs across postnatal development in mouse brain. Scientific Reports, 6(1).

Gibbs, R., Weinstock, G., Metzker, M., Muzny, D., Sodergren, E., Scherer, S., Scott, G., Steffen, D., Worley, K., Burch, P., Okwuonu, G., Hines, S., Lewis, L., DeRamo, C., Delgado, O., Dugan-Rocha, S., Miner, G., Morgan, M., Hawes, A., Gill, R., Holt, C., Adams, M., Amanatides, P., Baden-Tillson, H., Barnstead, M., Chin, S., Evans, C., Ferriera, S., Fosler, C., Glodek, A., Gu, Z., Jennings, D., Kraft, C., Nguyen, T., Pfannkoch, C., Sitter, C., Sutton, G., Venter, J., Woodage, T., Smith, D., Lee, H., Gustafson, E., Cahill, P., Kana, A., Doucette-Stamm, L., Weinstock, K., Fechtel, K., Weiss, R., Dunn, D., Green, E., Blakesley, R., Bouffard, G., de Jong, P., Osoegawa, K., Zhu, B., Marra, M., Schein, J., Bosdet, I., Fjell, C., Jones, S., Krzywinski, M., Mathewson, C., Siddiqui, A., Wye, N., McPherson, J., Zhao, S., Fraser, C., Shetty, J., Shatsman, S., Geer, K., Chen, Y., Abramzon, S., Nierman, W., Gibbs, R., Weinstock, G., Havlak, P., Chen, R., James Durbin, K., Egan, A., Ren, Y., Song, X., Li, B., Liu, Y., Qin, X., Cawley, S., Weinstock, G., Worley, K., Cooney, A., Gibbs, R., D'Souza, L., Martin, K., Qian Wu, J., Gonzalez-Garay, M., Jackson, A., Kalafus, K., McLeod, M., Milosavljevic, A., Virk, D., Volkov, A., Wheeler, D.,

Zhang, Z., Bailey, J., Eichler, E., Tuzun, E., Birney, E., Mongin, E., Ureta-Vidal, A., Woodwark, C., Zdobnov, E., Bork, P., Suyama, M., Torrents, D., Alexandersson, M., Trask, B., Young, J., Smith, D., Huang, H., Fechtel, K., Wang, H., Xing, H., Weinstock, K., Daniels, S., Gietzen, D., Schmidt, J., Stevens, K., Vitt, U., Wingrove, J., Camara, F., Mar Albà, M., Abril, J., Guigo, R., Smit, A., Dubchak, I., Rubin, E., Couronne, O., Poliakov, A., Hübner, N., Ganten, D., Goesele, C., Hummel, O., Kreitler, T., Lee, Y., Monti, J., Schulz, H., Zimdahl, H., Himmelbauer, H., Lehrach, H., Jacob, H., Bromberg, S., Gullings-Handley, J., Jensen-Seaman, M., Kwitek, A., Lazar, J., Pasko, D., Tonellato, P., Twigger, S., Ponting, C., Duarte, J., Rice, S., Goodstadt, L., Beatson, S., Emes, R., Winter, E., Webber, C., Brandt, P., Nyakatura, G., Adetobi, M., Chiaromonte, F., Elnitski, L., Eswara, P., Hardison, R., Hou, M., Kolbe, D., Makova, K., Miller, W., Nekrutenko, A., Riemer, C., Schwartz, S., Taylor, J., Yang, S., Zhang, Y., Lindpaintner, K., Andrews, T., Caccamo, M., Clamp, M., Clarke, L., Curwen, V., Durbin, R., Eyras, E., Searle, S., Cooper, G., Batzoglou, S., Brudno, M., Sidow, A., Stone, E., Craig Venter, J., Payseur, B., Bourque, G., López-Otín, C., Puente, X., Chakrabarti, K., Chatterji, S., Dewey, C., Pachter, L., Bray, N., Yap, V., Caspi, A., Tesler, G., Pevzner, P., Haussler, D., Roskin, K., Baertsch, R., Clawson, H., Furey, T., Hinrichs, A., Karolchik, D., Kent, W., Rosenbloom, K., Trumbower, H., Weirauch, M., Cooper, D., Stenson, P., Ma, B., Brent, M., Arumugam, M., Shteynberg, D., Copley, R., Taylor, M., Riethman, H., Mudunuri, U., Peterson, J., Guyer, M., Felsenfeld, A., Old, S., Mockrin, S. and Collins, F., 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature, 428(6982), pp.493-521.

Gibson, B. and Kraus, W., 2012. New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. Nature Reviews Molecular Cell Biology, 13(7), pp.411-424.

Gilbert, N., Lutz, S., Morrish, T. and Moran, J., 2005. Multiple Fates of L1 Retrotransposition Intermediates in Cultured Human Cells. Molecular and Cellular Biology, 25(17), pp.7780-7795.

Glazko, G., 2003. Estimation of Divergence Times for Major Lineages of Primate Species. Molecular Biology and Evolution, 20(3), pp.424-434.

Goodier, J., Mandal, P., Zhang, L. and Kazazian, H., 2010. Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. Human Molecular Genetics, 19(9), pp.1712-1725.

Goodier, J., Ostertag, E., Du, K. and Kazazian, H., 2001. A Novel Active L1 Retrotransposon Subfamily in the Mouse. Genome Research, 11(10), pp.1677-1685.

Goodier, J., Zhang, L., Vetter, M. and Kazazian, H., 2007. LINE-1 ORF1 Protein Localizes in Stress Granules with Other RNA-Binding Proteins, Including Components of RNA Interference RNA-Induced Silencing Complex. Molecular and Cellular Biology, 27(18), pp.6469-6483.

Götz, M. and Huttner, W., 2005. The cell biology of neurogenesis. Nature Reviews Molecular Cell Biology, 6(10), pp.777-788.

Gould, S. and Vrba, E., 1982. Exaptation—a Missing Term in the Science of Form. Paleobiology, 8(1), pp.4-15.

Gray, P., Fu, H., Luo, P., Zhao, Q., Yu, J., Ferrari, A., Tenzen, T., Yuk, D., Tsung, E., Cai, Z., Alberta, J., Cheng, L., Liu, Y., Stenman, J., Valerius, M., Billings, N., Kim, H., Greenberg, M., McMahon, A., Rowitch, D., Stiles, C. and Ma, Q., 2004. Mouse Brain Organization Revealed Through Direct Genome-Scale TF Expression Analysis. Science, 306(5705), pp.2255-2257.

Gregersen, L., Schueler, M., Munschauer, M., Mastrobuoni, G., Chen, W., Kempa, S., Dieterich, C. and Landthaler, M., 2014. MOV10 Is a 5′ to 3′ RNA Helicase Contributing to UPF1 mRNA Target Degradation by Translocation along 3′ UTRs. Molecular Cell, 54(4), pp.573-585.

Greig, L., Woodworth, M., Galazo, M., Padmanabhan, H. and Macklis, J., 2022. Molecular logic of neocortical projection neuron specification, development and diversity.

Grimaldi, G., Skowronski, J. and Singer, M., 1984. Defining the beginning and end of KpnI family segments. The EMBO Journal, 3(8), pp.1753-1759.

Guler, G., Tindell, C., Pitti, R., Wilson, C., Nichols, K., KaiWai Cheung, T., Kim, H., Wongchenko, M., Yan, Y., Haley, B., Cuellar, T., Webster, J., Alag, N., Hegde, G., Jackson, E., Nance, T., Giresi, P., Chen, K., Liu, J., Jhunjhunwala, S., Settleman, J., Stephan, J., Arnott, D. and Classon, M., 2017. Repression of Stress-Induced LINE-1 Expression Protects Cancer Cell Subpopulations from Lethal Drug Exposure. Cancer Cell, 32(2), pp.221-237.e13.

Guo, H., Chitiprolu, M., Gagnon, D., Meng, L., Perez-Iratxeta, C., Lagace, D. and Gibbings, D., 2014. Autophagy supports genomic stability by degrading retrotransposon RNA. Nature Communications, 5(1).

Guo, Y., Zhao, S. and Wang, G., 2021. Polycomb Gene Silencing Mechanisms: PRC2 Chromatin Targeting, H3K27me3 'Readout', and Phase Separation-Based Compaction. Trends in Genetics, 37(6), pp.547-565.

Guy, J., Cheval, H., Selfridge, J. and Bird, A., 2011. The Role of MeCP2 in the Brain. Annual Review of Cell and Developmental Biology, 27(1), pp.631-652.

Hanashima, C., Li, S., Shen, L., Lai, E. and Fishell, G., 2004. Foxg1 Suppresses Early Cortical Cell Fate. Science, 303(5654), pp.56-59.

Hanashima, C., Shen, L., Li, S. and Lai, E., 2002. Brain Factor-1Controls the Proliferation and Differentiation of Neocortical Progenitor Cells through Independent Mechanisms. The Journal of Neuroscience, 22(15), pp.6526-6536.

Hannenhalli, S. and Kaestner, K., 2009. The evolution of Fox genes and their role in development and disease. Nature Reviews Genetics, 10(4), pp.233-240.

Harris, C., DeWan, A., Zupnick, A., Normart, R., Gabriel, A., Prives, C., Levine, A. and Hoh, J., 2009. p53 responsive elements in human retrotransposons. Oncogene, 28(44), pp.3857-3865.

Hartfuss, E., Galli, R., Heins, N. and Götz, M., 2001. Characterization of CNS Precursor Subtypes and Radial Glia. Developmental Biology, 229(1), pp.15-30.

Hata, K. and Sakaki, Y., 1997. Identification of critical CpG sites for repression of L1 transcription by DNA methylation. Gene, 189(2), pp.227-234.

Haubensak, W., Attardo, A., Denk, W. and Huttner, W., 2004. Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: A major site of neurogenesis. Proceedings of the National Academy of Sciences, 101(9), pp.3196-3201.

