Scuola Internazionale di Studi Superiori Avanzati

phD Course in Functional and Structural Genomics

In vivo analysis of L1 retrotransposition in Huntington's disease mouse brain

Thesis submitted for the degree of "Philosophiæ Doctor"

Candidate
Floreani Lavinia

Supervisor
prof. Gustincich Stefano

Co-Supervisor
prof. Persichetti Francesca

Academic Year 2015/2016

This page intentionally left blank

TABLE OF CONTENTS

This page intentionally left blank

ABSTRACT

Transposable elements (TEs) are mobile genetic elements that constitute a large fraction of eukaryotic genomes. TEs co-evolved with their host genomes, providing powerful tools of genome plasticity and regulation.

Long Interspersed Nuclear Elements 1 (L1) are the most numerous TEs in mouse and human genomes. They mobilize via a "copy and paste" mechanism that requires an RNA intermediate. Although most L1s have lost their activity during evolution, the remaining subset continues to move both in the germline and in adult somatic tissues.

Mounting evidence suggests that L1s are active in somatic cells of the mammalian brain and that dysregulated activation of L1s is associated with neuropathology, such as schizophrenia, Rett syndrome and Ataxia telengectasia.

Huntington's disease (HD) is an autosomal dominant disorder that manifests in mid-life and is characterized by neuronal loss, prominently in striatum and deep layers of the cerebral cortex. Typical HD-associated phenotypes include somatic genomic instability, epigenetic and transcriptional dysregulations, impaired neurogenesis and altered DNA damage response. At the same time, in HD, the origin and the role of many genetic and epigenetic modifiers acting on disease onset and progression remain largely unknown.

In this scenario, I investigated whether L1 retrotransposition might be altered in HD and if it could have a role in HD pathogenesis.

To address this question, using a novel *Taqman* qPCR technique, I characterized endogenous L1 retrotransposition events in the brains of a precise genetic mouse model of HD, considering both pre-symptomatic and symptomatic developmental stages. From this study, I showed that similar levels of L1 genomic copies are present between HD and control brains. Moreover, differences in full length L1 transcript levels have been reported in HD brains. Interestingly, in HD striatum, at 12 months of age, expression of full length L1s was consistently impaired, whereas in the cortex, L1 mRNA levels were increased in HD mice at 3 months and 24 months of age.

The dysregulation of L1 expression in the striatum of 12 months old mice did not appear to be linked to differential deposition of H3K4me3, H3K9me3, H3K27me3 and MeCP2 on L1 promoter in HD conditions. Nonetheless, L1 transcriptional alterations might involve a piRNA-mediated regulation. Indeed, in both cortex and striatum of adult HD and control mice I detected appreciable levels of MILI protein, the crucial

factor of piRNA biogenesis, suggesting its role not only in the germline but also in adult mammalian brains, as recently proposed by other two independent works.

Additionally, I showed that in a subset of neurons of the adult mouse pre-frontal cortex, endogenous L1 transcription is accompanied by expression of L1-encoded ORF2 protein.

Finally, by characterizing the transcription of active murine full length L1 elements in a broad range of developmental stages (from E10 up to 24 months), I described the expression profiles of endogenous L1 elements during the entire mouse development. From this study, I showed that a wave of L1 transcription takes place between E12 and P0 in both striatum and cerebral cortex and it is concomitant with telencephalic neurogenesis. In post-natal stages, L1 transcription is maintained at basal levels.

ABBREVIATIONS

ASP	Antisense Promoter
ChIP	Chromatin Immunoprecipitation
CNV	Copy Number Variation
DSB	Double Strand Break
dsDNA	Double Strand DNA
E	Embryonal day
EN	Endonuclease
gDNA	genomic DNA
H3K4me3	Tri-methylated Lys4 on Histone 3
H3K9me3	Tri-methylated Lys9 on Histone 3
H3K27me3	Tri-methylated Lys27 on Histone 3
HD	Huntington's disease
Htt	Huntingtin
IHC	Immunohistochemistry
KI	Knock-in
L1	LINE1, Long Interspersed Nuclear Element 1
LTR	Long Terminal Repeat
MeCP2	Methyl-CpG binding protein 2
MILI	murine Piwi-like protein 2
MIWI2	murine Piwi-like protein 4
mHtt	Mutant Huntingtin
mo	months
ORF	Open Reading Frame
ORF2p	ORF2 protein
PCR	Polymerase Chain Reaction
piRNA	Piwi-interacting RNA
poly-Q	poly-Glutamine
PRC2	Polycomb Repressive Complex 2
qPCR	quantitative PCR
RISC	RNA-Induced Silencing Complex
RNP	Ribonucleoprotein Particle
RT	Reverse Transcriptase

SINE	Short Interspersed Nuclear Element
TE	Transposable Element
TPRT	Target Primed Reverse Transcription
TSD	Target Site Duplication
UTR	Untranslated Region
WB	Western Blot
WT	Wild- Type

INTRODUCTION

1. TRANSPOSABLE ELEMENTS

Transposable elements (TEs), also known as *jumping genes*, are repetitive genomic sequences able to move from site to site within the genome. Apart from rare exceptions, TEs have been identified in all organisms and comprise a high proportion of a species' genome. For example, TEs occupy up to 85% of the *Zea mays* corn genome (figure 1) (Huang et al., 2012) and indeed it was in this specie that TEs were first discovered in 1950 by geneticist Barbara McClintock. In her seminal work, she revealed that distinct coloration of maize kernels was due to the excision of a transposon from a region encoding for an enzyme involved in pigmentation (McCLINTOCK, 1950).

Owing to their mutagenic properties, TEs were believed to represent *selfish DNA parasites,* that exist merely to propagate themselves, representing a potential threat to genome integrity (Fedoroff, 2012). Nonetheless, the waves of transposon mobilization that occurred throughout evolution and the accumulation of TE-derived sequences in the genome suggest that TEs could provide an inheritable and powerful source of genomic plasticity and regulation, as proposed since the beginning by McClintock. To further support this idea, over the last two decades, increasing evidence has shown that, although most transposable elements have lost their activity during evolution, the remaining subset continues to move both in the germline and in adult somatic tissues (Richardson et al., 2015). Although functional roles and mobilization mechanisms remain poorly understood, it is now clear that TEs coevolved with their host genomes. Indeed cells developed defenses from TE uncontrolled genomic expansion but, at the same time, took advantage of TE domestication.

Overall, ongoing TE mobilization is still impacting structure and functions of most eukaryotic genomes, thus creating important levels of intra- and inter-individual genomic variability.

According to their mechanism of mobilization, termed transposition, mammalian transposable elements are classified in two main groups: class I or retrotransposons, which mobilize via a RNA intermediate, using a *copy-and-paste* mechanism and class II or DNA transposons, which mobilize via a DNA intermediate, using a *cut-and-paste* mechanism (Wicker et al., 2007).

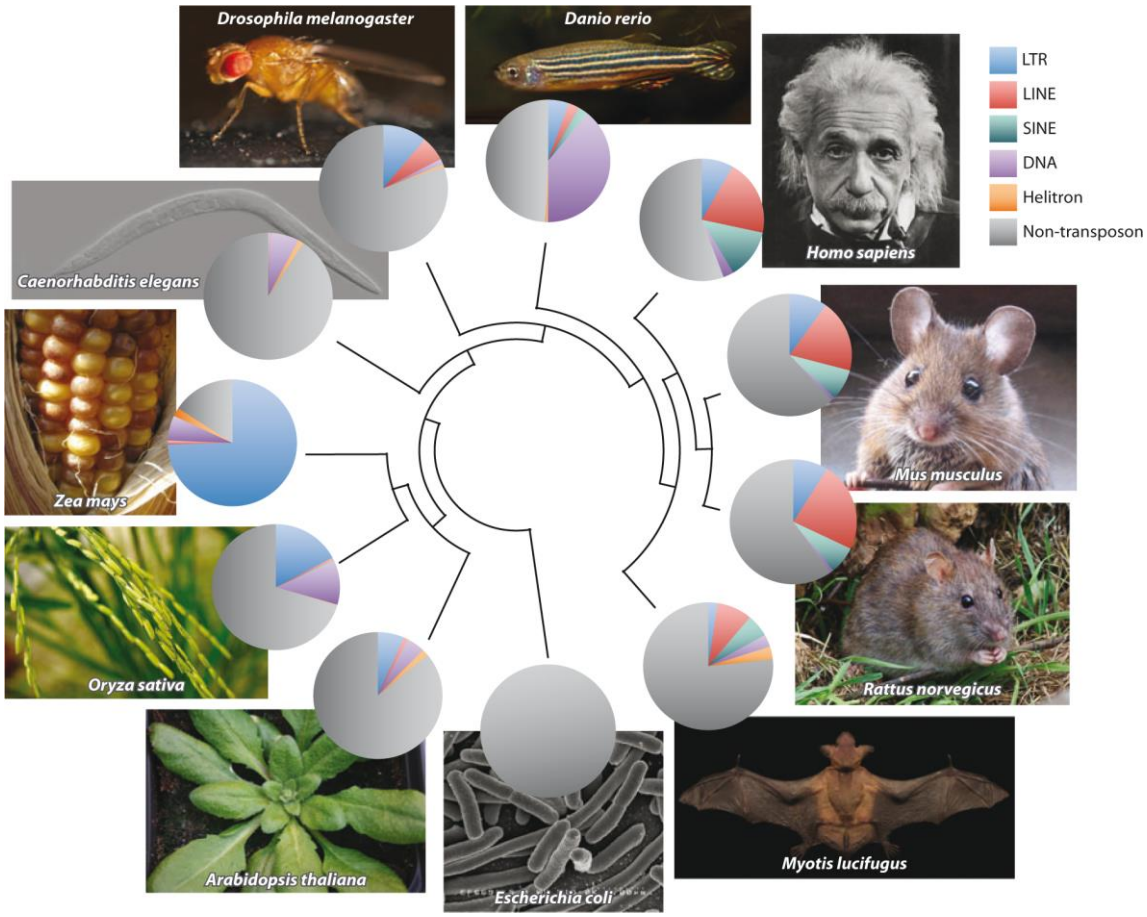


Figure 1. Transposons composition in different eukaryotic genomes. Pie charts show the abundance of each transposon type within the genome. The phylogenetic tree in the center describes evolutionary relationshops among species (from Huang et al.,2012).

1.1 DNA transposons

DNA transposons move within the genome by generating a DNA intermediate through the so called *cut-and-paste* mobilization mechanism: they are first excised from their original location and then reinserted elsewhere in the genome. Typically DNA transposons are composed of an open reading frame (ORF), encoding for the transposase enzyme, and are flanked by inverted terminal repeats (ITRs) (figure 2). In a round of transposition, the transposase recognizes ITRs and then mediates double strand excision and integration of DNA transposon into the new genomic location (Ivics et al., 1997). For their mobilization, autonomous DNA transposons rely on their own transposase enzyme, whereas non autonomous elements take advantage of the transposase encoded by autonomous transposons. Upon insertion, target site duplications (TSDs) are generated at both ends of the "moved" DNA transposon and the size of the TSDs results to be unique for each different class of DNA transposon. Thus the "*cut and paste*" mobilization process consists of a non-replicative mechanism where

the DNA transposon is moved to a new location but no transposon copies are generated, although few exceptions are represented by *Helitron* and *Maverick* transposons which are replicated and do not induce double strand breaks during their insertion (Muñoz-López and García-Pérez, 2010). DNA transposons are active in many species (i.e. P-elements in *Drosophila*, Activator-Dissociation elements in maize and PiggyBac from cabbage looper moth) (Richardson et al., 2015), however, in mammals, with the exception of bats, they had a limited life span owing to their mutagenic potential. Indeed during evolution they have largely accumulated mutations which rendered them immobile. Computational analyses indicated that DNA transposons activity ceased in the primate lineage at least 37 million years ago (Pace and Feschotte, 2007). Approximately 3% of a typical mammalian genome is composed by DNA transposons (Beck et al., 2011; Keane et al., 2014; Lander et al., 2001; Mouse Genome Sequencing Consortium et al., 2002).

Figure 2. Structure of DNA transposons. DNA transposons have flanking inverted terminal repeats (ITRs) and a unique open reading frame encoding for transposase (Martens et al., 2005).

1.2 Retrotransposons

Retrotransposons represent the predominant class of TEs in most mammalian genomes, accounting for up to 40% of human and mouse genomes (Lander et al., 2001; Mouse Genome Sequencing Consortium et al., 2002). They duplicate themselves via a RNA intermediate using a *copy-and-paste* mechanism: retrotransposons are first transcribed to an RNA copy and then reverse transcribed and inserted in a new genomic location. In this way, a second novel insertion is created, while the original transposon is preserved. Unlike DNA transposons, particular subsets of retrotransposons retained their transposition activity in mammalian genomes.

Depending on the presence of long terminal repeats (LTRs) at their extremities, retrotransposons can be classified in LTR and non-LTR retrotransposons (Crichton et al., 2014). As for DNA transposons, retrotransposons can self-mobilize using the transposition machinery autonomously encoded (autonomous retrotransposons) or they can be mobilized relying upon the proteins encoded by other elements (non autonomous elements) (Cordaux and Batzer, 2009).

LTR retrotransposons

Approximately 10% of a typical mammalian genome is made of long terminal repeat (LTR) retrotransposons (Lander et al., 2001; Mouse Genome Sequencing Consortium et al., 2002). In terms of structure these elements share features similar to those of retroviruses but they have lost their inter-cellular mobilization capability, therefore they are also called endogenous retroviruses (ERVs). A canonical full length LTR retrotransposon presents 5' and 3' LTRs along with *gag*, *env* and *pol* sequences (figure 3). In general, LTR retrotransposons contain functional *gag* and *pol* genes that encode for proteins and enzymes involved in retrotransposition, but lack a functional envelope (*env*) gene. Sometimes, the recombination between LTRs eliminate completely *env* gene, generating "solo LTRs" (Mager and Stoye, 2015). The absence of a functional *env* gene relegate ERVs to an intracellular life.

Most mammalian LTR retrotransposons underwent a massive accumulation of mutations that made them immobile, especially in the human genome. Nonetheless in mouse, ERV insertions are responsible for almost 10% of spontaneously occurring mutations. Murine active ERVs include both autonomous elements, such as intracisternal A-particles (IAPs) and MusD elements, and non-autonomous retrotransposons, such as early transposons (ETn) and mammalian apparent LTR retrotransposons (MALRs) (Maksakova et al., 2006).

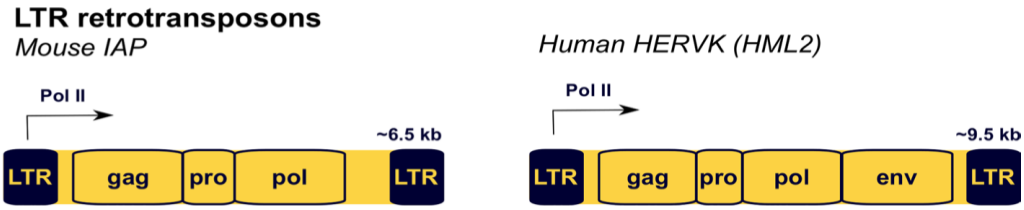


Figure 3. Structure of the major mammalian LTR retrotransposons. Main classes of LTR retrotransposons present in human and mouse genomes are given by IAP and HERV elements. Transcriptional regulatory regions are indicated with filled rectangles. Transcription start sites are depicted with arrows. Main protein coding regions are shown with open rectangles. LTR, Long terminal repeats (Crichton et al.,2014).

Non-LTR retrotransposons

Non autonomous elements

Non autonomous elements mobilize by using in *trans* the retrotransposition machinery encoded by autonomous elements (generally LINE1s). The most abundant class of non autonomous elements is represented by Short Interspersed Nuclear Elements (SINE) that constitute about 10% of the sequenced mammalian genome. They originate from small RNApol III transcripts, including tRNAs and the 7SLRNA component of the signal recognition particle (Crichton et al., 2014). Since their insertions can occur both in gene deserts and in gene rich regions, SINE elements represent a major source of genomic variability and provide regulatory elements for gene expression (Carrieri et al., 2012; Elbarbary et al., 2016). However in the less fortunate cases, SINE insertions can be associated with diseases (Hancks and Kazazian, 2012).

In humans, the most abundant SINEs are *SVA* and *Alu* elements. *SINE-R/VNTR/Alu* elements or *SVA* (~2 kb sequence) are the youngest human retrotransposons and account for about 0.2% of the human genome (Hancks and Kazazian, 2012). They are composed of a hexameric repeat region, followed sequentially by an inverted *Alu*-like sequence, a variable number of tandem repeats region, a HERV-K10-like region (SINE-R) and a polyA tail of variable length (figure 4). Apparently *SVA* elements do not contain internal promoter; nevertheless they are transcribed by RNA polymerase II and the resultant RNA transcript is presumably trans-mobilized by the L1-encoded proteins (Beck et al., 2011,Cordaux and Batzer, 2009). *Alu* elements (~300bp) represent almost 11% of the human genome and typically exhibit a dimeric structure with a left monomer separated from the right monomer by an A-rich linker region (figure 4). The 5'end contains an internal RNA polymerase III promoter (A and B boxes) and the element ends with a polyA tail of variable length. As *Alu* elements do not contain any RNA polymerase III terminator, their transcripts extend into the downstream flanking sequence until a terminator is found (Richardson et al., 2015). The counterpart of Alu elements in mice is given by B1 elements (~140 bp) that accounts for roughly 550,000 copies in the mouse genome (approximately 2,7% of mouse DNA) (Ponicsan et al., 2010). Like *Alu*s, they derive from a 7SL RNA precursor and present an internal RNA pol III promoter composed of A and B boxes and terminate with a polyA tract that is flanked by genomic DNA containing an RNA pol III terminator; but unlike *Alu*s, mouse B1s have a monomeric structure (figure 4).

B2 elements (~200 bp) are a second class of tRNA-derived murine SINEs that account for about 300,000 copies in mouse genome. Their structure is similar to B1s and they are transcribed by RNA pol III, as well (figure 4).

In addition, the transposition machinery encoded by autonomous elements occasionally trans-mobilizes mature cellular RNAs, like housekeeping genes mRNAs, ribosomal protein encoding genes, uracil-rich small nuclear RNAs and small nucleolar RNAs (Garcia-Perez et al., 2007; Weber, 2006). This process leads to the formation of processed pseudogenes that are scattered throughout the genome (figure 4). In some cases, processed pseudogenes can develop new cellular functions providing cells with new adaptive tools.

Autonomous elements

Long interspersed nuclear elements (LINE) are autonomous non-LTR autonomous retrotransposons that are highly abundant in mammalian genomes. They include LINE1, LINE2 and LINE3 subfamilies but among these, only LINE1 (or L1 elements) maintained mobilization capabilities in humans and mice (Lee et al., 2007). Being currently active in the mammalian genome and being able to mobilize both autonomous and non autonomous retrotransposons, L1s can significantly impact many biological processes of a L1-harboring cell.

The abundance, structure and properties of mammalian L1 elements will be extensively described in the next paragraph, with a particular focus on murine L1s.

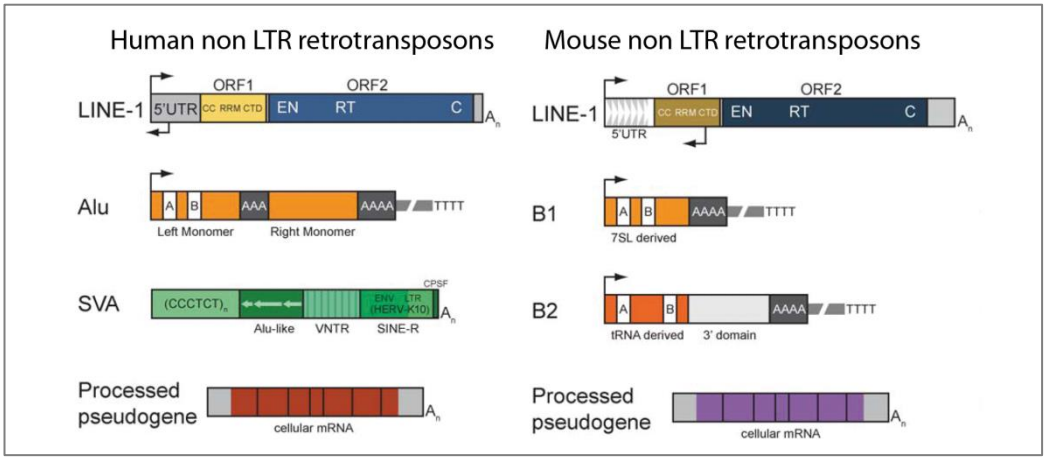


Figure 4. Structure of the main human and mouse non LTR retrotransposons. Each class of non LTR in human (left) and mouse (right) genome is listed with its name. Arrows depict transcription start sites. Structural components of each element are enclosed in rectangles. Detailed description of every retrotransposon type is in text (Richardson et al.,2015).

1.3 L1 retrotransposons

L1s constitute approximately 17% of the human genome and 19% of the mouse genome (Mandal and Kazazian, 2008). Most L1s are retrotransposition-defective owing to 5' end truncation, internal rearrangements or point mutations that disrupt their open reading frames. Nonetheless, a portion of L1s in both human and mouse genomes are full length elements, potentially able to mobilize. Among the 5,000 L1 full length copies present in the human genome, roughly 80-100 elements contain two intact open reading frames and are retrotransposition-competent (Beck et al., 2011). By comparison, the mouse genome contains about 11,000 L1 full length copies, of which at least 3,000 are still active (Goodier et al., 2001).

Full length L1 elements are ~6kb sequences composed by a 5' untranslated region (UTR), two open reading frames (ORF1 and ORF2) and a 3'UTR followed by a polyadenylation signal (AATAAAA) (figure 4).

The 5'UTR region houses an internal RNA polymerase II sense strand promoter that in humans also display antisense properties (Speek, 2001, Mätlik et al., 2006). 5'UTR is followed by ORF1 and ORF2 that encode for the proteins required for L1 retrotransposition (Moran et al., 1996). ORF1 encodes for a 40kDa protein (ORF1p) having RNA-binding and chaperone activities (Martin, 2006). ORF1p contains a highly conserved C-terminal region, an RNA binding motif and a less-conserved N-terminal α-helical domain (Ostertag and Kazazian, 2012). ORF2 encodes for a 150 kDa protein (ORF2p) with three conserved domains, an N-terminal endonuclease (EN) domain, a central reverse-transcriptase domain (RT) and a C-terminal zinc knuckle-domain (Doucet et al., 2010). The L1 EN domain is responsible for the dsDNA nick at the target insertion site, whereas the RT activity generates the cDNA copy of L1 to be inserted into the new genomic location (Feng et al., 1996; Mathias et al., 1991, Beck et al. 2011; Ostertag and Kazazian, 2001). The 3'UTR of L1 contains a functional but weak RNApol II polyadenylation signal near the 3'end. Signs of promoter activity residing on the 3'UTR of both human and murine L1s have been recently reported (Faulkner et al., 2009).

Mouse L1 elements

The sequence length of a full length mouse L1 element can vary between 6kb up to 8kb (Sookdeo et al., 2013). These differences are due to the presence of a variable region located in the 5'UTR of mouse L1s, which contains tandem repeats of ~200 bp, named monomers, followed by a short non-monomeric linker region right upstream of the ORF1 sequence (figure 5). The monomers at the 5'UTR differ in their number and sequence among different L1s. According to the organization of monomers at the 5'UTR, murine L1s can be divided into subclasses. Phylogenetically, all mouse L1 subtypes seem to derive from a common ancestor but during evolution older L1 elements accumulate more mutations and are more divergent than young L1s (Adey et al., 1994). At least three groups of monomers have been identified: *V*, *F* and *A* subtypes (Goodier et al., 2001). The most ancient *V* family is not very abundant and lacks the 5'end therefore accounting only for inactive elements (Adey et al., 1994). The *A* family members contain monomer repeats of 208bp at their 5' region and include about 6,500 full length elements. A small subset of them (~900) are supposed to be active because they contain intact ORF1 and ORF2 and they are transcribed (Severynse et al., 1992). The *F*-lineage was recently been discovered to comprise new young, transcriptionally active L1 families: the *Tf* and *Gf* L1 subtypes. Likely, these two families derive from a common ancestor, but during evolution, they have diverged in their 5'UTR monomeric sequence. In terms of abundance, *Tf* subfamily has 1800 potentially active elements among 3000 full length members (Naas et al., 1998), whereas the most recently discovered *Gf* type includes 400 active elements among 1500 full length members (Goodier et al., 2001).

In vitro experiments aimed to test the activity of the mouse *Tf* 5'UTR monomers have revealed that the promoter activity lies within the monomers and it seems to be enhanced proportionally to the number of monomers (DeBerardinis and Kazazian, 1999). Hence, the acquisition of new promoter sequences seems to lead to evolutionary success.

Concerning murine L1 ORF1 sequence, it displays trimeric organization having a coil coiled (CC) domain at the 5'end, followed by RNA-recognition motif (RRM) and a C-terminal domain (CTD). RRM and CTD domains are well conserved across L1 subfamilies whereas the CC domain, involved in trimerization of ORF1p, presents relevant levels of structural variation. Notably, a functional antisense RNA pol II

promoter, residing in mouse L1 ORF1 sequence, has been recently identified (Li et al., 2014).

The second open reading frame, ORF2, is the most conserved region among the different murine L1 subfamilies, showing very few amino acid changes, indicator of strong purifying selection on this sequence. By contrast, the 3'UTR includes the highest number of variable sites within murine L1 regions (Sookdeo et al., 2013).

Within the mouse genome, L1s are non-randomly distributed. Tipically, L1s are more abundant in AT-rich and gene-poor regions of the genome (Boyle et al., 1990), suggesting that possible deleterious effects of L1 insertion within gene bodies have been strongly discouraged during evolution. Nonetheless, when L1 insertions occur close to genes or even within them, the host gene expression can be consequently modulated (Han et al.,2004, Muotri et al, 2005). Notably, L1 localization depends also on L1 age, with younger L1s found in closer proximity to genes than older ones. Moreover, recent insertions tend to fall in preferential genomic regions, arising the concept of "hotspot" of retrotransposition within the genome (Jachowicz and Torres-Padilla, 2016). Additionally,intergenic L1 distribution are more frequent in autosomes and X chromosomes, whereas a greater intragenic density was reported in Y chromosome (Ngamphiw et al., 2014).

Importantly, a certain degree of variability in genomic L1 repertoire between inbred mouse strains has been described. A comprehensive study of intermediate length structural variants evidenced a high degree of polymorphisms at sites of integration of both L1s and non-autonomous retrotransposons. Moreover this study showed that L1 integrants can be either sense or antisense oriented (Akagi et al., 2008). A complete and updated summary of genomic rearrangements in mouse species is provided by MouseIndelDB (Akagi et al., 2010).

Altogether, this evidence shows the presence of ongoing endogenous L1 mobilization in mice. This phenomenon possibly leads to important inter-individual differences in L1 copy number and transcriptional variations in murine genomes.

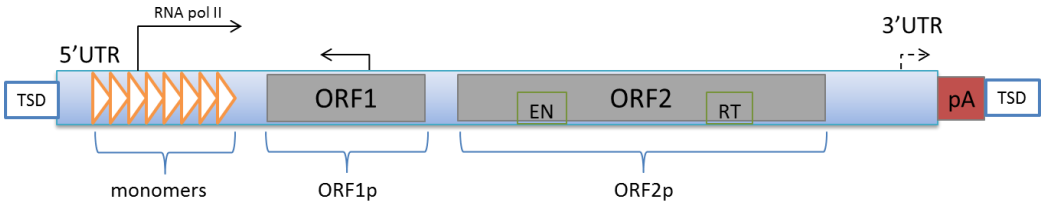


Figure 5. Structure of a full length murine L1. Mouse L1 is a ~7kb sequence composed of a 5'untraslated region (UTR), two open reading frames, a 3'UTR and a polyadenylation (pA) tail. L1 element is flanked by target site duplications (TSD) of the insertion site. The 5'UTR region is composed of a variable number of ~200 bp monomers, depicted with orange triangles. ORF1 encodes for the RNA-binding protein ORF1p. ORF2 encodes for ORF2p protein, having endonuclease (EN) and reverse-transcriptase (RT) activities. Black arrows indicate sense and antisense promoters. Dashed arrows show additional putative promoter in the 3'UTR.

2. L1 RETROTRANSPOSITION

L1 elements have shaped and greatly expanded mammalian genomes through at least two mechanisms: their self-retrotransposition and their ability to provide the machinery for the transposition of other mobile elements. The study of L1 structure has elucidated many aspects of the mechanism of L1 retrotransposition, a multi-step process that ultimately results in a *de novo* insertion.

2.1 Molecular mechanism of retrotransposition

The L1 mobilization process starts with the transcription by RNA polymerase II of a genomic L1 sequence from its own internal promoter that generates a bicistronic mRNA. The L1 mRNA contains a polyA tail which can be encoded by its own polyadenylation signal or a downstream signal. Whether L1 transcripts contain a 7-methylguanosine cap remains unknown (Ostertag and Kazazian, 2001). After transcription the L1 mRNA exits the nucleus through an unidentified mechanism and, once in the cytoplasm, ORF1 and ORF2 proteins are translated and bind to an mRNA molecule (figure 6). These two proteins show strong *cis*-preference, so that they preferentially associate with their own L1 mRNA, to form a ribonucleoprotein particle (RNP) (Doucet et al., 2010; Hohjoh and Singer, 1996; Kulpa and Moran, 2006; Martin, 1991). The formation of the complex is believed to involve multiple copies of ORF1p and only few copies of ORF2p (Beck et al., 2011). Interestingly, ORF1p, ORF2p and L1 RNA have been shown to accumulate in cytoplasmic foci that are closely associated with stress granules proteins (Doucet et al., 2010; Goodier et al., 2007).

Then, through a mechanism yet poorly understood, which might involve active transport or nuclear membrane breakdown, the RNP is transported back to the nucleus. Here the integration of the L1 copy into the genome likely occurs via a process called target-site primed reverse transcription (TPRT) (figure 6). During TPRT the L1 endonuclease recognizes the consensus sequence 5'-TTTT/A-3' and catalyzes a single-strand endonucleolytic nick in genomic DNA at the target site. The exposed 3'OH group is then used as a primer by the L1 reverse transcriptase ORF2p to synthesize a copy DNA (cDNA) of the L1 mRNA. Whether the L1mRNA acts only as a template or plays additional roles during TPRT remains still unclear (Feng et al.,1996, (Cost et al., 2002). Then, the second-strand of the target-site is cleaved and used to prime the synthesis of the cDNA second strand, through a mechanism still poorly understood. Typical hallmarks of the integration process are the target-site duplications (TSDs), that are ~ 7-20 bp long sequences generated at each end of the L1 copy; the 3' dA-rich tail of variable length, likely generated by the pairing of the A-rich sequence at the 3'end of L1 with the T-rich primer formed at the insertion site (Cordaux and Batzer, 2009, Sen et al., 2007).

Notably, the vast majority of L1 insertion events lead to 5'end truncation on the L1 copy, probably due to the inability of L1 reverse transcriptase to copy the entire L1-mRNA or to the action of cellular RNAse H which competes with L1 reverse-transcriptase (Ostertag and Kazazian, 2001). As a consequence, the newly inserted L1 copy lacks its promoter and is therefore unable to remobilize. The high frequency of 5' end truncations resulting from new L1 retrotransposition events remains enigmatic but might reflect a first line of host defense against uncontrolled L1 mobilization (Beck et al., 2011). On average, in humans, the length of an inserted L1 is about one sixth of a full length element (Ostertag and Kazazian, 2001).

Alternatively, to the canonical TPRT, L1 elements can also rely on a endonuclease independent (ENi) mechanism for their integration. This pathway likely represents a reminiscence of an RNA-mediated DNA repair mechanism and occurs when L1 uses pre-existing genomic double strand lesions to initiate TPRT (Eickbush, 2002, Morrish et al.,2002). L1 elements integrated via the ENi mechanism display structural features distinct from those typical of TPRT-mediated L1 insertions, in that they generally lack TSDs, they are frequently 5' and 3' truncated and are often accompanied by the deletion of genomic DNA at the integration site (Morrish et al., 2002). L1 integrants derived from ENi events have been observed in telomeres (Viollet

et al., 2014) but this process seems to be strongly discouraged *in vivo* (Babushok et al., 2006)(Morrish et al., 2002).

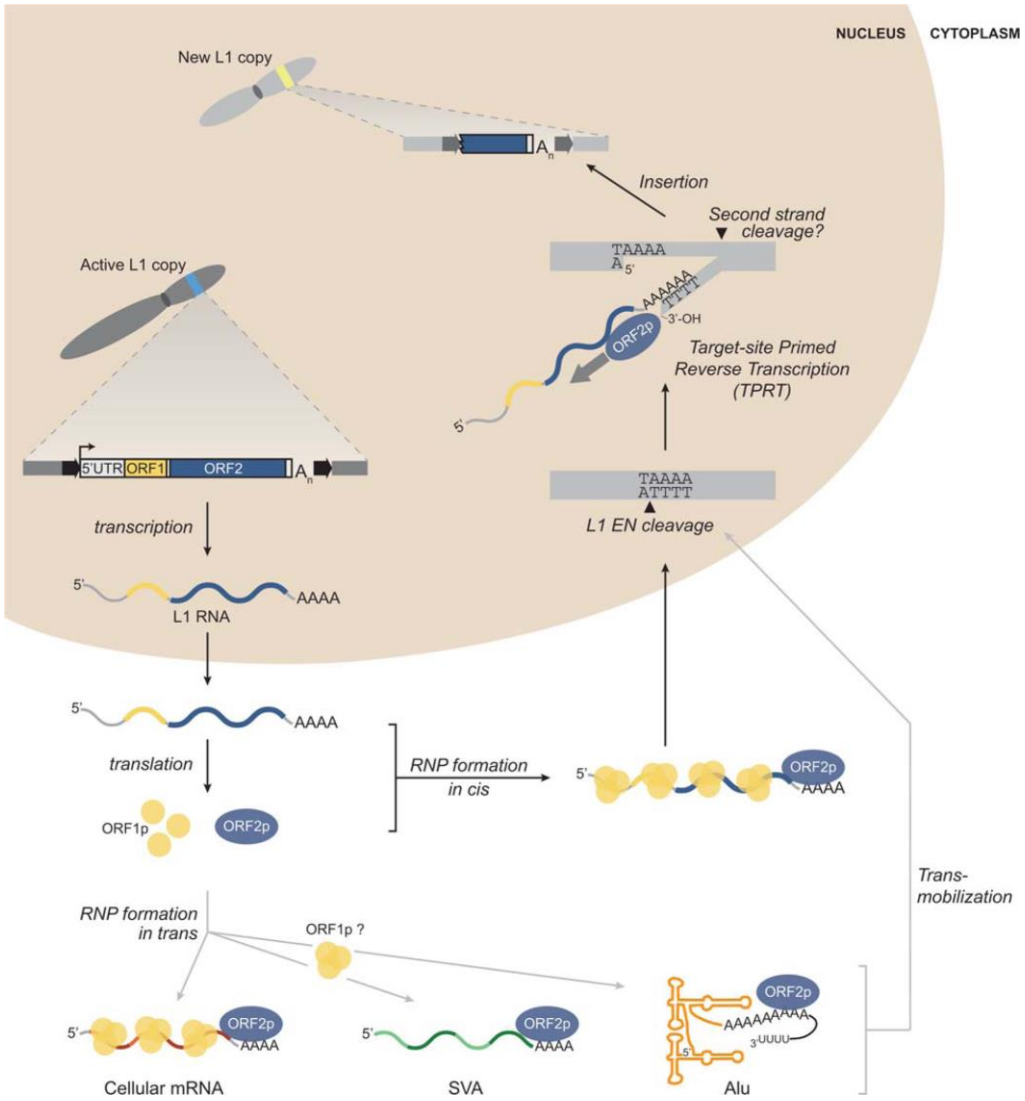


Figure 6. L1 retrotransposition cycle. A full length active L1 element is transcribed from its internal promoter to produce a bicitronic mRNA. L1-mRNA is then exported to the cytoplasm where ORF1 and ORF2 proteins are translated. ORF1 and ORF2 proteins can bind in *cis* to L1 mRNA or *in trans* to other cellular mRNAs to form the ribonucleoprotein (RNP) complex. The RNP complex then gains access to the nucleus, where the L1-RNA undergoes target site primed reverse transcription (TPRT). In TPRT a single strand nick is first made at the insertion site, then L1-mRNA is reverse transcribed and integrated into the new genomic location. Generally TPRT leads to 5'end truncation of the new L1 copy (Richardson et al.,2015).

2.2 Molecular mechanisms influencing L1 retrotransposition

Since L1 insertions and L1-mediated insertions might represent a significant threat to genome stability, host cells have developed a variety of mechanisms, at both transcriptional and post-transcriptional levels, to restrict L1 activity. Inhibition of

retrotransposition can occur through inter-correlated pathways, including epigenetic mechanisms (mainly DNA methylation and post-translational histone modifications on L1 promoter); premature polyadenylation; alternative splicing; small interfering RNAs and the activity of various cellular factors, some of which are typically involved in cellular defense against virus infection.

Epigenetic silencing of L1 promoter

Methylation of 5' cytosine on CpG dinucleotides is a common mechanism of DNA methylation that leads to transcriptional silencing of the downstream gene. The presence of CpG rich regions in the L1-5'UTR sequence and their consequent methylation is associated with repression of L1 mRNA expression (Yoder et al., 1997). Typically, DNA methylation patterns are established in primordial germ cells and are then maintained throughout the entire life of an organism. *De novo* DNA methylation on L1 promoter can be carried out by DNA methyltransferases in germ cells. The deletion of *de novo* methyltransferase 3L (DNMT3L) gene in mouse germline leads to meiotic catastrophe, concomitant with extensive demethylation of L1 promoter region and aberrant overexpression of L1s and other transposons (Bourc'his and Bestor, 2004). Nonetheless, DNA methylation is thought to control L1 expression also in somatic tissues. In particular, demethylation of L1 5'UTR promoter and consequent L1 activation can be provoked by the loss of methyl-CpG-binding protein 2 (MeCP2), a protein involved in global DNA methylation, primarily active in nervous tissues (Yu et al., 2001). MeCP2 knockout neural precursor cells and mice brains showed increased susceptibility for L1 retrotransposition. Accordingly, L1 promoter hypomethylation and abnormal L1 activity were reported in patients affected by Rett syndrome (RTT), a neurological disorder caused by a mutation in the MeCP2 gene (Muotri et al., 2010, 2005).

An alternative mechanism affecting the accessibility to L1 promoter during transcription is given by epigenetic modification on histone tails. Depending on which histone residue is modified, the output can be an increase or decrease of chromatin compactness around L1 promoter. Trimethylated lysine 9 on histone 3 (H3K9me3), an histone mark typically associated with repressed transcription, is largely present on L1 promoter. This leads to a constitutive repression of full length L1 expression in normal conditions (Fadloun et al., 2013; Pezic et al., 2014). In mice, H3K9me3 is found across the entire L1-5'UTR sequence, including both monomeric and non-monomeric regions,

and co-localizes with the histone variant H2A.Z and the heterochromatin-binding protein HP1α (Rangasamy, 2013).

Interestingly, during early embryogenesis (2-cell and 8-cell embryos), it has been shown that repressive H3K9me3 marks are consitutively present on L1 elements, whereas the deposition of active H3K4me3 can vary. The loss of H3K4me3 in 8-cell embryos consequently leads to L1 repression. This suggests that activation or silencing of L1 elements might be achieved by the modulation of activating marks, such as H3K4me3, rather than affecting constitutive repressive H3K9me3 (Fadloun et al., 2013). Also H3K9me2 is associated with L1 sequences primarily in germ cells and its deposition is developmentally regulated (Di Giacomo et al., 2013).

L1-RNA alternative splicing and premature polyadenylation

L1 transcript maturation can be inhibited by alternative splicing. Indeed, human full length L1 RNA molecules contain a conserved donor splice site within the 5'UTR sequence. This splice site was revealed to be functional and its use leads to the generation of shorter L1 transcripts, that likely lost their mobilization properties (Belancio et al., 2008, 2006). Additionally, the interaction between hnRNPL, a protein involved in alternative splicing, and mouse L1 RNA (Peddigari et al., 2013) or human L1 RNPs (Goodier et al., 2013), suggest a further evidence for L1 retrotransposition modulation mediated by alternative splicing.

Alternatively, L1 transcription can be influenced by premature polyadenylation. L1 elements are characterized by A-rich strands of variable length at their 3'region. This sequence contains both canonical and non-canonical polyadenylation (polyA) signals that can generate truncated L1 transcripts with compromised retrotransposition activity (Perepelitsa-Belancio and Deininger, 2003). Noteworthy, the presence of multiple polyadenylation signals might also contribute to fine–tune the expression of genes containing full length L1s (Han et al., 2004).

Small interfering RNA-mediated L1 silencing

Silencing of L1 retrotransposition by means of small interfering RNAs, particularly endogenous small-interfering RNAs (endo-siRNAs) and piRNAs, gives rise to another powerful cellular defense against L1 activation. Regarding endo-siRNAs, ongoing studies aim to elucidate the mechanisms by which they are generated and how they work to inhibit transposons. So far, it is known that endo-siRNAs are able to

repress L1 mRNA at a post-transcriptional level by promoting the disruption of the transposon transcript. They originate from a precursor dsRNA which is processed by Dicer proteins to form 21-24nt endo-siRNAs; these endo-siRNAs are then loaded by Argonaute proteins to form the RNA-induced silencing complex (RISC) that ultimately targets L1 mRNA for their cleavage (Levin and Moran, 2011). Precursor dsRNAs generating L1-specific endo-siRNAs can derive from the antisense promoter activity of L1-5'UTR (Yang and Kazazian, 2006) or from L1 mRNA itself. Indeed it was reported that L1-mRNA can be loaded by Microprocessor/Drosha siRNA machinery *in vitro* and this promote L1 RNA cleavage (Heras et al., 2013).

Alternatively, L1 silencing can be induced by a specific class of small non-coding RNAs, the 26-31 nucleotide RNAs termed PIWI-interacting RNAs (piRNAs). PIWI (P-element induced wimpy testes) proteins are a subclade of the Argonaute family, first discovered in *Drosophila*, that specifically binds to piRNAs in order to protect the genome of germ cells from the invasion of potential "genomic parasites", such as transposons (Malone et al., 2009). Primary piRNAs originate from piRNA clusters, that are intergenic genomic loci encoding long precursor RNAs, and bind to PIWI proteins in the cytoplasm, in order to be processed and guided to their target transcripts. Once the piRNA recognizes the complementary L1-mRNA, two outcomes can follow: (i) disruption of the L1-mRNA similarly to what endo-siRNAs do, or (ii) production of a secondary piRNA, sense-oriented to the L1-mRNA that, in turn, targets antisense L1-mRNA for the production of a new secondary antisense piRNA, able again to bind sense L1-mRNA for silencing. This gives rise to a cycle of rapid amplification of piRNAs known as "ping-pong cycle" (Ishizu et al., 2012).

In mice, the PIWI clade proteins are represented by MILI, MIWI2 and MIWI (Kuramochi-Miyagawa et al., 2008). MILI binds primary piRNAs in the cytoplasm and targets them to the transposon mRNA to promote its cleavage for silencing or for the initiation of the "ping-pong" cycle (figure 7). MIWI2 loads secondary piRNAs in the cytoplasm and can either stoke the amplification cycle or translocate the piRNA-MIWI2 complex to the nucleus to trigger DNA methylation on L1 transposon (Pillai and Chuma, 2012) (figure 7). Indeed, MILI and MIWI2 deficiencies in mouse spermatogonia induce the activation of L1 elements and this is accompanied with extensive demethylation on retrotransposon sequences, suggesting a role of these two factors in L1 promoter methylation (Aravin et al., 2008; Kuramochi-Miyagawa et al.,

2008). MIWI contributes to transposon silencing by inducing L1 mRNA cleavage or translational repression (figure 7).

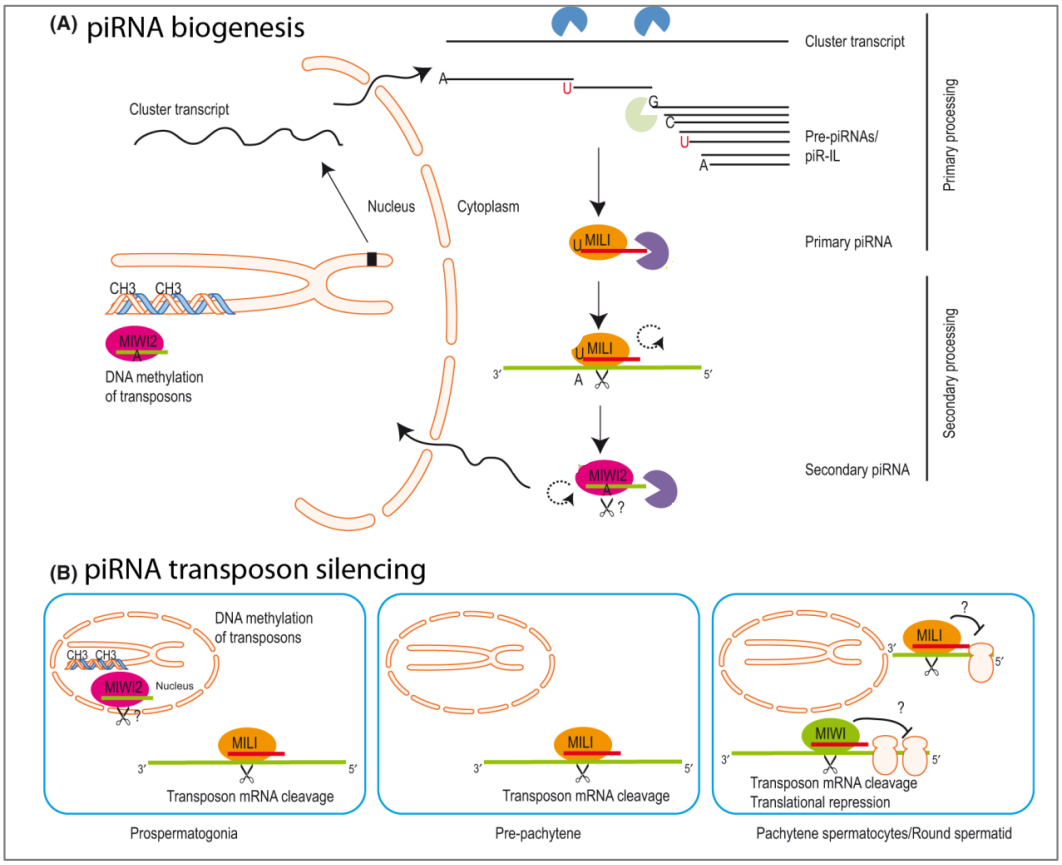


Figure 7. MILI, MIWI2 and MIWI functions in piRNA pathway in germ cells. A) In piRNA biogenesis, MILI (orange) participates in the primary processing of primary piRNAs and then guides piRNAs to target mRNAs (green) to promote their cleavage or to initiate the "ping pong" amplification cycle (dashed circle arrow). MIWI2 (purple) binds secondary piRNAs during "ping pong" cycle to induce their maturation and/or translocate them into the nucleus for DNA methylation on target transposons **B)** piRNAs bound to MILI induce transposon silencing by cleavage of the transposon mRNA or potentially by inhibiting the translation of the transposon mRNA. piRNAs associated to MIWI2 promotes DNA methylation on transposon promoters for their silencing. MIWI (green), the third class of PIWI proteins, participates in transposon silencing by promoting mRNA cleavage and/or translation repression. (adapted from Pillai and Chuma, 2012).

MILI and MIWI2 are particularly crucial during spermatogenesis, although they seem to play independent roles in piRNA biogenesis and DNA methylation. Indeed, in MIWI2 deficient mice L1 TE are derepressed, as expected, but "ping-pong" cycle activation can anyway take place. Moreover MILI is likely to control the expression of a larger subset of transposable elements than MIWI2, since MILI deficiency induces the expression of a larger repertoire of transposable elements and, at the same time,

promotes the DNA methylation on many more different transposons than MIWI2 deficiency (Manakov et al., 2015).

Surprisingly, very recent reports showed that piRNA pathway is functional not only in the germline. Indeed, piRNAs and MILI expression has been detected in adult brains of rodents and Aplysia (Ghosheh et al., 2016, Rajasethupathy et al., 2012). Moreover MILI deficient mice display hypomethylation of intergenic regions and L1 promoters in the brain genome, suggesting that the roles of piRNAs in nervous tissues might be similar to those carried out in testes (Nandi et al., 2016; Pillai and Chuma, 2012; Rajasethupathy et al., 2012)).

Other cellular L1 restriction factors

A number of transcription factors and cellular proteins involved in nucleic acid metabolism and DNA repair can contribute to restrict L1 retrotransposition. Among transcription factors able to bind L1 5'UTR and then regulate L1 transcription there are members of Sox and KRAB-zinc finger proteins. In particular, Sox2 can bind to Sox-binding sites located in the L1-5'UTR promoter region and its expression inversely correlates with that of L1 elements (Muotri et al., 2005). The level of Sox2 association to L1-5'UTR correlates with the decrease of Sox2 expression across neuronal differentiation (Coufal et al., 2009). Alternatively, KRAB-associated protein 1 (KAP1) repressive complex can be recruited on L1-5'UTR by KRAB zinc finger proteins to repress L1 expression (Castro-Diaz et al., 2014; Jacobs et al., 2014). Notably, the ribosylation of KAP1, promoted by binding of SIRT6 to L1-5'UTR, facilitates KAP1 interaction with HP1α and this leads to a heterochromatization of L1 promoter (Van Meter et al., 2014).

Frequently, cells limit retrotransposon activity taking advantage of the innate anti-viral immune response mechanisms. For example, deaminase proteins, tipically involved in innate response to retroviral infections, can act to inhibit also L1 retrotransposons. Members of the APOBEC3 family as well as ADAR-1 deaminase are shown to robustly suppress L1 retrotransposition in cultured cells. Interestingly, ADAR-1 directly interacts with L1 RNPs. However it remains yet poorly understood if L1 inhibition requires or not the deaminase activity, that might influence for instance cDNA synthesis during L1-TPRT (Bogerd et al., 2006; Muckenfuss et al., 2006; Orecchini et al., 2016). Alternatively, overexpression of *Trex-1*, a DNA exonuclease, and SAMHD1, a tryphosphohydrolase that can decrease intracellular dNTPs

availability, is accompanied by repression of L1 retrotransposition in cultured cells (Stetson et al., 2008; Zhao et al., 2013). Moreover MOV10, an RNA helicase and component of RISC, able to prevent retrovirus replication, can inhibit retrotransposon activity, including human L1, Alu, SVA and mouse IAP elements and it can directly interact with L1-RNA in the L1 RNP, in cultured cells (Goodier et al., 2012). Similarly, in cultured human cells, RNAseL, an endoribonuclease that cleaves single-stranded RNAs, seems to be activated by double-stranded regions existing within L1-RNA or that originates from annealing to a complementary transcript generated by antisense L1 promoter activity (Zhang et al., 2014).

Finally, since L1 retrotransposition events via TPRT are prone to generate DNA double strand breaks (Gasior et al., 2006), it is reasonable that DNA repair pathways might influence L1 retrotransposition success. In this context, recent studies demonstrated that Ataxia telengectasia mutated (ATM) protein, a Ser/Thr kinase involved in DNA damage signaling, and excision repair cross-complementation group 1 (ERCC1) proteins, a component of nucleotide excision repair (NER), can modulate L1 retrotransposition (Coufal et al., 2011; Gasior et al., 2008).

The extensive use of proteomics that is ongoing in these years allowed to identify a long list of host factors interacting with L1. However the effects on L1 retrotransposition of most of these factors have not been determined yet (Goodier et al., 2013; Moldovan and Moran, 2015; Taylor et al., 2013).

3. L1 MOBILIZATION EFFECTS

Fine-tuned L1 regulation is crucial for organismal biology. Several lines of evidence show that while elevated transposon expression is detrimental, some elements have also been recruited as new transcripts and/or proteins to benefit organismal function. From here derives a complex interaction between transposons and the host genome. Here it will be briefly described, from a molecular point of view, how L1-retrotransposition events can create diversity through structural as well as functional changes in genome and transcriptome.

3.1 L1-insertional effects

Genomic structural alterations

L1-mediated retrotransposition events induce a variety of structural modifications in genome structure, including insertional mutagenesis, genomic rearrangements, mobilization of L1 flanking sequences and chromatin remodeling.

- Insertional mutagenesis. When L1s are inserted in exons they induce interruption of the coding sequence. Owing to the immediate phenotypic effect, *de novo* L1 insertions causing insertional mutagenesis were the first to be detected and let us understand that transposable elements cannot be dismissed as mere "junk DNA". In 1988 the group of Kazazian demonstrated for the first time that two independent mutagenic L1 insertions in the *factor VIII* gene caused hemophilia A in unrelated boys (Kazazian et al., 1988). The full length L1 precursor from which the disease-causing-L1 derived was then recovered from the genome of the mother of one of the two hemophilic patients (Dombroski et al., 1991). Since then L1s started to be studied with more attention.
- Structural variations at the insertion site. Upon integration, L1 retrotransposition events can lead to genomic rearrangements at the target site through a variety of mechanisms. For example, the resolution of TPRT intermediates by single-strand annealing or synthesis-dependent single strand annealing can lead to the formation of L1-mediated deletions or duplication at the insertion site. These processes likely involve the cellular repair machinery which recognizes and tries to repair the dsDNA nick caused by L1 during TPRT (Gilbert et al., 2005, Gilbert et al., 2002, Symer et al., 2002). From a genomic large-scale analysis it turned out that L1 insertions are frequently associated with structural rearrangements in human genomes (Beck et al., 2011). For example, the characterization of a deletion of 46kb in the *PDHX* gene, causing pyruvate dehydrogenase complex deficiency, led to the discovery of a full-length L1 insertion between exon 2 and 10 of the gene, corresponding to the deleted region (Miné et al., 2007). Very recently, L1-insertion mediated deletions were observed also in healthy adult human brains (Erwin et al., 2016).
- Nonallelic homologous recombination. Dispersed distribution of LINE and other repetitive sequences in the genome can lead to mispairing and consequent

nonallelic homologous recombination (NAHR) between them, thus producing structural variants. Alignment analysis of human and chimpanzee genomes has identified 55 cases of human-specific L1-NAHR associated deletions. In this mechanism, two pre-recombination L1s present in the chimpanzee genome have recombined via NAHR to generate the chimeric L1 in human genome (Han et al., 2005). The amount of structural variation caused by NAHR is significant and accounts for more than 0,3% of human genetic diseases (Belancio et al., 2008).

- 3' and 5' transduction. Because in L1 elements canonical polyadenylation site is often bypassed on behalf of downstream signals, sequences flanking L1 3' ends may be mobilized during L1 retrotransposition, a process termed 3'transduction. Transduction of 3'-L1 flanking sequences has been detected both in mouse and humans: through GenBank L1 database analysis, Goodier and colleagues have shown that 23% of 66 uncharacterized L1 sequences carried 3'flanking DNA with an average in length of 207 nucleotides and 3'transduction was more frequent in full-length L1 copies (Goodier et al., 2000). Similarly, 5' transduction occurs when the promoter activity of a transcript initiating upstream to a L1 element is used to transcribe. 5' transduction can be detected only by examining full length L1s, therefore its influence appears to be much less common than 3' transduction (Beck et al., 2011). Through these two mechanisms, non-retrotransposon DNA located downstream or upstream of active L1s can be shuffled into new sites, thereby creating new genes.
- Heterochromatization. L1 insertions seem to be able to alter chromatin state. This phenomenon was first hypothesized by M. Lyon after the observation of L1 accumulation on X-chromosome, which led him to speculate that L1 might act as booster elements to promote the spreading of heterochromatin formation during X-inactivation (Lyon, 1998). This hypothesis was then demonstrated by experiments in ES cells that showed that L1s could have a dual role in nucleating heterochromatin formation on the inactive X chromosome. First, silent L1s, which are tightly packaged in heterochromatin, facilitate nucleation of a silent, heterochromatic compartment into which genes are recruited; second, a subset of active L1s, expressed during X-chromosome inactivation (XCI), participate in local propagation of XCI to certain genes that otherwise would be prone to escape silencing (Chow et al., 2010).

Alterations of cellular transcriptome

Through their insertions, L1 elements can also contribute to modulating the transcriptome of the cell by providing new splice sites, adenylation signals and new promoters that can finally generate new reorganized transcription units. For example, when a L1 integrant falls into an acceptor splice site, it disrupts it and provokes exon skipping. On the other hand, if a new functional donor or acceptor splice site is added by L1 insertion, this leads to L1-exonization. Clearly, depending on the position of the splice site inside the L1 sequence, insertion of exonized L1s of different length might occur. In addition, truncated or rearranged L1s and antisense oriented splice sites can provide further sources of alternative splicing. As a consequence, the combinatorial usage of the numerous potential splice sites generated through L1 insertions adds complexity to the alternative splicing scenario of the cell, as confirmed by a screening analysis of a mouse cDNA library (Zemojtel et al., 2007).

Otherwise, cellular transcriptome might also be affected by alternative promoters provided by L1s, highlighting a role for L1s as "portable promoters". Functional sequences can derive from the sense canonical L1-5'UTR promoter of full length L1s and from 5'end truncated L1 fragments. In this context, Faulkner and colleagues unveiled a novel promoter sequence in the 3'UTR of L1s in both humans and mice. Moreover they showed that the canonical 5'UTR promoter of full length L1 elements is active in a tissue specific fashion, with higher activity in developmental and cancerous tissues (Faulkner et al., 2009).

In humans, the antisense promoter (ASP) activity lying in the L1-5'UTR also plays a role in modulating cellular transcriptome. This emerged from studies in human embryonic stem cells and from analyses of expression profiles of chimeric mRNAs in different human tissues showing that L1-ASP is functional and can give rise to new transcription start sites for genes upstream of L1s on the antisense strand. Antisense promoter activity correlates with the one of the native promoter gene or can act in a tissue-specific manner, thus contributing to tissue-specific control of expression (Macia et al., 2011, Mätlik et al., 2006).

Additionally, it has been shown that expressed L1s originate mostly from intronic L1s, suggesting that expression and insertion of L1s can act on transcriptional regulatory regions, such as enhancers or 3'UTRs, to modulate the expression of the hosting gene (Faulkner et al., 2009).

Recently, a new ORF has been discovered in the primate L1s, termed ORF0. This ORF is located upstream to L1-5'UTR in antisense orientation; it can be found in almost 3000 loci within human and chimpanzee genomes and it is functional. Additionally, since it contains two splice donor sites, ORF0 can also form fusion proteins with nearby exons. ORF0 seems favoring L1 mobility (Denli et al., 2015).

Intriguingly, L1 insertions represent a major force in shaping the non-coding transcriptome in humans and mice. In these species more than 40% of long non-coding RNAs contain embedded transposable elements, including L1-derived sequences (Kapusta et al., 2013).

3.2 L1 expression-mediated effects

The retrotransposition of L1 elements can influence genomic stability and function upon insertion and through the expression of L1-mRNA and L1-encoded proteins. Overexpression of L1s has been linked to apoptosis, DNA damage and repair, cellular plasticity and stress (Goodier, 2016). For example, the production of L1-ORF2p, not necessarily linked to a retrotranspositional event, can trigger DNA damage and provide a functional RT to cellular RNAs, thanks to its intrinsic EN and RT activities. As shown by immunostaining for γ-H2AX foci, L1 overexpression seems to induce large amounts of DNA double strand breaks in cultured cells (DSBs) (Erwin et al., 2016; Gasior et al., 2006). The high levels of DSBs exert cellular toxicities that eventually lead to cell cycle arrest, apoptosis and cellular senescence of L1 overexpressing cells (Belancio et al., 2010; Belgnaoui et al., 2006). The induction of DNA DSBs can be clearly ascribed to the endonuclease domain of ORF2p. However, although to a lesser extent than compared to EN domain, also mutations in the RT domain of ORF2p can modulate cell viability (Wallace et al., 2008). How the RT domain might be detrimental for cells remains largely unclear. According to a model, the interaction of L1 RT and random genomic breaks, occurring for instance during endonuclease-independent integrations (Morrish et al., 2002, Sen et al.,2007), could inhibit DNA repair. Noteworthy, the mobilization of SINE elements do not necessarily require an entire L1 RNP but can rely on the production of only a functional L1-ORF2p (Dewannieux et al., 2003). Therefore, the expression of both full length and truncated L1 retrotransposons can affect the host genome.

Concerning L1 transcripts, as already described previously, L1-mRNA might be the source of cellular piRNAs and endo-siRNAs, therefore contributing to gene

transcriptional regulation and chromatin remodeling. Notably, from the analysis of CAGE tags of both human and mouse transcriptomes, it emerged that L1-derived transcripts display a prevalent nuclear localization, therefore suggesting their potential role in association with nuclear non-coding RNAs (Faulkner et al., 2009). Furthermore it has been proposed that actively transcribed L1s might help the spreading of repressive Xist-dependent domain during X chromosome inactivation (Chow et al.,2010).

Hence, whether L1 expression has a role for the host organism or represents merely "transcriptional noise" seems to strongly depend on the context in which it takes place and it is nowadays under intense investigation. As it will be described in the next chapters, L1 expression seems to occur in precise temporal and spatial windows. This suggests that this phenomenon requires a strict control by the host cell, since it can provide evolutionary advantages for the host organism, but it can also entail pathology onset.

4. GERMLINE L1 RETROTRANSPOSITION

The accumulation of L1 elements in the genome during evolution of eukaryotes is a clear sign of heritable retrotransposition events occurring in germ cells during gametogenesis (figure 8). Plenty of data shows that L1 retrotransposition largely occurs in the germline. High levels of L1 expression was found in these cells in mammals: increased levels of full length L1 mRNA and L1 proteins were originally detected in leptotene and zygotene spermatocytes and this expression seems to be lost during development (Trelogan and Martin, 1995, Branciforte and Martin, 1994). Additionally, by taking advantage of L1 transgenic mice harboring an engineered human L1 element under the control of its endogenous promoter, L1 retrotransposition events were detected in male germ cells while L1 expression is restricted to testis and ovary (Ostertag et al., 2002). Furthermore, expression of L1-RNA was shown to be high in mouse full grown oocytes (Peaston et al., 2004) and L1 reverse transcriptase activity was detected in mature spermatozoa (Vitullo et al., 2012).

Germline L1 activation reflects the distinctive pattern of DNA hypomethylation typical of germ cells, where epigenetic patterns has to be flexible to permit cellular specification and differentiation (Smallwood and Kelsey, 2012). Not surprisingly, the tight control of L1 expression in germ cells is achieved by multiple epigenetic

pathways, including *de novo* DNA methylation, piRNA pathway and chromatin remodeling (Bourc'his and Bestor 2004, DiGiacomo et al.,2013).

Interestingly, a recent study conferred a functional role to L1 retrotransposons in the process of fetal oocyte attrition (FOA) in mice. FOA is the process of elimination of most meiotic prophase I (MPI) oocytes before birth. Increased levels of L1 expression and nuclear accumulation of L1-ORF1p associate with enhanced MPI defects, FOA, oocyte aneuploidy and embryonic lethality. On the other hand, reverse transcriptase inhibitors considerably affect FOA dynamics and meiotic recombination, suggesting an RT-dependent pathway for L1 implication in oocyte elimination in MPI. They propose that in FOA mechanism the oocytes with less L1 activity are favored, since they represent a lower risk for future generations (Malki et al., 2014).

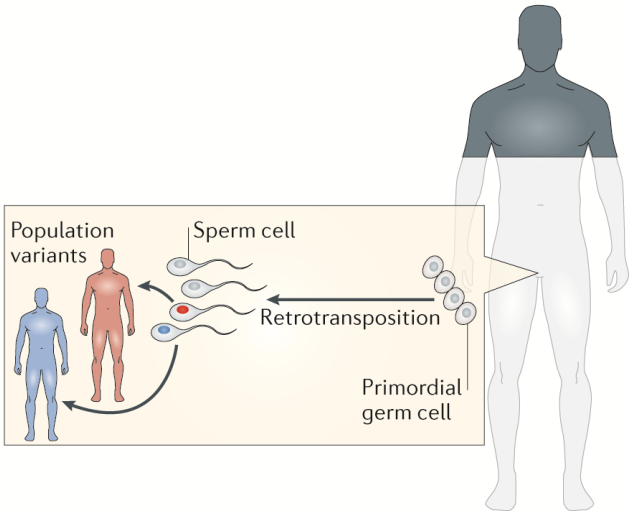


Figure 8. Effects of germline retrotransposition. Retrotransposition of L1 elements occurs in the germline and thus creates population variants that are present in every cell of an individual's body and are passed on future generations (Erwin et al.,2014)

5. SOMATIC L1 RETROTRANSPOSITION

Although originally thought to occur only in the germline, L1 retrotransposition is emerging having a major role in shaping genomes and reconfiguring gene expression networks in somatic cells. Notably, L1 integrants resulting from somatic retrotransposition events are not incorporated into germ cells, so they are not heritable and will not accumulate in the genome of all cells of an individual. This leads to the formation of genomic diversity between the cells of a somatic tissues, an event known as somatic mosaicism. The best characterized example of somatic mosaicism is the one creating VDJ recombination in B lymphocytes, aimed at providing a defense against the

variety of natural pathogens. Afterwards, the use of whole genome sequencing screenings allowed detecting many other cases of genomic variants in somatic tissues of an individual. Notably, a substantial portion of these variations arose from L1 retrotransposition events (Lupski, 2013). So far, early embryos, cancer and the brain have been described as particularly fertile soils for somatic L1 activity, thus representing typical examples of L1-derived genomic mosaics. Here will be described up-to-date knowledge of somatic L1 retrotransposition in early embryos and in the brain.

5.1 L1 retrotransposition in the early embryo

In 2009, Kano and colleagues demonstrated for the first time that L1 retrotransposition can occur in early embryos. In particular, they detected high levels of L1-mRNA in mouse pre-implantation embryos and showed that L1 integration is likely favored during embryogenesis. To demonstrate this activity, they created a transgenic mouse model harboring a L1 retrotransposition cassette where the L1 element is controlled by its endogenous promoter and is interrupted by an intronic sequence. According to this strategy, PCRs using intron-flanking primers allowed to distinguish the L1 transgene from a L1 retrotransposed copy since intron removal can occur only upon L1 retrotransposition. Interestingly, the comparison of L1 integrants in different developmental stages revealed that L1 retrotransposition events are more frequent in embryos (blastocyst and $E_{9.5}$) than in spermatogenic fractions (Kano et al., 2009).

Consistent with previous observations, another study showed that mouse early embryos physiologically contain abundant L1 mRNAs and dynamic expression of L1 retrotransposons occurs during embryo maturation (Fadloun et al., 2013).

Moreover a significant increase in L1 copies was registered in genomic DNAs derived from mouse zygotes and 2-cell embryos as compared to sperm and oocytes, suggesting active integration in this developmental window (Vitullo et al., 2012). Accordingly, the same group showed that L1-RT plays a crucial role in early embryo development: inhibition of RT leads to an arrest of development at the two- and four-cell stage (Beraldi et al., 2006; Pittoggi et al., 2003).

In humans, a X-linked L1 insertion responsible for choroideremia eye pathology displays somatic and germ-line mosaicism in the mother of the affected patient, suggesting that L1 insertion can occur very early in human embryonic development (van den Hurk et al., 2007).

Indication of a good tolerance of L1 retrotransposition in early embryos is suggested also by the high expression of L1s and the accumulation of new L1 insertion reported in human pluripotent stem cells and embryo stem cell lines (Garcia-Perez et al., 2007; Klawitter et al., 2016; Marchetto et al., 2013; Wissing et al., 2012, 2011).

Clearly, the fact that the *de novo* L1 insertion is incorporated in the pluripotent stem cell genome, before or after germline specification, defines its hereditability (figure 9). However, the frequency and the specific timing of endogenous L1 retrotransposition across development have not been fully described.

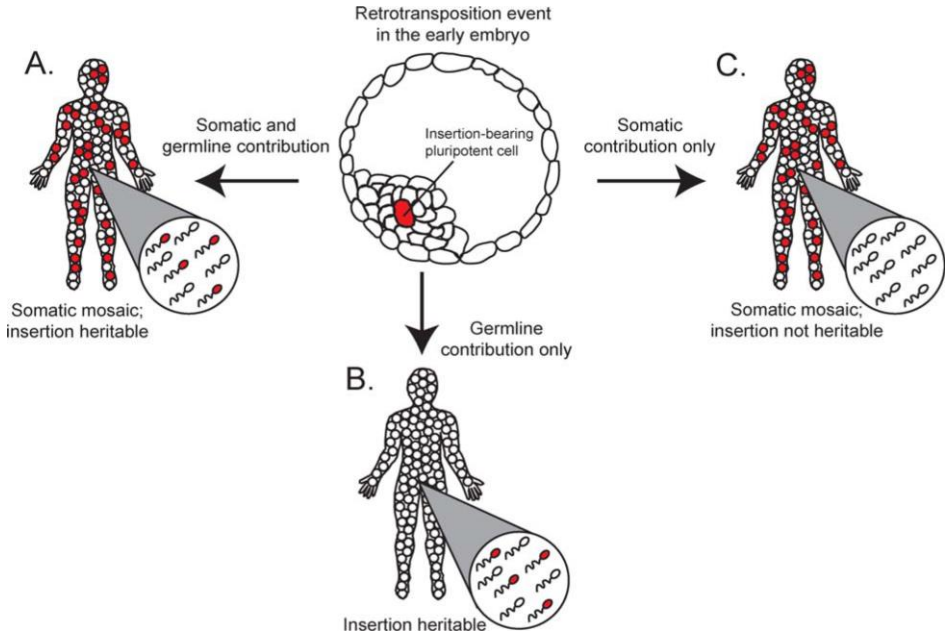


Figure 9. Effects of retrotransposition in early embryos. Retrotransposition of L1 elements occurring in pluripotent cells of the early embryo can have three outcomes. If pluripotent cells harboring *de novo* L1 insertions differentiate into germ cells as well as somatic cells (A.), or only into germ cells (B.) or only into somatic cells (C.), they will contribute, respectively, to germline together with somatic mosaicism (A.), only germline (B.) or only somatic mosaicism (C.) within an individual. Germline insertions can be consequently passed on future generations (Richardson et al.,2015).

5.2 L1 retrotransposition in the brain

The brain is composed of a heterogeneous population of cells having peculiar functional and structural features, even within cells of the same subtype. Plenty of data now shows that part of this variability might derive from L1 mobilization events occurring in this organ (figure 10). Both engineered and endogenous L1s have been shown to mobilize in the mammalian brain. The first evidence for this came from the Gage laboratory when it was shown that the differentiation of adult rat hippocampal neural stem cells into neuronal precursor cells (NPCs) and neurons leads to an increase

in L1 transcript levels and that an engineered L1 element can undergo retrotransposition in cultured rat NPCs (Muotri et al., 2005). Later, L1 retrotransposition was reported also in embryonic stem cells derived from neuronal progenitors (Coufal et al., 2009). More recently, it has been shown that also mature neuronal cells differentiated from human embryonic stem cells (hESC)-derived NPCs contain high levels of L1 mRNA and can support L1 genomic integration of an engineered L1 element (Macia et al., 2016).

As previously mentioned, Sox2 can regulate L1 transcription and consequent retrotransposition in rat hippocampal neural stem cells. Alternatively, activation of L1 retrotransposition in neural precursor cells could arise from DNA methylation permissive patterns, where MeCP2 plays a crucial role, or from an alteration in DNA repair cellular machinery, as shown in the case of reduced kinase activity of Ataxia Telengectasia Mutated (ATM) (Muotri et al. 2005, Coufal et al.,2009, Macia et al.,2016, Coufal et al.,2011). All these data were then confirmed according to a model of L1 retrotranposition control where the activation of L1 retrotransposons corresponds to the progression from neural stem cell to neural progenitor (Erwin et al., 2014; Muotri et al., 2010).

In line with *in vitro* experiments, L1 retrotransposition events were also reported *in vivo*. These studies took advantage of L1-EGFP-transgene where the EGFP reporter is interrupted by an intron and is controlled by CMV promoter and inserted in opposite orientation in the 3'UTR region of a human L1. By this approach, L1 retrotransposition events were revealed in the brain occurring during the first phases of neurogenesis, particularly between $E_{8.5}$ and $E_{10.5}$ (Muotri et al. 2005). In addition, it was recently reported that L1 copy number is increased in the adult mouse hippocampus, a preferential site of adult neurogenesis in mammalian brains (Ueno et al., 2016).