Hayashi, K., Chuva de Sousa Lopes, S., Kaneda, M., Tang, F., Hajkova, P., Lao, K., O'Carroll, D., Das, P., Tarakhovsky, A., Miska, E. and Surani, M., 2008. MicroRNA Biogenesis Is Required for Mouse Primordial Germ Cell Development and Spermatogenesis. PLoS ONE, 3(3), p.e1738.

He, J., Fu, X., Zhang, M., He, F., Li, W., Abdul, M., Zhou, J., Sun, L., Chang, C., Li, Y., Liu, H., Wu, K., Babarinde, I., Zhuang, Q., Loh, Y., Chen, J., Esteban, M. and Hutchins, A., 2019. Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells. Nature Communications, 10(1).

Hébert, J. and McConnell, S., 2000. Targeting of cre to the Foxg1 (BF-1) Locus Mediates loxP Recombination in the Telencephalon and Other Developing Head Structures. Developmental Biology, 222(2), pp.296-306.

Herrmann, A., Wittmann, S., Thomas, D., Shepard, C., Kim, B., Ferreirós, N. and Gramberg, T., 2018. The SAMHD1-mediated block of LINE-1 retroelements is regulated by phosphorylation. Mobile DNA, 9(1).

Hervouet, E., Vallette, F. and Cartron, P., 2010. Dnmt1/Transcription Factor Interactions: An Alternative Mechanism of DNA Methylation Inheritance. Genes & Cancer, 1(5), pp.434-443.

Hervouet, E., Vallette, F. and Cartron, P., 2010. Dnmt1/Transcription Factor Interactions: An Alternative Mechanism of DNA Methylation Inheritance. Genes & Cancer, 1(5), pp.434-443.

Hevner, R., 2019. Intermediate progenitors and Tbr2 in cortical development. Journal of Anatomy, 235(3), pp.616-625.

Hoffmann, A., Zimmermann, C. and Spengler, D., 2015. Molecular epigenetic switches in neurodevelopment in health and disease. Frontiers in Behavioral Neuroscience, 9.

Horn, A., Celic, I., Dong, C., Martirosyan, I. and Han, J., 2017. A conserved role for the ESCRT membrane budding complex in LINE retrotransposition. PLOS Genetics, 13(6), p.e1006837.

Hou, P., hAilín, D., Vogel, T. and Hanashima, C., 2020. Transcription and Beyond: Delineating FOXG1 Function in Cortical Development and Disorders. Frontiers in Cellular Neuroscience, 14.

Hou, P., hAilín, D., Vogel, T. and Hanashima, C., 2022. Transcription and Beyond: Delineating FOXG1 Function in Cortical Development and Disorders.

Huang, X., Liu, J., Ketova, T., Fleming, J., Grover, V., Cooper, M., Litingtung, Y. and Chiang, C., 2010. Transventricular delivery of Sonic hedgehog is essential to cerebellar ventricular zone development. Proceedings of the National Academy of Sciences, 107(18), pp.8422-8427.

Huh, S., Hatini, V., Marcus, R., Li, S. and Lai, E., 1999. Dorsal–Ventral Patterning Defects in the Eye of BF-1-Deficient Mice Associated with a Restricted Loss of shh Expression. Developmental Biology, 211(1), pp.53-63.

Hunter, R., Murakami, G., Dewell, S., Seligsohn, M., Baker, M., Datson, N., McEwen, B. and Pfaff, D., 2012. Acute stress and hippocampal histone H3 lysine 9 trimethylation, a retrotransposon silencing response. Proceedings of the National Academy of Sciences, 109(43), pp.17657-17662.

Hwang, C., Simeone, A., Lai, E. and Wu, D., 2009. Foxg1is required for proper separation and formation of sensory cristae during inner ear development. Developmental Dynamics, 238(11), pp.2725-2734.

Ichiyanagi, K. and Okada, N., 2008. Mobility Pathways for Vertebrate L1, L2, CR1, and RTE Clade Retrotransposons. Molecular Biology and Evolution, 25(6), pp.1148-1157.

Inoue, K., Khajavi, M., Ohyama, T., Hirabayashi, S., Wilson, J., Reggin, J., Mancias, P., Butler, I., Wilkinson, M., Wegner, M. and Lupski, J., 2004. Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. Nature Genetics, 36(4), pp.361-369.

Ittner, L. and Götz, J., 2010. Amyloid-β and tau — a toxic pas de deux in Alzheimer's disease. Nature Reviews Neuroscience, 12(2), pp.67-72.

Ivancevic, A., Kortschak, R., Bertozzi, T. and Adelson, D., 2016. LINEs between Species: Evolutionary Dynamics of LINE-1 Retrotransposons across the Eukaryotic Tree of Life. Genome Biology and Evolution, 8(11), pp.3301-3322.

Jacob-Hirsch, J., Eyal, E., Knisbacher, B., Roth, J., Cesarkas, K., Dor, C., Farage-Barhom, S., Kunik, V., Simon, A., Gal, M., Yalon, M., Moshitch-Moshkovitz, S., Tearle, R., Constantini, S., Levanon, E., Amariglio, N. and Rechavi, G., 2018. Whole-genome sequencing reveals principles of brain retrotransposition in neurodevelopmental disorders. Cell Research, 28(2), pp.187-203.

Januszyk, K., Li, P., Villareal, V., Branciforte, D., Wu, H., Xie, Y., Feigon, J., Loo, J., Martin, S. and Clubb, R., 2007. Identification and Solution Structure of a Highly Conserved C-terminal Domain within ORF1p Required for Retrotransposition of Long Interspersed Nuclear Element-1. Journal of Biological Chemistry, 282(34), pp.24893-24904.

Jeon, S., Kim, S., Oh, M., Liang, P., Tang, W. and Han, K., 2021. A comprehensive analysis of gorilla-specific LINE-1 retrotransposons. Genes & Genomics, 43(10), pp.1133-1141.

Johnson, J., Kim, E. and Meredith, D., 2010. [S6.3]: bHLH transcription complexes couple neural differentiation and cell-type specification. International Journal of Developmental Neuroscience, 28(8), pp.650-650.

Jönsson, M., Ludvik Brattås, P., Gustafsson, C., Petri, R., Yudovich, D., Pircs, K., Verschuere, S., Madsen, S., Hansson, J., Larsson, J., Månsson, R., Meissner, A. and Jakobsson, J., 2019. Activation of neuronal genes via LINE-1 elements upon global DNA demethylation in human neural progenitors. Nature Communications, 10(1).

Kajikawa, M., Yamaguchi, K. and Okada, N., 2012. A new mechanism to ensure integration during LINE retrotransposition: A suggestion from analyses of the 5′ extra nucleotides. Gene, 505(2), pp.345-351.

Kapusta, A. and Suh, A., 2016. Evolution of bird genomes-a transposon's-eye view. Annals of the New York Academy of Sciences, 1389(1), pp.164-185.

Karayiorgou, M., Morris, M., Morrow, B., Shprintzen, R., Goldberg, R., Borrow, J., Gos, A., Nestadt, G., Wolyniec, P. and Lasseter, V., 1995. Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. Proceedings of the National Academy of Sciences, 92(17), pp.7612-7616.

Kemp, T., 2004. The Origin and Evolution of Mammals.

Kessler, R., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K., Rush, A., Walters, E. and Wang, P., 2003. The Epidemiology of Major Depressive Disorder. JAMA, 289(23), p.3095.

Khazina, E. and Weichenrieder, O., 2009. Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. Proceedings of the National Academy of Sciences, 106(3), pp.731-736.

Kim, K., 2019. PIWI Proteins and piRNAs in the Nervous System. Mol Cells, 42(12), pp.828–835.

Kim, K., Tang, N., Andrusiak, M., Wu, Z., Chisholm, A. and Jin, Y., 2018. A Neuronal piRNA Pathway Inhibits Axon Regeneration in C. elegans. Neuron, 97(3), pp.511-519.e6.

Kim, S., Günesdogan, U., Zylicz, J., Hackett, J., Cougot, D., Bao, S., Lee, C., Dietmann, S., Allen, G., Sengupta, R. and Surani, M., 2014. PRMT5 Protects Genomic Integrity during Global DNA Demethylation in Primordial Germ Cells and Preimplantation Embryos. Molecular Cell, 56(4), pp.564-579.

Kishi, N. and Macklis, J., 2005. Dissecting MECP2 Function in the Central Nervous System. Journal of Child Neurology, 20(8), pp.753-759.

Kolosha, V. and Martin, S., 2003. High-affinity, Non-sequence-specific RNA Binding by the Open Reading Frame 1 (ORF1) Protein from Long Interspersed Nuclear Element 1 (LINE-1). Journal of Biological Chemistry, 278(10), pp.8112-8117.

Komada, M., Saitsu, H., Kinboshi, M., Miura, T., Shiota, K. and Ishibashi, M., 2008. Hedgehog signaling is involved in development of the neocortex. Development, 135(16), pp.2717-2727.

Kon, E., Cossard, A. and Jossin, Y., 2017. Neuronal Polarity in the Embryonic Mammalian Cerebral Cortex. Frontiers in Cellular Neuroscience, 11.

Konkel, M., Walker, J. and Batzer, M., 2010. LINEs and SINEs of primate evolution. Evolutionary Anthropology: Issues, News, and Reviews, 19(6), pp.236-249.

Kopera, H., Moldovan, J., Morrish, T., Garcia-Perez, J. and Moran, J., 2011. Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. Proceedings of the National Academy of Sciences, 108(51), pp.20345-20350.

Korada, S., Zheng, W., Basilico, C., Schwartz, M. and Vaccarino, F., 2002. Fibroblast Growth Factor 2 Is Necessary for the Growth of Glutamate Projection Neurons in the Anterior Neocortex. The Journal of Neuroscience, 22(3), pp.863-875.