Similarly, increases of genomic L1 copies that likely derived from somatic L1 retrotransposition were detected also in human adult hippocampus. Using a *Taqman*-qPCR approach on genomic DNA extracted from bulk tissue, Coufal et al. performed a relative quantification of L1-ORF2 sequences in different somatic tissues, showing that approximately 80 more L1 copies were present in the human hippocampus when compared to other non-nervous tissues, like heart and liver, taken from the same individual. Notably this experiment evidenced significant levels of genomic variability between different brain regions within an individual and between different individuals (Coufal et al., 2009).

In line with these results, the use of next-generation sequencing further confirmed the occurrence of somatic L1 retrotransposition in human brain genomes. Baillie et al. developed a retrotransposon-capture sequencing (RC-Seq) technique, which is able to enrich sequencing libraries for fragments containing L1-genomic junctions. With this technology, they demonstrated that 7743 putative new L1 somatic insertions were present in the hippocampus and caudate nucleus of three individuals, supporting the idea of somatic mosaicism in the brain (Baillie et al., 2011).

Consistent with these data, high levels of L1 mRNA were detectable in brain tissue samples compared to several other somatic tissues (Belancio et al., 2010) and the analysis of CAGE (Cap-analysis gene expression) tags on human and mouse transcriptome revealed that L1s are extensively expressed in the brain (Faulkner et al.,2009).

However, the analysis of bulk tissue, though being highly informative, carries limitations in the detection of L1 retrotransposition events occurring in a small subset of cells within the entire tissue. Therefore, the employment of single cell approaches is now trying to better assess the extent of neuronal L1 retrotransposition. A first single-cell genomic study performed on whole cell amplified genomic DNA from 300 single neurons of cerebral cortex and caudate nucleus estimated the frequency of unique somatic L1 insertion around 0.6 per neuron (Evrony et al., 2012). Two subsequent studies evaluated different rates of genomic L1 insertion in neurons of human hippocampus as well as frontal cortex: Upton and colleagues showed an average of ~13 L1 insertion events per hippocampal neuron, whereas Erwin and colleagues estimated 0.58-1 somatic L1 insertion events per neuron in hippocampus and frontal cortex (Erwin et al., 2016; Upton et al., 2015).

Overall, although the exact L1 retrotranspositional rate per neuron needs to be refined and likely varies in different brain regions and cell types, it is now clear that the mammalian brain represents a somatic mosaic resulting from active L1 retrotransposition.

Notably, the mobility of transposable elements has been also detected in the *Drosophila* brain (Li et al., 2013; Perrat et al., 2013), indicating its conserved feature throughout evolution and raising the hypothesis that mobilization of TEs could have a functional role in normal brain physiology.

Conversely, the perturbation of this mechanism can be associated with neuropathology, as described with some examples in the next paragraph.

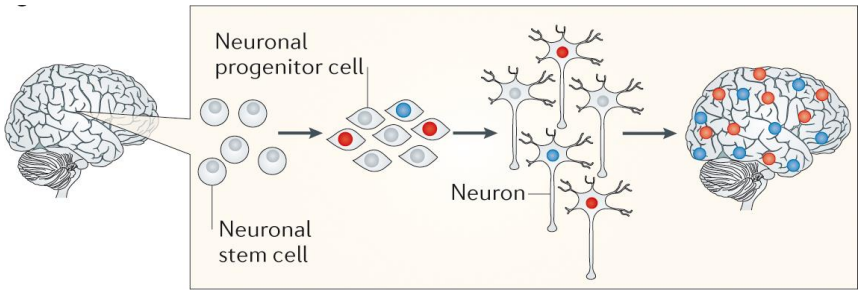


Figure 10. Effects of somatic L1 retrotransposition in adut brain. Somatic retrotransposition increases as neural stem cells differentiate into neurons and results in neurons with unique genomes (blu and red). Variability exists between the rates of retrotransposition and regions in which it occurs between individuals. High rates of retrotransposition events seem to occur in the hippocampus and frontal cortex in humans (Erwin et al.,2014).

L1 retrotransposition and neurological diseases

The insertion of L1 elements in intragenic regions with the consequent disruption of the targeted gene function has been associated with several diseases, as extensively described in many reviews by Kazazian and Hancks (Hancks and Kazazian, 2016, 2012). However, several pathologies, including many neurological diseases, correlate with deregulation of L1 retrotransposition.

In particular, it has been shown that iPS cells from patients affected by Rett Syndrome (RTT), a neurodegenerative disorder caused by a mutation in the X-linked MeCP2 gene, present an increased susceptibility to L1 retrotransposition while MeCP2 knockout mice display enhanced L1 transcription and retrotransposition (Muotri et al., 2010).

Additionally, unregulated L1 mobilization has been reported in Ataxia telengectasia, a neurodegenerative disorder caused by mutations in the ATM gene that is involved in DNA double strand break signaling. Cells lacking ATM show increased retrotransposition of an engineered human L1 element and postmortem brains of ATM patients present an increased genomic L1 content (Coufal et al., 2011).

Furthermore, a higher L1 copy number was found in the pre-frontal cortex neurons and iPS cells of patients affected by schizophrenia (Bundo et al., 2014).

Altogether, these findings highlight a correlation between enhanced L1 activation and neuropathology. The triggering event and the consequences of L1 dysregulations remain largely unknown. Especially in the case of neurodegenerative disorders, it must be considered that disease-associated alterations might promote L1 retrotransposition, but at the same time, normal aging and stress can influence transposon mobilization. In mice, the suppression of L1s mediated by SIRT6, has been shown to be lost during

aging and in DNA damage response, leading to a consequent activation of L1 elements (Van Meter et al., 2014). Additionally, in *Drosophila*, it has been extensively described that the activation of LINE-like retrotransposons can occur during aging in the brain (Li et al.,2013) and in concomitance with stress response (Specchia et al., 2010).

A complete study of L1 retrotransposition in different neurodegenerative contexts might significantly help in understanding the common and distinctive features of L1 alterations in brain diseases, in turn providing hints of L1 retrotransposition functions in brain physiology.

The aim of the present study was to contribute to the understanding of L1 retrotransposition events in Huntington's disease.

6. HUNTINGTON'S DISEASE (HD)

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder first described in 1872 by George Huntington in his paper "On Chorea", thereafter termed Huntington's disease. In 1983 the scientists of the Huntington's Disease Collaborative Research Group found out that the genomic sequence linked to the disease was mapping on chromosome 4 (Gusella et al., 1983) and ten years later the gene was identified ("The Huntington's Disease Collaborative Research Group.," 1993). The disease-causing mutation was shown to be a polymorphic trinucleotide (CAG) repeat in the 5' region of the IT15 gene that can expand beyond physiological thresholds.

Since then, many aspects of the molecular basis of the pathology have been elucidated; however, no cure is currently available for this devastating disease. HD affects 4 to 10 cases per 100,000 in Western Europe populations.

6.1 Clinical manifestations of HD

Clinical diagnosis of HD is made in the presence of movement disorders typical of HD in the context of a family history and genetic test positivity to HD mutation (Williams et al., 2007). In addition to movement disorders, typically consisting of choreic movements and loss of coordination, HD patients show many other clinical manifestations. Psychiatric symptoms are common in HD and they can arise about a decade before the onset of motor disturbances. They include mood and personality changes, apathy and depression, alterations in memory, learning and planning abnormalities and sleep disturbances (Rosenblatt, 2007). As the disease progresses, patients develop overt choreiform movements of the whole body, that worsens with time, together with facial grimacing and twisting of the trunk and limbs (Bachoud-Lévi et al., 2001). Weight loss also appears and could be due to dysphagia or degeneration of hypothalamic orexin neurons (Aziz et al., 2008).

In HD, symptoms typically appear in mid-life, on average around 40 years old, but can span between 2 years old, in juvenile HD cases, up to 80-90 years old. The age at onset inversely correlates with the extent of the CAG expansion in the HD gene as will be explained later. The most common causes of death in HD patients are cardiovascular and pneumonic problems deriving from incessant choreic movements,

injuries related to falls, poor nutrition and infection. Choking, caused by aspiration of food, and suicide are also relatively common causes of death (Di Maio et al., 1993).

6.2 Neuropathology of the disease

HD is characterized by a general shrinkage of the brain and degeneration of the striatum (caudate and putamen), with specific and gradual loss of efferent medium spiny neurons (MSNs) (Reiner et al., 1988). Although the striatum appears to be the most affected region of the brain, a specific thinning of cortical ribbon was also found in HD patients (Rosas et al., 2002). The loss of cortical mass is an early event in HD pathology that worsens with disease progression and mostly occurs at the expenses of the striatum-projecting neurons in cortical layer V and VI (Hedreen et al., 1991). Atrophy and neuronal loss in the hypothalamus have also been reported in HD brains (Kremer et al., 1991; Petersén and Björkqvist, 2006).

Studies performed on post-mortem brains revealed the presence of a significant *microglia* activation in the brain regions affected by neurodegeneration, particularly striatum, frontal cortex and globus pallidus (Sapp et al., 2001). Other non-neurological signs have been also described in HD, including diabetes, peripheral muscle weakness, stress-induced apoptotic cell death in lymphocytes and endocrine changes (Ross et al., 2014).

6.3 Genetics of HD

HD is caused by mutations in exon 1 of the IT15 (HTT) gene that encodes for huntingtin, a ubiquitously expressed protein (Huntington's Disease collaborative research group 1993). HTT gene is located in chromosome 4p16.3, it contains 67 exons and spans over more than 200kb. The 5'end of the gene contains a highly polymorphic stretch of CAG trinucleotides that, in wild-type conditions, ranges from 6 to 35 repeats, whereas in HD mutant alleles is expanded from 36 to more than 250 CAG repeats (Gusella and MacDonald, 2002).

There is a strong inverse correlation between the number of CAG repeats and the age at onset of symptoms: larger CAG expansion relates to earlier ages at onset (Andrew et al., 1993). HD alleles carrying CAG repeats between 35 and 40 might show late disease onset or might not be penetrant, whereas individuals with more than 60 CAG typically develop juvenile forms of HD. However the CAG repeat number only

partially explains (~70% of cases) the variability in age at onset of the disease, that seems to be affected also by other genetic and environmental factors (Neto et al., 2016; Rosenblatt et al., 2001). Concerning the genetic background, it has been reported that CAG repeats are highly unstable during germline transmission with a tendency to increase in size particularly when coming from male germline (Trottier et al., 1994). As a result, the age at onset of HD symptoms tends to decrease in successive generations.

Of note, in addition to the one reported in the germline, somatic tissues of HD patients present CAG instability. In HD individuals, higher number of CAG repeats have been detected in disease-affected regions as compared to other non-affected tissues, including cerebellum, blood, lung, kidney, muscle, spleen, stomach, prostate and colon (De Rooij et al., 1995; Telenius et al., 1994). Somatic instability was found to be a significant predictor of onset age, with larger somatic expansions in HD-target brain regions (primarily striatum and cortex) associated with earlier disease onset (Swami et al., 2009, Lee et al.,2012). Similarly to humans, HD mouse models show somatic CAG expansion in pathologically affected brain regions and also in kidney and liver, whereas it is always lacking in the cerebellum (Mangiarini et al., 1997). This phenomenon advances in an age-dependent but not sex-dependent manner (Ishiguro et al., 2001; Wheeler et al., 1999). Moreover, CAG repeat size is larger in neurons of the striatum and triplets are particularly unstable in terminally differentiated neurons (Gonitel et al., 2008).

Notably, in mouse striatum, repeat instability is differentially modulated by the genetic background of mice (Lloret et al., 2006). In humans, HD monozygotic twins have been reported to show different clinical symptoms (Georgiou et al., 1999). This suggests that other cellular or epigenetic factors or tissue-specific variation in CAG repeats might influence the disease. In this context, it has been shown that an important contribution to somatic CAG instability derives from DNA handling and mismatch repair protein members, including Msh2 (Manley et al., 1999; Wheeler et al., 2003), Msh3 (Tomé et al., 2013), Mlh1 and Mlh3 (Pinto et al., 2013). Accordingly, a huge genome wide association study on HD affected families with different age at onset revealed that increased levels of CAG expansion (both germinal and somatic) associates with polymorphic variants of DNA repair genes (Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium, 2015).

Altogether, these findings show that the brain seems to be particularly susceptible to somatic CAG instability in HD, with consequences on the age at onset and

progression of the disease. This phenomenon might be modulated by many genetic and environmental factors, most of which are still unknown.

6.4 Huntingtin protein

Human huntingtin (Htt) is a ~ 350kDa protein containing a polyglutamine (polyQ) sequence, encoded by the CAG stretch at the N-terminus. Across the entire protein multiple HEAT motif important for protein-protein interactions have been identified (Takano and Gusella, 2002). Htt localizes mainly in the cytoplasm, where it frequently associates with vesicle membranes, but it can also be found in the nucleus. The N-terminal portion contains the nuclear localization signal (NLS) and interacts with nuclear pore proteins, whereas the C-terminus includes the nuclear export sequence.

Htt is ubiquitously expressed in humans and rodents, with highest levels in the neurons of the central nervous system (DiFiglia et al., 1995). Htt is particularly enriched in striatal medium spiny neurons (MSNs) and corticostriatal neurons (Fusco et al., 1999). Since its discovery as the protein responsible for HD, many efforts have been made to understand Htt functions and many roles have been described so far. Here I present a brief summary of the most studied ones.

Physiological functions

Since Htt knockout mice die around day 8.5, right before the formation of the nervous system, it seems that Htt expression is required for the normal early embryonic development (Nasir et al., 1995; Zeitlin et al., 1995). Interestingly, Htt is crucial in neurogenesis: reduced levels of Htt in embryo stem cells severely compromise the specification, self-renewal and proliferation potential of primitive neuronal stem cells during the process of neural induction (Nguyen et al., 2013) and, in mice, decreased amounts of Htt lead to aberrant brain development and perinatal lethality (White et al., 1997).

Wild-type Htt can also act as a protein scaffold: Htt can interact with β-tubulin and dynein/dynactin complex and binds to microtubules, therefore participating in intracellular trafficking processes (Caviston et al., 2007; Hoffner et al., 2002). Htt has been also shown to localize to spindle pole during mitosis, controlling orientation in mouse neuronal cells (Godin et al., 2010).

Furthermore, Htt seems to exert a transcriptional regulatory function, as extensively described for BDNF. In this case, Htt sequesters and inhibits in the cytoplasm the REST/NRSF factor, a repressor of BDNF transcription, therefore stimulating BDNF expression (Zuccato et al., 2003). It has been also shown that nuclear Htt interacts with MeCP2 in mouse and cellular models of HD and that this interaction modulates the expression of MeCP2-target genes like BDNF (McFarland et al., 2014). Moreover Htt can enhance the activity of the polycomb repressive complex 2 (PRC2), an epigenetic silencer that promotes chromatin remodeling, thus regulating the expression of several genes involved in stem cell differentiation and embryonal development (Seong et al., 2010).

Involvement of Htt in synaptic connectivity has also been proposed. Htt associates with synaptic vesicles and seems to be required for the correct formation of cortical and striatal excitatory synapses, as shown by huntingtin silencing in developing mouse cortex that leads to an excessive excitatory synapses formation in cortex and striatum, followed by gliosis (McKinstry et al., 2014).

Notably, wild-type Htt protein seems to have antiapoptotic properties as shown both in striatal cells and *in vivo*, where the overexpression of Htt is neuroprotective to various apoptotic stimuli (Cattaneo et al., 2001; Leavitt et al., 2006, 2001; Rigamonti et al., 2000).

6.5 Mutant Huntingtin

CAG expansion leads to the lengthening of the poly-glutamine (polyQ) tract at the N-terminus of huntingtin protein. The polyQ expansion in mutant huntingtin (mHtt) is likely to confer novel toxic gain of function properties to the protein. However, since deletion or inactivation of the wild-type protein also leads to neurodegeneration, the hypothesis of a toxicity mediated by the loss of normal Htt functions cannot be excluded (Dragatsis et al., 2000; O'Kusky et al., 1999).

Like other polyQ-expanded disorders, protein aggregates are found in both HD post mortem brains and HD mouse models, thus representing a pathological hallmark of the disease. Aggregates are composed of misfolded mHtt together with many other proteins including ubiquitin (Becher et al., 1998; DiFiglia et al., 1997), proteasome subunits and chaperones (Cummings et al., 1998; Warrick et al., 1999), transcription factors (Huang et al., 1998; Steffan et al., 2000) or even wild-type huntingtin (Busch et al., 2003).

The identification of N-terminal Htt fragments in neuronal intranuclear inclusion (NIIs) in HD brains suggested proteolytic cleavage of Htt as a toxic mechanism in HD (DiFiglia et al., 1997). These fragments can derive from the proteolysis by caspases, calpains and other proteases (Gafni et al., 2004; Graham et al., 2006; Hermel et al., 2004; Kim et al., 2003, 2001; Wellington et al., 1998). Although both wild-type and mutant Htt can be cleaved, it seems that mHtt fragments is associated with increased toxicity and this might be due to their higher propensity to form nuclear aggregates, instead of cytoplasmic aggregates that are less toxic (Hackam et al., 1998; Kim et al., 1999; Lunkes et al., 2002). The relevance of aggregate toxicity in HD remains controversial. The positive correlation between the number of polyQ and the rate of aggregation and earlier disease onset suggests a toxic role (Martindale et al., 1998), however the inverse correlation between the presence of aggregates and cell death suggests a potential protective role of these inclusions, that might sequester toxic soluble species (Arrasate et al., 2004)

6.6 HD pathogenesis

HD pathogenesis can be viewed as a progressive disease cascade of molecular alterations, where the trigger event takes place far before the onset of neurological dysfunctions and affects specific neuronal populations (MacDonald et al., 2003). HD alterations are primarily detected in striatal neurons, suggesting that HD toxicity starts in this brain area. However, in HD brains, loss of cortical mass has been reported in in early stages of the pathology (Rosas et al., 2002) and alterations in corticostriatal synaptic connectivity in the presence of mHtt seem to occur early in HD mouse models, thus preceding striatal toxicity (McKinstry et al., 2014). Accordingly, the BDNF deficit in striatal neurons derives from defects in axonal transport of BDNF from the cortex to the striatum through the corticostriatal pathway (Gauthier et al., 2004).

As summarized in figure 11, pathophysiological mechanisms in HD include, in addition to aggregate formation, several other processes as diverse as transcriptional dysregulation underlying epigenetic alterations, deregulated protein homeostasis leading to apoptosis, mitochondrial dysfunction, altered synaptic plasticity, axonal transport defects and neuroglia activation. These mechanisms can act both in a synergistic and in a sequential way in HD pathogenesis.

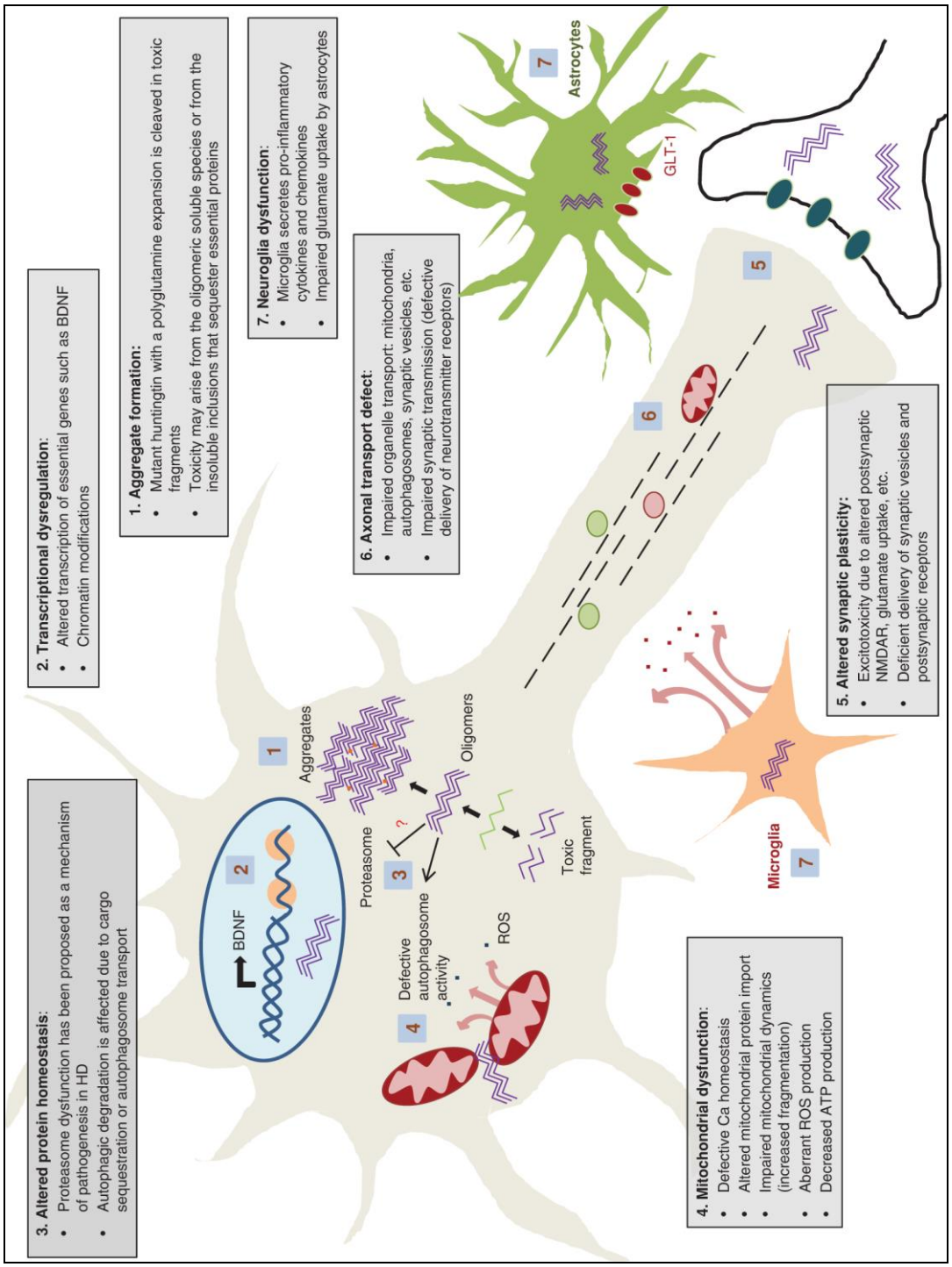


Figure 11. Mechanisms of neuropathogenesis in Huntington's disease. Summary of the main pathogenic mechanisms in HD. BDNF, Brain derived neurotrophic factor. ROS, reactive oxygen species. NMDAR, N-methyl-D-aspartate receptor (Jimenez-Sanchez et al., 2016).

Neurogenesis and neuronal maturation impairments in HD

As previously explained, huntingtin protein plays a crucial role in neurogenesis. Consequently, the presence of the mutant protein can lead to neurogenesis abnormalities. In fact, impairments in the neurogenic niches of the adult brain have been described in HD mouse models: two independent studies showed a decrease in cell proliferation rates of NeuN-positive neurons in the hippocampus of HD adult mice (Gil et al., 2005; Lazic et al., 2004). However, dysfunctions on both neurogenesis and neuronal specification in the presence of mHtt have been also reported brain regions which are not strictly neurogenic but severely affected by HD. As shown by Molero and colleagues, HD knock-in mice exhibit delayed acquisition of cytoarchitecture and altered expression of markers of MSNs neurogenesis in the striatum. Additionally, HD striatal progenitors show an abnormal profile of neurogenesis with an increase in the pool of intermediate progenitors; an altered cell cycle progression with delayed cell-cycle exit between $E_{11.5}$ and $E_{15.5}$ and an enlarged fraction of cells in an abnormal cell cycle state, together with an overexpression of the core pluripotency factor Sox2. Furthermore, in HD conditions neural stem cells display impaired lineage restriction, reduced proliferative potential and deregulated MSN subtype specification (Molero et al., 2009).

Interestingly, it has been shown that Htt in embryos is able to interact with subunits of the epigenetic silencer polycomb repressive complex 2 (PRC2), enhancing its activity in a polyQ dependent manner (Seong et al., 2010). In turn, PRC2 is crucial for the suppression of transcriptional programs that otherwise would be detrimental for neuronal function and survival. As recently reported, PRC2-deficient striatal neurons loose de-repression of target genes involved in transcription regulation, neuron specification and cell death, with a consequent progressive neurodegeneration (von Schimmelmann et al., 2016).

Altogether, these data show that in HD conditions neuronal specification and maturation programs are perturbed and affect particularly striatal neurons, thus rendering MSNs more vulnerable to later stressors.

Remarkably, all these events take place very early in HD brains, suggesting that the first pathological alterations triggered by mutant huntingtin arise already during embryonal neurogenesis.

Transcriptional dysregulation

Transcriptional dysregulation has emerged as a major and early pathogenic event in HD. Microarray studies revealed that expression profiles of both coding and non-coding RNAs are altered in HD (Luthi-Carter et al., 2002; Sipione et al., 2002). Gene expression anomalies have been reported in HD patients (Hodges et al., 2006), are recapitulated in HD mouse models (Kuhn et al., 2007). In many cases, these alterations affect mainly striatal neurons, explaining part of the enhanced toxicity of mHtt in this tissue. A variety of mechanisms has been proposed to explain how mHtt dysregulates transcription patterns, including inhibition of positive regulators or loss of inhibition of transcription's negative regulators. Htt can bind to many transcription regulators both in its insoluble and soluble state. Many transcription factors have been reported to accumulate in the aggregates and this could suggest that their sequestration might limit the ability to bind and regulate their target promoters. On the other hand, many transcription factors are able to interact with the N-term of Htt in its soluble state. These bindings are altered in case of expanded polyQ and lead to a dysregulation of transcription factor activity in a polyQ-dependent fashion (Seredenina and Luthi-Carter, 2012).

An additional line of evidence proposes that mHtt affects the levels of active transcription factors by influencing their degradation by the proteasome (Cong et al., 2005). Moreover, Htt can shuttle from the nucleus to the cytoplasm and since the nuclear localization signal resides in the N-terminus of the protein, the expanded polyQ tract can disrupt the balance of Htt distribution inside cell compartments and promote mHtt accumulation in the nucleus (Benn et al., 2005; Wheeler et al., 2002). This is particularly relevant if we consider that Htt, in both wild-type and mutant forms, can directly bind promoters and several other intronic and intergenic regions, likely leading to chromatin remodeling and eventually affecting transcription (Benn et al., 2008).

One of the most studied evidence of transcriptional alterations in HD is the one related to BDNF, a pro-survival factor produced in cortex and fundamental for striatal neuron survival. Decreased transcription of BDNF occurs in the in the presence of the mutated protein and this leads to degeneration specifically of striatal neurons (Zuccato et al., 2001).

Epigenetic alterations and noncoding RNAs deregulation

Gene expression alterations in HD revealed that the epigenetic landscape and the non-coding RNA transcriptome can be severely affected in HD.

Significant changes in DNA methylation have been detected in striatal cells expressing mHtt. DNA methylation in promoter regions, that can result either in gene activation or silencing, was changed in many genes altered in HD. Importantly, Sox2 binding sites are targets of altered methylation and, in case of increased methylation, loss of Sox2 binding occurs on neurogenesis genes (Ng et al., 2013). However, how mHtt triggers DNA methylation, remains largely unknown. Interestingly, it has been reported that the presence of polyQ-expanded tract on Htt reinforces its binding capacity to MeCP2, and these interactions are stronger in the nucleus compared to the cytoplasm. Furthermore in the presence of mHtt, there is an increased binding of MeCP2 to target genes, like BDNF, and this leads to the suppression of target gene transcription (McFarland et al., 2014).

In HD, extensive deregulation of histone modifications has also been reported. A first hint of this phenomenon came from the study of CBP, a transcriptional co-activator with histone acetyltransferase (HAC) activity. The binding of mHtt to the HAC domain disrupts its function. Likewise, histone deacetylase (HDAC) inhibitors prevent neurodegeneration in cells, *Drosophila* and mouse models of HD (Ferrante et al., 2003; McCampbell et al., 2001; Steffan et al., 2001).

As for histone acetylation, altered histone methylation levels were detected in HD. Differential distribution of H3K9me3 on gene promoters with consequent altered gene expression was observed in HD striatal cell lines (Lee et al., 2013). In HD patients and transgenic HD mice increased levels of ESET, a H3K9 methyltransferase, were found affecting cellular H3K9 trimethylation (Ryu et al., 2006). Moreover, distinctive signatures of H3K4me3 have been reported on dysregulated promoters in HD brains and in cortex and striatum of 12 months old HD mice. Furthermore, increased levels of H3K27me3 are present in HD embryos, embryo stem cells and neural precursors (Seong et al., 2010). The role of mHtt on the deposition of these histone marks remains largely unclear (Dong et al., 2015; Vashishtha et al., 2013). The modulation of H3K27me3 and H3K4me3 has been proposed to strongly depend on PRC2 activity that, in turn, can be promoted by Htt (Seong et al., 2010). Interestingly, in embryo stem cells and neural precursor cells, normal Htt affects both H3K27me3 levels and distribution, acting on

"bivalent loci". In contrast, mHtt is mainly accompanied by altered H3K4me3 at active loci (Biagioli et al., 2015).

Distinctive chromatin signatures are therefore established in HD conditions through mechanisms still largely unknown. This highlights a role of epigenetic modifiers, in addition to genetic modifiers, on pathological mechanisms and progression.

In addition to these effects, HD-associated abnormalities in miRNA biogenesis have also been described. Altered miRNA expression were registered in HD mice at symptomatic stages and in HD patients (Johnson et al., 2008; Packer et al., 2008). Notably, in some cases altered expression of miRNAs is accompanied by altered expression of key miRNA processing enzymes, as well (Lee and Kim, 2011). Moreover polyQ-induced toxicity seems to be modulated by some miRNAs while in a *Drosophila* model of SCA3 (a polyQ disorder), it seems to require an intact miRNA machinery processing (Bilen et al., 2006). All this associates with the detection of huntingtin in RNA structures such as P-bodies (Savas et al., 2008), stress granules (Ratovitski et al., 2012) and dendritic RNA granules (Savas et al., 2010).

DNA damage in HD

An important contribution to neurodegeneration in HD is provided by apoptosis driven by mHtt-induced DNA damage. Its first evidence came from mHtt-expressing PC12 cells and human HD fibroblasts where the production of ROS (reactive oxygen species) correlates with the activation of a DNA damage response (Giuliano et al., 2003). Importantly, in cells and mice, mHtt expression associates with accumulation of DNA lesions, both single-strand and double-strand breaks. Consequently, the activation of DNA damage response leads to phosphorylation of H2AX at the damaged sites and stabilization of p53, which in turn promote cytotoxicity (Illuzzi et al., 2009). This phenotype is an early event in pathology progression, by far preceding protein aggregate formation, and can be modulated by other factors such as Pin1 (Agostoni et al., 2016; Illuzzi et al., 2009; Wheeler et al., 2002). Conversely, DNA damage can influence huntingtin toxicity in neurons, by altering the state of protein phosphorylation, a process long being shown to be crucial for neuronal cell death (Anne et al., 2007).

6.7 Hdh $^{Q7/Q111}$ mouse model

A large contribution to the elucidation of the various mechanisms involved in HD derived from studies performed using mouse models of the disease. Knock-in mice carrying the murine Htt protein with expanded polyQ are close models to the HD genetic state, since they express mHtt at endogenous levels. Several models of HD knock-in mice have been generated (Pouladi et al., 2013), including the Hdh $^{Q7/Q111}$ knock-in mouse model that I used in this study.

The Hdh $^{Q7/Q111}$ knock-in mouse model has been generated by the insertion of 109 CAG repeats in the murine homolog HD gene (Hdh) (White et al., 1997). While mouse lifespan is not compromised, the first signs of neurodegeneration involve early molecular changes that are predictive of later neuropathological phenotypes. These alterations hit the striatum as well as the cortex and then, at later stages, also progress to other brain regions (MacDonald et al., 2003). For this reason, it represents an ideal model to investigate pre-symptomatic HD pathways.

Importantly, these mice display germline CAG instability together with signs of somatic instability particularly in striatum and liver (Wheeler et al., 1999).

As depicted in figure 12, a progressive disease cascade of pathological events is recapitulated in this mouse model. In the first months of life, mice display several molecular striatum-specific alterations: progressive accumulation of mHtt in the nucleus of striatal neurons starts around 2.5 months (mo) and first neuronal intranuclear inclusions are detected at 6 mo of age. ER stress (Carnemolla et al., 2009); somatic CAG repeat expansion (Lloret et al., 2006), reduction of cAMP and decreased levels of BDNF and its TrkB receptor (Gines et al., 2003; Ginés et al., 2006) are observed from the age of 3 mo. As soon as 10 weeks of age, Hdh $^{Q7/Q111}$ mice manifest behavioral abnormalities, including hypoactivity, decreased anxiety, motor learning, long term recognition, spatial memory and coordination deficits and impaired olfactory discrimination (Giralt et al., 2012; Hölter et al., 2013). Importantly, epigenetic alterations triggered by mHtt are recapitulated in these mice: impaired CBP histone acetylase activity and increased levels of H3K27me3 in embryoid bodies have been reported (Giralt et al., 2012, Molero et al., 2009). Signs of neurodegeneration, reactive gliosis and neuronal loss are displayed at older age, from 18mo (Grison et al., 2011; Wheeler et al., 2002).

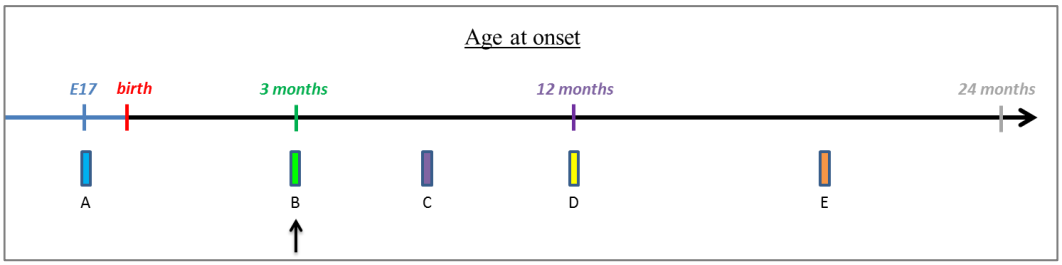


Figure 12. Temporal disease cascade in Hdh$^{Q7/Q111}$ mice. A. Huntingtin facilitates PRC2 histone H3K27 methyltransferase complex (Seong et al., 2010). B. Nuclear accumulation of mutant Htt in MSNs (Wheeler et al., 2002); ER stress (Carnemolla et a.,2009); somatic CAG expansion (Lloret et al.,2006); reduction of cAMP and decreased BDNF production and transport from cortex to striatum (Gines et al., 2003, Zuccato et al,2001). C. Mutant Htt intranuclear inclusions in MSNs. D. Methylation pattern alterations and transcriptional dysregulation in cortex and striatum (Ng et al., 2013. McFarland et al., 2014. Seong et al., 2010); N-terminal inclusions and DNA damage in MSNs (Wheeler et al.,2002). E. Neurodegeneration, reactive gliosis and neuronal loss (Grison et al., 2011; Wheeler et al., 2002). Black arrow indicate onset of early behavioral abnormalities. E17, embryonal day 17.

AIM

TEs are mobile genetic elements that constitute a large fraction of eukaryotic genomes. TEs co-evolved with their host genomes providing tools for genome plasticity and regulation. Mounting evidence shows that autonomous L1 transposable elements are active in somatic cells of the mammalian brain and that dysregulated activation of L1s is associated with neuropathology, such as schizophrenia, Rett syndrome and Ataxia telengectasia.

Likewise, in HD, a series of pathogenic phenotypes that potentially might affect L1 retrotransposition has been reported. These alterations include somatic DNA instability in disease-affected tissues, transcriptional and DNA methylation dysregulations, neurogenesis and DNA repair impairments.

On the other hand, several genetic and epigenetic modifiers able to modulate HD onset and progression have still to be identified, thus suggesting that L1 retrotransposition might play a role in this scenario.

The aim of the present study was to investigate whether L1 retrotransposition could be altered in HD-affected brains.

In this work, taking advantage of a novel *Taqman* qPCR-based approach, we provided a complete characterization of L1 retrotransposition events, either at genomic and transcriptional levels, in a genetically precise HD mouse model, following the entire pathology progression. Subsequently, we described L1-ORF2p expression patterns in adult mouse brains, specifically in striatal and cortical brain regions. Finally, we interrogated potential regulatory pathways affecting L1 retrotransposition that could be altered in mutant huntingtin expressing brains.

MATERIALS AND METHODS

Mouse tissue dissection

In this study, heterozygous ($Hdh^{Q7/Q111}$) knock-in mice in C57BL/6 background were used (Lloret et al., 2006). Animals were provided by prof. M. MacDonald (Massachusetts General Hospital, Boston, MA, USA). Animal care, handling and subsequent procedures were performed in accordance with the European Community Council Directive of November 24, 1986 (86/609EEC) and following SISSA Ethical Committee permissions.