Kordiš, D., Lovšin, N. and Gubenšek, F., 2006. Phylogenomic Analysis of the L1 Retrotransposons in Deuterostomia. Systematic Biology, 55(6), pp.886-901.

Kowalczyk, T., Pontious, A., Englund, C., Daza, R., Bedogni, F., Hodge, R., Attardo, A., Bell, C., Huttner, W. and Hevner, R., 2009. Intermediate Neuronal Progenitors (Basal Progenitors) Produce Pyramidal–Projection Neurons for All Layers of Cerebral Cortex. Cerebral Cortex, 19(10), pp.2439-2450.

Kriegstein, A. and Alvarez-Buylla, A., 2009. The Glial Nature of Embryonic and Adult Neural Stem Cells. Annual Review of Neuroscience, 32(1), pp.149-184.

Kubo, S., Seleme, M., Soifer, H., Perez, J., Moran, J., Kazazian, H. and Kasahara, N., 2006. L1 retrotransposition in nondividing and primary human somatic cells. Proceedings of the National Academy of Sciences, 103(21), pp.8036-8041.

Kulpa, D. and Moran, J., 2005. Ribonucleoprotein particle formation is necessary but not sufficient for LINE-1 retrotransposition. Human Molecular Genetics, 14(21), pp.3237-3248.

Kumamoto, T. and Hanashima, C., 2017. Evolutionary conservation and conversion of Foxg1 function in brain development. Development, Growth & Differentiation, 59(4), pp.258-269.

Küppers, R. and Dalla-Favera, R., 2001. Mechanisms of chromosomal translocations in B cell lymphomas. Oncogene, 20(40), pp.5580-5594.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T., Hata, K., Li, E., Matsuda, Y., Kimura, T., Okabe, M., Sakaki, Y., Sasaki, H. and Nakano, T., 2008. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes & Development, 22(7), pp.908-917.

Kurzynska-Kokorniak, A., Jamburuthugoda, V., Bibillo, A. and Eickbush, T., 2007. DNA-directed DNA Polymerase and Strand Displacement Activity of the Reverse Transcriptase Encoded by the R2 Retrotransposon. Journal of Molecular Biology, 374(2), pp.322-333.

Kuwabara, T., Hsieh, J., Muotri, A., Yeo, G., Warashina, M., Lie, D., Moore, L., Nakashima, K., Asashima, M. and Gage, F., 2009. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nature Neuroscience, 12(9), pp.1097-1105.

Kuwabara, T., Hsieh, J., Muotri, A., Yeo, G., Warashina, M., Lie, D., Moore, L., Nakashima, K., Asashima, M. and Gage, F., 2009. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nature Neuroscience, 12(9), pp.1097-1105.

Land, P., 2003. Expression of the Transcription Factor, tailless, Is Required for Formation of Superficial Cortical Layers. Cerebral Cortex, 13(9), pp.921-931.

Lander, E., Linton, L., Birren, B., Nusbaum, C., Zody, M., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R., Wilson, R., Hillier, L., McPherson, J., Marra, M., Mardis, E., Fulton, L., Chinwalla, A., Pepin, K., Gish, W., Chissoe, S., Wendl, M., Delehaunty, K., Miner, T., Delehaunty, A., Kramer, J., Cook, L., Fulton, R., Johnson, D., Minx, P., Clifton, S., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R., Muzny, D., Scherer, S., Bouck, J., Sodergren, E., Worley, K., Rives, C., Gorrell, J., Metzker, M., Naylor, S., Kucherlapati, R., Nelson, D., Weinstock, G., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Smith, D., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H., Dubois, J., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R., Federspiel, N., Abola, A., Proctor, M., Roe, B., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W., de la Bastide, M., Dedhia, N., Blöcker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D., Burge, C., Cerutti, L., Chen, H., Church, D., Clamp, M., Copley, R., Doerks, T., Eddy, S., Eichler, E., Furey, T., Galagan, J., Gilbert, J., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L., Jones, T., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W., Kitts, P., Koonin, E., Korf, I., Kulp, D., Lancet, D., Lowe, T., McLysaght, A., Mikkelsen, T., Moran, J., Mulder, N., Pollara, V., Ponting, C., Schuler, G., Schultz, J., Slater, G., Smit, A., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y., Wolfe, K., Yang, S., Yeh, R., Collins, F., Guyer, M., Peterson, J., Felsenfeld, A., Wetterstrand, K., Myers, R., Schmutz, J., Dickson, M., Grimwood, J., Cox, D., Olson, M., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G., Athanasiou, M., Schultz, R., Patrinos, A. and Morgan, M., 2001. Initial sequencing and analysis of the human genome. Nature, 409(6822), pp.860-921.

Laski, F., Rio, D. and Rubin, G., 1986. Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. Cell, 44(1), pp.7-19.

Le Guen, T., Bahi-Buisson, N., Nectoux, J., Boddaert, N., Fichou, Y., Diebold, B., Desguerre, I., Raqbi, F., Daire, V., Chelly, J. and Bienvenu, T., 2010. A FOXG1 mutation in a boy with congenital variant of Rett syndrome. neurogenetics, 12(1), pp.1-8.

Lee, E., Banerjee, S., Zhou, H., Jammalamadaka, A., Arcila, M., Manjunath, B. and Kosik, K., 2011. Identification of piRNAs in the central nervous system. RNA, 17(6), pp.1090-1099.

Lee, H., Farhanullah, Lee, J., Jajoo, R., Kong, S., Shin, J., Kim, J., Lee, J., Lee, J. and Kim, H., 2015. Discovery of a Small Molecule that Enhances Astrocytogenesis by Activation of STAT3, SMAD1/5/8, and ERK1/2 via Induction of Cytokines in Neural Stem Cells. ACS Chemical Neuroscience, 7(1), pp.90-99.

Lee, J., Cordaux, R., Han, K., Wang, J., Hedges, D., Liang, P. and Batzer, M., 2007. Different evolutionary fates of recently integrated human and chimpanzee LINE-1 retrotransposons. Gene, 390(1-2), pp.18-27.

Lee, J., Cordaux, R., Han, K., Wang, J., Hedges, D., Liang, P. and Batzer, M., 2007. Different evolutionary fates of recently integrated human and chimpanzee LINE-1 retrotransposons. Gene, 390(1-2), pp.18-27.

Lee, J., Wendelken, C., Bunge, S. and Ghetti, S., 2015. A Time and Place for Everything: Developmental Differences in the Building Blocks of Episodic Memory. Child Development, 87(1), pp.194-210.

Lee, S., Tang, W., Liang, P. and Han, K., 2019. A comprehensive analysis of chimpanzee (Pan troglodytes)-specific LINE-1 retrotransposons. Gene, 693, pp.46-51.

Leighton, L., Wei, W., Marshall, P., Ratnu, V., Li, X., Zajaczkowski, E., Spadaro, P., Khandelwal, N., Kumar, A. and Bredy, T., 2019. Disrupting the hippocampal Piwi pathway enhances contextual fear memory in mice. Neurobiology of Learning and Memory, 161, pp.202-209.

Li, J., Dai, S., Chen, X., Liang, X., Qu, L., Jiang, L., Guo, M., Zhou, Z., Wei, H., Zhang, H., Chen, Z., Chen, L. and Chen, Y., 2021. Mechanism of forkhead transcription factors binding to a novel palindromic DNA site. Nucleic Acids Research, 49(6), pp.3573-3583.

Li, J., Kannan, M., Trivett, A., Liao, H., Wu, X., Akagi, K. and Symer, D., 2014. An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. Nucleic Acids Research, 42(7), pp.4546-4562.

Li, M., Zhu, M., Feng, F., Cai, F., Fan, K., Jiang, H., Wang, Z. and Linghu, E., 2014. Long interspersed nucleotide acid element-1 ORF-1 protein promotes proliferation and invasion of human colorectal cancer LoVo cells through enhancing ETS-1 activity. Genetics and Molecular Research, 13(3), pp.6981-6994.

Li, M., Zhu, M., Feng, F., Cai, F., Fan, K., Jiang, H., Wang, Z. and Linghu, E., 2014. Long interspersed nucleotide acid element-1 ORF-1 protein promotes proliferation and invasion of human colorectal cancer LoVo cells through enhancing ETS-1 activity. Genetics and Molecular Research, 13(3), pp.6981-6994.

Li, W., Cogswell, C. and LoTurco, J., 1998. Neuronal Differentiation of Precursors in the Neocortical Ventricular Zone Is Triggered by BMP. The Journal of Neuroscience, 18(21), pp.8853-8862.

Lindblad-Toh, K., Wade, C., Mikkelsen, T., Karlsson, E., Jaffe, D., Kamal, M., Clamp, M., Chang, J., Kulbokas, E., Zody, M., Mauceli, E., Xie, X., Breen, M., Wayne, R., Ostrander, E., Ponting, C., Galibert, F., Smith, D., deJong, P., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C., Cook, A., Cuff, J., Daly, M., DeCaprio, D., Gnerre, S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K., Parker, H., Pollinger, J., Searle, S., Sutter, N., Thomas, R., Webber, C. and Lander, E., 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature, 438(7069), pp.803-819.

Liu, E., Russ, J., Cali, C., Phan, J., Amlie-Wolf, A. and Lee, E., 2019. Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. Cell Reports, 27(5), pp.1409-1421.e6.

Liu, J., Fang, J., Wang, Z., Rong, P., Hong, Y., Fan, Y., Wang, X., Park, J., Jin, Y., Liu, C., Zhu, B. and Kong, J., 2016. Transcutaneous vagus nerve stimulation modulates amygdala functional connectivity in patients with depression. Journal of Affective Disorders, 205, pp.319-326.

Loeb, D., Padgett, R., Hardies, S., Shehee, W., Comer, M., Edgell, M. and Hutchison, C., 1986. The sequence of a large L1Md element reveals a tandemly repeated 5' end and several features found in retrotransposons. Molecular and Cellular Biology, 6(1), pp.168-182.

Logan, C. and Nusse, R., 2004. THE WNT SIGNALING PATHWAY IN DEVELOPMENT AND DISEASE. Annual Review of Cell and Developmental Biology, 20(1), pp.781-810.

LoTurco, J. and Bai, J., 2006. The multipolar stage and disruptions in neuronal migration. Trends in Neurosciences, 29(7), pp.407-413.

Luan, D., Korman, M., Jakubczak, J. and Eickbush, T., 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. Cell, 72(4), pp.595-605.

Luchetti, A., Plazzi, F. and Mantovani, B., 2017. Evolution of Two Short Interspersed Elements in Callorhinchus milii (Chondrichthyes, Holocephali) and Related Elements in Sharks and the Coelacanth. Genome Biology and Evolution, 9(6), pp.1406-1417.

MacDuff, D., Demorest, Z. and Harris, R., 2009. AID can restrict L1 retrotransposition suggesting a dual role in innate and adaptive immunity. Nucleic Acids Research, 37(6), pp.1854-1867.

Macia, A., Muñoz-Lopez, M., Cortes, J., Hastings, R., Morell, S., Lucena-Aguilar, G., Marchal, J., Badge, R. and Garcia-Perez, J., 2011. Epigenetic Control of Retrotransposon Expression in Human Embryonic Stem Cells. Molecular and Cellular Biology, 31(2), pp.300-316.

Macia, A., Widmann, T., Heras, S., Ayllon, V., Sanchez, L., Benkaddour-Boumzaouad, M., Muñoz-Lopez, M., Rubio, A., Amador-Cubero, S., Blanco-Jimenez, E., Garcia-Castro, J., Menendez, P., Ng, P., Muotri, A., Goodier, J. and Garcia-Perez, J., 2016. Engineered LINE-1 retrotransposition in nondividing human neurons. Genome Research, 27(3), pp.335-348.

Makałowski, W., Gotea, V., Pande, A. and Makałowska, I., 2019. Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics. Methods in Molecular Biology, pp.177-207.

Malatesta, P., Hack, M., Hartfuss, E., Kettenmann, H., Klinkert, W., Kirchhoff, F. and Götz, M., 2003. Neuronal or Glial Progeny. Neuron, 37(5), pp.751-764.

Malatesta, P., Hartfuss, E. and Gotz, M., 2000. Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neuronal lineage. Development, 127(24), pp.5253-5263.

Mall, E., Herrmann, D. and Niemann, H., 2017. Murine pluripotent stem cells with a homozygous knockout of Foxg1 show reduced differentiation towards cortical progenitors in vitro. Stem Cell Research, 25, pp.50-60.

Mallamaci, A. and Stoykova, A., 2006. Gene networks controlling early cerebral cortex arealization. European Journal of Neuroscience, 23(4), pp.847-856.

Mallamaci, A., 2011. Molecular bases of cortico-cerebral regionalization. Progress in Brain Research, pp.37-64.

Manuel, M., Martynoga, B., Yu, T., West, J., Mason, J. and Price, D., 2010. The transcription factor Foxg1 regulates the competence of telencephalic cells to adopt subpallial fates in mice. Development, 137(3), pp.487-497.

Marçal, N., Patel, H., Dong, Z., Belanger-Jasmin, S., Hoffman, B., Helgason, C., Dang, J. and Stifani, S., 2005. Antagonistic Effects of Grg6 and Groucho/TLE on the Transcription Repression Activity of Brain Factor 1/FoxG1 and Cortical Neuron Differentiation. Molecular and Cellular Biology, 25(24), pp.10916-10929.

Martin, S., Branciforte, D., Keller, D. and Bain, D., 2003. Trimeric structure for an essential protein in L1 retrotransposition. Proceedings of the National Academy of Sciences, 100(24), pp.13815-13820.

Martin, S., Cruceanu, M., Branciforte, D., Wai-lun Li, P., Kwok, S., Hodges, R. and Williams, M., 2005. LINE-1 Retrotransposition Requires the Nucleic Acid Chaperone Activity of the ORF1 Protein. Journal of Molecular Biology, 348(3), pp.549-561.

Martynoga, B., Morrison, H., Price, D. and Mason, J., 2005. Foxg1 is required for specification of ventral telencephalon and region-specific regulation of dorsal telencephalic precursor proliferation and apoptosis. Developmental Biology, 283(1), pp.113-127.

Maside, X., Bartolomé, C. and Charlesworth, B., 2002. S-element Insertions Are Associated with the Evolution of the Hsp70 Genes in Drosophila melanogaster. Current Biology, 12(19), pp.1686-1691.

Mateo, J., van den Berg, D., Haeussler, M., Drechsel, D., Gaber, Z., Castro, D., Robson, P., Lu, Q., Crawford, G., Flicek, P., Ettwiller, L., Wittbrodt, J., Guillemot, F. and Martynoga, B., 2014. Characterization of the neural stem cell gene regulatory network identifies OLIG2 as a multifunctional regulator of self-renewal. Genome Research, 25(1), pp.41-56.

Mathelier, A., Fornes, O., Arenillas, D., Chen, C., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., Zhang, A., Parcy, F., Lenhard, B., Sandelin, A. and Wasserman, W., 2015. JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Research, 44(D1), pp.D110-D115.

McClintock, B., 1950. The origin and behavior of mutable loci in maize. Proceedings of the National Academy of Sciences, 36(6), pp.344-355.

McDonald, J., 1998. Transposable elements, gene silencing and macroevolution. Trends in Ecology & Evolution, 13(3), pp.94-95.

Mencarelli, M., Spanhol-Rosseto, A., Artuso, R., Rondinella, D., De Filippis, R., Bahi-Buisson, N., Nectoux, J., Rubinsztajn, R., Bienvenu, T., Moncla, A., Chabrol, B., Villard, L., Krumina, Z., Armstrong, J., Roche, A., Pineda, M., Gak, E., Mari, F., Ariani, F. and Renieri, A., 2009. Novel FOXG1 mutations associated with the congenital variant of Rett syndrome. Journal of Medical Genetics, 47(1), pp.49-53.

Metcalfe, C., Filee, J., Germon, I., Joss, J. and Casane, D., 2012. Evolution of the Australian Lungfish (Neoceratodus forsteri) Genome: A Major Role for CR1 and L2 LINE Elements. Molecular Biology and Evolution, 29(11), pp.3529-3539.

Meyer, A., Schloissnig, S., Franchini, P., Du, K., Woltering, J., Irisarri, I., Wong, W., Nowoshilow, S., Kneitz, S., Kawaguchi, A., Fabrizius, A., Xiong, P., Dechaud, C., Spaink, H., Volff, J., Simakov, O., Burmester, T., Tanaka, E. and Schartl, M., 2021. Giant lungfish genome elucidates the conquest of land by vertebrates. Nature, 590(7845), pp.284-289.

Meyer, G., 2010. Building a human cortex: the evolutionary differentiation of Cajal-Retzius cells and the cortical hem. Journal of Anatomy, 217(4), pp.334-343.

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., WKinzler, K., Vogelstein, B. and Nakamura, Y., 1992. Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. Cancer Res . 1992 Feb 1;52(3):643-5., 52(3), pp.643-5.

Mikkelsen, T., Ku, M., Jaffe, D., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T., Koche, R., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. and Bernstein, B., 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature, 448(7153), pp.553-560.

Mishra, S., Gray, C., Miyamoto, S., Bers, D. and Brown, J., 2011. Location Matters. Circulation Research, 109(12), pp.1354-1362.

Mita, P., Wudzinska, A., Sun, X., Andrade, J., Nayak, S., Kahler, D., Badri, S., LaCava, J., Ueberheide, B., Yun, C., Fenyö, D. and Boeke, J., 2018. LINE-1 protein localization and functional dynamics during the cell cycle. eLife, 7.

Miyata, T., Kawaguchi, A., Saito, K., Kawano, M., Muto, T. and Ogawa, M., 2004. Asymmetric production of surface-dividing and non-surface-dividing cortical progenitor cells. Development, 131(13), pp.3133-3145.

Miyoshi, G. and Fishell, G., 2012. Dynamic FoxG1 Expression Coordinates the Integration of Multipolar Pyramidal Neuron Precursors into the Cortical Plate. Neuron, 74(6), pp.1045-1058.

Miyoshi, T., Makino, T. and Moran, J., 2019. Poly(ADP-Ribose) Polymerase 2 Recruits Replication Protein A to Sites of LINE-1 Integration to Facilitate Retrotransposition. Molecular Cell, 75(6), pp.1286-1298.e12.

Mnatzakanian, G., Lohi, H., Munteanu, I., Alfred, S., Yamada, T., MacLeod, P., Jones, J., Scherer, S., Schanen, N., Friez, M., Vincent, J. and Minassian, B., 2004. A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. Nature Genetics, 36(4), pp.339-341.

Mokrani, H., Sharaf el dein, O., Mansuroglu, Z. and Bonnefoy, E., 2006. Binding of YY1 to the Proximal Region of the Murine Beta Interferon Promoter Is Essential To Allow CBP Recruitment and K8H4/K14H3 Acetylation on the Promoter Region after Virus Infection. Molecular and Cellular Biology, 26(22), pp.8551-8561.

Molecular Biology and Evolution, 1993. L1 A-monomer tandem arrays have expanded during the course of mouse L1 evolution.

Molyneaux, B., Arlotta, P., Hirata, T., Hibi, M. and Macklis, J., 2005. Fezl Is Required for the Birth and Specification of Corticospinal Motor Neurons. Neuron, 47(6), pp.817-831.