Mice were sacrificed by cervical dislocation. Brains were dissected out and immediately transferred to a Petri dish on ice, dropping some PBS on the tissue chop.

In mice embryos, tissue dissection of the entire cerebral cortex was performed after removal of meninges. From E_{12}, striatal tissue samples were collected from the region of basal ganglia underlying hippocampus.

In adult mice, bulk cortex tissue samples were taken from either left and right hemispheres, including all cortical regions. Striatum tissue samples were dissected from both hemispheres and pulled together for subsequent extractions.

After dissection, all tissue samples were immediately frozen in liquid nitrogen and then stored at -80°C.

Genomic DNA extraction and quantification

Phenol/Chloroform DNA extraction

Dissected tissues (almost 50 mg) were homogenized at room temperature in 2mL of lysis buffer (Tris 100mM; EDTA 5mM; SDS 0,5%; NaCl 150mM) using a glass-Teflon potter. Homogenates were then treated with RNAseA (40µg/mL) (Sigma) at 37°C for 60 minutes in order to remove any residual RNA. After that, proteinase K (Roche) was added at a final concentration of 10µg/mL and samples were incubated at 37°C, overnight (O/N).

The day after, genomic DNA extraction was performed using a standard phenol/chloroform/isoamyl alcohol method: one volume of phenol (water-saturated, pH 8-Sigma) was added to each sample, followed by a centrifugation at 10000rpm for 20 minutes. The aqueous upper phase was collected and then added with 1 volume of

phenol : chloroform-isoamyl alcohol (25: 24:1), followed by a centrifugation at 10000rpm for 10 minutes. Again, the aqueous upper phase was collected in a new tube and added with 1 volume of chloroform : isoamyl alcohol (24:1), followed by a centrifugation at 10000rpm for 10 minutes. From the resulting aqueous phase DNA was finally precipitated by adding two volumes of 100% ethanol so that a visible white flake formed. Flakes were then transferred into fresh tubes, containing 70% ethanol and centrifuged at 12000 rpm for 15 minutes at 4°C in order to wash and gradually hydrate DNA. The DNA pellets were dried from ethanol and then dissolved in 300 µL of Tris 10mM O/N at RT. Genomic DNA quality was then checked by gel electrophoresis using a 0.9% EtBr agarose gel.

Genomic DNA quantification using Quant-iT™ PicoGreen® dsDNA kit (Invitrogen)

In order to detect very small differences within very abundant genomic repetitive sequences, precisely quantified small amount of genomic DNA has to be used for L1 copy number variation analysis. According to the literature (Coufal et al., 2009), the optimal amount of genomic DNA for this type of analysis is 80 picograms corresponding approximately to 12 genomes. To obtain such an accurate DNA quantification, we took advantage of Quant-iT™ PicoGreen® dsDNA kit (Invitrogen), an ultrasensitive double strand DNA (dsDNA) quantification method able to quantify DNA amounts ranging from 25 pg/mL up to 1000 ng/mL.

First of all, DNA concentrations of each sample were measured by standard spectrophotometric techniques, using Nanodrop 2000 (ThermoFisher). According to the initial concentration, DNA samples were diluted in TE buffer (10mM Tris-HCl, 1 mM EDTA, pH 7,5) to a final concentration of 80 pg/µL. A bacteriophage λ dsDNA (provided by the kit) was used to prepare a five-points standard curve, at the following concentrations: 0,1 ng/mL, 0,5 ng/mL, 1 ng/mL, 5 ng/mL, 10 ng/mL. Each point of the standard curve, the DNA samples and the blank were diluted in TE buffer to a final volume of 100 µL, loaded in duplicate in a microtiter plate and added with 100µL of working solution (Quant-iT™ PicoGreen® dsDNA reagent diluted 200-fold in TE buffer). The plate was gently mixed and incubated for 5 minutes at room temperature, protected from light. After incubation, samples' fluorescence was measured at standard fluorescein wavelengths using an EnSpire spectrofluorometer (Perkin Elmer). Sample concentrations were calculated in respect to the standard curve: by plotting fluorescence versus DNA concentration of λ DNA, a standard curve was generated and its equation

was used to assess the DNA concentration of each sample. DNA concentrations ranging from 60 to 100 pg/µL were accepted.

Total RNA extraction and RT PCR

Total RNA was extracted from 30 mg of striatum and cerebral cortex tissue samples using Trizol reagent according to manufacturer's instructions (Ambion). To remove any residual DNA, RNA samples were treated with DNAse I (Ambion) using 2 U every 10 ug of RNA, added with SUPERase RNAse inhibitor (Invitrogen) in a reaction volume of 50uL and incubated at 37°C for 30min. After this treatment the enzyme was inactivated and RNA was purified using Cleanup RNeasy® Mini kit according to manufacturer's instructions (QIAGEN). The RNA concentration and purity were determined by spectrophotometric measurement using Nanodrop 2000 (ThermoFisher) and by running denaturing formaldehyde agarose gel, assessing the ratio between 28S and 18S rRNA subunits, approximately near to 2. After that, reverse transcription (RT) was performed using iScript cDNA Synthesis kit according to manufacturer's instructions (Biorad). In reverse transcription reactions, 0.5 µg of total RNA were used as template in a final reaction volume of 20 µL. For each sample both RT+ and RT- reactions were run by, respectively, adding reverse transcriptase enzyme or replacing it with an equivalent volume of RNAse free water.

Quantitative Real-Time PCR (qPCR)

Quantitative Real-Time PCR (qPCR) is a method of amplification and detection of nucleic acids, which allows determining the quantity of a given template with accuracy, specificity and high sensitivity over a dynamic wide range. Unlike conventional PCR where the amplification product, or amplicon, is detected by an end-point method, with qPCR it is possible to detect the amplicon at each step of amplification, as the reaction progresses. The detection of the amplicon relies on the use of fluorescent molecules, such as DNA-binding dyes (i.e. SYBR Green) or fluorescently labeled sequence-specific probes (i.e. *Taqman* probes), whose fluorescent signal is proportional to the amount of DNA amplified at each cycle.

In this study, we performed *Taqman*-qPCR in L1 specific assays and SYBR Green qPCR in MILI expression experiments. qPCRs were performed using CFX96 Touch™ Real-Time PCR Detection System (Biorad), in a final volume of 20 µL.

Taqman-qPCR experiments were performed in duplex using FAM-labeled *Taqman* probes specific for the gene of interest together with VIC-labeled *Taqman* probe specific for the reference gene. PCR mixes were prepared by adding 10x iQ^{TM} Multiplex Powermix (Biorad), 80pg of genomic DNA or 3ng of cDNA, 0.3 µM of each primer, 0.1 µM of each probe, and RNAse free water up to 20uL. PCR cycling conditions used were: 1)95°C for 20sec, 2) 95°C for 10 sec, 3)59°C for 30 sec. Steps 2) and 3) were repeated 40 times.

For SYBR Green qPCR experiments, master mixes were prepared by adding 2x SYBR Green (Biorad), 10ng of cDNA, 0.2 µM of each primer, and RNAse free water up to 20uL. In expression qPCR experiments, we loaded both RT+ and RT- samples to control residual genomic DNA contamination.

Primer and probes sequences are listed in table 1.

For each qPCR experiment we constructed a standard curve and calculated the primer and probe amplification efficiencies $E=10^{(-1/slope)}$. The amplification efficiencies of the target gene and of the reference gene were accepted between 90-100%.

Relative quantification was assessed by normalizing Ct values of the target gene with those of a reference gene, by using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). In detail, for each sample, the average Ct values of the triplicate for both target and reference genes were first calculated and then the average value of the reference gene was subtracted to that of the gene of interest, obtaining the ΔCt value:

ΔCt = average Ct test probe – average Ct reference probe

Then for each replica, the samples with the highest value of ΔCt was chosen as internal calibrator to normalize ΔCt values between samples within a plate and between samples run in different plates. In this way, ΔΔCt was calculated:

ΔΔCt = ΔCt sample – ΔCt calibrator

Finally, the second derivative value was calculated as follows:

$2^{-\Delta\Delta Ct}$ = normalized amount of target

By doing this normalization, any difference in the amount of input DNA between samples was compensated and an accurate relative quantification was performed.

ORF2-L1 Taqman copy number variation assay

In our laboratory, a novel *Taqman*-qPCR based assay was developed in order to assess L1 copy number variation in mouse genomic DNA, based on that published by Coufal and colleagues in humans (Coufal et al., 2009). Our *Taqman* qPCR strategy was

planned considering that probes complementary to the murine L1-ORF2 sequence are able to detect the entire L1 repertoire in the mouse genome, including both full length and truncated forms, which correspond to both original and retrotransposed L1 copies (Figure 13). The L1-ORF2 probe that we used was labeled with FAM dye at the 5'end and with MGB quencher at the 3'end according to provider's instructions (Applied Biosystem).

In order to perform a relative quantification of the genomic ORF2 content, we designed a *Taqman* probe against the centromeric microsatellite sequence or MICSAT (about 1 million copies), to be used as a reference gene with high number of copies in a duplex qPCR reaction. MICSAT-probe was labeled with VIC dye at the 5'end and with MGB quencher at the 3'end according to provider's instructions (Applied Biosystem). Sequences of primers and probes are listed in table 1.

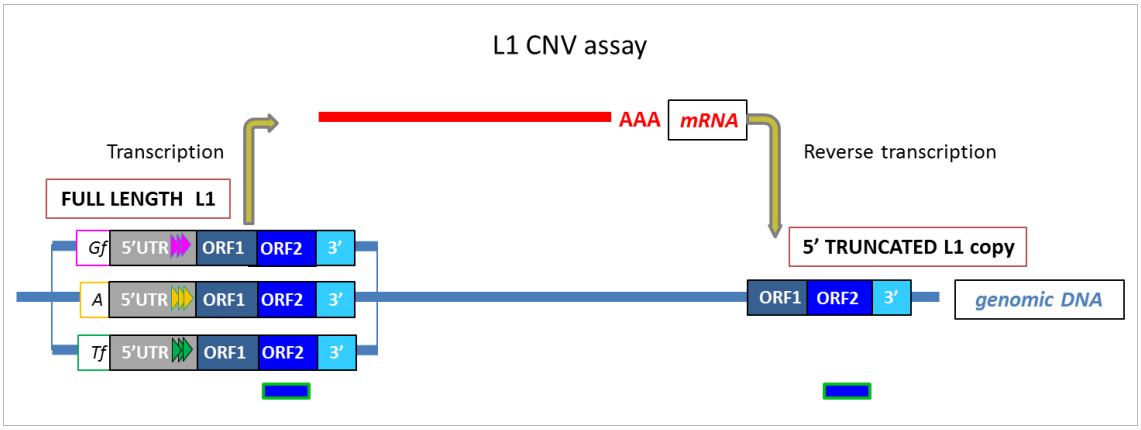


Figure 13. Schematic representation of ORF2-L1 *Taqman* assay. Blue boxes represent *Taqman* probes specific for L1 ORF2-sequence and able to detect both full length L1 and 5'truncated L1 copies present in the genome.

5'UTR-L1 Taqman expression assay

In order to detect endogenous L1 expression in mouse tissues, we developed new *Taqman* assays able to target only full length active L1 elements. Since the murine L1-5'UTR is composed by different repetitive monomers whose organization characterizes different L1 subfamilies (Goodier et al., 2001), we designed *Taqman* probes and primer pairs specific for active mouse L1 subtypes, corresponding to *A, Tf* and *Gf* L1 5'UTRs. In this way, we were able to follow the expression patterns of each L1 subfamily (figure 14). L1-5'UTR *Taqman* probes were labeled with FAM dye at the 5'end and with MGB quencher at the 3'end, according to provider's instructions (Applied Biosystem). In order to obtain a relative quantification of L1 5'UTR mRNA levels against a reference

housekeeping gene, we performed a GeNorm analysis in order to understand which was the most stable housekeeping gene among our set of samples, taking into account the developmental stage, the tissue and the HD genotype. To this aim we tested by GeNorm beta actin, GAPDH, Hmbs, Pgk1, UbC and Ywhaz in a representative number of samples of either cortex and striatum, at E14, P0 and 3mo, including both HD and WT mice. At the end, Ubc turned out to be the most suitable housekeeping gene for our experimental conditions, having a GeNorm M value < 0.5 in all samples analyzed. UbC *Taqman* probe was labeled with VIC dye at the 5'end and with MGB quencher at the 3'end according to provider's instructions (Applied Biosystem). Primers and probes used are listed in table 1.

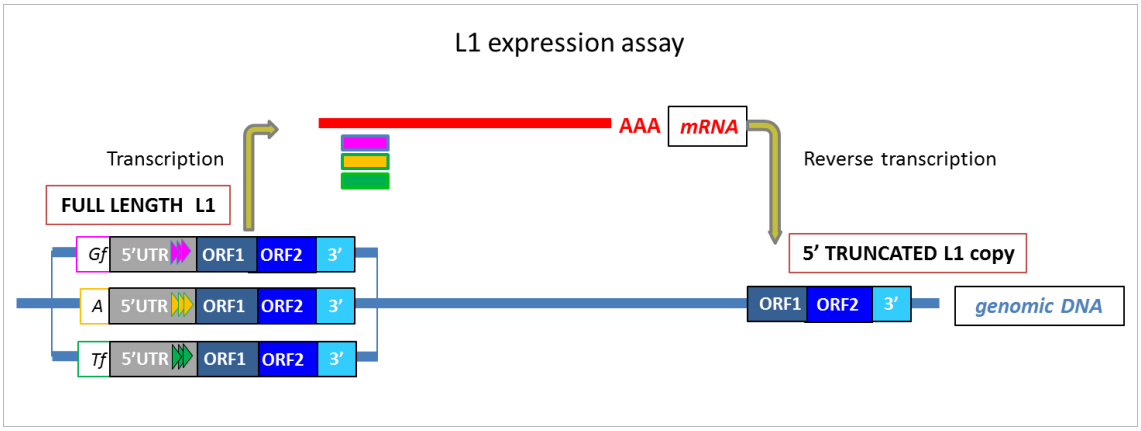


Figure 14. Schematic representation of 5'UTR L1 *Taqman* assay on full length L1s. Triangles represent 5'UTR monomers present on full length mouse L1 elements. Purple, yellow and green boxes indicate, respectively, Gf, A and Tf specific *Taqman* probes, able to detect only full length L1 mRNAs.

Statistical analysis

For each assay, at least three independent qPCR technical replicates were performed on all samples.

In scatter plots, each dot represents the average value for each biological replica derived from three qPCR replicates. Horizontal black line indicates the mean value of at least six independent biological replicas. Error bars indicate standard deviations.

Bar plots report mean value of at least six independent biological replicas. Error bars represent standard error.

In Line graphs, each dot represents the mean value of at least six independent biological replicas. Error bars represent standard error.

Statistical analyses were performed by means of two tailed, Mann Whitney non-parametric test, and taking advantage of GraphPad Prism 5 Statistics Toolbox. * p value<0.05, ** p value<0.01.

MILI and MIWI2 PCR

In order to detect MILI and MIWI2 expression in adult cortex and striatum samples, we performed end point PCR on the same cDNAs previously used for L1 qPCR experiments that displayed no genomic contamination. PCRs were performed using CFX96 Touch™ Real-Time PCR Detection System (Biorad) machine, in a final volume of 50 µL, using ExTaq DNA polymerase (Takara).

Primers used are listed in table 1.

Table 1. List of primers and probes used for qPCR and PCR experiments

oligo name	5'-3' sequence
ORF2 FOR	CCCTCAACAGAGGAATGGAT
ORF2 REV	CCATCCATTTGGCTAGGAAT
ORF2 probe	AAATGTGGTACATCTACACAATGGA
MICSAT FOR	GAACATATTAGATGAGTGAGTTAC
MICSAT REV	GTTCTACAAATCCCGTTTCCAA
MICSAT probe	ACTGAAAAACACATTCG
5'UTR Gf FOR	CCAAACACCAGATAACTGTACACC
5'UTR Gf REV	CGTGGGAGACAAGCTCTCTT
5'UTR Gf probe	TGAAAGAGGAGAGCTTGCCT
5'UTR A FOR	TGCCCACTGAAACTAAGGAGA
5'UTR A REV	GCTTGTTCTTCAGGTGACTCTGT
5'UTR A probe	TGCTACCCTCCAGGTCTGCT
5'UTR Tf FOR	TGAGCACTGAAACTCAGAGGAG
5'UTR Tf REV	GATTGTTCTTCTGGTGATTCTGTTA
5'UTR Tf probe	GAATCTGTCTCCCAGGTCTG
UbC FOR	ACAGACGTACCTTCCTCACC
UbC REV	CCCCATCACACCCAAGAACA
UbC probe	AAAAAGAGCCCTCCTTGTGC
MILI FOR	AAGGACAGAGAAGAACCCCG
MILI REV	ATACTACTGGCTGCTCGTCC
MIWI2 FOR	CCCGACTCGTGGATGACAT
MIWI2 REV	GGCTTCACAGACCAGTCAATG

Protein extraction and western blot

Total protein lysates from cerebral cortex and striatum were obtained using Trizol reagent according to manufacturer's instructions (Ambion) after RNA extraction and then sonicated for complete lysis. Protein lysates concentrations were determined by Bradford (BIORAD) assay using a calibration curve built with serial dilutions of bovine serum albumin (BSA). 80 µg of cortex and striatum lysates and 20 µg of testes lysates were mixed with 2X SDS sample buffer. Samples were boiled at 95°C for 5 min and separated by SDS-PAGE electrophoresis. Proteins were then transferred to nitrocellulose membrane (GE Healthcare) by blotting at 100V for 1 hour. Membranes were blocked with 5% non-fat milk in Tris Buffer Saline Tween20 (TBST) and then incubated at 4°C overnight with primary antibody diluted in 5% milk-TBST, with agitation. The day after, membranes were washed three times in TBST and then incubated with secondary HRP-conjugated antibody diluted in 5% milk-TBST at RT for 1 hour. Proteins bands were developed using ECL developing system (GE Healthcare).

The following primary antibodies were used: anti MILI (1:1000, MABE363, Millipore), anti MIWI2 (1:500, Abcam), anti beta actin (1:20000, Sigma).

Immunohistochemistry

Mice were transcardially perfused with 4% paraformaldehyde in PBS (pH 7.4). The brain was rapidly removed from the skull, post-fixed in PFA 4% o/n at 4°C and then transferred in 30% sucrose/PBS until it sunk. Brain samples were included in Killik medium (Bio-Optica) and then flash frozen by immersion in liquid nitrogen. Coronal cryosections of 16µM thickness were prepared cutting throughout the striatum (+3mm : -3mm from Bregma).

For immunohistochemistry, slices were first rinsed in Tris Buffered Saline (TBS) solution for 10 min at room temperature (RT), then added with 0.1 M glycine/TBS for 10 min at RT to quench autofluorescence and then quickly rinsed three times in TBS. After that, slices were incubated for 1 hour at RT in blocking solution (1% BSA, 10% FBS,1% fish gelatin in TBS) and then at RT overnight with primary antibody in TBS added with 0.1% fish gelatin, 0.3% Triton X-100, 1% BSA. The next day, sections were incubated with secondary antibodies: first with biotinylated antibody for 1h30 min at RT, and then with streptavidin Alexa 488 conjugated antibody for 1h30 min at RT. After each antibody incubation, slices were washed 3 times in TBS. For nuclei staining,

DAPI was added during the streptavidin A488 incubation at a dilution of 1:500. Finally slices were rinsed 2 additional times in H_2O and slides were mounted using Vectashield mounting medium. Images were captured using Leica TCSSP2 Confocal microscope.

The following primary antibodies were used: anti LINE-1 (M300) (1:100, Santa Cruz Biotechnology), anti MILI (1:10, Millipore).

Chromatin immunoprecipitation

Striatum tissue samples (~50mg) were first minced into small pieces with minimal thawing, then added with 1 mL of ice cold PBS 1X and completely homogenized using a glass-Teflon potter. For crosslinking, homogenates were incubated with 1% PFA in a final volume of 2mL at RT for 20 minutes, rocking. PFA was quenched by incubation with 125mM of glycine at RT for 5 minutes, rocking. Lysates were then collected by centrifugation at 5000 rpm, at RT for 5 minutes and then washed two times with 500µL of ice-cold PBS 1X to remove any residual PFA and glycine.

Then, cell lysis was performed in 500uL of cell lysis buffer (10mM HEPES pH 8, 85mM KCl, 0.5% NP-40) added with 25X proteinase inhibitors. Samples were homogenized by douncing, incubated on ice for 10min and finally collected at 5000rpm, at RT for 5 minutes. After that, lysates were resuspended in 300 uL of nuclei lysis buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH8) added with 25X proteinase inhibitors and incubated on ice for 10min. Samples were then transferred into sonication tubes and sonicated using Bioruptor Sonicator NGS (Diagenode) using the following settings: high power, 30 seconds ON and 30 seconds OFF per each cycle, 30 cycles. After sonication, 1/30 of sonicated DNA fraction was saved, decrosslinked, DNA was extracted, quantified using Nanodrop 2000 (ThermoFisher) in order to have an estimate of sonicated chromatin concentration and run on a 1% EtBr agarose gel to assess chromatin shearing to be around 200bp.

In the meanwhile, antibody-magnetic Dynabeads® M-280 Sheep anti-Rabbit IgG (Invitrogen) complexes were prepared by first washing beads twice in blocking solution (0.5% BSA in 1x PBS) and then by incubating beads with IgG or 5ug of antibody per each immunoprecipitation in blocking solution at 4°C, overnight, rocking. The day after, beads-antibody complexes were washed again 3 times in blocking solution to remove any unbound antibody and resuspended in 100uL of blocking solution, ready to be used for immunoprecipitation.

The sonicated DNA, after checking the chromatin fragmentation size, was centrifuged at 20000g, at 4°C for 30 minutes to pellet cell and nuclear debris and diluted 1:10 in ChIP Dilution buffer (0.01% SDS, 0.275% Triton X-100, 1mM EDTA, 16.7mM Tris-HCl pH 8, 167 mM NaCl) added with 25X proteinase inhibitors. After having removed a control aliquot to be used as INPUT sample, sonicated chromatin was split in equivalent portions per each immunoprecipitation (IP) (approximately 20ug per IP) and incubated with antibody-beads complexes, previously prepared, at 4°C, overnight, rocking.

Chromatin-antibody complexes were then sequentially washed with increasing stringency salt buffers as follows: 2 times in low salt buffer (20mM Tris-HCl pH 8, 150mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), 2 times in high salt buffer (20mM Tris-HCl pH 8, 500mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA), 2 times in LiCl buffer (10mM Tris-HCl pH 8, 250mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1mM EDTA). At each wash, samples were incubated for 5 minutes, at 4°C, rocking and then collected using pre-chilled magnetic stand. Finally samples were washed once in TE buffer (10mM Tris-HCl pH 8, 1mM EDTA) and spin at 1000 x g for 3 min at 4°C. Immunoprecipitated chromatin was then eluted in elution buffer (1% SDS, 100mM NaHCO3, 5mM DTT), at 65°C for 15 minutes, shaking. An aliquot of eluted IP sample was saved to check immunoprecipitation specificity by Western Blot analysis. The remaining eluted IP chromatin was collected and decrosslinked at 65°C overnight, together with INPUT chromatin sample. After that, proteinase K (1mg/mL) (Roche) treatment was performed at 37°C for 2 hours followed by Phenol/Chloroform DNA extraction. IP and INPUT DNA samples were then treated with RNAse A (80ug/uL) (Sigma) and purified using MinElute kit (QIAGEN) according to manufacturer's protocol. Finally, DNA concentration was measured using Quant-iT™ PicoGreen® dsDNA kit (Invitrogen).

Four independent biological replicas were performed per each chIP experiment.

Immunoprecipitation specificity with the expected antibody was assessed by Western Blot analysis on IP and IgG samples for each replica.

ChIP enrichments were assessed by SYBR Green-qPCR using CFX96 Touch™ Real-Time PCR Detection System (Biorad), loading 100 pg of ChIP DNA and an equal amount of un-enriched INPUT DNA. Enrichments were calculated from 4 independent biological replicas. For chIP experiments on histone modifications, a gene desert region located on mouse chromosome 6 (Active Motif) was used as internal negative control

for normalization of chIP DNA enrichments. DNA enrichments were assessed as follows: chIP DNA enrichment for L1 5'UTR = [(2^ (-ΔCt))*100] with ΔCt = [(avg Ct IP-Avg Ct input (for L1 5'UTR)) - (avg Ct IP- Avg Ct input (for chr6 gene desert region)].

For chromatin immunoprecipitation the following antibody were used: anti H3K4me3 (Millipore), anti H3K9me3 (Millipore), anti H3K27me3 (Millipore) and anti-MeCP2 (Abcam)

RESULTS

Preliminary data

In humans, it is known that several regions of adult brain, in particular neurogenic compartments like hippocampus and olfactory epithelium, and frontal cortex display an increased number of genomic L1 copies as compared to other somatic non nervous tissues, such as heart and liver (Coufal et al., 2009, Baillie et al., 2011).

We asked ourselves whether this pattern of L1 copy number in different tissues might be present also in mice. To this aim, in a previous study performed in our laboratory, the rate of L1 retrotransposition in different mouse somatic tissues was assessed, with a particular interest on the comparison between brain and non-nervous tissues.

As explained in detail in the methods section, L1 CNV analysis in mouse genome was performed with a *Taqman* qPCR assay developed in our laboratory by adapting the procedure described for human samples by Coufal and colleagues. In brief, in order to account for the entire repertoire of L1 elements present in the mouse genome (more than 100000 copies) we used a *Taqman* probe complementary to the ORF2 sequence of a murine L1, able to detect both full length L1 elements and 5'truncated L1 copies that underwent retrotransposition. Then, in order to normalize the amount of ORF2p sequences on a reference gene, we designed an assay against the high copy number invariant microsatellite centromeric sequence, MICSAT (about 500000 copies). With this technology, a relative quantification of the total number of genomic L1s could be made, thus giving an estimate of the rate of L1 retrotransposition events in a given tissue.

Similarly, to the experiment performed in humans, we analyzed genomic DNA extracted from tissue samples of hippocampus, cortex and olfactory epithelium, together with two non nervous organs corresponding to liver and kidney. Samples derived from 5 wild-type C57Black mice at 3 months of age. Three independent qPCR replica were performed.

From this study, consistent with what observed in humans, a marked alteration in total L1 copy number was observed in the brain regions compared to non-brain tissues (figure 15). In particular, higher amounts of genomic L1 elements were found in adult hippocampus, olfactory epithelium and cortex relative to liver and kidney. A paired *t*-

test comparing the different brain samples to liver and kidney showed that the increase in ORF2 content in the brain was statistically significant. In detail, we observed that ORF2 copies in olfactory epithelium and cortex were increased of 14% (p-value = 0.001) and 12% (p-value = 0.01), respectively, compared to liver and of 22% (p-value = 0.006) and 19% (p-value = 0.02) compared to kidney. The difference in relative L1-ORF2 copies in hippocampus was less evident when compared to liver (hippocampus contains on average 10% more ORF2 than liver), but it became significant when compared to kidney (18% (p-value = 0.04) more L1 copies in hippocampus than kidney).

Overall, since increased L1-ORF2 copies result from active L1 retrotransposition, we may conclude that, in adult mouse, somatic endogenous L1 retrotransposition is higher in the brain than in non-brain tissues, thus confirming data reported in human hippocampus and frontal cortex (Coufal et al., 2009, Baillie et al.,2011).

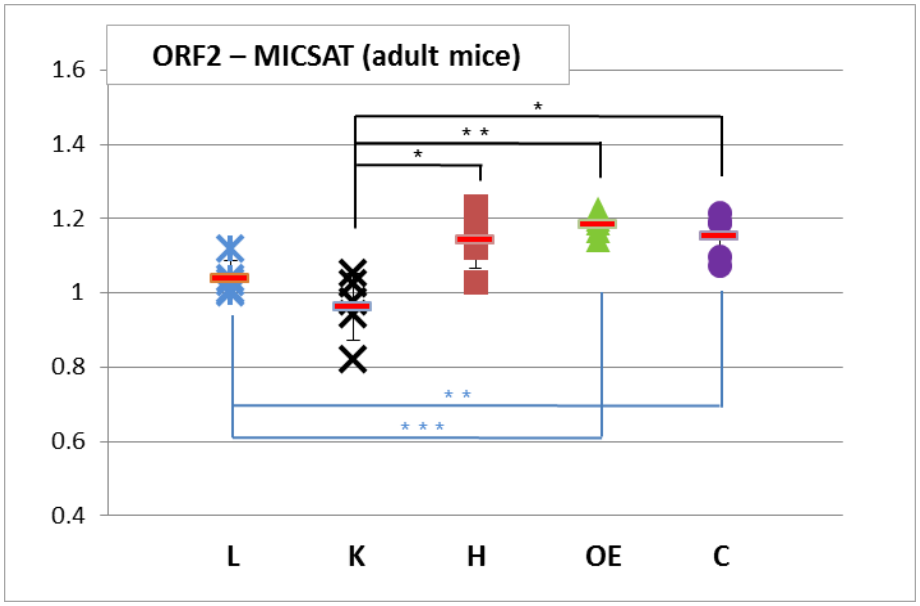


Figure 15. The copy number of genomic L1s is increased in mouse adult brain regions as compared to non-nervous tissues. Relative quantification of total number of genomic L1 elements obtained by qPCR using the ORF2-MICSAT *Taqman* assay. L, liver; K, kidney; H, Hippocampus; OE, olfactory epithelium; C, cortex. The amount of L1 ORF2 sequences is significantly higher in all brain regions respect to liver and kidney. Scattered graphs represent L1 ORF2 relative amount in 5 adult mice. Each symbol represents the average of 3 qPCR independent replica. Red lines represent the average of all samples. Error bars indicate standard deviations. *P<0.05; **P<0.01; ***P<0.001 resulting from *t-student* paired test.

1. Characterization of L1 retrotransposition events in a mouse model of HD

In the last decade, several works, mostly taking advantage of engineered L1 reporter assays and L1 CNV analysis, showed that L1s are active during neurogenesis and that dysregulated activation of L1s is associated with neuropathology, such as schizophrenia, Rett syndrome and ataxia telengectasia (ATM) (Bundo et al., 2014; Coufal et al., 2011; Muotri et al., 2010).

In Huntington's disease (HD), pathogenic mechanisms including somatic CAG instability, altered neurogenesis, transcriptional dysregulation, epigenetic alterations and DNA damage may potentially alter L1 retrotransposition activity.

At the same time, several genetic and epigenetic modifiers, able to modulate age at onset and progression of the disease, remains largely uncharacterized, suggesting that L1 retrotransposition might play a role in this scenario.

The first step that we made towards an understanding of L1 dynamics in HD was the characterization of L1 retrotransposition events in a precise genetic mouse model of HD.

1.1. L1 copy number variation analysis in WT/KI compared to WT/WT mice

First, we assessed the number of genomic L1 copies in HD mice as compared to controls. To this aim, we performed L1 copy number variation (CNV) analysis in cortex and striatum, the most affected tissues in HD, of Hdh $^{Q7/Q111}$ (WT/KI) and Hdh $^{Q7/Q7}$ (WT/WT) mice. To this purpose, we took advantage of the ORF2-MICSAT *Taqman* assay able to account for the entire repertoire of L1 copies in the mouse genomes, including both full length and truncated forms.

It has been demonstrated that mutant huntingtin can trigger early alterations during the embryonic development (Molero et al., 2009, Seong et al., 2010). At the same time, somatic retrotransposition of an engineered L1 has been reported during early embryonal neurogenesis, in particular between E8.5 and E10.5 (van den Hurk et al., 2007, Muotri et al., 2005). For these reasons, we decided to analyze L1 CNV in both embryonic and adult stages. In detail, L1 CNV analyses were performed at $\underline{E_{12}}$, right after the bulk of L1 retrotransposition reported by Muotri et al. and corresponding to the beginning of telencephalic neurogenesis; at $\underline{P_0}$, corresponding to the end of telencephalic neurogenesis; at three progressive adult ages when HD-specific

phenotypes are manifest in affected neurons: 3 months of age (3mo), when mutant huntingtin accumulates in the nucleus of medium spiny neurons and a reduced production and transport of BDNF from the cortex to the striatum is observed (Wheeler et al., 2002), (Zuccato et al., 2001); 12 months of age (12mo), when medium spiny neurons display Htt nuclear inclusions and DNA damage, and cortex and striatum are affected by alterations in methylation patterns and transcriptional dysregulation (McFarland et al., 2014; Ng et al., 2013; Seong et al., 2010); 24 months of age (24mo) when neurodegeneration of medium spiny neurons and reactive gliosis is detected (Wheeler et al., 2002).

For each assay at least three independent technical replicas were performed. For each developmental stage and genotype, we analyzed at least 6 independent biological replicas.

Therefore summarizing, by assessing relative ORF2-L1 content, we compared L1 CNV in the striatum and in the cerebral cortex of WT/WT and WT/KI mice, at the stages of E_{12}, P_0, 3mo, 12 mo and 24 mo.

As shown in figure 16, L1 CNV analysis in the striatum revealed no significant differences between WT/WT and WT/KI mice in relative ORF2 genomic content. This was consistent for all the developmental stages analyzed, both at pre-natal and post-natal ages (figure 16). As indicated by each dot of the scatter plot, a certain level of inter-individual variability was observed at E12 and P0. However, in post-natal stages the deviations among biological replicas were very small.

In the cerebral cortex, similarly to the striatum, we observed equivalent levels of genomic L1 elements in WT/WT and WT/KI mice, at all developmental stages analyzed (figure 17). As for the striatum, also in the cortical tissue samples, high levels of variability between mice were detected at E12 and P0. However, in the case of cortex, a certain level of inter-individual variability was registered also in post-natal stages, particularly at 3mo of age.

Altogether these data show that both in the striatum and in the cortex of E12, P0 and adult mice, equivalent levels of L1 genomic copies are present in Hdh $^{Q111/Q7}$ and Hdh $^{Q7/Q7}$ mice.

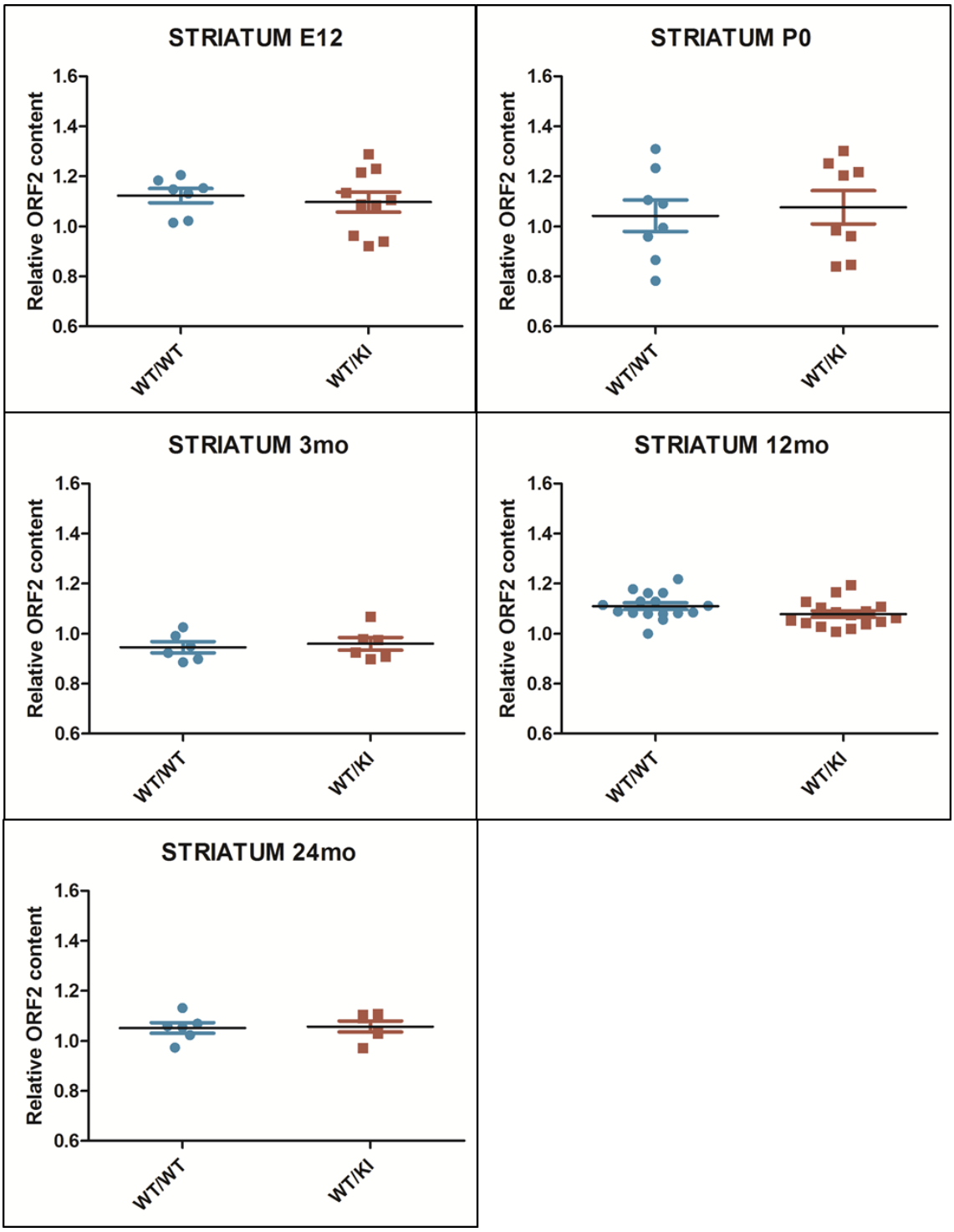


Figure 16. Genomic L1 copy number in striatum of WT/WT and WT/KI mice. Relative quantification of total genomic L1 elements using ORF2-MICSAT Taqman assay. E12, embryonal day 12; P0, date of birth; 3mo, 3 months of age; 12mo,12months of age; 24mo, 24 months of age. Similar levels of L1-ORF2 sequence are detected between WT/WT and WT/KI mice. Each dot of the scatter plot represents the average of 3 independent qPCR replica. The average value between biological replicas is shown with black horizontal line. Error bars indicate standard deviation.