Molyneaux, B., Arlotta, P., Menezes, J. and Macklis, J., 2007. Neuronal subtype specification in the cerebral cortex. Nature Reviews Neuroscience, 8(6), pp.427-437.

Moore, J. and Haber, J., 1996. Capture of retrotransposon DNA at the sites of chromosomal double-strand breaks. Nature, 383(6601), pp.644-646.

Morrish, T., Gilbert, N., Myers, J., Vincent, B., Stamato, T., Taccioli, G., Batzer, M. and Moran, J., 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nature Genetics, 31(2), pp.159-165.

Moyon, S., Huynh, J., Dutta, D., Zhang, F., Ma, D., Yoo, S., Lawrence, R., Wegner, M., John, G., Emery, B., Lubetzki, C., Franklin, R., Fan, G., Zhu, J., Dupree, J. and Casaccia, P., 2016. Functional Characterization of DNA Methylation in the Oligodendrocyte Lineage. Cell Reports, 15(4), pp.748-760.

Muotri, A., Chu, V., Marchetto, M., Deng, W., Moran, J. and Gage, F., 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature, 435(7044), pp.903-910.

Muotri, A., Marchetto, M., Coufal, N., Oefner, R., Yeo, G., Nakashima, K. and Gage, F., 2010. L1 retrotransposition in neurons is modulated by MeCP2. Nature, 468(7322), pp.443-446.

Muotri, A., Zhao, C., Marchetto, M. and Gage, F., 2009. Environmental influence on L1 retrotransposons in the adult hippocampus. Hippocampus, 19(10), pp.1002-1007.

Muratcioglu, S., Presman, D., Pooley, J., Grøntved, L., Hager, G., Nussinov, R., Keskin, O. and Gursoy, A., 2015. Structural Modeling of GR Interactions with the SWI/SNF Chromatin Remodeling Complex and C/EBP. Biophysical Journal, 109(6), pp.1227-1239.

Muratcioglu, S., Presman, D., Pooley, J., Grøntved, L., Hager, G., Nussinov, R., Keskin, O. and Gursoy, A., 2015. Structural Modeling of GR Interactions with the SWI/SNF Chromatin Remodeling Complex and C/EBP. Biophysical Journal, 109(6), pp.1227-1239.

Muzio, L., Di Benedetto, B., Stoykova, A., Boncinelli, E., Gruss, P. and Mallamaci, A., 2002. Emx2 and Pax6 Control Regionalization of the Pre-neuronogenic Cortical Primordium. Cerebral Cortex, 12(2), pp.129-139.

Myers, J., Vincent, B., Udall, H., Watkins, W., Morrish, T., Kilroy, G., Swergold, G., Henke, J., Henke, L., Moran, J., Jorde, L. and Batzer, M., 2002. A Comprehensive Analysis of Recently Integrated Human Ta L1 Elements. The American Journal of Human Genetics, 71(2), pp.312-326.

Myers, J., Vincent, B., Udall, H., Watkins, W., Morrish, T., Kilroy, G., Swergold, G., Henke, J., Henke, L., Moran, J., Jorde, L. and Batzer, M., 2002. A Comprehensive Analysis of Recently Integrated Human Ta L1 Elements. The American Journal of Human Genetics, 71(2), pp.312-326.

Nature, 2002. Initial sequencing and comparative analysis of the mouse genome. 420(6915), pp.520-562.

Naufer, M., Furano, A. and Williams, M., 2019. Protein-nucleic acid interactions of LINE-1 ORF1p. Seminars in Cell & Developmental Biology, 86, pp.140-149.

Nguyen, S., Meletis, K., Fu, D., Jhaveri, S. and Jaenisch, R., 2007. Ablation of de novo DNA methyltransferase Dnmt3a in the nervous system leads to neuromuscular defects and shortened lifespan. Developmental Dynamics, 236(6), pp.1663-1676.

Nicetto, D., Donahue, G., Jain, T., Peng, T., Sidoli, S., Sheng, L., Montavon, T., Becker, J., Grindheim, J., Blahnik, K., Garcia, B., Tan, K., Bonasio, R., Jenuwein, T. and Zaret, K., 2019. H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. Science, 363(6424), pp.294-297.

Nigumann, P., Redik, K., Mätlik, K. and Speek, M., 2002. Many Human Genes Are Transcribed from the Antisense Promoter of L1 Retrotransposon. Genomics, 79(5), pp.628-634.

Noctor, S., Flint, A., Weissman, T., Dammerman, R. and Kriegstein, A., 2001. Neurons derived from radial glial cells establish radial units in neocortex. Nature, 409(6821), pp.714-720.

Noctor, S., Martínez-Cerdeño, V., Ivic, L. and Kriegstein, A., 2004. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. Nature Neuroscience, 7(2), pp.136-144.

Nomura, Y., 2005. Early behavior characteristics and sleep disturbance in Rett syndrome. Brain and Development, 27, pp.S35-S42.

Ohshima, K., 2012. Parallel Relaxation of Stringent RNA Recognition in Plant and Mammalian L1 Retrotransposons. Molecular Biology and Evolution, 29(11), pp.3255-3259.

Okano, M., Bell, D., Haber, D. and Li, E., 1999. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. Cell, 99(3), pp.247-257.

O'Leary, D. and Terashima, T., 1988. Cortical axons branch to multiple subcortical targets by interstitial axon budding: Implications for target recognition and “waiting periods”. Neuron, 1(10), pp.901-910.

Ooi, S., O'Donnell, A. and Bestor, T., 2009. Mammalian cytosine methylation at a glance. Journal of Cell Science, 122(16), pp.2787-2791.

Orgel, L. and Crick, F., 1980. Selfish DNA: the ultimate parasite. Nature, 284(5757), pp.604-607.

Ostertag, E. and Kazazian Jr, H., 2001. Biology of Mammalian L1 Retrotransposons. Annual Review of Genetics, 35(1), pp.501-538.

Pacaud, R., Sery, Q., Oliver, L., Vallette, F., Tost, J. and Cartron, P., 2014. DNMT3L interacts with transcription factors to target DNMT3L/DNMT3B to specific DNA sequences: Role of the DNMT3L/DNMT3B/p65-NFκB complex in the (de-)methylation of TRAF1. Biochimie, 104, pp.36-49.

Pacaud, R., Sery, Q., Oliver, L., Vallette, F., Tost, J. and Cartron, P., 2014. DNMT3L interacts with transcription factors to target DNMT3L/DNMT3B to specific DNA sequences: Role of the DNMT3L/DNMT3B/p65-NFκB complex in the (de-)methylation of TRAF1. Biochimie, 104, pp.36-49.

Padgett, R., Hutchison, C. and Edgell, M., 1988. The F-type 5′ motif of mouse L1 elements: a major class of L1 termini similar to the A-type in organization but unrelated in sequence. Nucleic Acids Research, 16(2), pp.739-749.

Pelka, G., Watson, C., Christodoulou, J. and Tam, P., 2005. Distinct expression profiles of Mecp2 transcripts with different lengths of 3'UTR in the brain and visceral organs during mouse development. Genomics, 85(4), pp.441-452.

Perepelitsa-Belancio, V. and Deininger, P., 2003. RNA truncation by premature polyadenylation attenuates human mobile element activity. Nature Genetics, 35(4), pp.363-366.

Perera, B., Tsai, Z., Colwell, M., Jones, T., Goodrich, J., Wang, K., Sartor, M., Faulk, C. and Dolinoy, D., 2019. Somatic expression of piRNA and associated machinery in the mouse identifies short, tissue-specific piRNA. Epigenetics, 14(5), pp.504-521.

Piskareva, O., Ernst, C., Higgins, N. and Schmatchenko, V., 2013. The carboxy-terminal segment of the human LINE-1 ORF2 protein is involved in RNA binding. FEBS Open Bio, 3(1), pp.433-437.

Platt, R., Chen, S., Zhou, Y., Yim, M., Swiech, L., Kempton, H., Dahlman, J., Parnas, O., Eisenhaure, T., Jovanovic, M., Graham, D., Jhunjhunwala, S., Heidenreich, M., Xavier, R., Langer, R., Anderson, D., Hacohen, N., Regev, A., Feng, G., Sharp, P. and Zhang, F., 2014. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. Cell, 159(2), pp.440-455.

Price, N. and Born, R., 2010. Timescales of Sensory- and Decision-Related Activity in the Middle Temporal and Medial Superior Temporal Areas. Journal of Neuroscience, 30(42), pp.14036-14045.

Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character, 1922. Croonian Lecture:—On the mechanism of heredity. 94(659), pp.162-197.

Proekt, A., Wong, J., Zhurov, Y., Kozlova, N., Weiss, K. and Brezina, V., 2008. Predicting Adaptive Behavior in the Environment from Central Nervous System Dynamics. PLoS ONE, 3(11), p.e3678.

Protasova, M., Andreeva, T. and Rogaev, E., 2021. Factors Regulating the Activity of LINE1 Retrotransposons. Genes, 12(10), p.1562.

Protasova, M., Gusev, F., Grigorenko, A., Kuznetsova, I., Rogaev, E. and Andreeva, T., 2017. Quantitative analysis of L1-retrotransposons in Alzheimer's disease and aging. Biochemistry (Moscow), 82(8), pp.962-971.

Puelles, L., Ayad, A., Alonso, A., Sandoval, J., MartÍnez-de-la-Torre, M., Medina, L. and Ferran, J., 2015. Selective early expression of the orphan nuclear receptorNr4a2identifies the claustrum homolog in the avian mesopallium: Impact on sauropsidian/mammalian pallium comparisons. Journal of Comparative Neurology, 524(3), pp.665-703.

Raballo, R., Rhee, J., Lyn-Cook, R., Leckman, J., Schwartz, M. and Vaccarino, F., 2000. Basic Fibroblast Growth Factor (Fgf2) Is Necessary for Cell Proliferation and Neurogenesis in the Developing Cerebral Cortex. The Journal of Neuroscience, 20(13), pp.5012-5023.

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

Rakic, P., 1974. Neurons in Rhesus Monkey Visual Cortex: Systematic Relation between Time of Origin and Eventual Disposition. Science, 183(4123), pp.425-427.

Rakic, P., 2003. Elusive radial glial cells: Historical and evolutionary perspective. Glia, 43(1), pp.19-32.

Rash, B., Tomasi, S., Lim, H., Suh, C. and Vaccarino, F., 2013. Cortical Gyrification Induced by Fibroblast Growth Factor 2 in the Mouse Brain. Journal of Neuroscience, 33(26), pp.10802-10814.

Ravel-Godreuil, C., Znaidi, R., Bonnifet, T., Joshi, R. and Fuchs, J., 2021. Transposable elements as new players in neurodegenerative diseases. FEBS Letters, 595(22), pp.2733-2755.

Reiner, A., Perkel, D., Bruce, L., Butler, A., Csillag, A., Kuenzel, W., Medina, L., Paxinos, G., Shimizu, T., Striedter, G., Wild, M., Ball, G., Durand, S., Gütürkün, O., Lee, D., Mello, C., Powers, A., White, S., Hough, G., Kubikova, L., Smulders, T., Wada, K., Dugas-Ford, J., Husband, S., Yamamoto, K., Yu, J., Siang, C. and Jarvis, E., 2004. Revised nomenclature for avian telencephalon and some related brainstem nuclei. The Journal of Comparative Neurology, 473(3), pp.377-414.

Rekvig, O., 2022. The Anti-DNA Antibodies: Their Specificities for Unique DNA Structures and Their Unresolved Clinical Impact—A System Criticism and a Hypothesis. Frontiers in Immunology, 12.

Richardson, S., Doucet, A., Kopera, H., Moldovan, J., Garcia-Perez, J. and Moran, J., 2015. The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. Microbiology Spectrum, 3(2).

Richardson, S., Doucet, A., Kopera, H., Moldovan, J., Garcia-Perez, J. and Moran, J., 2015. The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. Microbiology Spectrum, 3(2).

Rojas-Ríos, P. and Simonelig, M., 2018. piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. Development, 145(17).

Rosser, J., 2012. L1 expression and regulation in humans and rodents. Frontiers in Bioscience, E4(1), p.2203.

Roth, G. and Dicke, U., 2013. Evolution of Nervous Systems and Brains. Neurosciences - From Molecule to Behavior: a university textbook, pp.19-45.

Rowe, T., 2017. The Emergence of Mammals. Evolution of Nervous Systems, pp.1-52.

Roy, N., Nakano, T., Keyoung, H., Windrem, M., Rashbaum, W., Alonso, M., Kang, J., Peng, W., Carpenter, M., Lin, J., Nedergaard, M. and Goldman, S., 2004. Telomerase immortalization of neuronally restricted progenitor cells derived from the human fetal spinal cord. Nature Biotechnology, 22(3), pp.297-305.

Ryu, S., Kang, K., Yoo, T., Joe, C. and Chung, J., 2011. Transcriptional repression of repeat-derived transcripts correlates with histone hypoacetylation at repetitive DNA elements in aged mice brain. Experimental Gerontology, 46(10), pp.811-818.

Saksouk, N., Simboeck, E. and Déjardin, J., 2015. Constitutive heterochromatin formation and transcription in mammals. Epigenetics & Chromatin, 8(1).

Sanchez-Luque, F., Kempen, M., Gerdes, P., Vargas-Landin, D., Richardson, S., Troskie, R., Jesuadian, J., Cheetham, S., Carreira, P., Salvador-Palomeque, C., García-Cañadas, M., Muñoz-Lopez, M., Sanchez, L., Lundberg, M., Macia, A., Heras, S., Brennan, P., Lister, R., Garcia-Perez, J., Ewing, A. and Faulkner, G., 2019. LINE-1 Evasion of Epigenetic Repression in Humans. Molecular Cell, 75(3), pp.590-604.e12.

Sanchez-Luque, F., Richardson, S. and Faulkner, G., 2016. Retrotransposon Capture Sequencing (RC-Seq): A Targeted, High-Throughput Approach to Resolve Somatic L1 Retrotransposition in Humans. Methods in Molecular Biology, pp.47-77.

Sasaki, T., Nishihara, H., Hirakawa, M., Fujimura, K., Tanaka, M., Kokubo, N., Kimura-Yoshida, C., Matsuo, I., Sumiyama, K., Saitou, N., Shimogori, T. and Okada, N., 2008. Possible involvement of SINEs in mammalian-specific brain formation. Proceedings of the National Academy of Sciences, 105(11), pp.4220-4225.

Schlappi, M., 1994. Epigenetic regulation of the maize Spm transposable element: Novel activation of a methylated promoter by TnpA. Cell, 77(3), pp.427-437.

Schlenke, T. and Begun, D., 2004. Strong selective sweep associated with a transposon insertion in Drosophila simulans. Proceedings of the National Academy of Sciences, 101(6), pp.1626-1631.

Seidl, M. and Thomma, B., 2017. Transposable Elements Direct The Coevolution between Plants and Microbes. Trends in Genetics, 33(11), pp.842-851.

Seoane, J., Le, H., Shen, L., Anderson, S. and Massagué, J., 2004. Integration of Smad and Forkhead Pathways in the Control of Neuroepithelial and Glioblastoma Cell Proliferation. Cell, 117(2), pp.211-223.

Seoane, J., Le, H., Shen, L., Anderson, S. and Massagué, J., 2004. Integration of Smad and Forkhead Pathways in the Control of Neuroepithelial and Glioblastoma Cell Proliferation. Cell, 117(2), pp.211-223.

Servant, G., Streva, V., Derbes, R., Wijetunge, M., Neeland, M., White, T., Belancio, V., Roy-Engel, A. and Deininger, P., 2017. The Nucleotide Excision Repair Pathway Limits L1 Retrotransposition. Genetics, 205(1), pp.139-153.

Shahbazian, M., Young, J., Yuva-Paylor, L., Spencer, C., Antalffy, B., Noebels, J., Armstrong, D., Paylor, R. and Zoghbi, H., 2002. Mice with Truncated MeCP2 Recapitulate Many Rett Syndrome Features and Display Hyperacetylation of Histone H3. Neuron, 35(2), pp.243-254.

Shao, F., Han, M. and Peng, Z., 2019. Evolution and diversity of transposable elements in fish genomes. Scientific Reports, 9(1).

Shimamura, K. and Rubenstein, J., 1997. Inductive interactions direct early regionalization of the mouse forebrain. Development, 124(14), pp.2709-2718.

Shoichet, S., Kunde, S., Viertel, P., Schell-Apacik, C., von Voss, H., Tommerup, N., Ropers, H. and Kalscheuer, V., 2005. Haploinsufficiency of novel FOXG1B variants in a patient with severe mental retardation, brain malformations and microcephaly. Human Genetics, 117(6), pp.536-544.

Shreyer, D. and Jones, E., 1988. Topographic sequence of outgrowth of corticospinal axons in the rat: a study using retrograde axonal labeling with Fast blue. Developmental Brain Research, 38(1), pp.89-101.

Siegenthaler, J., Tremper-Wells, B. and Miller, M., 2007. Foxg1 Haploinsufficiency Reduces the Population of Cortical Intermediate Progenitor Cells: Effect of Increased p21 Expression. Cerebral Cortex, 18(8), pp.1865-1875.

Silva, J., Shabalina, S., Harris, D., Spouge, J. and Kondrashov, A., 2003. Conserved fragments of transposable elements in intergenic regions: evidence for widespread recruitment of MIR- and L2-derived sequences within the mouse and human genomes. Genetical Research, 82(1), pp.1-18.

Singer, T., McConnell, M., Marchetto, M., Coufal, N. and Gage, F., 2010. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes?. Trends in Neurosciences, 33(8), pp.345-354.

Siomi, M., Sato, K., Pezic, D. and Aravin, A., 2011. PIWI-interacting small RNAs: the vanguard of genome defence. Nature Reviews Molecular Cell Biology, 12(4), pp.246-258.

Skalka, A., 2014. Retroviral DNA Transposition: Themes and Variations. Microbiology Spectrum, 2(5).

Smart, I. and McSherry, G., 1982. Growth patterns in the lateral wall of the mouse telencephalon. II. Histological changes during and subsequent to the period of isocortical neuron production. Journal of Anatomy, 134(Pt 3), pp.415-42.

Smart, I., 1973. Proliferative characteristics of the ependymal layer during the early development of the mouse neocortex: a pilot study based on recording the number, location and plane of cleavage of mitotic figures. Journal of Anatomy, 116(Pt1), pp.67–91.

Smit, A., Tóth, G., Riggs, A. and Jurka, J., 1995. Ancestral, Mammalian-wide Subfamilies of LINE-1 Repetitive Sequences. Journal of Molecular Biology, 246(3), pp.401-417.

Smith, K., Budde, A., Dill-Macky, R., Rasmusson, D., Schiefelbein, E., Steffenson, B., Wiersma, J., Wiersma, J. and Zhang, B., 2013. Registration of 'Quest' Spring Malting Barley with Improved Resistance to Fusarium Head Blight. Journal of Plant Registrations, 7(2), pp.125-129.

Sohn, E., Jo, Y. and Park, H., 2019. Downregulation MIWI-piRNA regulates the migration of Schwann cells in peripheral nerve injury. Biochemical and Biophysical Research Communications, 519(3), pp.605-612.

Soifer, H., 2006. Do Small RNAs Interfere With LINE-1?. Journal of Biomedicine and Biotechnology, 2006, pp.1-8.