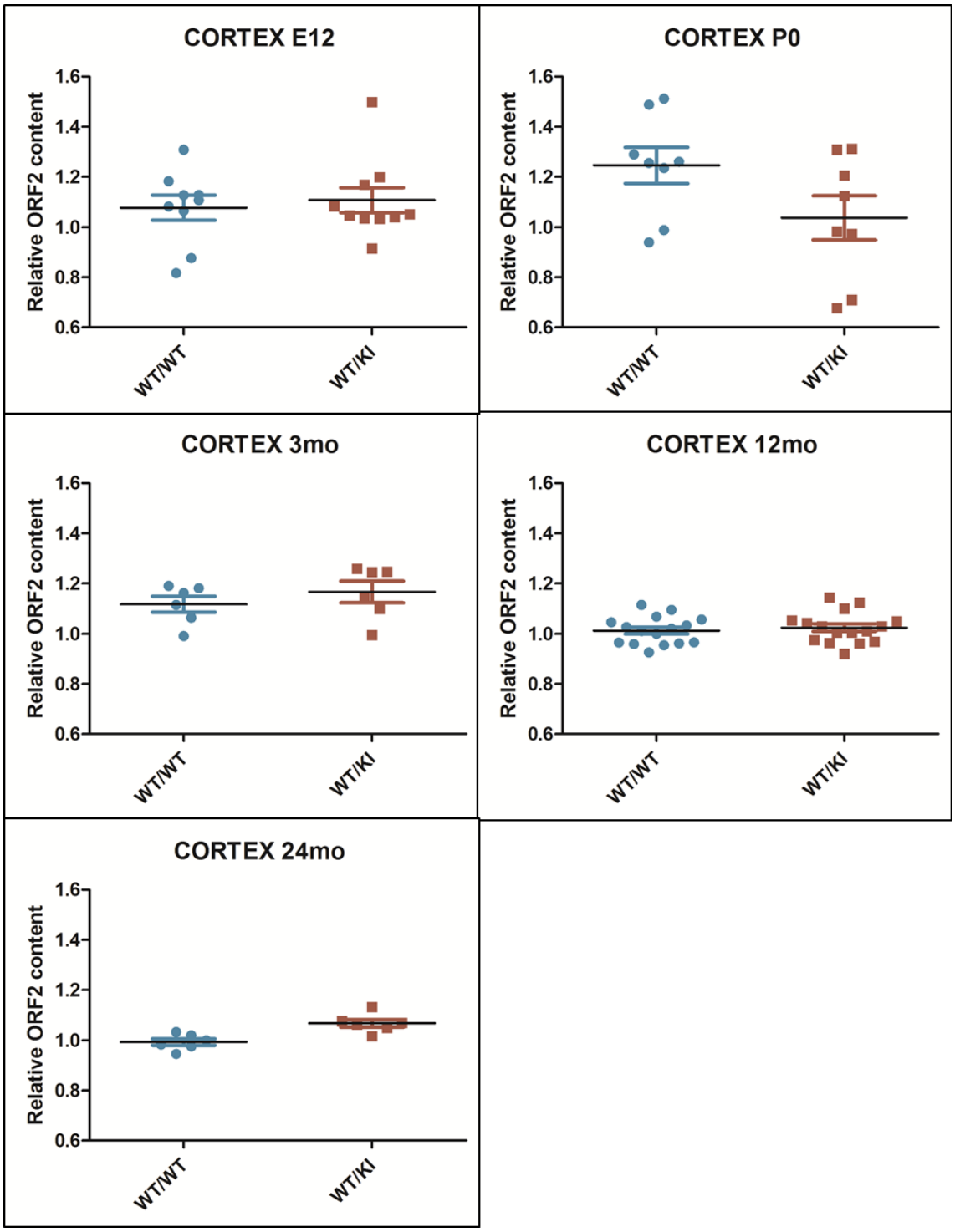


Figure 17. Genomic L1 copy number in cortex of WT/WT and WT/KI mice. Relative quantification of total genomic L1 elements using ORF2-MICSAT Taqman assay. E12, embryonal day 12; P0, date of birth; 3mo, 3 months of age; 12mo,12months of age; 24mo, 24 months of age. Similar levels of L1-ORF2 sequence are detected between WT/WT and WT/KI mice. Each dot of the scatter plot represents the average of at least 3 independent qPCR replica. The average value between biological replicas is shown with black horizontal line. Error bars indicate standard deviation.

1.2. L1 expression analysis in WT/KI compared to WT/WT mice

In parallel to L1 CNV analysis, in order to investigate whether L1 transcription can be altered by mutant huntingtin, we also performed L1 expression analysis. To this purpose, we compared endogenous L1 transcript levels of WT/KI ($Hdh^{Q7/Q111}$) and WT/WT ($Hdh^{Q7/Q7}$) mice, using RNA extracted from striatum and cerebral cortex.

To carry out endogenous L1 expression analysis we developed a *Taqman* qPCR-based approach able to detect only full length active L1 elements, transcribed from the endogenous 5'UTR sense promoter. Since the murine genome is scattered with several L1 classes that differ for the monomeric organization at their 5'UTR, we designed *Taqman* probes specific for each 5'UTR-L1 subtype. In this way, we were able to discriminate between the different subfamilies of active L1s, that are Gf, A and Tf. Transcript levels of each L1 5'UTR were normalized with the housekeeping gene ubiquitin C (UbC), since it resulted to be the most stable housekeeping gene for our experimental conditions, according to the GeNorm analysis that considered the different stages, tissues and genotypes.

As previously mentioned, neurogenesis alterations and early epigenetic changes in the presence of mutant huntingtin can be detected in HD mice already during embryonal stages (Seong et al., 2010, Molero et al., 2009). Therefore, we performed L1 expression analysis starting from embryonal day 10 up to 24 months of age. In detail, we considered four embryonal stages: E10, E12, E14 and E17; the date of birth (P0) and the same adult stages considered for L1 CNV: 3 months, 12 months and 24 months.

Similarly to L1 CNV, per each stage, we analyzed both striatum and cortex. At E10 striatum and cortex are not yet differentiated and therefore, at this embryonal stage, we analyzed only the common precursor, here simply referred as "cortex". For each assay at least three independent technical replicas were performed.

For each assay at least three independent technical replicas were performed. For each developmental stage and genotype, we analyzed at least 6 independent biological replicas.

L1 expression in embryos and at P0

The expression analysis of full length Gf, A and Tf L1 elements in the striatum of E12, E14, E17 and P0 mice revealed similar levels of L1 mRNAs between WT/WT and WT/KI mice (figure 18) at all the developmental stages analyzed.

Bar plots summarize the results obtained for each L1 subfamily and show that expression patterns are similar for L1 5'UTR Gf, L1 5'UTR A and L1 5'UTR Tf subfamilies.

Detailed scatter plots showing distribution of all biological replicas are shown in Appendix.

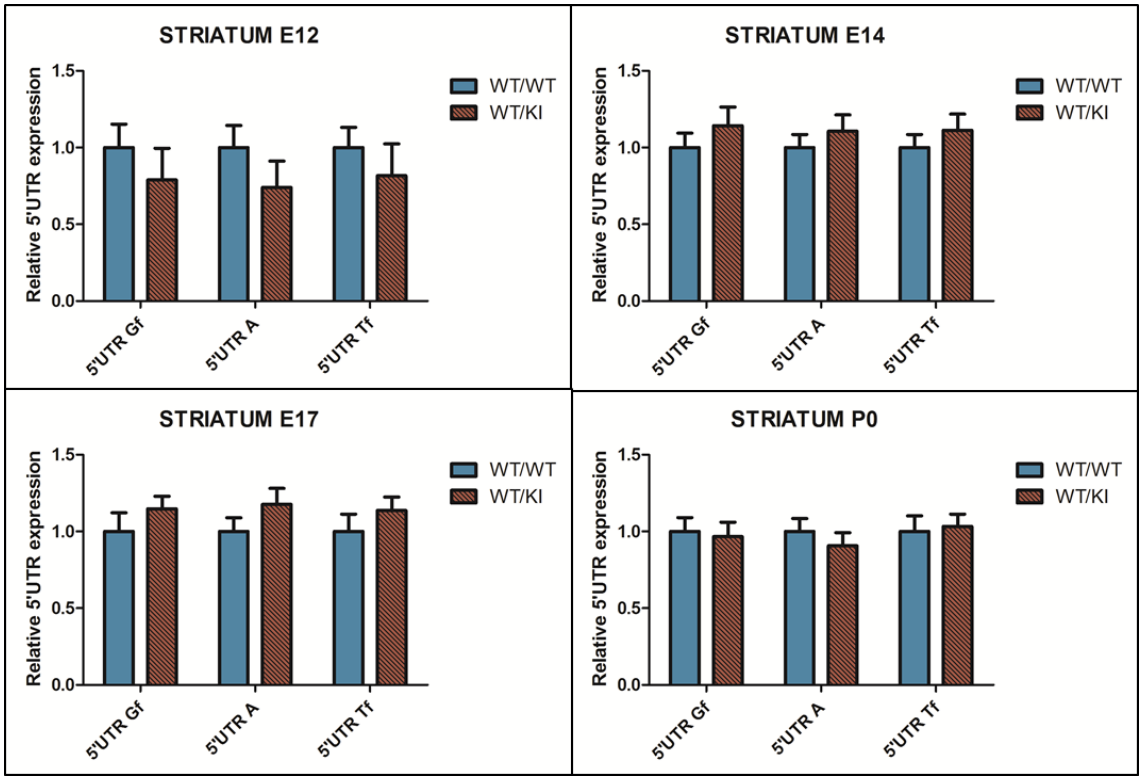


Figure 18. Full length L1 mRNA levels are similar in the striatum of WT/WT and WT/KI mice in late embryos and at P0. Relative quantification of 5'UTR L1 mRNA levels in the striatum during late embryonal development. E12, embryonal day 12; E14, embryonal day 14; E17, embryonal day 17; P0, date of birth. 5'UTR Gf, 5'UTR A, 5'UTR Tf refer to full length Gf, A and Tf L1 subfamilies present in the mouse genome. Relative 5'UTR L1 mRNA levels are normalized with the housekeeping gene UbC. L1-5'UTR transcript levels are equivalent between WT/WT (light blue) and WT/KI (dark red) mice. 3 independent qPCR replicas were performed on each biological sample. Each bar represents average value of six independent biological replicas (n=6). Error bars indicate standard error.

In the cerebral cortex (figure 19), similarly to what shown in the striatum, we observed comparable levels of full length L1 transcripts between WT/WT and WT/KI mice. This was true for all the three L1 subfamilies considered and in all embryonal stages as well as at P0.

Detailed scatter plots showing distribution of all biological replicas are shown in Appendix.

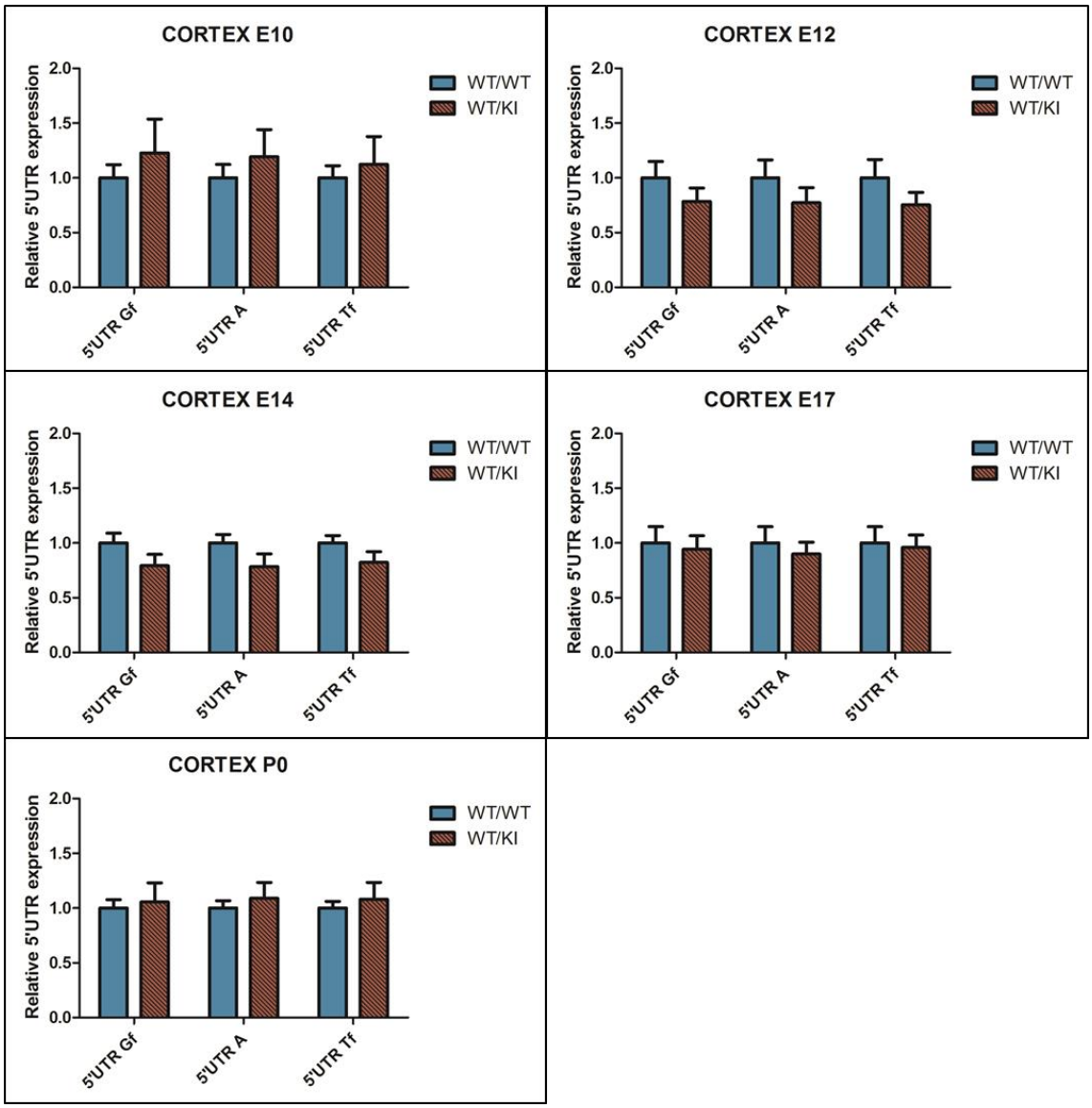


Figure 19. Full length L1 mRNA levels are similar in the cortex of WT/WT and WT/KI mice in late embryos and at P0. Relative quantification of 5'UTR L1 mRNA levels in the cortex during late embryonal development. E10, embryonal day 10; E12, embryonal day 12; E14, embryonal day 14; E17, embryonal day 17; P0, date of birth. 5'UTR Gf, 5'UTR A, 5'UTR Tf refer to full length Gf, A and Tf L1 subfamilies present in the mouse genome. Relative 5'UTR L1 mRNA levels are normalized with the housekeeping gene UbC. L1-5'UTR transcript levels are equivalent between WT/WT (light blue) and WT/KI (dark red) mice. 3 independent qPCR replicas were performed on each biological sample. Each bar represents average value of six independent biological replicas (n=6). Error bars indicate standard error.

L1 expression in adult stages

To provide a complete characterization of L1 expression throughout the lifespan of the HD mice, in addition to embryos, we analyzed the same adult stages considered for L1 CNV (3mo, 12mo and 24mo mice) following the progressive manifestation of HD phenotypes that specifically occurs at the level of striatum and cortex.

In the the striatum, at 3mo of age, we have not found significant alterations in WT/WT compared to WT/KI mice, for none of the three L1 5'UTR subtypes (figure 20a). However, at 12mo of age (figure 20b), interestingly, we observed a significant reduction (about 20%) of L1 transcript levels in WT/KI mice compared to WT/WT. Notably, this decrease was conserved between all the three L1 subfamilies. Later on, at 24 mo of age, equivalent levels of L1 mRNAs could be detected between WT/KI and WT/WT mice (figure 20c).

Detailed scatter plots showing distribution of all biological replicas are shown in Appendix.

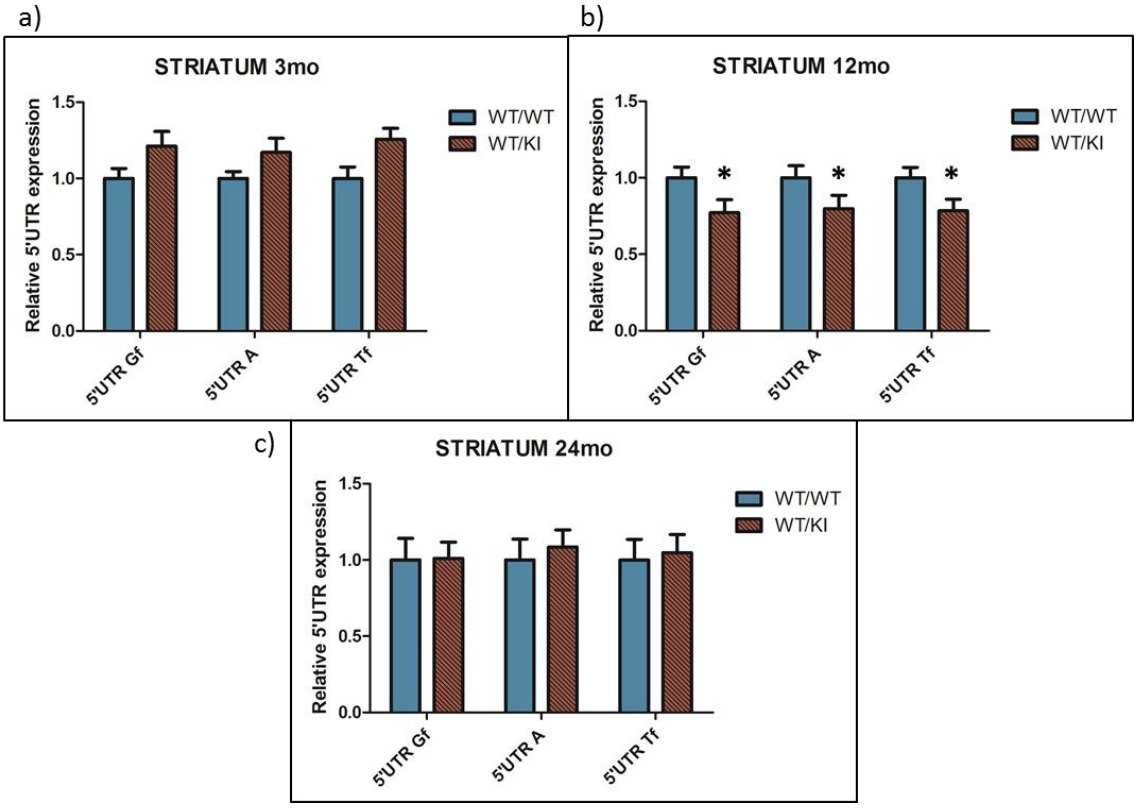


Figure 20. Full length L1 mRNA levels are decreased in the striatum of 12months old WT/KI mice. Relative quantification of 5'UTR L1 mRNA levels in the striatum of adult mice. 3mo, 3 months of age; 12mo, 12 months of age; 24mo, 24 months of age. 5'UTR Gf, 5'UTR A, 5'UTR Tf refer to full length Gf, A and Tf L1 subfamilies present in the mouse genome. Relative 5'UTR L1 mRNA levels are normalized with the housekeeping gene UbC. At 12mo of age, L1-5'UTR expression of all three L1 subfamilies is significantly decreased in WT/KI (dark red) mice as compared to controls (WT/WT, light blue). 3 independent qPCR replicas were performed on each biological sample. Each bar represents the average value of twelve independent biological replicas (n=12). Error bars indicate standard error. *P<0.05 resulting from *Mann Whitney* unpaired test.

As for the striatum, also from the analysis of L1 expression in adult cerebral cortices, we could conclude some original observations. On one hand, at 3mo of age, we observed higher levels (~ 25% increase) of L1 5'Gf mRNA in WT/KI mice compared to

controls. This increase was not shared with the other two L1 subfamilies (A and Tf) (figure 21a). On the other hand, no differences between WT/WT and WT/KI mice were detected at 12 mo of age for none L1 5'UTR classes (figure 21b). But interestingly, at 24 mo, similarly to what observed at 3mo, WT/KI mice displayed a strong activation of L1 transcription. Notably, the L1 mRNA overexpression in 24mo WT/KI mice was consistent in Gf, A and Tf L1 subtypes (figure 21c). Detailed scatter plots showing distribution of all biological replicas are shown in Appendix.

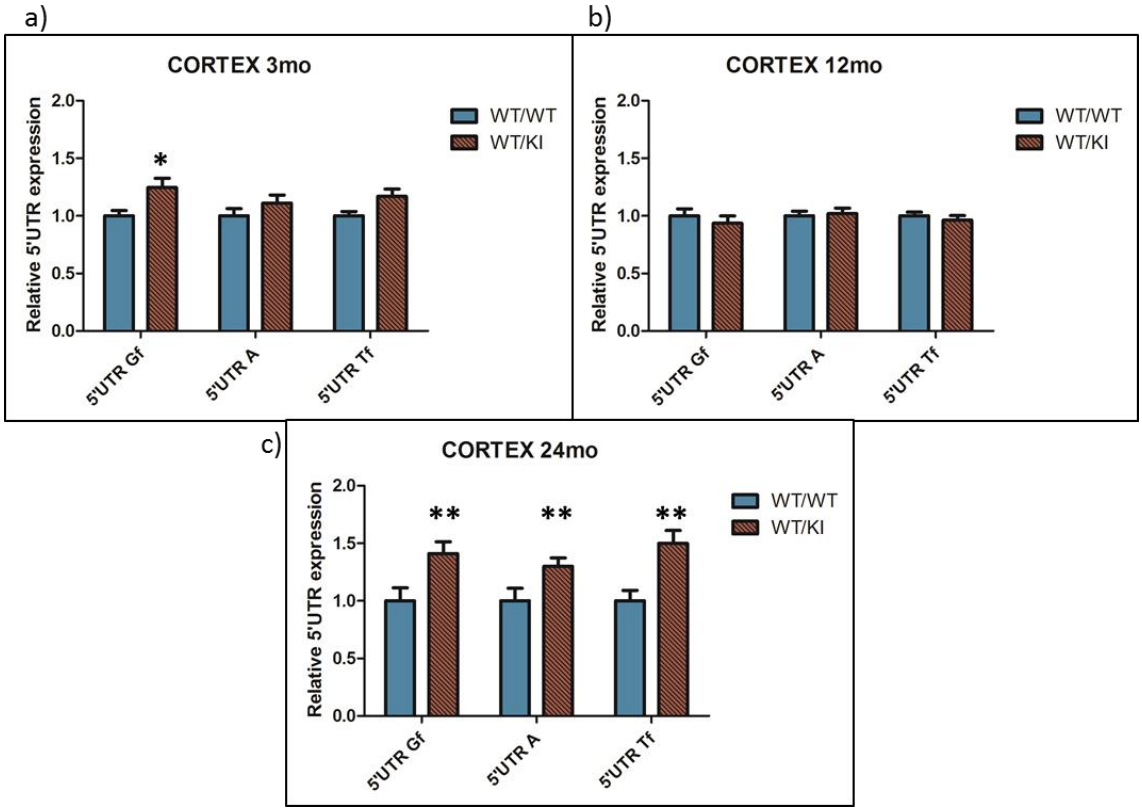


Figure 21. Full length L1 mRNA levels are increased in the cortex of 3 months and 24 months old WT/KI mice. Relative quantification of 5'UTR L1 mRNA levels in the cortex of adult mice. 3mo, 3 months of age; 12mo, 12 months of age; 24mo, 24 months of age. 5'UTR Gf, 5'UTR A, 5'UTR Tf refer to full length Gf, A and Tf L1 subfamilies present in the mouse genome. Relative 5'UTR L1 mRNA levels are normalized with thw housekeeping gene UbC. At 3mo of age, relative L1-5'UTR Gf expression is significantly higher in WT/KI (dark red) mice as compared to controls (WT/WT, light blue). At 24 mo of age, L1 5'UTR mRNA levels of all three L1 subfamilies are consistently increased in WT/KI mice relative to WT/WT mice. 3 independent qPCR replicas were performed on each biological sample. Each bar represents average value of twelve independent biological replicas (n=12). Error bars indicate standard error. *P<0.05, **P<0.01 resulting from *Mann Whitney* unpaired test.

1.3. Endogenous L1 expression profile during brain development

High levels of L1 mRNAs, leading in some cases to retrotransposition events, have been extensively described during the early stages of embryo development

(Fadloun et al., 2013; Kano et al., 2009; Vitullo et al., 2012). Moreover, neural precursor cells (NPCs) have been demonstrated to be able to support L1 retrotransposition of an engineered L1 reporter construct (Muotri et al., 2005). However little is known about the physiological L1 transcription dynamics during embryonal neurogenesis and in the adult differentiated brain.

Here, in addition to compare L1 transcript levels between WT and HD mice, the analysis of L1 expression in four late embryonal stages (E10, E12, E14, E17), at P0 and in adult mice (3mo, 12mo, 24mo), allowed us to follow the transcription dynamics of full length L1 elements during development. By plotting the average value of relative L1 mRNA levels at each developmental stage, we described the expression profiles of each L1 subfamily in the developmental window going from embryonal stage E10 until 24 mo of age, in striatum and cortex.

L1 expression profiles in striatum

In the striatum, we evidenced a wave of L1 transcription during late embryonal development. As shown in figure 22, in both WT/WT and WT/KI mice, L1 expression only slowly increased during the first phases of striatal differentiation between E10 and E12, but then this trend significantly increased between E12 and E14 up to E17 when L1 transcript levels reached their maximum during embryogenesis. After E17, L1 transcription decreased until P0 and this slowdown was more marked in WT/KI mice. After birth, between P0 and 3mo no significant changes were observed, but from 3 mo to 12 mo, surprisingly, a burst of L1 transcription, severely impaired in WT/KI mice, was detected in WT/WT mice. At later stages, L1 transcripts returned to basal levels in both WT/WT and WT/KI mice.

Interestingly, for all the three L1 subfamilies we described almost overlapping L1 expression profiles (figure 22 a,b,c).

In conclusion, these findings show that in the striatum, excluding the stage of 12mo, L1 expression levels and profiles are similar between WT/WT and WT/KI mice and L1s undergo a wave of transcription during late embryonal development.

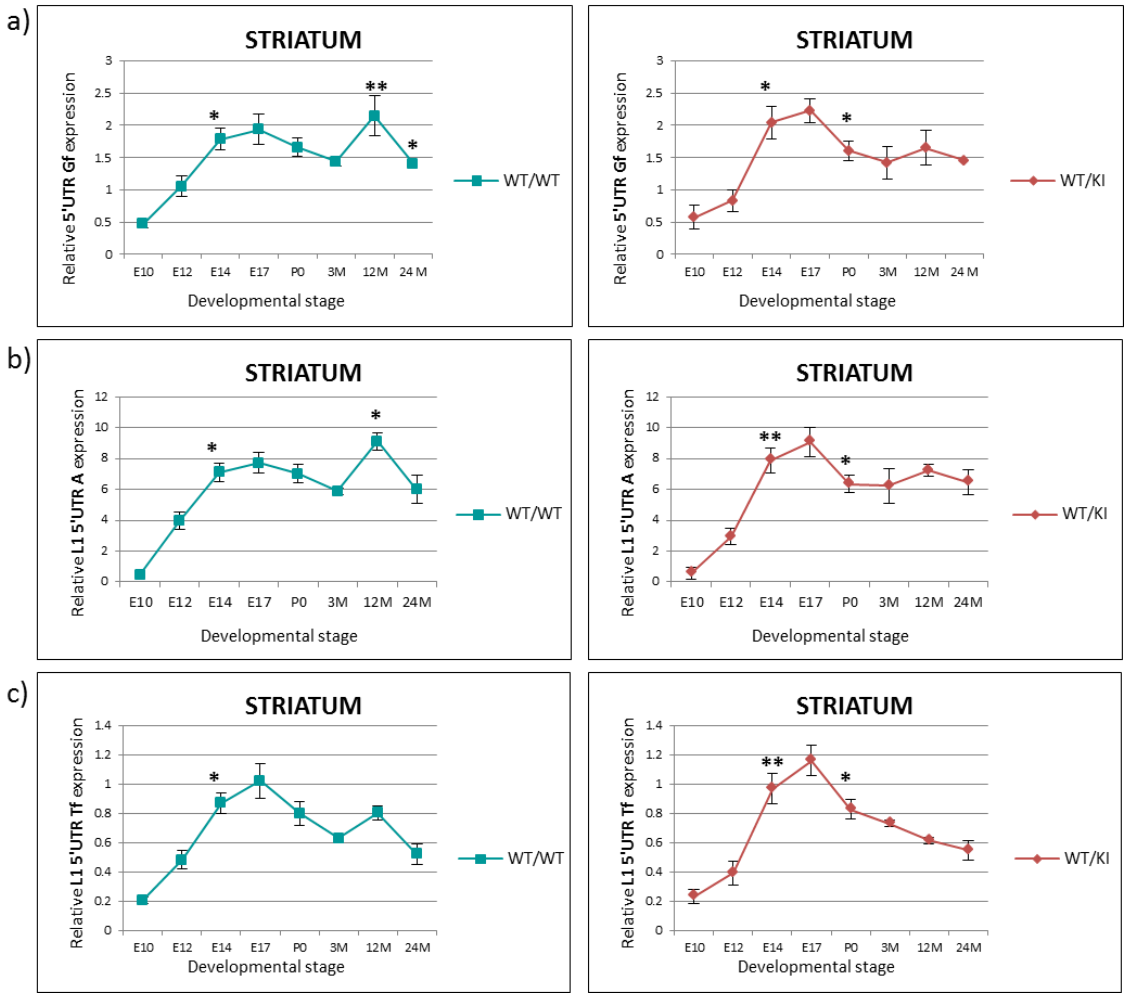


Figure 22. Endogenous L1 expression profiles during development of striatum. a) Expression analysis of full length L1 *Gf* elements in WT/WT (left, light blue) and WT/KI (right, dark red) mice. **b)** Expression analysis of full length L1 *A* elements **c)** Expression analysis of full length L1 *Tf* elements. E10, embryonal day 10; E12, embryonal day 12; E14, embryonal day 14; E17, embryonal day 17; P0, date of birth; 3M, 3 months, 12M, 12months; 24M, 24months. Relative quantification of L1 5'UTR mRNA were normalized with the housekeeping gene UbC. Each dot represents average value of 6 independent biological replicas. Error bars represent standard error. *P<0.05, **P<0.01 resulting from *Mann Whitney* unpaired test. Statistical significance is relative to the previous developmental stage.

L1 expression profiles in cerebral cortex

In cerebral cortex, similarly to the striatum, L1 elements appeared to be actively transcribed and underwent a wave of transcription during embryonal development.

In embryonal stages, as shown in figure 23, we could evidence marked similarities in L1 expression profiles of WT/WT and WT/KI mice. In detail, L1 expression increased from E10 to E14, especially between E10 and E12. Then, at E14, the peak of L1 expression during neuronal embryogenesis was reached in both WT/WT and WT/KI mice. Until P0, L1 mRNA levels progressively decreased.

Later on, at post-natal stages, in WT/WT mice (figure 23, left panels), L1 transcription increased from P0 to 3mo at statistically significant levels for 5'A and 5'Tf elements. Between 3mo and 12mo, mRNA levels remained stable for 5'Gf, but decreased for 5'A and increased for 5'Tf subtype. Then, along with senescence (from 12 mo to 24 mo) L1 mRNA expression of all three subfamilies substantially decreased.

On the other hand, in adult WT/KI mouse brains (figure 23, right panels), L1 expression profiles showed a marked, statistically significant, increase of L1 elements between P0 and 3mo and this was consistent in all three L1 5'UTR subfamilies. However, with time (from 3 mo to 24 mo) L1 activation tended to decrease.

As previously described, in post-natal stages of WT/KI cortices, apart from 12mo, L1 mRNA levels remained higher when compared to WT/WT mice (see also figure 21).

Noteworthy, also in the case of cerebral cortex, we observed similar expression profiles for Gf, A and Tf L1 elements. This was true particularly during embryonal development (figure 21 a,b,c).

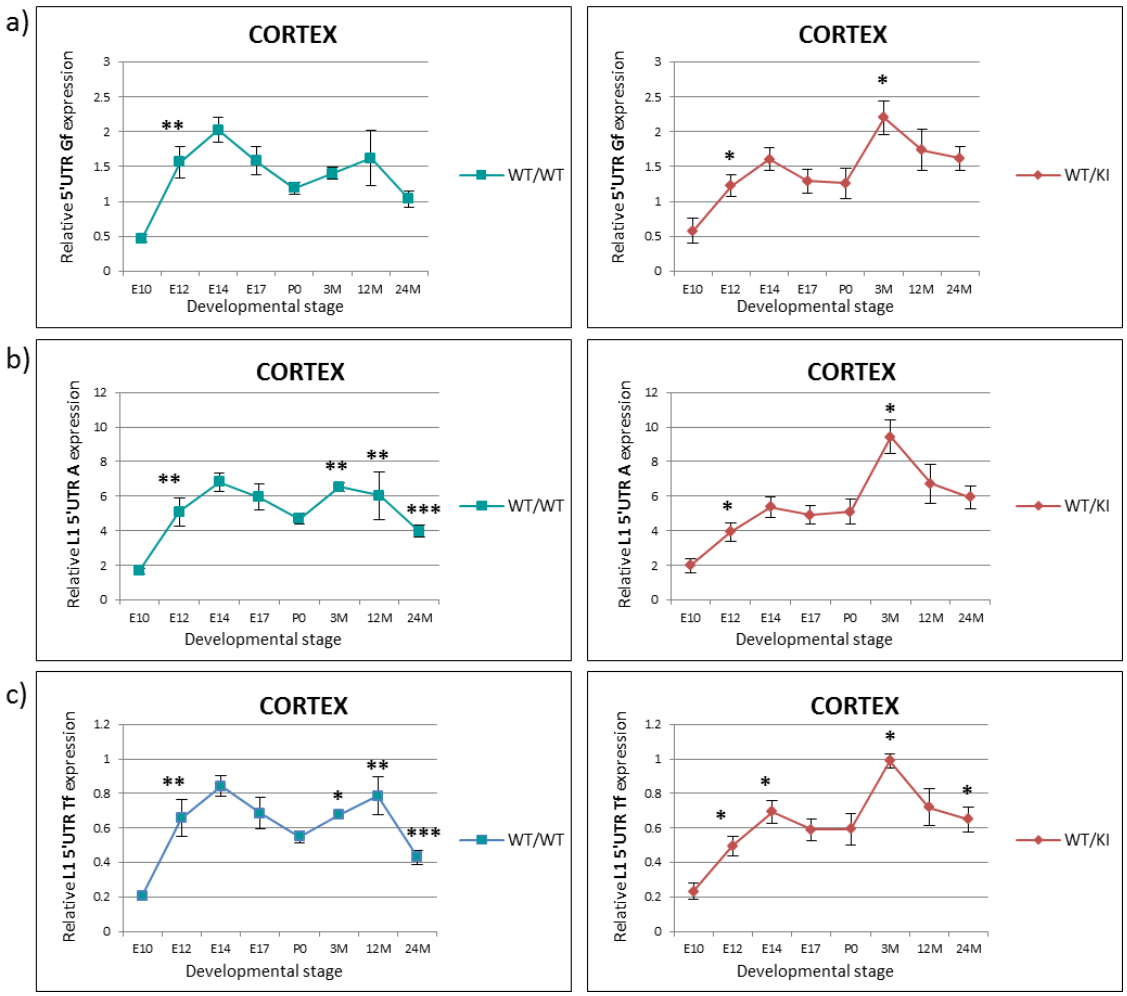


Figure 23. Endogenous L1 expression profiles during development of cerebral cortex. a) Expression analysis of full length L1 *Gf* elements in WT/WT (left, light blue) and WT/KI (right, dark red) mice. **b)** Expression analysis of full length L1 *A* elements. **c)** Expression analysis of full length L1 *Tf* elements. E10, embryonal day 10; E12, embryonal day 12; E14, embryonal day 14; E17, embryonal day 17; P0, date of birth; 3M, 3 months, 12M, 12months; 24M, 24months. Relative quantification of L1 5'UTR mRNA were normalized the housekeeping gene UbC. Each dot represents average value of 6 independent biological replica. Error bars represent standard error. *P<0.05, **P<0.01, ***P<0.001 resulting from *Mann Whitney* unpaired test. Statistical significance is relative to the previous developmental stage.

1.4. ORF2 protein expression in adult mouse brain

Full length L1 mRNA is a bicistronic transcript that encodes for ORF1 and ORF2 proteins (ORF1p and ORF2p). ORF1p is an RNA binding protein with chaperone activity (Martin, 2006), whereas ORF2p has both endonuclease and reverse transcriptase activities (Ostertag and Kazazian, 2001). Together these two proteins bind to L1 mRNAs or to non-autonomous retrotransposons mRNAs to form a ribonucleoprotein complex that mediates reverse transcription (TPRT) and genomic re-integration of the retrotransposon.

We next decided to investigate whether the L1 mRNAs endogenously transcribed in the adult brain can lead to detectable levels of ORF2p and, if so, whether this is a common feature among striatal and cortical neurons or if it is restricted to a subset of cells. To this purpose, using immunofluorescence analysis, we interrogated the expression and localization of ORF2 protein in striatum and cortex of WT/WT and WT/KI adult mice (3mo, 12mo and 24mo).

As shown in figure 24, we couldn't detect any ORF2p staining in the striatum of WT/WT or WT/KI mice at 3mo, 12mo and 24mo of age. This might suggest that the protein is not expressed in this tissue or that levels of expression are under the detection threshold.

However, we could evidence the presence of ORF2p-positive cells in some regions of the cerebral cortex. Expression levels were overall low. ORF2p-expressing neurons localized particularly in outer cortical layers of the pre-frontal cortex, in those areas corresponding to primary motor cortex (highlighted with a red square in the representative bright-field image of mouse brain coronal section on the top of figure 24).

The staining for ORF2p in prefrontal cortex cells display prevalently a cytoplasmic signal even though some ORF2 positivity was detected in few nuclei (figure 24). ORF2p appears to be expressed in a subset of cells of the region analyzed.

Similar patterns of ORF2p expression were observed in 3mo, 12 mo and 24mo cortices, in both WT/WT and WT/KI mice.

In summary, this analysis revealed that the endogenous ORF2p protein can be detected in the pre-frontal cortex of adult mice. Noteworthy, ORF2p expression is not a common feature of pre-frontal cortical neurons but is limited to a small subset of them.

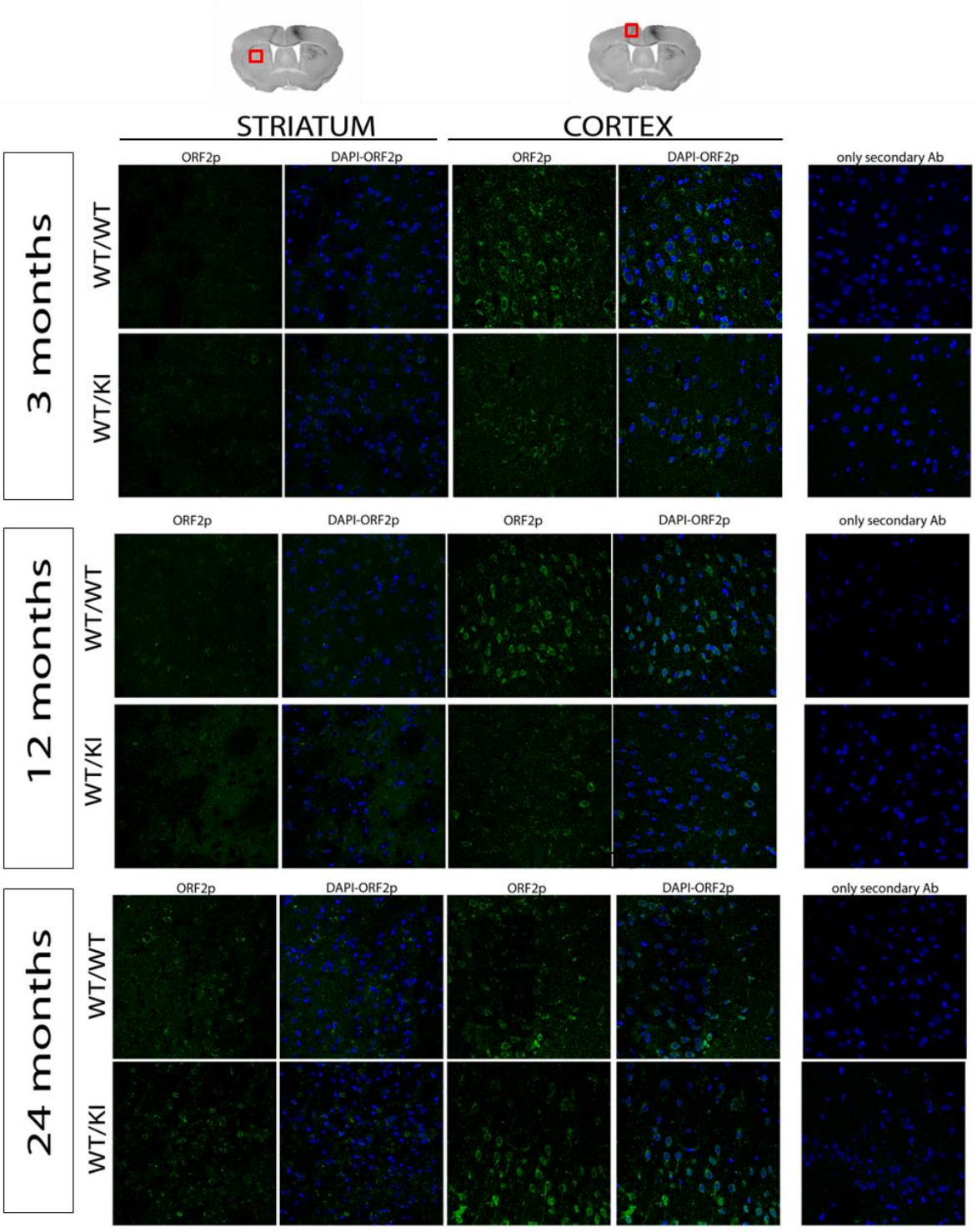


Figure 24. ORF2 protein expression in adult mouse brains Representative figure of L1 ORF2p immunostaining in coronal brain sections (0.3 mm from Bregma) of 3months, 12 months, 24 months old WT/WT and WT/KI mice (n=4). Nuclear staining with DAPI is shown in blue. L1ORF2p signal is shown in green. Immunostaining using only secondary antibody is used as negative control.

2. Study of regulatory pathways of L1 retrotransposition in HD mouse brains

2.1. Analysis of L1 transcriptional mechanisms

During evolution, cells developed surveillance mechanisms to prevent L1 transposition, acting both at transcriptional and post-transcriptional levels. The control of L1 mRNA at a transcription level can occur mainly through 1) binding of transcription factors; 2) DNA methylation at the CpG rich region on L1 5'UTR, mediated mainly by de novo methyltransferases in germline (Bourch'his et al.,2004), and by methyl-CpG-binding protein 2 (MeCP2) in somatic tissues (Muotri et al. 2010); 3) epigenetic modifications on histone tails able to remodel chromatin status on L1 promoter. In particular, L1s are constitutively repressed by H3K9me3 in somatic cells, but their expression can be modulated by loss or acquisition of diverse repressive or active histone marks (i.e. H3K4me3), that finally lead to enhanced L1 activation or repression, as observed especially during spermatogenesis (Fadloun et al.2013, (Pezic et al., 2014) Rangasamy et al.,2013).

In HD, several epigenetic alterations have been reported, including alterations of methylation on histone tails and aberrant interaction between mutant huntingtin and MeCP2, leading to transcriptional dysregulation of target genes (McFarland et al.,2014, Ryu et al.,2006, Vashishtha et al 2013, Biagioli et al.,2015).

Therefore, we asked whether the reduced L1 expression detected in the striatum of 12mo WT/KI mice could be associated to some of the epigenetic alterations described in HD brains. To this purpose, we interrogated the methylation status of histone tails on L1 promoter together with the level of MeCP2 association to L1 5'UTR.

H3K4me3, H3K9me3, H3K27me3 association to L1-5'UTR

Elevated levels of trimethylation at H3K9 (H3K9me3) are detected in HD striatum (Ryu et al., 2006). Moreover, mutant huntingtin enhances polycomb repressive complex 2 (PRC2) activity by direct interaction, which, in turn, affects the deposition of H3K27me3 and H3K4me3 on bivalent loci of genes crucial for neuronal identity and survival (Biagioli et al., 2015; Seong et al., 2010; von Schimmelmann et al., 2016). Therefore, as a first step, we checked whether these three histone modifications (H3K9me3, H3K4me3 and H3K27me3), known to be altered in HD, might account for

the reduced L1 expression in 12mo HD striatum. To this purpose, using qPCR, we compared the levels of L1 5'UTR enrichment in WT/WT and WT/KI mice after chromatin immunoprecipitation with specific antibodies for each histone mark.

For each chromatin immunoprecipitation experiment, we first assessed the level of enrichment obtained for the immunoprecipitated (IP) sample compared to the relative IgG. According to the literature (Chen-Plotkin et al., 2006), if the enrichment for IP versus IgG was lower than 1.5, samples were discarded. After that, to overcome the huge variability detected between replicas, and to compare the levels of L1 5'UTR enrichment in WT/WT and WT/KI mice, we normalized the IP/input fraction obtained for L1 5'UTR on that obtained for a gene desert region not associated to H3K9me3, H3K27me3 and H3K4me3. With this approach, we can adjust the background level derived from aspecific immunoprecipitated DNA.

As shown in figure 25, from ChIP analysis on H3K4me3 histone mark, typically associated with active chromatin, we obtained acceptable levels of IP enrichment as compared to relative IgG for all the striatal samples (4 WT/WT and 4 WT/KI) (figure 25a). However, when we compared the L1 5'UTR enrichment between striatum of 12 mo WT/WT and WT/KI mice, we could not appreciate any significant quantitative difference (figure 25b). Scatter plots show the result of four independent biological replicas and suggest the presence of a high level of variability between replicas in both WT/WT and WT/KI mice, even after the normalization to the gene desert region.

Similar results were obtained from the analysis of H3K9me3 and H3K27me3 repressive histone marks. In particular, for H3K9me3 we discarded samples nr. 2 and nr. 3 (corresponding to 1 WT/WT and 1 WT/KI mice) because of unsuccessful immunoprecipitation compared to the relative IgG (figure 26a). The L1 5'UTR relative enrichment analysis revealed a high interindividual variability, particularly in WT/KI samples. Overall, after immunoprecipitation, we detected comparable levels of L1 5'UTR enrichment in WT/WT and WT/KI mice (figure 26b).

Chromatin immunoprecipitations using H3K27me3, indicator of silenced chromatin, showed high levels of IP versus IgG enrichment in all the 8 samples analyzed, suggesting high efficiencies of chromatin immunoprecipitation (figure 27a). However, as well as for H3K4me3 and H3K9me3 ChIPs, we could not detect quantitative differences of L1 5'UTR association levels to H3K27me3 in the striatum of WT/WT compared to WT/KI mice (figure 27b).

Overall, from our chIP analyses on histone marks, we were not able to show altered methylation status of histone tails on L1 5'UTR that could explain the reduced L1 expression in the striatum of 12mo HD mice.

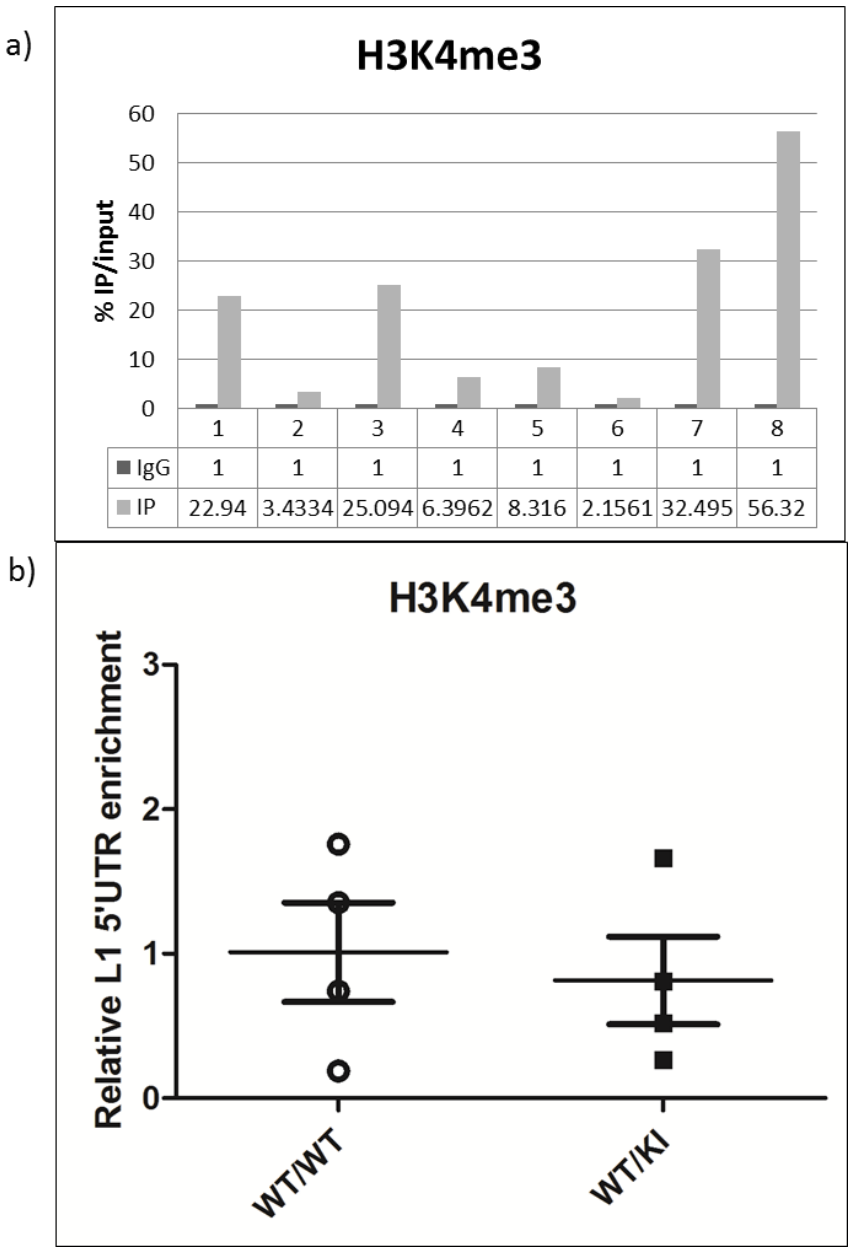
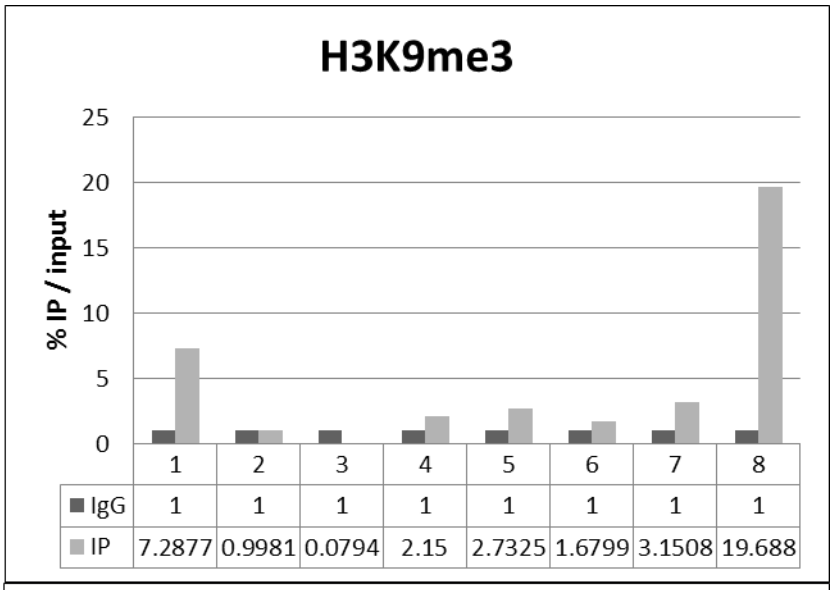


Figure 25. H3K4me3 association to L1-5'UTR. a) ChIP raw data for L1-5'UTR Gf in IP samples (light grey) compared to respective IgG (dark grey). The data are calculated as % of IP/input and numeric values of each replica are indicated in the bottom table. X-axis values label the number of biological replica. **b)** Enrichment of L1-5'UTR Gf after immunoprecipitation with H3K4me3 antibody in the striatum of 12 months old WT/WT compared to WT/KI mice. Values for each IP sample were first adjusted to respective INPUT sample. Then adjusted 5'UTR values were normalized on those of a gene desert sequence used as negative internal normalizer. Main horizontal black line indicates the average value of 4 biological replicas. Error bars indicate standard deviation.

a)



	1	2	3	4	5	6	7	8
IgG	1	1	1	1	1	1	1	1
IP	7.2877	0.9981	0.0794	2.15	2.7325	1.6799	3.1508	19.688

b)

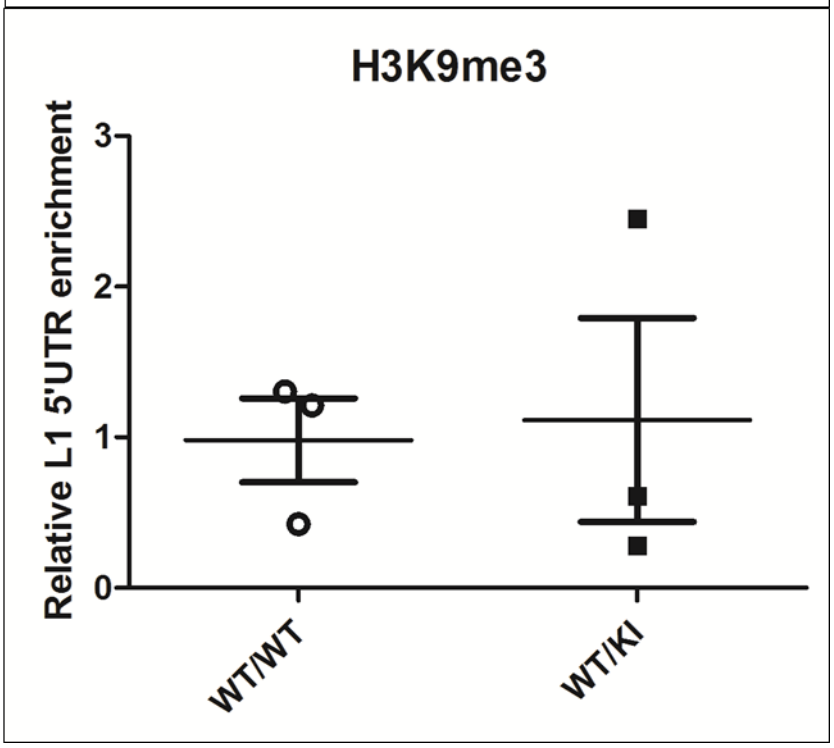


Figure 26. H3K9me3 association to L1-5'UTR. a) ChIP raw data for L1-5'UTR Gf in IP samples (light grey) compared to respective IgG (dark grey). The data are calculated as % of IP/input and numeric values of each replica are indicated in the bottom table. X-axis values label the number of biological replica. **b)** Enrichment of L1-5'UTR Gf after immunoprecipitation with H3K9me3 antibody in the striatum of 12 months old WT/WT as compared to WT/KI mice. Values for each IP sample were first adjusted on respective INPUT sample. Then adjusted 5'UTR values were normalized on those of a gene desert sequence used as negative internal normalizer. Main horizontal black line indicates the average value of 3 biological replicas. Error bars indicate standard deviation.

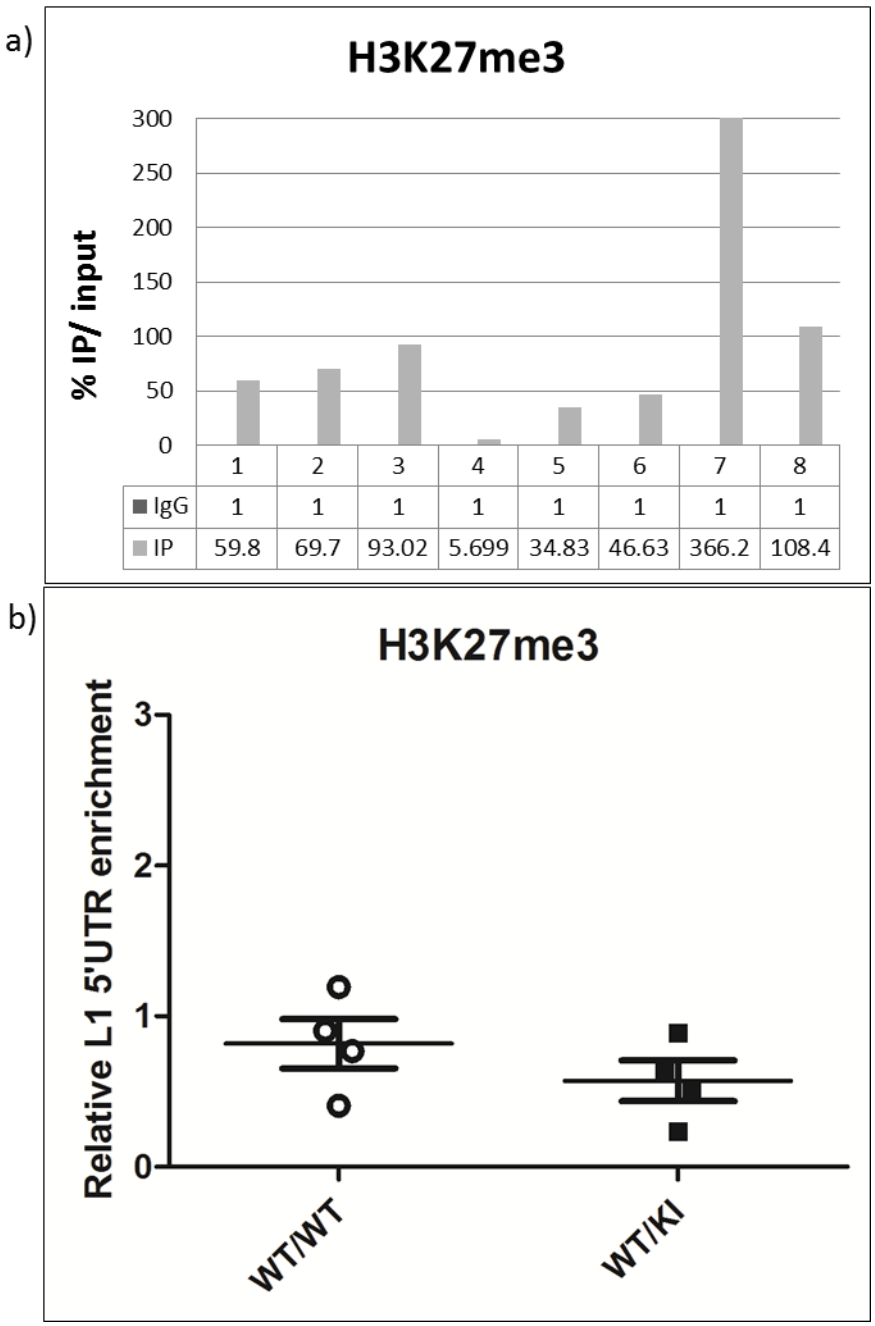


Figure 27. H3K27me3 association to L1-5'UTR. a) ChIP raw data for L1-5'UTR Gf in IP samples (light grey) compared to respective IgG (dark grey). The data are calculated as % of IP/input and numeric values of each replica are indicated in the bottom table. X-axis values label the number of biological replica. **b)** Enrichment of L1-5'UTR Gf after immunoprecipitation with H3K27me3 antibody in the striatum of 12 months old WT/WT as compared to WT/KI mice. Values for each IP sample were first adjusted on respective INPUT sample. Then adjusted 5'UTR values were normalized on those of a gene desert sequence used as negative internal normalizer. Main horizontal black line indicates the average value of 4 biological replicas. Error bars indicate standard deviation.

MeCP2 association to L1-5'UTR

In addition to the study of histone modifications on L1 promoter, using a similar approach, we interrogated MeCP2 binding to L1 5'UTR in the striatum of 12mo mice. MeCP2 promotes L1 silencing by binding to L1 5'UTR (Muotri et al., 2010) and it is able to interact with huntingtin in a polyQ-dependent fashion with a consequent increase in MeCP2 silencing on target genes (McFarland et al., 2014).

Therefore, we performed ChIP-qPCR analysis in order to assess the association levels of MeCP2 to 5'UTR L1 sequences in the striatum of 12 mo HD mice compared to controls.

As shown in figure 28, similarly to what observed with the histone marks, a high degree of variability between replicas was detected with MeCP2 ChIP experiments. We obtained reliable enrichments for L1 5'UTR sequences in 6 out of 8 IP samples, so that we discarded samples nr. 1 and nr. 5 in the final analysis (figure 28a).

When we compared the 5'UTR enrichment levels of MeCP2-precipitated chromatin fractions between WT/WT and WT/KI mice, we observed a slight increase in the striatum of WT/KI mice, but this difference was not statistically significant (figure 28b), suggesting that additional replicas are needed to overcome the interindividual variability.

Taken together these data suggest that HD-specific epigenetic alterations exerted by H3K4me3, H3K9me3, H3K27me3 and MeCP2 do not directly affect L1 5'UTR transcription in the striatum of 12 mo Hdh $^{Q7/Q111}$ mice.

However, in all these ChIP experiments we highlighted huge degrees of inter-replica variability, suggesting that ChIP followed by L1 5'UTR qPCR probably cannot represent an enough sensitive technique for the detection of small differences of highly abundant genomic repeats with such ubiquitous chromatin binding proteins.

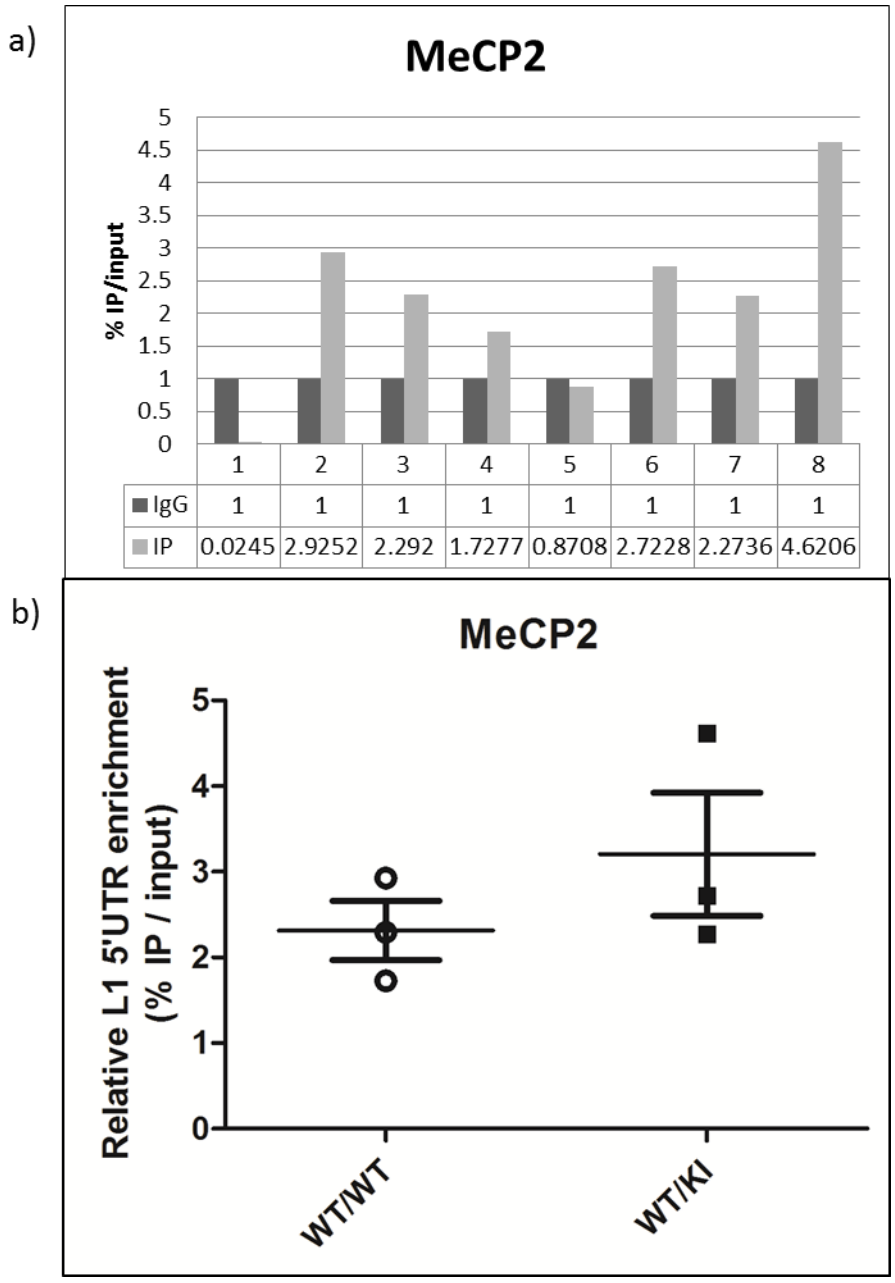


Figure 28. MeCP2 binding to L1-5'UTR. a) ChIP raw data for L1-5'UTR Gf in IP samples (light grey) compared to respective IgG (dark grey). The data are calculated as % of IP/input and numeric values of each replica are indicated in the bottom table. X-axis values label the number of the biological replica. **b)** Enrichment of L1-5'UTR Gf after immunoprecipitation with MeCP2 antibody in the striatum of 12 months old WT/WT compared to WT/KI mice. Results are calculated as % of IP/input. Main horizontal black line indicates the average value of 3 biological replicas. Error bars indicate standard deviation.

2.2. Analysis of L1 post-transcriptional mechanisms

The piRNA pathway is one of the main sources of post-transcriptional L1 silencing. First discovered in *Drosophila melanogaster*, this pathway is active primarily in the germline. MILI and MIWI2 are the murine homologs of Drosophila PIWI proteins and represent the two main factors of piRNA biogenesis, playing crucial roles in promoting de novo DNA methylation on L1 promoters (Kuramochi-Miyagawa et al. 2008, Aravin et al. 2008). Very recent reports identified and proposed potential activity of piRNA pathway also in the mouse brain (Ghosheh et al., 2016, Nandi et al. 2016), suggesting the existence of a conserved mechanism that regulates retrotransposon silencing among eukaryotes. In line and simultaneously to these two works, we interrogated the potential involvement of the piRNA pathway in the regulation of L1 expression in the adult brain of our mouse model. To this purpose, we analyzed the endogenous expression of MIWI2 and MILI in the striatum and the cortex of adult WT/WT and WT/KI mice.

MILI expression in adult mouse brain

First, using RT-PCR, we assessed the presence of MILI transcript in the cortex and striatum of adult WT/WT and WT/KI mice, using testis sample as a positive control. As shown in figure 29a, appreciable levels of endogenous MILI mRNA were detected in both nervous tissues analyzed, in WT/WT as well as in WT/KI mice and in all the three adult ages that we considered.

Accordingly, MILI protein expression was confirmed by western blot analysis in both striatum and cortex of 3mo and 12 mo mice. As expected, MILI protein levels in brain tissues were much more reduced as compared to those expressed in testis (figure 29b).

Additionally, by immunohistochemistry on coronal brain sections of 12 mo mice, we detected MILI positive neurons in cerebral cortex and striatum with a prevalent cytoplasmic staining (figure 29c), consistent with what published by Nandi et al., 2016.

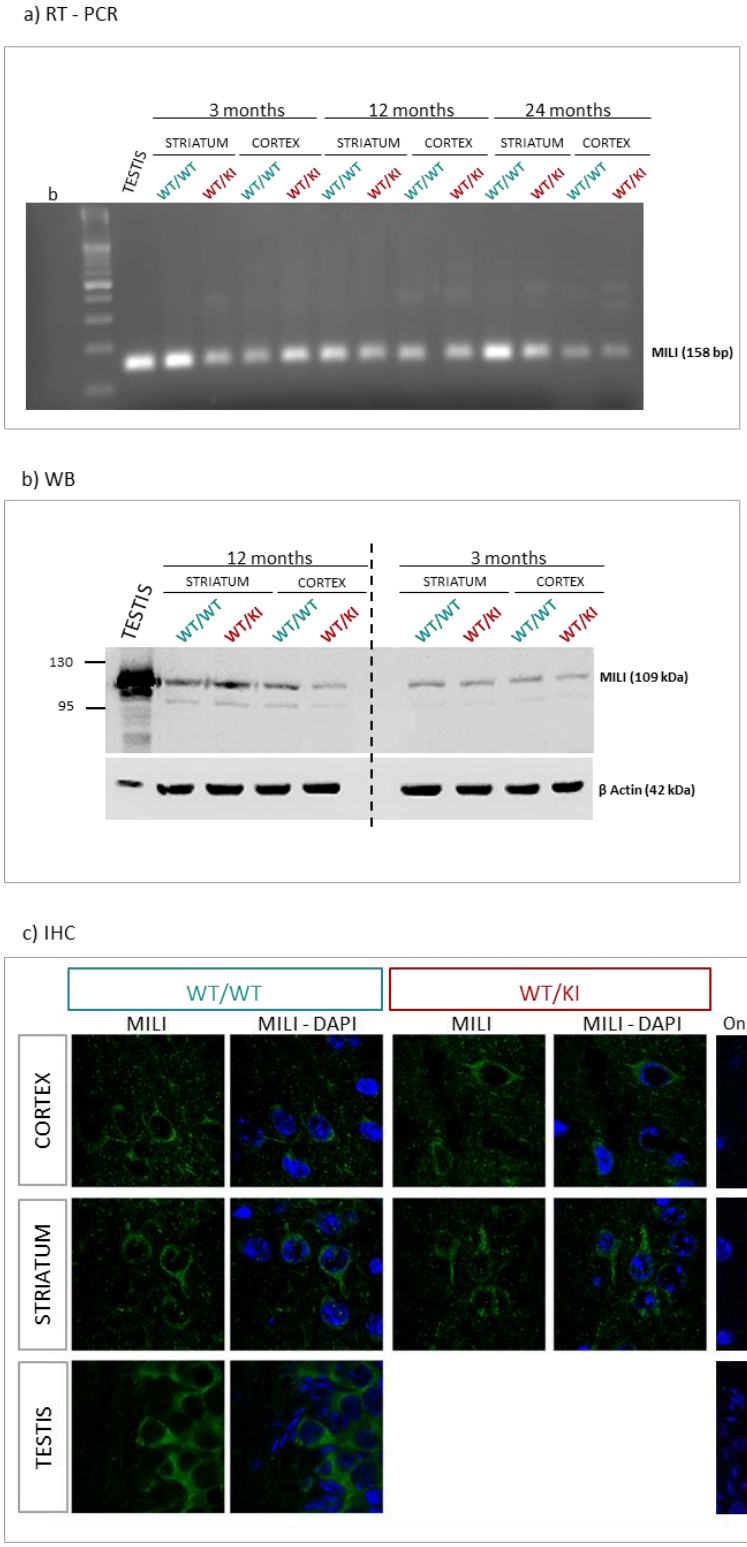


Figure 29. MILI expression in striatum and cortex of adult mouse brain. a) RT-PCR of MILI on cDNAs obtained from striatum, cortex and testis samples of adult WT/WT and WT/KI mice. Testis is used as positive control sample. **b)** Representative western blot showing MILI protein expression in striatum and cortex of 3 months and 12 months old mice (n=4). β-actin is used as loading control. **c)** Immunofluorescence staining of MILI protein in testis (positive control) and 16uM coronal brain sections of 12mo mice. Nuclear staining with DAPI is shown in blue. MILI staining is shown in green. Immunostaining using only secondary antibody is used as negative control.