Sookdeo, A., Hepp, C., McClure, M. and Boissinot, S., 2013. Revisiting the evolution of mouse LINE-1 in the genomic era. Mobile DNA, 4(1).

Sookdeo, A., Hepp, C., McClure, M. and Boissinot, S., 2013. Revisiting the evolution of mouse LINE-1 in the genomic era. Mobile DNA, 4(1).

Sotero-Caio, C., Platt, R., Suh, A. and Ray, D., 2017. Evolution and Diversity of Transposable Elements in Vertebrate Genomes. Genome Biology and Evolution, 9(1), pp.161-177.

Spruston, N., 2008. Pyramidal neurons: dendritic structure and synaptic integration. Nature Reviews Neuroscience, 9(3), pp.206-221.

Squillaro, T., Alessio, N., Cipollaro, M., Melone, M., Hayek, G., Renieri, A., Giordano, A. and Galderisi, U., 2012. Reduced expression ofMECP2affects cell commitment and maintenance in neurons by triggering senescence: new perspective for Rett syndrome. Molecular Biology of the Cell, 23(8), pp.1435-1445.

Stage, D. and Eickbush, T., 2009. Origin of nascent lineages and the mechanisms used to prime second-strand DNA synthesis in the R1 and R2 retrotransposons of Drosophila. Genome Biology, 10(5), p.R49.

Stetson, D., Ko, J., Heidmann, T. and Medzhitov, R., 2008. Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity. Cell, 134(4), pp.587-598.

Stetson, D., Ko, J., Heidmann, T. and Medzhitov, R., 2008. Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity. Cell, 134(4), pp.587-598.

Striano, P., Paravidino, R., Sicca, F., Chiurazzi, P., Gimelli, S., Coppola, A., Robbiano, A., Traverso, M., Pintaudi, M., Giovannini, S., Operto, F., Vigliano, P., Granata, T., Coppola, G., Romeo, A., Specchio, N., Giordano, L., Osborne, L., Gimelli, G., Minetti, C. and Zara, F., 2011. West syndrome associated with 14q12 duplications harboring FOXG1. Neurology, 76(18), pp.1600-1602.

Striedter, G., 2005. Principles of Brain Evolution. Sunderland, MA.

Sturtevant, A., 1913. The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of association. Journal of Experimental Zoology, 14(1), pp.43-59.

Suarez, N., Macia, A. and Muotri, A., 2017. LINE-1 retrotransposons in healthy and diseased human brain. Developmental Neurobiology, 78(5), pp.434-455.

Sun, W., Samimi, H., Gamez, M., Zare, H. and Frost, B., 2018. Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. Nature Neuroscience, 21(8), pp.1038-1048.

Swergold, G., 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. Molecular and Cellular Biology, 10(12), pp.6718-6729.

Swergold, G., 1990. Identification, characterization, and cell specificity of a human LINE-1 promoter. Molecular and Cellular Biology, 10(12), pp.6718-6729.

Szak, S., Pickeral, O., Makalowski, W., Boguski, M., Landsman, D. and Boeke, J., 2002. Genome Biology, 3(10), pp.research0052.1.

Takahashi, T., Nowakowski, R. and Caviness, V., 1996. The Leaving or Q Fraction of the Murine Cerebral Proliferative Epithelium: A General Model of Neocortical Neuronogenesis. The Journal of Neuroscience, 16(19), pp.6183-6196.

Tang, W. and Liang, P., 2020. Identification and characterization of retro-DNAs, a new type of retrotransposons originated from DNA transposons, in primate genomes.

Taylor, M., Altukhov, I., Molloy, K., Mita, P., Jiang, H., Adney, E., Wudzinska, A., Badri, S., Ischenko, D., Eng, G., Burns, K., Fenyö, D., Chait, B., Alexeev, D., Rout, M., Boeke, J. and LaCava, J., 2018. Dissection of affinity captured LINE-1 macromolecular complexes. eLife, 7.

Taylor, M., LaCava, J., Mita, P., Molloy, K., Huang, C., Li, D., Adney, E., Jiang, H., Burns, K., Chait, B., Rout, M., Boeke, J. and Dai, L., 2013. Affinity Proteomics Reveals Human Host Factors Implicated in Discrete Stages of LINE-1 Retrotransposition. Cell, 155(5), pp.1034-1048.

Tchenio, T., 2000. Members of the SRY family regulate the human LINE retrotransposons. Nucleic Acids Research, 28(2), pp.411-415.

Tchenio, T., 2000. Members of the SRY family regulate the human LINE retrotransposons. Nucleic Acids Research, 28(2), pp.411-415.

Telley, L., Govindan, S., Prados, J., Stevant, I., Nef, S., Dermitzakis, E., Dayer, A. and Jabaudon, D., 2016. Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. Science, 351(6280), pp.1443-1446.

Thomas, C., 2012. LINE-1: creators of neuronal diversity. Frontiers in Bioscience, E4(1), p.1663.

Thomas, C., Paquola, A. and Muotri, A., 2012. LINE-1 Retrotransposition in the Nervous System. Annual Review of Cell and Developmental Biology, 28(1), pp.555-573.

Thomas, C., Tejwani, L., Trujillo, C., Negraes, P., Herai, R., Mesci, P., Macia, A., Crow, Y. and Muotri, A., 2017. Modeling of TREX1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation. Cell Stem Cell, 21(3), pp.319-331.e8.

Thomas, C., Tejwani, L., Trujillo, C., Negraes, P., Herai, R., Mesci, P., Macia, A., Crow, Y. and Muotri, A., 2017. Modeling of TREX1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation. Cell Stem Cell, 21(3), pp.319-331.e8.

Tigani, W., Rossi, M., Artimagnella, O., Santo, M., Rauti, R., Sorbo, T., Ulloa Severino, F., Provenzano, G., Allegra, M., Caleo, M., Ballerini, L., Bozzi, Y. and Mallamaci, A., 2020. Foxg1 Upregulation Enhances Neocortical Activity. Cerebral Cortex, 30(9), pp.5147-5165.

Tilgner, H., Knowles, D., Johnson, R., Davis, C., Chakrabortty, S., Djebali, S., Curado, J., Snyder, M., Gingeras, T. and Guigó, R., 2012. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. Genome Research, 22(9), pp.1616-1625.

Tohyama, J., Yamamoto, T., Hosoki, K., Nagasaki, K., Akasaka, N., Ohashi, T., Kobayashi, Y. and Saitoh, S., 2011. West syndrome associated with mosaic duplication ofFOXG1in a patient with maternal uniparental disomy of chromosome 14. American Journal of Medical Genetics Part A, 155(10), pp.2584-2588.

Tóth, K., Pezic, D., Stuwe, E. and Webster, A., 2015. The piRNA Pathway Guards the Germline Genome Against Transposable Elements. Non-coding RNA and the Reproductive System, pp.51-77.

Trojer, P. and Reinberg, D., 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature?. Molecular Cell, 28(1), pp.1-13.

Tsankova, N., Renthal, W., Kumar, A. and Nestler, E., 2007. Epigenetic regulation in psychiatric disorders. Nature Reviews Neuroscience, 8(5), pp.355-367.

Tsujimura, K., Abematsu, M., Kohyama, J., Namihira, M. and Nakashima, K., 2009. Neuronal differentiation of neural precursor cells is promoted by the methyl-CpG-binding protein MeCP2. Experimental Neurology, 219(1), pp.104-111.

Upton, K., Gerhardt, D., Jesuadian, J., Richardson, S., Sánchez-Luque, F., Bodea, G., Ewing, A., Salvador-Palomeque, C., van der Knaap, M., Brennan, P., Vanderver, A. and Faulkner, G., 2015. Ubiquitous L1 Mosaicism in Hippocampal Neurons. Cell, 161(2), pp.228-239.

Venkatesh, B., Lee, A., Ravi, V., Maurya, A., Lian, M., Swann, J., Ohta, Y., Flajnik, M., Sutoh, Y., Kasahara, M., Hoon, S., Gangu, V., Roy, S., Irimia, M., Korzh, V., Kondrychyn, I., Lim, Z., Tay, B., Tohari, S., Kong, K., Ho, S., Lorente-Galdos, B., Quilez, J., Marques-Bonet, T., Raney, B., Ingham, P., Tay, A., Hillier, L., Minx, P., Boehm, T., Wilson, R., Brenner, S. and Warren, W., 2014. Elephant shark genome provides unique insights into gnathostome evolution. Nature, 505(7482), pp.174-179.

Venter, J., Adams, M., Myers, E., Li, P., Mural, R., Sutton, G., Smith, H., Yandell, M., Evans, C., Holt, R., Gocayne, J., Amanatides, P., Ballew, R., Huson, D., Wortman, J., Zhang, Q., Kodira, C., Zheng, X., Chen, L., Skupski, M., Subramanian, G., Thomas, P., Zhang, J., Gabor Miklos, G., Nelson, C., Broder, S., Clark, A., Nadeau, J., McKusick, V., Zinder, N., Levine, A., Roberts, R., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T., Higgins, M., Ji, R., Ke, Z., Ketchum, K., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G., Milshina, N., Moore, H., Naik, A., Narayan, V., Neelam, B., Nusskern, D., Rusch, D., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali,

F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J., Guigó, R., Campbell, M., Sjolander, K., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y., Coyne, M., Dahlke, C., Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A. and Zhu, X., 2001. The Sequence of the Human Genome. Science, 291(5507), pp.1304-1351.

Vezzali, R., Weise, S., Hellbach, N., Machado, V., Heidrich, S. and Vogel, T., 2016. The FOXG1/FOXO/SMAD network balances proliferation and differentiation of cortical progenitors and activates Kcnh3 expression in mature neurons. Oncotarget, 7(25), pp.37436-37455.

Villanueva-Cañas, J., Rech, G., Cara, M. and González, J., 2017. Beyond <scp>SNP</scp> s: how to detect selection on transposable element insertions. Methods in Ecology and Evolution, 8(6), pp.728-737.

Visel, A., 2004. GenePaint.org: an atlas of gene expression patterns in the mouse embryo. Nucleic Acids Research, 32(90001), pp.552D-556.

Voliva, C., Jahn, C., Comer, M., Hutchison, C. and Edgell, M., 1983. The L1Md long interspersed repeat family in the mouse: almost all examples are truncated at one end. Nucleic Acids Research, 11(24), pp.8847-8859.

Wallace, N., Wagstaff, B., Deininger, P. and Roy-Engel, A., 2008. LINE-1 ORF1 protein enhances Alu SINE retrotransposition. Gene, 419(1-2), pp.1-6.

Wang, L., Hou, S. and Han, Y., 2016. Hedgehog signaling promotes basal progenitor expansion and the growth and folding of the neocortex. Nature Neuroscience, 19(7), pp.888-896.

Wang, N., Finegold, M., Bradley, A., Ou, C., Abdelsayed, S., Wilde, M., Taylor, L., Wilson, D. and Darlington, G., 1995. Impaired Energy Homeostasis in C/EBPα Knockout Mice. Science, 269(5227), pp.1108-1112.

Wang, N., Finegold, M., Bradley, A., Ou, C., Abdelsayed, S., Wilde, M., Taylor, L., Wilson, D. and Darlington, G., 1995. Impaired Energy Homeostasis in C/EBPα Knockout Mice. Science, 269(5227), pp.1108-1112.

Ware, M., 1999. Coexistence of Widespread Clones and Large Radial Clones in Early Embryonic Ferret Cortex. Cerebral Cortex, 9(6), pp.636-645.

Warren, W., Hillier, L., Marshall Graves, J., Birney, E., Ponting, C., Grützner, F., Belov, K., Miller, W., Clarke, L., Chinwalla, A., Yang, S., Heger, A., Locke, D., Miethke, P., Waters, P., Veyrunes, F., Fulton, L., Fulton, B., Graves, T., Wallis, J., Puente, X., López-Otín, C., Ordóñez, G., Eichler, E., Chen, L., Cheng, Z., Deakin, J., Alsop, A., Thompson, K., Kirby, P., Papenfuss, A., Wakefield, M., Olender, T., Lancet, D., Huttley, G., Smit, A., Pask, A., Temple-Smith, P., Batzer, M., Walker, J., Konkel, M., Harris, R., Whittington, C., Wong, E., Gemmell, N., Buschiazzo, E., Vargas Jentzsch, I., Merkel, A., Schmitz, J., Zemann, A., Churakov, G., Kriegs, J., Brosius, J., Murchison, E., Sachidanandam, R., Smith, C., Hannon, G., Tsend-Ayush, E., McMillan, D., Attenborough, R., Rens, W., Ferguson-Smith, M., Lefèvre, C., Sharp, J., Nicholas, K., Ray, D., Kube, M., Reinhardt, R., Pringle, T., Taylor, J., Jones, R., Nixon, B., Dacheux, J., Niwa, H., Sekita, Y., Huang, X., Stark, A., Kheradpour, P., Kellis, M., Flicek, P., Chen, Y., Webber, C., Hardison, R., Nelson, J., Hallsworth-Pepin, K., Delehaunty, K., Markovic, C., Minx, P., Feng, Y., Kremitzki, C., Mitreva, M., Glasscock, J., Wylie, T., Wohldmann, P., Thiru, P., Nhan, M., Pohl, C., Smith, S., Hou, S., Nefedov, M., de Jong, P., Renfree, M., Mardis, E. and Wilson, R., 2008. Genome analysis of the platypus reveals unique signatures of evolution. Nature, 455(7210), pp.256-256.

Wei, W., Gilbert, N., Ooi, S., Lawler, J., Ostertag, E., Kazazian, H., Boeke, J. and Moran, J., 2001. Human L1 Retrotransposition: cis Preference versus trans Complementation. Molecular and Cellular Biology, 21(4), pp.1429-1439.

Weichenrieder, O., Repanas, K. and Perrakis, A., 2004. Crystal Structure of the Targeting Endonuclease of the Human LINE-1 Retrotransposon. Structure, 12(6), pp.975-986.

Weick, E. and Miska, E., 2014. piRNAs: from biogenesis to function. Development, 141(18), pp.3458-3471.

Weill, L., Shestakova, E. and Bonnefoy, E., 2003. Transcription Factor YY1 Binds to the Murine Beta Interferon Promoter and Regulates Its Transcriptional Capacity with a Dual Activator/Repressor Role. Journal of Virology, 77(5), pp.2903-2914.

Weimann, J., Zhang, Y., Levin, M., Devine, W., Brûlet, P. and McConnell, S., 1999. Cortical Neurons Require Otx1 for the Refinement of Exuberant Axonal Projections to Subcortical Targets. Neuron, 24(4), pp.819-831.

Weise, S., Arumugam, G., Villarreal, A., Videm, P., Heidrich, S., Nebel, N., Dumit, V., Sananbenesi, F., Reimann, V., Craske, M., Schilling, O., Hess, W., Fischer, A., Backofen, R. and Vogel, T., 2018. FOXG1 Regulates PRKAR2B Transcriptionally and Posttranscriptionally via miR200 in the Adult Hippocampus. Molecular Neurobiology, 56(7), pp.5188-5201.

Wellmann, U., Letz, M., Herrmann, M., Angermüller, S., Kalden, J. and Winkler, T., 2005. The evolution of human anti-double-stranded DNA autoantibodies. Proceedings of the National Academy of Sciences, 102(26), pp.9258-9263.

Wessler, S., 2006. Transposable elements and the evolution of eukaryotic genomes. Proceedings of the National Academy of Sciences, 103(47), pp.17600-17601.

Wong, E., Glasgow, S., Trigiani, L., Chitsaz, D., Rymar, V., Sadikot, A., Ruthazer, E., Hamel, E. and Kennedy, T., 2019. Spatial memory formation requires netrin-1 expression by neurons in the adult mammalian brain. Learning & Memory, 26(3), pp.77-83.

Wong, L., Singh, S., Wang, H., Hsu, C., Hu, S. and Lee, W., 2019. FOXG1-Related Syndrome: From Clinical to Molecular Genetics and Pathogenic Mechanisms. International Journal of Molecular Sciences, 20(17), p.4176.

Wong, L., Wu, Y., Hsu, C., Weng, W., Tsai, W. and Lee, W., 2019. Cognition and Evolution of Movement Disorders of FOXG1-Related Syndrome. Frontiers in Neurology, 10.

Xu, M., You, Y., Hunsicker, P., Hori, T., Small, C., Griswold, M. and Hecht, N., 2008. Mice Deficient for a Small Cluster of Piwi-Interacting RNAs Implicate Piwi-Interacting RNAs in Transposon Control1. Biology of Reproduction, 79(1), pp.51-57.

Xuan, S., Baptista, C., Balas, G., Tao, W., Soares, V. and Lai, E., 1995. Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres. Neuron, 14(6), pp.1141-1152.

Yang, A., Estecio, M., Garcia-Manero, G., Kantarjian, H. and Issa, J., 2003. Comment on "Chromosomal Instability and Tumors Promoted by DNA Hypomethylation" and "Induction of Tumors in Mice by Genomic Hypomethylation." Science, 302(5648), pp.1153-1153.

Yang, N. and Kazazian, H., 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nature Structural & Molecular Biology, 13(9), pp.763-771.

Yang, W., Lee, J. and Nowotny, M., 2006. Making and Breaking Nucleic Acids: Two-Mg2+-Ion Catalysis and Substrate Specificity. Molecular Cell, 22(1), pp.5-13.

Yao, J., Lai, E. and Stifani, S., 2001. The Winged-Helix Protein Brain Factor 1 Interacts with Groucho and Hes Proteins To Repress Transcription. Molecular and Cellular Biology, 21(6), pp.1962-1972.

Ypsilanti, A. and Rubenstein, J., 2015. Transcriptional and epigenetic mechanisms of early cortical development: An examination of how Pax6 coordinates cortical development. Journal of Comparative Neurology, 524(3), pp.609-629.

Yu, F., 2001. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Research, 29(21), pp.4493-4501.

Zemojtel, T., Penzkofer, T., Schultz, J., Dandekar, T., Badge, R. and Vingron, M., 2007. Exonization of active mouse L1s: a driver of transcriptome evolution?. BMC Genomics, 8(1), p.392.

Zhang, P., Kang, J., Gou, L., Wang, J., Xue, Y., Skogerboe, G., Dai, P., Huang, D., Chen, R., Fu, X., Liu, M. and He, S., 2015. MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. Cell Research, 25(2), pp.193-207.

Zhao, P., Yao, M., Chang, S., Gou, L., Liu, M., Qiu, Z. and Yuan, X., 2015. Novel function of PIWIL1 in neuronal polarization and migration via regulation of microtubule-associated proteins. Molecular Brain, 8(1).

Zingler, N., Willhoeft, U., Brose, H., Schoder, V., Jahns, T., Hanschmann, K., Morrish, T., Löwer, J. and Schumann, G., 2005. Analysis of 5′ junctions of human LINE-1 and Alu retrotransposons suggests an alternative model for 5′-end attachment requiring microhomology-mediated end-joining. Genome Research, 15(6), pp.780-789.

Acknowledgments

We thank Osvaldo Basilio Artimagnella for helping with the E16 culture and RNA immunoprecipitation (RIP) assay experiments (Fig. 6; Fig. 14).

We thank Laura Rigoldi for contributing to the genotyping of male and female CD1-wt mice used to refine DNA isolation protocols necessary for quantification of *LINE1*-DNA content (Fig. 11B 2,3,4).