Then, we asked if any correlation between MILI expression and L1 mRNA reduction could be evidenced in the striatum of HD mice at 12mo of age. To this aim, we performed qPCR experiments to assess MILI transcript levels in WT/KI relative to WT/WT mice.

Unfortunately, as shown in figure 30, qPCR analysis did not reveal significant (p-value=0.068) quantitative differences in MILI expression between WT/KI and WT/WT mice.

Nonetheless we must consider that in this study we focused on MILI, thus giving no information about the types of piRNAs expressed in HD striatum, that represent the direct regulators of L1 expression. Therefore, it would be interesting to explore whether selected piRNA subsets are differentially expressed in WT/WT compared to WT/KI mice.

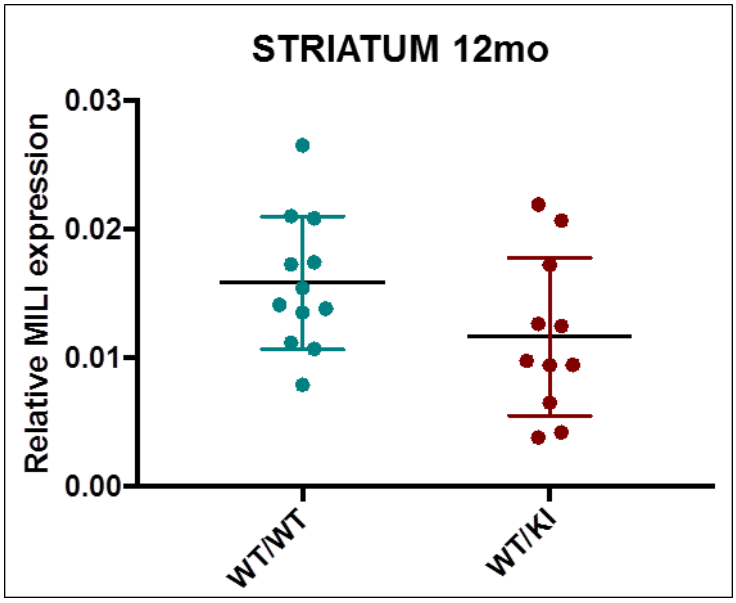


Figure 30. MILI expression in striatum of 12mo WT/WT as compared to WT/KI mice. RT-qPCR analysis revealed no significant differences in MILI transcript levels between WT/WT and WT/KI mice. MILI transcript levels are normalized with housekeeping gene UbC. Horizontal black lines indicate average value of 12 independent biological replica (n=12). Error bars display standard deviation.

MIWI2 expression in adult mouse brain

At the same time, we also tested the expression of the second player of the piRNA pathway: MIWI2. To this purpose, we assessed MIWI2 expression by RT-PCR and WB analyses in adult brain tissues.

As expected from the literature, we detected high levels of MIWI2 in testis but undetectable levels of MIWI2 expression in the two nervous tissues analyzed (cortex and striatum). This was confirmed by both RT-PCR and Western blot analyses (figure 31).

Therefore, taken together, these data provide an additional independent evidence of MILI expression and MIWI2 absence of expression in the adult mouse brain. This result supports the idea of a functional role of MILI not only in the germline but also in mammalian adult brain and consequently suggests a function of MILI associated piRNAs outside germ cells, as proposed by Nandi and colleagues.

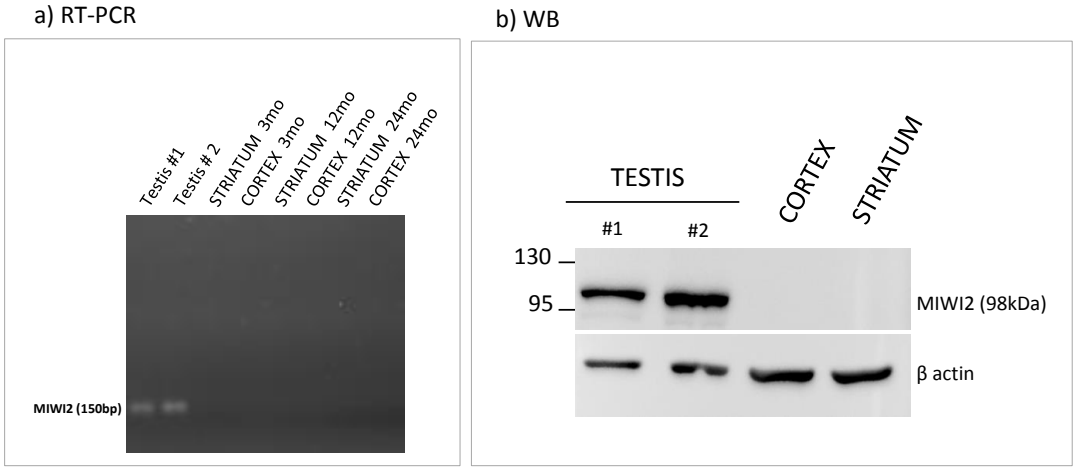


Figure 31. MIWI2 expression in striatum and cortex of adult mice. a) RT-PCR of MIWI2 on cDNAs obtained from striatum, cortex and testis samples of adult WT/WT mice. Testes are used as positive control samples **b)** Representative Western Blot showing MIWI2 protein expression in testis samples but not in striatum and cortex of 12 months old mice (n=3). β-actin is used as loading control.

DISCUSSION

"Insertional mutagens" and "regulatory elements", with these words geneticist Barbara McClintock first described a class of "mutant loci" that she discovered be responsible for altered kernels' pigmentation in maize. After her seminal work, these "mutant loci" were identified as transposable elements (TEs), a class of repetitive genomic sequences able to mobilize from one site to another within the genome. Owing to their mobilization capabilities, TEs accumulated over time in almost all eukaryotic genomes and equipped host genomes with useful tools of plasticity and evolutionary adaptation. Overall, TEs compose almost 40% of human and mouse genomes.

L1s are far and away the most active TEs in mammalian genomes. They compose 17% and 19% of human and mouse genomes, respectively, and mobilize via the L1 retrotransposition process, a "copy and paste" mechanism where a L1-RNA intermediate is first "copied" and then reverse-transcribed and "pasted" in a new genomic location. L1 retrotransposons can mediate mobilization of their own transcripts and of other non-L1 retrotransposons mRNAs, thus significantly affecting genomes' structure and function. By promoting new retrotransposons insertions, L1s can lead to genomic rearrangements and remodeling of transcriptional units. At the same time, L1 mRNA and L1-encoded proteins, even without resulting in a new genomic transposon insertion, can perturb cellular functions by providing cells with additional sources of endonucleases, reverse transcriptases and potentially non-coding RNAs. Consequently, to prevent potential deleterious effects of L1 abnormal activity, cells have developed several mechanisms to safeguard and fine-tune L1 retrotransposition. At the end, the balance between evolutionary advantages of L1 mobilization and L1 regulation generates L1 permissive environments fundamental for organismal biology. In particular, germ cells experience the highest levels of L1 activity due to their flexible methylation signatures and therefore giving the opportunity to new favorable L1 insertions to be passed on future generations. Similarly, the first phases of embryonal development seem to be particularly tolerant to L1 retrotransposition. Interestingly, in the last decade, an increasing number of evidence is showing that also somatic cells of the adult brain can support L1 mobilization. L1 retrotransposition events have been extensively described during neurogenesis and recently in terminally differentiated neurons. Single cell genomic studies estimated that at least 0.6 L1 retrotransposition events per neuron occur in adult human hippocampus and frontal cortex. This feature is

evolutionary conserved from *Drosophila* to mammals and represent a major source of somatic mosaicism in adult central nervous system, suggesting its crucial role for normal brain physiology.

Conversely, unregulated activation of L1s associates with neuropathology, as observed in Rett syndrome, Ataxia telengectasia and schizophrenia.

Considering these recent findings, we asked whether L1 retrotransposition could be altered in Huntington's disease (HD) and, in turn, if L1 retrotransposition could affect HD pathogenesis.

HD is a late-onset, autosomal dominant disorder caused by a CAG expansion in the IT15 gene that leads to the production of a mutated toxic form of huntingtin protein. Since the discovery of the gene responsible for the pathology, an impressive amount of data on the molecular basis of HD neurodegeneration accumulated, however relatively little is known about the genetic and epigenetic factors that could be crucial in modulating disease onset and pathology progression. For instance, the basis of somatic CAG expansion leading to a consequent anticipation of symptoms remains largely unclear. On the other hand, the presence in HD brains of extensive epigenetic alterations, transcriptional dysregulation, neuronal specification impairments and DNA damage could have important effects on several cellular processes, including L1 retrotransposition.

Here, using a novel qPCR *Taqman* assay we assessed the expression of full length active murine L1 elements during mouse brain development (from E_{10} to 24 mo) in a genetically precise mouse model of HD. With these analyses, we showed that HD mice display an altered expression of L1 mRNAs. Additionally, we observed that the endogenous transcription of L1 elements can lead to the expression of ORF2 protein in a particular subset of neurons of the adult frontal cortex, both in WT and HD conditions. Moreover, we highlighted that L1 elements undergo a wave of transcription during cortical and striatal embryonic development which is then maintained at basal levels after birth. Finally, we provide additional evidence of a functional piRNA pathway in the adult mammalian brain.

1. L1 expression dysregulations in HD mouse brain

In this study, taking advantage of Hdh $^{Q7/Q111}$ knock-in (WT/KI) mice, for the first time we assessed L1 retrotransposition in the cerebral cortex and striatum of a precise genetic model of HD. In order to monitor the entire disease progression, we performed L1 retrotransposition characterization in the adult mouse as well as in embryonal stages, at both genomic and transcriptional levels.

From a genomic point of view, by performing L1 copy number variation analysis using a qPCR-based approach targeting L1-ORF2 sequences, we were not able to evidence differences in genomic L1 content between HD and control mice. This was true both in striatum and cortex and in all the developmental stages analyzed. This result was very informative, but at the same time, as already proposed by others (Goodier 2016), it might be biased by the fact that qPCR analyses on bulk tissues likely underestimate L1 insertion events that are private for a small subset of cells within a tissue. Therefore, to finely characterize the genomic content of L1 elements in a tissue and to accurately compare it between different samples, it would be useful to perform relative L1 copy number quantification after the sorting of a selected number of cells, a procedure that is becoming a gold standard in this field after its first application (Bundo et al.,2014).

Nonetheless, even if the genomic L1 content seems to be equivalent in HD and WT mice, we decided to study endogenous L1 expression in HD adult mouse brains. To this purpose, we first performed a relative quantification analysis of full length L1 transcripts, distinguishing between the three active murine L1 subfamilies, in striatum and cortex of HD mice as compared to controls. Then we assessed the endogenous expression of ORF2 protein that is autonomously encoded by full length L1 elements. These experiments highlighted L1 expression alterations in HD brains, affecting both cortex and striatum.

In the cerebral cortex, we showed that L1 Gf elements are overexpressed in 3 mo old HD mice and that all three L1 subfamilies are consistently overexpressed at 24 mo of age, while no alterations are detected at 12mo of age, in late embryos and at P0. Moreover, in both HD and control adult mice, from the age of 3mo up to 24 mo, endogenous levels of ORF2p were detected by immunohistochemistry in a small subset of neurons of the pre-frontal cortex. ORF2p staining was mostly cytoplasmatic, as expected from the literature (Gualtieri et al., 2013). Although we did not reveal significant differences in ORF2p staining between WT and HD mice, the detection of

an endogenous signal for ORF2p in a non-pathological tissue represented an important finding, considering the great efforts generally required to detect this protein (Doucet et al., 2010). Therefore, we might conclude that in a subpopulation of neurons of the pre-frontal cortex, endogenous L1 mRNAs are translated to produce appreciable levels of ORF2p. This would be in agreement with the neuronal L1 retrotransposition that has been reported in somatic cells of the adult frontal cortex in mammalian brains (Erwin et al., 2016).

On the other hand, the overexpression of L1 elements detected at 3 mo and 24 mo in mutant huntingtin-expressing brains, might represent a novel pathological feature of HD, arising already at pre-symptomatic stages (3mo). The mechanisms underlying these alterations remain largely unclear and could be related to aberrant transcriptional or epigenetic programs, previously described in HD cortical neurons (McKinstry et al., 2014, Gauthier et al., 2004).

Concerning very old 24 mo mice, we speculate that HD conditions, in the cortex, might somehow exacerbate the process of L1 activation normally occurring in aging brains (Van Meter et al.,2014, Li et al.,2013), in turn rendering HD-affected cortical neurons more susceptible to neurodegeneration.

From the same analysis performed in the striatum of HD mice, we could not detect appreciable levels of ORF2p expression. Additionally, no differences have been evidenced in L1 expression between WT and HD mice during embryonal development, at P0, at 3mo and at 24 mo. However, at 12mo, we interestingly observed a significant decrease in L1 transcript levels in the striatum of HD mice that was consistent in all the three murine L1 subfamilies. Notably, this finding represented a strong evidence of L1 transcriptional dysregulation at the age and in the brain region where the major HD-pathological phenotypes accumulate. Therefore, we then asked whether L1 repression in the striatum of 12mo HD mice could relate to altered L1 epigenetic regulation. Noteworthy, at 12mo of age, several molecular alterations triggered by mutant huntingtin have been described in the striatum of Hdh $^{Q7/Q111}$ mice (Wheeler et al., 2002). These observations include altered methylation of DNA and of histone tails. Therefore, we assessed the levels of L1-5'UTR enrichment after chromatin immunoprecipitation with different chromatin remodelers known to be associated with L1 promoter and to be altered in HD pathology. We specifically chose H3K9me3, H3K27me3, H3K4me3 histone marks and MeCP2 for our analysis. Unfortunately, none of these regulators could be linked to altered L1 expression in HD striatum, suggesting

that other factors could underlie this phenomenon or that the technique that we employed was not sensitive enough in our experimental conditions.

We then decided to investigate the potential involvement of the piRNA pathway in the regulation of L1 expression in HD striatum. Indeed, it has been recently demonstrated that piRNAs could be functional in adult mouse brains (Nandi et al., 2016) and abnormal miRNA expression and miRNA biogenesis have been reported in HD brains (Johnson et al., 2008; Packer et al., 2008, Lee and Kim, 2011). To this aim, we interrogated in our samples the expression of MILI, the protein able to bind piRNAs and target them to transposons mRNAs for their silencing. Surprisingly, we detected appreciable levels of MILI expression in adult brains but it did not correlate with L1 deregulation in the striatum of 12mo HD mice. We have to consider however that, by addressing MILI expression, we do not obtain any information about piRNAs expression. Therefore, it would be interesting to assess whether piRNAs are differentially expressed in HD mice as compared to WT. This in turn could explain altered regulation of L1 mRNAs in HD brains.

An alternative mechanism of modulation of retrotransposons' activity in a context of HD, particularly in the striatum, could be mediated by PRC2, whose activity is known to be influenced by huntingtin protein in a polyQ-dependent fashion (Seong et al.,2010). In fact, PRC2 plays a crucial role in suppressing transcriptional programs that otherwise would be detrimental for striatal neurons' function and survival (VonSchimmelmann et al.,2016). Moreover, PRC2 can repress expansion of parasitic DNA, like endogenous retroviruses (Leeb et al., 2010).

In alternative to the "aberrant regulatory" hypothesis, to explain L1 downregulation in 12mo HD striatum, we can hypothesize a selective neuronal loss of striatal neurons expressing L1 transcripts. In this case, L1 expression could represent a susceptibility factor for enhanced neurodegeneration in HD striatum. Notably, at 12 months of age, this phenomenon would be accompanied by extensive DNA damage (Wheeler et al., 2002).

In summary, these findings provide the first description of L1 retrotransposition in a defined genetic mouse model of HD. Interestingly, our analysis shows that altered expression of L1 elements occurs in both striatum and cerebral cortex of HD mice. In cortex, L1 overexpression appears both at early and late stages, whereas in the striatum, L1 downregulation occurs later, when many pathological HD phenotypes are manifest. In both cases, L1 retrotransposition might represent a previously unrecognized

modulator of HD age at onset and progression. However further studies are needed to better understand if L1 activity is driver or bystander of HD neuropathogenesis.

2. Endogenous L1 expression dynamics during striatal and cortical development

The study of L1 expression in striatum and cortex of E10, E12, E14, E17, P0, 3mo, 12mo and 24 mo old mice, allowed us to describe the expression profiles of full length L1 elements in WT as well as HD mice. With these experiments, we provided the first complete characterization of endogenous L1 expression in striatum and cortex during late embryonal development (between E10 and P0) and the entire adulthood of murine life. Expression dynamics were outlined for all the active L1 subfamilies present in the mouse genome: Gf, A and Tf L1s.

We interestingly observed that, during late embryonal development, L1 mRNA of all the three L1 subtypes are actively transcribed in both cortex and striatum. L1 elements undergo a wave of transcription between E10 and P0, reaching their highest levels between E14 and E17 in both tissues. Notably, L1 transcription dynamics are almost overlapping in WT and HD mice, thus highlighting a common feature of mouse cortical and striatal embryogenesis.

Interestingly, we noticed an extraordinary parallelism between L1 expression profiles and the wave of neurogenesis that takes place during differentiation of embryonal telencephalon, the brain structure comprising cerebral cortex and striatum (Götz and Huttner, 2005). When the highest number of newly born neurons is reached, the bulk of L1 transcription is registered. On the other hand, as neurons in excess start to physiologically die, a reduction of L1 expression follows. These findings seem to be in agreement with the idea that L1 overexpression associates with neuronal death. Therefore, L1 activation seems to be a mechanism with a dual role, strictly related to the context in which it takes place: in some cases it can take part to developmentally crucial functions, as described in embryonal neurogenesis and fetal oocyte attrition (Malki et al., 2014); in other cases, it can be toxic for the cell, potentially leading to an increased susceptibility to neurodegeneration, as hypothesized for the striatum of 12mo HD mice.

After birth, L1 expression profiles show that L1 mRNAs are generally maintained at basal levels, although, from 3mo to 12 mo, we unexpectedly observed an increase in L1 transcription in physiological conditions in the striatum. This event is consistent for the three L1 subfamilies. At present, its biological significance remains unclear.

By contrast, the analysis of adult cerebral cortex revealed less conserved expression profiles between L1 Gf, A and Tf subfamilies. We think that this could reflect the neuronal heterogeneity typical of the cerebral cortex that is not present in other brain regions, like striatum or cerebellum.

In summary, this study allowed us to provide a complete depiction of endogenous L1 expression dynamics during the development of central nervous system, with particular attention to striatal and cortical brain regions. With this approach, we interestingly unveiled a wave of L1 transcription mirroring embryonal telencephalic neurogenesis.

3. MILI expression in adult mouse brain

In view of recent findings showing piRNAs expression in adult mammalian brain (Nandi et al.,2016; (Ghosheh et al., 2016)) and considering previous RNA-seq data obtained in our group showing expression of MILI in the striatum of 12mo mice, we begun to explore the expression of piRNA biogenesis factors in our brain samples. From this analysis, consistent with the literature, we did not find evidence of MIWI2 expression in the brain of adult mice. However, as expected, the presence of MILI in adult striatum and cerebral cortex was confirmed by RT-PCR, Western Blot and immunohistochemistry. In the mouse brain, MILI seems to be expressed throughout the entire adulthood, thus indicating a common feature of adult mouse brain, not restricted to particular developmental stages. Furthermore, by histological studies, we showed that MILI is expressed mainly outside the nucleus, therefore suggesting that in the brain MILI probably participate in the load of cellular piRNAs or other endogenous siRNAs to promote post-transcriptional silencing of target mRNAs. Alternatively, we can expect that in the brain, similarly to the germline, MILI might give rise to the “ping-pong” amplification cycle. Since the initiation of this cycle relies on the presence of retrotransposon mRNAs, we suppose that neuronal retrotransposons transcripts could be the source of piRNAs that are then loaded and processed by MILI to exploit their silencing function. However, due to the lack of MIWI2 expression, MILI could have taken over the role of MIWI2 in its capacity to silence retrotransposons in mouse adult brains.

Our results, together with data published this year by Ghosheh et al., 2016 and Nandi et al., represent three independent evidence of expression of piRNAs and piRNA biogenesis factors in adult mouse brains, therefore robustly supporting the idea of a

functional role of piRNAs not confined to the germline. It will be interesting to investigate the expression of retrotransposons in adult brains of MILI null mice. With this experiment, we might understand whether the function of piRNA pathway in the adult brain is mainly a defense against retrotransposon propagation, as in the germline, or if new roles developed during evolution. We can for instance suppose that piRNAs in the brain could contribute to neuronal plasticity by acting on the cellular epigenome.

Overall, if we think to the waves of retrotransposition that occurred during evolution, it is not surprisingly that factors involved in L1 mobilization in germ cells are related to factors acting in the brain. Indeed, it is thanks to the inter-generational transmission of advantageous insertional events, that L1 retrotransposition was passed to future generations, even promoting in some cases sexual reproduction success (Barsoum et al., 2010). Therefore, mechanisms that originally allowed retrotransposition in the germline should be similar to those existing in somatic cells. In line with this model, *in silico* data reported parallelisms between human brain and testes gene expression patterns (Guo et al., 2003, 2005) and it has been proposed that similarities in expression in brain and testes might play important roles for human speciation (Wilda et al., 2000).

Altogether, our findings give additional evidence of active retrotransposition in adult mammalian brains, therefore strongly supporting the idea of L1 retrotransposons as sources of somatic mosaicism in this tissue. However, in this context many open questions need to be addressed. We know that eukaryotic genomes underwent important waves of retrotransposition during evolution and they have never lost their intrinsic patrimony of transposable elements, therefore displaying an inextricable link between L1s and host genomes. Accordingly, germline retrotransposition crucially contributed to the transmission of advantageous insertional events. But what is the meaning of L1 retrotransposition in somatic cells? Which are the functions and mechanisms of somatic L1 insertions? Do genomes strictly need so many copies of the same element and why? Does L1 retrotransposition interest other somatic tissues, apart from brain and cancer? These are just some of the many questions that need to be addressed in the next future to have a complete overview of such an outstanding phenomenon, so powerful in creating somatic genomic variability within cells of the some tissue but at the same time having important implications in pathological processes.

.

BIBLIOGRAPHY

A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group., 1993. . Cell 72, 971–983.

Adey, N.B., Schichman, S.A., Graham, D.K., Peterson, S.N., Edgell, M.H., Hutchison, C.A., 1994. Rodent L1 evolution has been driven by a single dominant lineage that has repeatedly acquired new transcriptional regulatory sequences. Mol Biol Evol 11, 778–789.

Agostoni, E., Michelazzi, S., Maurutto, M., Carnemolla, A., Ciani, Y., Vatta, P., Roncaglia, P., Zucchelli, S., Leanza, G., Mantovani, F., Gustincich, S., Santoro, C., Piazza, S., Del Sal, G., Persichetti, F., 2016. Effects of Pin1 Loss in Hdh(Q111) Knock-in Mice. Front Cell Neurosci 10, 110. doi:10.3389/fncel.2016.00110

Akagi, K., Li, J., Stephens, R.M., Volfovsky, N., Symer, D.E., 2008. Extensive variation between inbred mouse strains due to endogenous L1 retrotransposition. Genome Res 18, 869–880. doi:10.1101/gr.075770.107

Akagi, K., Stephens, R.M., Li, J., Evdokimov, E., Kuehn, M.R., Volfovsky, N., Symer, D.E., 2010. MouseIndelDB: a database integrating genomic indel polymorphisms that distinguish mouse strains. Nucleic Acids Res 38, D600–6. doi:10.1093/nar/gkp1046

Andrew, S.E., Goldberg, Y.P., Kremer, B., Telenius, H., Theilmann, J., Adam, S., Starr, E., Squitieri, F., Lin, B., Kalchman, M.A., 1993. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. Nat Genet 4, 398–403. doi:10.1038/ng0893-398

Anne, S.L., Saudou, F., Humbert, S., 2007. Phosphorylation of huntingtin by cyclin-dependent kinase 5 is induced by DNA damage and regulates wild-type and mutant huntingtin toxicity in neurons. J Neurosci 27, 7318–7328. doi:10.1523/JNEUROSCI.1831-07.2007

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., Hannon, G.J., 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell 31, 785–799. doi:10.1016/j.molcel.2008.09.003

Arrasate, M., Mitra, S., Schweitzer, E.S., Segal, M.R., Finkbeiner, S., 2004. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. Nature 431, 805–810. doi:10.1038/nature02998

Aziz, N.A., van der Marck, M.A., Pijl, H., Olde Rikkert, M.G.M., Bloem, B.R., Roos, R.A.C., 2008. Weight loss in neurodegenerative disorders. J Neurol 255, 1872–1880. doi:10.1007/s00415-009-0062-8

Babushok, D.V., Ostertag, E.M., Courtney, C.E., Choi, J.M., Kazazian, H.H., 2006. L1 integration in a transgenic mouse model. Genome Res 16, 240–250. doi:10.1101/gr.4571606

Bachoud-Lévi, A.C., Maison, P., Bartolomeo, P., Boissé, M.F., Dalla Barba, G., Ergis, A.M., Baudic, S., Degos, J.D., Cesaro, P., Peschanski, M., 2001. Retest effects and cognitive decline in longitudinal follow-up of patients with early HD. Neurology 56, 1052–1058.

Baillie, J.K., Barnett, M.W., Upton, K.R., Gerhardt, D.J., Richmond, T.A., De Sapio, F., Brennan, P.M., Rizzu, P., Smith, S., Fell, M., Talbot, R.T., Gustincich, S., Freeman, T.C., Mattick, J.S., Hume, D.A., Heutink, P., Carninci, P., Jeddeloh, J.A., Faulkner, G.J., 2011. Somatic retrotransposition alters the genetic landscape of the human brain. Nature 479, 534–537. doi:10.1038/nature10531

Barsoum, E., Martinez, P., Aström, S.U., 2010. Alpha3, a transposable element that promotes host sexual reproduction. Genes Dev 24, 33–44. doi:10.1101/gad.557310

Becher, M.W., Kotzuk, J.A., Sharp, A.H., Davies, S.W., Bates, G.P., Price, D.L., Ross, C.A., 1998. Intranuclear neuronal inclusions in Huntington's disease and dentatorubral and pallidoluysian atrophy: correlation between the density of inclusions and IT15 CAG triplet repeat length. Neurobiol Dis 4, 387–397. doi:10.1006/nbdi.1998.0168

Beck, C.R., Garcia-Perez, J.L., Badge, R.M., Moran, J.V., 2011. LINE-1 elements in structural variation and disease. Annu Rev Genomics Hum Genet 12, 187–215. doi:10.1146/annurev-genom-082509-141802

Belancio, V.P., Hedges, D.J., Deininger, P., 2006. LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic Acids Res 34, 1512–1521. doi:10.1093/nar/gkl027

Belancio, V.P., Hedges, D.J., Deininger, P., 2008. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome Res 18, 343–358. doi:10.1101/gr.5558208

Belancio, V.P., Roy-Engel, A.M., Pochampally, R.R., Deininger, P., 2010. Somatic expression of LINE-1 elements in human tissues. Nucleic Acids Res 38, 3909–3922. doi:10.1093/nar/gkq132

Belgnaoui, S.M., Gosden, R.G., Semmes, O.J., Haoudi, A., 2006. Human LINE-1 retrotransposon induces DNA damage and apoptosis in cancer cells. Cancer Cell Int 6, 13. doi:10.1186/1475-2867-6-13

Benn, C.L., Landles, C., Li, H., Strand, A.D., Woodman, B., Sathasivam, K., Li, S.-H., Ghazi-Noori, S., Hockly, E., Faruque, S.M.N.N., Cha, J.-H.J., Sharpe, P.T., Olson, J.M., Li, X.-J., Bates, G.P., 2005. Contribution of nuclear and extranuclear polyQ to neurological phenotypes in mouse models of Huntington's disease. Hum Mol Genet 14, 3065–3078. doi:10.1093/hmg/ddi340

Benn, C.L., Sun, T., Sadri-Vakili, G., McFarland, K.N., DiRocco, D.P., Yohrling, G.J., Clark, T.W., Bouzou, B., Cha, J.-H.J., 2008. Huntingtin modulates transcription, occupies gene promoters in vivo, and binds directly to DNA in a polyglutamine-dependent manner. J Neurosci 28, 10720–10733. doi:10.1523/JNEUROSCI.2126-08.2008

Beraldi, R., Pittoggi, C., Sciamanna, I., Mattei, E., Spadafora, C., 2006. Expression of LINE-1 retroposons is essential for murine preimplantation development. Mol Reprod Dev 73, 279–287. doi:10.1002/mrd.20423

Biagioli, M., Ferrari, F., Mendenhall, E.M., Zhang, Y., Erdin, S., Vijayvargia, R., Vallabh, S.M., Solomos, N., Manavalan, P., Ragavendran, A., Ozsolak, F., Lee, J.M., Talkowski, M.E., Gusella, J.F., Macdonald, M.E., Park, P.J., Seong, I.S., 2015. Htt CAG repeat expansion confers pleiotropic gains of mutant huntingtin function in chromatin regulation. Hum Mol Genet 24, 2442–2457. doi:10.1093/hmg/ddv006

Bilen, J., Liu, N., Burnett, B.G., Pittman, R.N., Bonini, N.M., 2006. MicroRNA pathways modulate polyglutamine-induced neurodegeneration. Mol Cell 24, 157–163. doi:10.1016/j.molcel.2006.07.030

Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O'Shea, K.S., Moran, J.V., Cullen, B.R., 2006. Cellular inhibitors of long interspersed element 1 and

Alu retrotransposition. Proc Natl Acad Sci U S A 103, 8780–8785. doi:10.1073/pnas.0603313103

Bourc'his, D., Bestor, T.H., 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature 431, 96–99. doi:10.1038/nature02886

Boyle, A.L., Ballard, S.G., Ward, D.C., 1990. Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization. Proc Natl Acad Sci U S A 87, 7757–7761.

Branciforte, D., Martin, S.L., 1994. Developmental and cell type specificity of LINE-1 expression in mouse testis: implications for transposition. Mol Cell Biol 14, 2584–2592.

Bundo, M., Toyoshima, M., Okada, Y., Akamatsu, W., Ueda, J., Nemoto-Miyauchi, T., Sunaga, F., Toritsuka, M., Ikawa, D., Kakita, A., Kato, M., Kasai, K., Kishimoto, T., Nawa, H., Okano, H., Yoshikawa, T., Kato, T., Iwamoto, K., 2014. Increased l1 retrotransposition in the neuronal genome in schizophrenia. Neuron 81, 306–313. doi:10.1016/j.neuron.2013.10.053

Busch, A., Engemann, S., Lurz, R., Okazawa, H., Lehrach, H., Wanker, E.E., 2003. Mutant huntingtin promotes the fibrillogenesis of wild-type huntingtin: a potential mechanism for loss of huntingtin function in Huntington's disease. J Biol Chem 278, 41452–41461. doi:10.1074/jbc.M303354200

Carnemolla, A., Fossale, E., Agostoni, E., Michelazzi, S., Calligaris, R., De Maso, L., Del Sal, G., MacDonald, M.E., Persichetti, F., 2009. Rrs1 is involved in endoplasmic reticulum stress response in Huntington disease. J Biol Chem 284, 18167–18173. doi:10.1074/jbc.M109.018325

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., Forrest, A.R.R., Carninci, P., Biffo, S., Stupka, E., Gustincich, S., 2012. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature 491, 454–457. doi:10.1038/nature11508

Castro-Diaz, N., Ecco, G., Coluccio, A., Kapopoulou, A., Yazdanpanah, B., Friedli, M., Duc, J., Jang, S.M., Turelli, P., Trono, D., 2014. Evolutionally dynamic L1 regulation in embryonic stem cells. Genes Dev 28, 1397–1409. doi:10.1101/gad.241661.114

Cattaneo, E., Rigamonti, D., Goffredo, D., Zuccato, C., Squitieri, F., Sipione, S., 2001. Loss of normal huntingtin function: new developments in Huntington's disease research. Trends Neurosci 24, 182–188.

Caviston, J.P., Ross, J.L., Antony, S.M., Tokito, M., Holzbaur, E.L.F., 2007. Huntingtin facilitates dynein/dynactin-mediated vesicle transport. Proc Natl Acad Sci U S A 104, 10045–10050. doi:10.1073/pnas.0610628104

Chen-Plotkin, A.S., Sadri-Vakili, G., Yohrling, G.J., Braveman, M.W., Benn, C.L., Glajch, K.E., DiRocco, D.P., Farrell, L.A., Krainc, D., Gines, S., MacDonald, M.E., Cha, J.-H.J., 2006. Decreased association of the transcription factor Sp1 with genes downregulated in Huntington's disease. Neurobiol Dis 22, 233–241. doi:10.1016/j.nbd.2005.11.001

Chow, J.C., Ciaudo, C., Fazzari, M.J., Mise, N., Servant, N., Glass, J.L., Attreed, M., Avner, P., Wutz, A., Barillot, E., Greally, J.M., Voinnet, O., Heard, E., 2010. LINE-1 activity in facultative heterochromatin formation during X chromosome inactivation. Cell 141, 956–969. doi:10.1016/j.cell.2010.04.042

Cong, S.-Y., Pepers, B.A., Evert, B.O., Rubinsztein, D.C., Roos, R.A.C., van Ommen, G.-J.B., Dorsman, J.C., 2005. Mutant huntingtin represses CBP, but not p300, by binding and protein degradation. Mol Cell Neurosci 30, 560–571.

Cordaux, R., Batzer, M.A., 2009. The impact of retrotransposons on human genome evolution. Nat Rev Genet 10, 691–703. doi:10.1038/nrg2640

Cost, G.J., Feng, Q., Jacquier, A., Boeke, J.D., 2002. Human L1 element target-primed reverse transcription in vitro. EMBO J 21, 5899–5910.

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Marchetto, M.C.N., Muotri, A.R., Mu, Y., Carson, C.T., Macia, A., Moran, J.V., Gage, F.H., 2011. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proc Natl Acad Sci U S A 108, 20382–20387. doi:10.1073/pnas.1100273108

Coufal, N.G., Garcia-Perez, J.L., Peng, G.E., Yeo, G.W., Mu, Y., Lovci, M.T., Morell, M., O'Shea, K.S., Moran, J.V., Gage, F.H., 2009. L1 retrotransposition in human neural progenitor cells. Nature 460, 1127–1131. doi:10.1038/nature08248

Crichton, J.H., Dunican, D.S., Maclennan, M., Meehan, R.R., Adams, I.R., 2014. Defending the genome from the enemy within: mechanisms of retrotransposon

suppression in the mouse germline. Cell Mol Life Sci 71, 1581–1605. doi:10.1007/s00018-013-1468-0

Cummings, C.J., Mancini, M.A., Antalffy, B., DeFranco, D.B., Orr, H.T., Zoghbi, H.Y., 1998. Chaperone suppression of aggregation and altered subcellular proteasome localization imply protein misfolding in SCA1. Nat Genet 19, 148–154. doi:10.1038/502

De Rooij, K.E., De Koning Gans, P.A., Roos, R.A., Van Ommen, G.J., Den Dunnen, J.T., 1995. Somatic expansion of the (CAG)n repeat in Huntington disease brains. Hum Genet 95, 270–274.

DeBerardinis, R.J., Kazazian, H.H., 1999. Analysis of the promoter from an expanding mouse retrotransposon subfamily. Genomics 56, 317–323. doi:10.1006/geno.1998.5729

Denli, A.M., Narvaiza, I., Kerman, B.E., Pena, M., Benner, C., Marchetto, M.C.N., Diedrich, J.K., Aslanian, A., Ma, J., Moresco, J.J., Moore, L., Hunter, T., Saghatelian, A., Gage, F.H., 2015. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell 163, 583–593. doi:10.1016/j.cell.2015.09.025

Dewannieux, M., Esnault, C., Heidmann, T., 2003. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet 35, 41–48. doi:10.1038/ng1223

Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A.J., O'Carroll, D., 2013. Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol Cell 50, 601–608. doi:10.1016/j.molcel.2013.04.026

Di Maio, L., Squitieri, F., Napolitano, G., Campanella, G., Trofatter, J.A., Conneally, P.M., 1993. Suicide risk in Huntington's disease. J Med Genet 30, 293–295.

DiFiglia, M., Sapp, E., Chase, K., Schwarz, C., Meloni, A., Young, C., Martin, E., Vonsattel, J.P., Carraway, R., Reeves, S.A., 1995. Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. Neuron 14, 1075–1081.

DiFiglia, M., Sapp, E., Chase, K.O., Davies, S.W., Bates, G.P., Vonsattel, J.P., Aronin, N., 1997. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. Science 277, 1990–1993.

Dombroski, B.A., Mathias, S.L., Nanthakumar, E., Scott, A.F., Kazazian, H.H., 1991. Isolation of an active human transposable element. Science 254, 1805–1808.

Dong, X., Tsuji, J., Labadorf, A., Roussos, P., Chen, J.-F., Myers, R.H., Akbarian, S., Weng, Z., 2015. The Role of H3K4me3 in Transcriptional Regulation Is Altered in Huntington's Disease. PLoS ONE 10, e0144398. doi:10.1371/journal.pone.0144398

Doucet, A.J., Hulme, A.E., Sahinovic, E., Kulpa, D.A., Moldovan, J.B., Kopera, H.C., Athanikar, J.N., Hasnaoui, M., Bucheton, A., Moran, J.V., Gilbert, N., 2010. Characterization of LINE-1 ribonucleoprotein particles. PLoS Genet 6. doi:10.1371/journal.pgen.1001150

Dragatsis, I., Levine, M.S., Zeitlin, S., 2000. Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. Nat Genet 26, 300–306. doi:10.1038/81593

Eickbush, T.H., 2002. Repair by retrotransposition. Nat Genet 31, 126–127. doi:10.1038/ng897

Elbarbary, R.A., Lucas, B.A., Maquat, L.E., 2016. Retrotransposons as regulators of gene expression. Science 351, aac7247. doi:10.1126/science.aac7247

Erwin, J.A., Marchetto, M.C., Gage, F.H., 2014. Mobile DNA elements in the generation of diversity and complexity in the brain. Nat Rev Neurosci 15, 497–506. doi:10.1038/nrn3730

Erwin, J.A., Paquola, A.C.M., Singer, T., Gallina, I., Novotny, M., Quayle, C., Bedrosian, T.A., Alves, F.I.A., Butcher, C.R., Herdy, J.R., Sarkar, A., Lasken, R.S., Muotri, A.R., Gage, F.H., 2016. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. Nat Neurosci 19, 1583–1591. doi:10.1038/nn.4388

Evrony, G.D., Cai, X., Lee, E., Hills, L.B., Elhosary, P.C., Lehmann, H.S., Parker, J.J., Atabay, K.D., Gilmore, E.C., Poduri, A., Park, P.J., Walsh, C.A., 2012. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell 151, 483–496. doi:10.1016/j.cell.2012.09.035

Fadloun, A., Le Gras, S., Jost, B., Ziegler-Birling, C., Takahashi, H., Gorab, E., Carninci, P., Torres-Padilla, M.-E., 2013. Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. Nat Struct Mol Biol 20, 332–338. doi:10.1038/nsmb.2495

Faulkner, G.J., Kimura, Y., Daub, C.O., Wani, S., Plessy, C., Irvine, K.M., Schroder, K., Cloonan, N., Steptoe, A.L., Lassmann, T., Waki, K., Hornig, N., Arakawa, T., Takahashi, H., Kawai, J., Forrest, A.R.R., Suzuki, H., Hayashizaki, Y.,

Hume, D.A., Orlando, V., Grimmond, S.M., Carninci, P., 2009. The regulated retrotransposon transcriptome of mammalian cells. Nat Genet 41, 563–571. doi:10.1038/ng.368

Fedoroff, N.V., 2012. McClintock's challenge in the 21st century. Proc Natl Acad Sci U S A 109, 20200–20203. doi:10.1073/pnas.1215482109

Feng, Q., Moran, J.V., Kazazian, H.H., Boeke, J.D., 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87, 905–916.

Ferrante, R.J., Kubilus, J.K., Lee, J., Ryu, H., Beesen, A., Zucker, B., Smith, K., Kowall, N.W., Ratan, R.R., Luthi-Carter, R., Hersch, S.M., 2003. Histone deacetylase inhibition by sodium butyrate chemotherapy ameliorates the neurodegenerative phenotype in Huntington's disease mice. J Neurosci 23, 9418–9427.

Fusco, F.R., Chen, Q., Lamoreaux, W.J., Figueredo-Cardenas, G., Jiao, Y., Coffman, J.A., Surmeier, D.J., Honig, M.G., Carlock, L.R., Reiner, A., 1999. Cellular localization of huntingtin in striatal and cortical neurons in rats: lack of correlation with neuronal vulnerability in Huntington's disease. J Neurosci 19, 1189–1202.

Gafni, J., Hermel, E., Young, J.E., Wellington, C.L., Hayden, M.R., Ellerby, L.M., 2004. Inhibition of calpain cleavage of huntingtin reduces toxicity: accumulation of calpain/caspase fragments in the nucleus. J Biol Chem 279, 20211–20220. doi:10.1074/jbc.M401267200

Garcia-Perez, J.L., Marchetto, M.C.N., Muotri, A.R., Coufal, N.G., Gage, F.H., O'Shea, K.S., Moran, J.V., 2007. LINE-1 retrotransposition in human embryonic stem cells. Hum Mol Genet 16, 1569–1577. doi:10.1093/hmg/ddm105

Gasior, S.L., Roy-Engel, A.M., Deininger, P.L., 2008. ERCC1/XPF limits L1 retrotransposition. DNA Repair (Amst) 7, 983–989. doi:10.1016/j.dnarep.2008.02.006

Gasior, S.L., Wakeman, T.P., Xu, B., Deininger, P.L., 2006. The human LINE-1 retrotransposon creates DNA double-strand breaks. J Mol Biol 357, 1383–1393. doi:10.1016/j.jmb.2006.01.089

Gauthier, L.R., Charrin, B.C., Borrell-Pagès, M., Dompierre, J.P., Rangone, H., Cordelières, F.P., De Mey, J., MacDonald, M.E., Lessmann, V., Humbert, S., Saudou, F., 2004. Huntingtin controls neurotrophic support and survival of

neurons by enhancing BDNF vesicular transport along microtubules. Cell 118, 127–138. doi:10.1016/j.cell.2004.06.018

Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium, 2015. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. Cell 162, 516–526. doi:10.1016/j.cell.2015.07.003

Georgiou, N., Bradshaw, J.L., Chiu, E., Tudor, A., O'Gorman, L., Phillips, J.G., 1999. Differential clinical and motor control function in a pair of monozygotic twins with Huntington's disease. Mov Disord 14, 320–325.

Ghosheh, Y., Seridi, L., Ryu, T., Takahashi, H., Orlando, V., Carninci, P., Ravasi, T., 2016. Characterization of piRNAs across postnatal development in mouse brain. Sci Rep 6, 25039. doi:10.1038/srep25039

Gil, J.M.A.C., Mohapel, P., Araújo, I.M., Popovic, N., Li, J.-Y., Brundin, P., Petersén, A., 2005. Reduced hippocampal neurogenesis in R6/2 transgenic Huntington's disease mice. Neurobiol Dis 20, 744–751. doi:10.1016/j.nbd.2005.05.006

Gilbert, N., Lutz, S., Morrish, T.A., Moran, J.V., 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. Mol Cell Biol 25, 7780–7795. doi:10.1128/MCB.25.17.7780-7795.2005

Gilbert, N., Lutz-Prigge, S., Moran, J.V., 2002. Genomic deletions created upon LINE-1 retrotransposition. Cell 110, 315–325.

Ginés, S., Bosch, M., Marco, S., Gavaldà, N., Díaz-Hernández, M., Lucas, J.J., Canals, J.M., Alberch, J., 2006. Reduced expression of the TrkB receptor in Huntington's disease mouse models and in human brain. Eur J Neurosci 23, 649–658. doi:10.1111/j.1460-9568.2006.04590.x

Gines, S., Seong, I.S., Fossale, E., Ivanova, E., Trettel, F., Gusella, J.F., Wheeler, V.C., Persichetti, F., MacDonald, M.E., 2003. Specific progressive cAMP reduction implicates energy deficit in presymptomatic Huntington's disease knock-in mice. Hum Mol Genet 12, 497–508. doi:10.1093/hmg/ddg046

Giralt, A., Puigdellívol, M., Carretón, O., Paoletti, P., Valero, J., Parra-Damas, A., Saura, C.A., Alberch, J., Ginés, S., 2012. Long-term memory deficits in Huntington's disease are associated with reduced CBP histone acetylase activity. Hum Mol Genet 21, 1203–1216. doi:10.1093/hmg/ddr552

Giuliano, P., De Cristofaro, T., Affaitati, A., Pizzulo, G.M., Feliciello, A., Criscuolo, C., De Michele, G., Filla, A., Avvedimento, E.V., Varrone, S., 2003. DNA damage

induced by polyglutamine-expanded proteins. Hum Mol Genet 12, 2301–2309. doi:10.1093/hmg/ddg242

Godin, J.D., Colombo, K., Molina-Calavita, M., Keryer, G., Zala, D., Charrin, B.C., Dietrich, P., Volvert, M.-L., Guillemot, F., Dragatsis, I., Bellaiche, Y., Saudou, F., Nguyen, L., Humbert, S., 2010. Huntingtin is required for mitotic spindle orientation and mammalian neurogenesis. Neuron 67, 392–406. doi:10.1016/j.neuron.2010.06.027

Gonitel, R., Moffitt, H., Sathasivam, K., Woodman, B., Detloff, P.J., Faull, R.L.M., Bates, G.P., 2008. DNA instability in postmitotic neurons. Proc Natl Acad Sci U S A 105, 3467–3472. doi:10.1073/pnas.0800048105

Goodier, J.L., 2016. Restricting retrotransposons: a review. Mob DNA 7, 16. doi:10.1186/s13100-016-0070-z

Goodier, J.L., Cheung, L.E., Kazazian, H.H., 2012. MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells. PLoS Genet 8, e1002941. doi:10.1371/journal.pgen.1002941

Goodier, J.L., Cheung, L.E., Kazazian, H.H., 2013. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic Acids Res 41, 7401–7419. doi:10.1093/nar/gkt512

Goodier, J.L., Ostertag, E.M., Du, K., Kazazian, H.H., 2001. A novel active L1 retrotransposon subfamily in the mouse. Genome Res 11, 1677–1685. doi:10.1101/gr.198301

Goodier, J.L., Ostertag, E.M., Kazazian, H.H., 2000. Transduction of 3'-flanking sequences is common in L1 retrotransposition. Hum Mol Genet 9, 653–657.

Goodier, J.L., Zhang, L., Vetter, M.R., Kazazian, H.H., 2007. LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. Mol Cell Biol 27, 6469–6483. doi:10.1128/MCB.00332-07

Götz, M., Huttner, W.B., 2005. The cell biology of neurogenesis. Nat Rev Mol Cell Biol 6, 777–788. doi:10.1038/nrm1739

Graham, R.K., Deng, Y., Slow, E.J., Haigh, B., Bissada, N., Lu, G., Pearson, J., Shehadeh, J., Bertram, L., Murphy, Z., Warby, S.C., Doty, C.N., Roy, S., Wellington, C.L., Leavitt, B.R., Raymond, L.A., Nicholson, D.W., Hayden, M.R., 2006. Cleavage at the caspase-6 site is required for neuronal dysfunction

and degeneration due to mutant huntingtin. Cell 125, 1179–1191. doi:10.1016/j.cell.2006.04.026

Grison, A., Mantovani, F., Comel, A., Agostoni, E., Gustincich, S., Persichetti, F., Del Sal, G., 2011. Ser46 phosphorylation and prolyl-isomerase Pin1-mediated isomerization of p53 are key events in p53-dependent apoptosis induced by mutant huntingtin. Proc Natl Acad Sci U S A 108, 17979–17984. doi:10.1073/pnas.1106198108

Gualtieri, A., Andreola, F., Sciamanna, I., Sinibaldi-Vallebona, P., Serafino, A., Spadafora, C., 2013. Increased expression and copy number amplification of LINE-1 and SINE B1 retrotransposable elements in murine mammary carcinoma progression. Oncotarget 4, 1882–1893. doi:10.18632/oncotarget.1188

Guo, J., Zhu, P., Wu, C., Yu, L., Zhao, S., Gu, X., 2003. In silico analysis indicates a similar gene expression pattern between human brain and testis. Cytogenet Genome Res 103, 58–62. doi:76290

Guo, J.H., Huang, Q., Studholme, D.J., Wu, C.Q., Zhao, Z., 2005. Transcriptomic analyses support the similarity of gene expression between brain and testis in human as well as mouse. Cytogenet Genome Res 111, 107–109. doi:10.1159/000086378

Gusella, J., MacDonald, M., 2002. No post-genetics era in human disease research. Nat Rev Genet 3, 72–79. doi:10.1038/nrg706

Gusella, J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., Sakaguchi, A.Y., 1983. A polymorphic DNA marker genetically linked to Huntington's disease. Nature 306, 234–238.

Hackam, A.S., Singaraja, R., Wellington, C.L., Metzler, M., McCutcheon, K., Zhang, T., Kalchman, M., Hayden, M.R., 1998. The influence of huntingtin protein size on nuclear localization and cellular toxicity. J Cell Biol 141, 1097–1105.

Han, J.S., Szak, S.T., Boeke, J.D., 2004. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. Nature 429, 268–274. doi:10.1038/nature02536

Han, K., Sen, S.K., Wang, J., Callinan, P.A., Lee, J., Cordaux, R., Liang, P., Batzer, M.A., 2005. Genomic rearrangements by LINE-1 insertion-mediated deletion in the human and chimpanzee lineages. Nucleic Acids Res 33, 4040–4052. doi:10.1093/nar/gki718

Hancks, D.C., Kazazian, H.H., 2012. Active human retrotransposons: variation and disease. Curr Opin Genet Dev 22, 191–203. doi:10.1016/j.gde.2012.02.006

Hancks, D.C., Kazazian, H.H., 2016. Roles for retrotransposon insertions in human disease. Mob DNA 7, 9. doi:10.1186/s13100-016-0065-9

Hedreen, J.C., Peyser, C.E., Folstein, S.E., Ross, C.A., 1991. Neuronal loss in layers V and VI of cerebral cortex in Huntington's disease. Neurosci Lett 133, 257–261.

Heras, S.R., Macias, S., Plass, M., Fernandez, N., Cano, D., Eyras, E., Garcia-Perez, J.L., Cáceres, J.F., 2013. The Microprocessor controls the activity of mammalian retrotransposons. Nat Struct Mol Biol 20, 1173–1181. doi:10.1038/nsmb.2658

Hermel, E., Gafni, J., Propp, S.S., Leavitt, B.R., Wellington, C.L., Young, J.E., Hackam, A.S., Logvinova, A.V., Peel, A.L., Chen, S.F., Hook, V., Singaraja, R., Krajewski, S., Goldsmith, P.C., Ellerby, H.M., Hayden, M.R., Bredesen, D.E., Ellerby, L.M., 2004. Specific caspase interactions and amplification are involved in selective neuronal vulnerability in Huntington's disease. Cell Death Differ 11, 424–438. doi:10.1038/sj.cdd.4401358

Hodges, A., Strand, A.D., Aragaki, A.K., Kuhn, A., Sengstag, T., Hughes, G., Elliston, L.A., Hartog, C., Goldstein, D.R., Thu, D., Hollingsworth, Z.R., Collin, F., Synek, B., Holmans, P.A., Young, A.B., Wexler, N.S., Delorenzi, M., Kooperberg, C., Augood, S.J., Faull, R.L.M., Olson, J.M., Jones, L., Luthi-Carter, R., 2006. Regional and cellular gene expression changes in human Huntington's disease brain. Hum Mol Genet 15, 965–977. doi:10.1093/hmg/ddl013

Hoffner, G., Kahlem, P., Djian, P., 2002. Perinuclear localization of huntingtin as a consequence of its binding to microtubules through an interaction with beta-tubulin: relevance to Huntington's disease. J Cell Sci 115, 941–948.

Hohjoh, H., Singer, M.F., 1996. Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. EMBO J 15, 630–639.

Hölter, S.M., Stromberg, M., Kovalenko, M., Garrett, L., Glasl, L., Lopez, E., Guide, J., Götz, A., Hans, W., Becker, L., Rathkolb, B., Rozman, J., Schrewed, A., Klingenspor, M., Klopstock, T., Schulz, H., Wolf, E., Wursta, W., Gillis, T., Wakimoto, H., Seidman, J., MacDonald, M.E., Cotman, S., Gailus-Durner, V., Fuchs, H., de Angelis, M.H., Lee, J.-M., Wheeler, V.C., 2013. A broad phenotypic screen identifies novel phenotypes driven by a single mutant allele in

Huntington's disease CAG knock-in mice. PLoS ONE 8, e80923. doi:10.1371/journal.pone.0080923

Huang, C.C., Faber, P.W., Persichetti, F., Mittal, V., Vonsattel, J.P., MacDonald, M.E., Gusella, J.F., 1998. Amyloid formation by mutant huntingtin: threshold, progressivity and recruitment of normal polyglutamine proteins. Somat Cell Mol Genet 24, 217–233.

Huang, C.R.L., Burns, K.H., Boeke, J.D., 2012. Active transposition in genomes. Annu Rev Genet 46, 651–675. doi:10.1146/annurev-genet-110711-155616

Illuzzi, J., Yerkes, S., Parekh-Olmedo, H., Kmiec, E.B., 2009. DNA breakage and induction of DNA damage response proteins precede the appearance of visible mutant huntingtin aggregates. J Neurosci Res 87, 733–747. doi:10.1002/jnr.21881

Ishiguro, H., Yamada, K., Sawada, H., Nishii, K., Ichino, N., Sawada, M., Kurosawa, Y., Matsushita, N., Kobayashi, K., Goto, J., Hashida, H., Masuda, N., Kanazawa, I., Nagatsu, T., 2001. Age-dependent and tissue-specific CAG repeat instability occurs in mouse knock-in for a mutant Huntington's disease gene. J Neurosci Res 65, 289–297. doi:10.1002/jnr.1153

Ishizu, H., Siomi, H., Siomi, M.C., 2012. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev 26, 2361–2373. doi:10.1101/gad.203786.112

Ivics, Z., Hackett, P.B., Plasterk, R.H., Izsvák, Z., 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91, 501–510.

Jachowicz, J.W., Torres-Padilla, M.-E., 2016. LINEs in mice: features, families, and potential roles in early development. Chromosoma 125, 29–39. doi:10.1007/s00412-015-0520-2

Jacobs, F.M.J., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A.D., Katzman, S., Paten, B., Salama, S.R., Haussler, D., 2014. An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. Nature 516, 242–245. doi:10.1038/nature13760

Jimenez-Sanchez, M., Licitra, F., Underwood, B.R., Rubinsztein, D.C., 2016. Huntington's Disease: Mechanisms of Pathogenesis and Therapeutic Strategies. Cold Spring Harb Perspect Med. doi:10.1101/cshperspect.a024240

Johnson, R., Zuccato, C., Belyaev, N.D., Guest, D.J., Cattaneo, E., Buckley, N.J., 2008. A microRNA-based gene dysregulation pathway in Huntington's disease. Neurobiol Dis 29, 438–445. doi:10.1016/j.nbd.2007.11.001

Kano, H., Godoy, I., Courtney, C., Vetter, M.R., Gerton, G.L., Ostertag, E.M., Kazazian, H.H., 2009. L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes Dev 23, 1303–1312. doi:10.1101/gad.1803909

Kapusta, A., Kronenberg, Z., Lynch, V.J., Zhuo, X., Ramsay, L., Bourque, G., Yandell, M., Feschotte, C., 2013. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. PLoS Genet 9, e1003470. doi:10.1371/journal.pgen.1003470

Kazazian, H.H., Wong, C., Youssoufian, H., Scott, A.F., Phillips, D.G., Antonarakis, S.E., 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332, 164–166. doi:10.1038/332164a0

Keane, T.M., Wong, K., Adams, D.J., Flint, J., Reymond, A., Yalcin, B., 2014. Identification of structural variation in mouse genomes. Front Genet 5, 192. doi:10.3389/fgene.2014.00192

Kim, M., Lee, H.S., LaForet, G., McIntyre, C., Martin, E.J., Chang, P., Kim, T.W., Williams, M., Reddy, P.H., Tagle, D., Boyce, F.M., Won, L., Heller, A., Aronin, N., DiFiglia, M., 1999. Mutant huntingtin expression in clonal striatal cells: dissociation of inclusion formation and neuronal survival by caspase inhibition. J Neurosci 19, 964–973.

Kim, M., Roh, J.-K., Yoon, B.W., Kang, L., Kim, Y.J., Aronin, N., DiFiglia, M., 2003. Huntingtin is degraded to small fragments by calpain after ischemic injury. Exp Neurol 183, 109–115.

Kim, Y.J., Yi, Y., Sapp, E., Wang, Y., Cuiffo, B., Kegel, K.B., Qin, Z.H., Aronin, N., DiFiglia, M., 2001. Caspase 3-cleaved N-terminal fragments of wild-type and mutant huntingtin are present in normal and Huntington's disease brains, associate with membranes, and undergo calpain-dependent proteolysis. Proc Natl Acad Sci U S A 98, 12784–12789. doi:10.1073/pnas.221451398

Klawitter, S., Fuchs, N.V., Upton, K.R., Muñoz-Lopez, M., Shukla, R., Wang, J., Garcia-Cañadas, M., Lopez-Ruiz, C., Gerhardt, D.J., Sebe, A., Grabundzija, I., Merkert, S., Gerdes, P., Pulgarin, J.A., Bock, A., Held, U., Witthuhn, A., Haase,

A., Sarkadi, B., Löwer, J., Wolvetang, E.J., Martin, U., Ivics, Z., Izsvák, Z., Garcia-Perez, J.L., Faulkner, G.J., Schumann, G.G., 2016. Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. Nat Commun 7, 10286. doi:10.1038/ncomms10286

Kremer, H.P., Roos, R.A., Dingjan, G.M., Bots, G.T., Bruyn, G.W., Hofman, M.A., 1991. The hypothalamic lateral tuberal nucleus and the characteristics of neuronal loss in Huntington's disease. Neurosci Lett 132, 101–104.

Kuhn, A., Goldstein, D.R., Hodges, A., Strand, A.D., Sengstag, T., Kooperberg, C., Becanovic, K., Pouladi, M.A., Sathasivam, K., Cha, J.-H.J., Hannan, A.J., Hayden, M.R., Leavitt, B.R., Dunnett, S.B., Ferrante, R.J., Albin, R., Shelbourne, P., Delorenzi, M., Augood, S.J., Faull, R.L.M., Olson, J.M., Bates, G.P., Jones, L., Luthi-Carter, R., 2007. Mutant huntingtin's effects on striatal gene expression in mice recapitulate changes observed in human Huntington's disease brain and do not differ with mutant huntingtin length or wild-type huntingtin dosage. Hum Mol Genet 16, 1845–1861. doi:10.1093/hmg/ddm133

Kulpa, D.A., Moran, J.V., 2006. Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nat Struct Mol Biol 13, 655–660. doi:10.1038/nsmb1107

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W., Hata, K., Li, E., Matsuda, Y., Kimura, T., Okabe, M., Sakaki, Y., Sasaki, H., Nakano, T., 2008. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev 22, 908–917. doi:10.1101/gad.1640708

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., et al., 2001. Initial sequencing and analysis of the human genome. Nature 409, 860–921. doi:10.1038/35057062

Lazic, S.E., Grote, H., Armstrong, R.J.E., Blakemore, C., Hannan, A.J., van Dellen, A., Barker, R.A., 2004. Decreased hippocampal cell proliferation in R6/1 Huntington's mice. Neuroreport 15, 811–813.

Leavitt, B.R., Guttman, J.A., Hodgson, J.G., Kimel, G.H., Singaraja, R., Vogl, A.W., Hayden, M.R., 2001. Wild-type huntingtin reduces the cellular toxicity of mutant huntingtin in vivo. Am J Hum Genet 68, 313–324. doi:10.1086/318207

Leavitt, B.R., van Raamsdonk, J.M., Shehadeh, J., Fernandes, H., Murphy, Z., Graham, R.K., Wellington, C.L., Raymond, L.A., Hayden, M.R., 2006. Wild-type huntingtin protects neurons from excitotoxicity. J Neurochem 96, 1121–1129. doi:10.1111/j.1471-4159.2005.03605.x

Lee, J., Cordaux, R., Han, K., Wang, J., Hedges, D.J., Liang, P., Batzer, M.A., 2007. Different evolutionary fates of recently integrated human and chimpanzee LINE-1 retrotransposons. Gene 390, 18–27. doi:10.1016/j.gene.2006.08.029

Lee, J., Hwang, Y.J., Shin, J.-Y., Lee, W.-C., Wie, J., Kim, K.Y., Lee, M.Y., Hwang, D., Ratan, R.R., Pae, A.N., Kowall, N.W., So, I., Kim, J.-I., Ryu, H., 2013. Epigenetic regulation of cholinergic receptor M1 (CHRM1) by histone H3K9me3 impairs Ca(2+) signaling in Huntington's disease. Acta Neuropathol 125, 727–739. doi:10.1007/s00401-013-1103-z

Lee, S.-T., Kim, M., 2011. MicroRNAs in Experimental Models of Movement Disorders. Journal of movement disorders 4, 55–59. doi:10.14802/jmd.11011

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., Wutz, A., 2010. Polycomb complexes act redundantly to repress genomic repeats and genes. Genes Dev 24, 265–276. doi:10.1101/gad.544410

Levin, H.L., Moran, J.V., 2011. Dynamic interactions between transposable elements and their hosts. Nat Rev Genet 12, 615–627. doi:10.1038/nrg3030

Li, J., Kannan, M., Trivett, A.L., Liao, H., Wu, X., Akagi, K., Symer, D.E., 2014. An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. Nucleic Acids Res 42, 4546–4562. doi:10.1093/nar/gku091

Li, W., Prazak, L., Chatterjee, N., Grüninger, S., Krug, L., Theodorou, D., Dubnau, J., 2013. Activation of transposable elements during aging and neuronal decline in Drosophila. Nat Neurosci 16, 529–531. doi:10.1038/nn.3368

Lloret, A., Dragileva, E., Teed, A., Espinola, J., Fossale, E., Gillis, T., Lopez, E., Myers, R.H., MacDonald, M.E., Wheeler, V.C., 2006. Genetic background modifies nuclear mutant huntingtin accumulation and HD CAG repeat instability in Huntington's disease knock-in mice. Hum Mol Genet 15, 2015–2024. doi:10.1093/hmg/ddl125

Lunkes, A., Lindenberg, K.S., Ben-Haïem, L., Weber, C., Devys, D., Landwehrmeyer, G.B., Mandel, J.-L., Trottier, Y., 2002. Proteases acting on mutant huntingtin

generate cleaved products that differentially build up cytoplasmic and nuclear inclusions. Mol Cell 10, 259–269.

Lupski, J.R., 2013. Genetics. Genome mosaicism--one human, multiple genomes. Science 341, 358–359. doi:10.1126/science.1239503

Luthi-Carter, R., Hanson, S.A., Strand, A.D., Bergstrom, D.A., Chun, W., Peters, N.L., Woods, A.M., Chan, E.Y., Kooperberg, C., Krainc, D., Young, A.B., Tapscott, S.J., Olson, J.M., 2002. Dysregulation of gene expression in the R6/2 model of polyglutamine disease: parallel changes in muscle and brain. Hum Mol Genet 11, 1911–1926.

Lyon, M.F., 1998. X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell Genet 80, 133–137.

MacDonald, M.E., Gines, S., Gusella, J.F., Wheeler, V.C., 2003. Huntington's disease. Neuromolecular Med 4, 7–20. doi:10.1385/NMM:4:1-2:7

Macia, A., Muñoz-Lopez, M., Cortes, J.L., Hastings, R.K., Morell, S., Lucena-Aguilar, G., Marchal, J.A., Badge, R.M., Garcia-Perez, J.L., 2011. Epigenetic control of retrotransposon expression in human embryonic stem cells. Mol Cell Biol 31, 300–316. doi:10.1128/MCB.00561-10

Macia, A., Widmann, T.J., Heras, S.R., Ayllon, V., Sanchez, L., Benkaddour-Boumzaouad, M., Munoz-Lopez, M., Rubio, A., Amador-Cubero, S., Blanco-Jimenez, E., Garcia-Castro, J., Menendez, P., Ng, P., Muotri, A.R., Goodier, J.L., Garcia-Perez, J.L., 2016. Engineered LINE-1 retrotransposition in non-dividing human neurons. Genome Res. doi:10.1101/gr.206805.116

Mager, D.L., Stoye, J.P., 2015. Mammalian Endogenous Retroviruses. Microbiology spectrum 3, MDNA3–0009. doi:10.1128/microbiolspec.MDNA3-0009-2014

Maksakova, I.A., Romanish, M.T., Gagnier, L., Dunn, C.A., van de Lagemaat, L.N., Mager, D.L., 2006. Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet 2, e2. doi:10.1371/journal.pgen.0020002

Malki, S., van der Heijden, G.W., O'Donnell, K.A., Martin, S.L., Bortvin, A., 2014. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. Dev Cell 29, 521–533. doi:10.1016/j.devcel.2014.04.027

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., Hannon, G.J., 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137, 522–535. doi:10.1016/j.cell.2009.03.040

Manakov, S.A., Pezic, D., Marinov, G.K., Pastor, W.A., Sachidanandam, R., Aravin, A.A., 2015. MIWI2 and MILI Have Differential Effects on piRNA Biogenesis and DNA Methylation. Cell Rep 12, 1234–1243. doi:10.1016/j.celrep.2015.07.036

Mandal, P.K., Kazazian, H.H., 2008. SnapShot: Vertebrate transposons. Cell 135, 192–192.e1. doi:10.1016/j.cell.2008.09.028

Mangiarini, L., Sathasivam, K., Mahal, A., Mott, R., Seller, M., Bates, G.P., 1997. Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. Nat Genet 15, 197–200. doi:10.1038/ng0297-197

Manley, K., Pugh, J., Messer, A., 1999. Instability of the CAG repeat in immortalized fibroblast cell cultures from Huntington's disease transgenic mice. Brain Res 835, 74–79.

Marchetto, M.C.N., Narvaiza, I., Denli, A.M., Benner, C., Lazzarini, T.A., Nathanson, J.L., Paquola, A.C.M., Desai, K.N., Herai, R.H., Weitzman, M.D., Yeo, G.W., Muotri, A.R., Gage, F.H., 2013. Differential L1 regulation in pluripotent stem cells of humans and apes. Nature 503, 525–529. doi:10.1038/nature12686

Martens, J.H.A., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., Jenuwein, T., 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. EMBO J 24, 800–812. doi:10.1038/sj.emboj.7600545

Martin, S.L., 1991. Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. Mol Cell Biol 11, 4804–4807.

Martin, S.L., 2006. The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. J Biomed Biotechnol 2006, 45621. doi:10.1155/JBB/2006/45621

Martindale, D., Hackam, A., Wieczorek, A., Ellerby, L., Wellington, C., McCutcheon, K., Singaraja, R., Kazemi-Esfarjani, P., Devon, R., Kim, S.U., Bredesen, D.E., Tufaro, F., Hayden, M.R., 1998. Length of huntingtin and its polyglutamine tract influences localization and frequency of intracellular aggregates. Nat Genet 18, 150–154. doi:10.1038/ng0298-150

Mathias, S.L., Scott, A.F., Kazazian, H.H., Boeke, J.D., Gabriel, A., 1991. Reverse transcriptase encoded by a human transposable element. Science 254, 1808–1810.

Mätlik, K., Redik, K., Speek, M., 2006. L1 antisense promoter drives tissue-specific transcription of human genes. J Biomed Biotechnol 2006, 71753. doi:10.1155/JBB/2006/71753

McCampbell, A., Taye, A.A., Whitty, L., Penney, E., Steffan, J.S., Fischbeck, K.H., 2001. Histone deacetylase inhibitors reduce polyglutamine toxicity. Proc Natl Acad Sci U S A 98, 15179–15184. doi:10.1073/pnas.261400698

McCLINTOCK, B., 1950. The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A 36, 344–355. doi:10.1073/pnas.36.6.344

McFarland, K.N., Huizenga, M.N., Darnell, S.B., Sangrey, G.R., Berezovska, O., Cha, J.-H.J., Outeiro, T.F., Sadri-Vakili, G., 2014. MeCP2: a novel Huntingtin interactor. Hum Mol Genet 23, 1036–1044. doi:10.1093/hmg/ddt499

McKinstry, S.U., Karadeniz, Y.B., Worthington, A.K., Hayrapetyan, V.Y., Ozlu, M.I., Serafin-Molina, K., Risher, W.C., Ustunkaya, T., Dragatsis, I., Zeitlin, S., Yin, H.H., Eroglu, C., 2014. Huntingtin is required for normal excitatory synapse development in cortical and striatal circuits. J Neurosci 34, 9455–9472. doi:10.1523/JNEUROSCI.4699-13.2014

Miné, M., Chen, J.-M., Brivet, M., Desguerre, I., Marchant, D., de Lonlay, P., Bernard, A., Férec, C., Abitbol, M., Ricquier, D., Marsac, C., 2007. A large genomic deletion in the PDHX gene caused by the retrotranspositional insertion of a full-length LINE-1 element. Hum Mutat 28, 137–142. doi:10.1002/humu.20449

Moldovan, J.B., Moran, J.V., 2015. The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLoS Genet 11, e1005121. doi:10.1371/journal.pgen.1005121

Molero, A.E., Gokhan, S., Gonzalez, S., Feig, J.L., Alexandre, L.C., Mehler, M.F., 2009. Impairment of developmental stem cell-mediated striatal neurogenesis and pluripotency genes in a knock-in model of Huntington's disease. Proc Natl Acad Sci U S A 106, 21900–21905. doi:10.1073/pnas.0912171106

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., Kazazian, H.H., 1996. High frequency retrotransposition in cultured mammalian cells. Cell 87, 917–927.

Morrish, T.A., Gilbert, N., Myers, J.S., Vincent, B.J., Stamato, T.D., Taccioli, G.E., Batzer, M.A., Moran, J.V., 2002. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat Genet 31, 159–165. doi:10.1038/ng898

Mouse Genome Sequencing Consortium, Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., et al., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562. doi:10.1038/nature01262

Muckenfuss, H., Hamdorf, M., Held, U., Perkovic, M., Löwer, J., Cichutek, K., Flory, E., Schumann, G.G., Münk, C., 2006. APOBEC3 proteins inhibit human LINE-1 retrotransposition. J Biol Chem 281, 22161–22172. doi:10.1074/jbc.M601716200

Muñoz-López, M., García-Pérez, J.L., 2010. DNA transposons: nature and applications in genomics. Curr Genomics 11, 115–128. doi:10.2174/138920210790886871

Muotri, A.R., Chu, V.T., Marchetto, M.C.N., Deng, W., Moran, J.V., Gage, F.H., 2005. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435, 903–910. doi:10.1038/nature03663

Muotri, A.R., Marchetto, M.C.N., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., Gage, F.H., 2010. L1 retrotransposition in neurons is modulated by MeCP2. Nature 468, 443–446. doi:10.1038/nature09544

Naas, T.P., DeBerardinis, R.J., Moran, J.V., Ostertag, E.M., Kingsmore, S.F., Seldin, M.F., Hayashizaki, Y., Martin, S.L., Kazazian, H.H., 1998. An actively retrotransposing, novel subfamily of mouse L1 elements. EMBO J 17, 590–597. doi:10.1093/emboj/17.2.590

Nandi, S., Chandramohan, D., Fioriti, L., Melnick, A.M., Hébert, J.M., Mason, C.E., Rajasethupathy, P., Kandel, E.R., 2016. Roles for small noncoding RNAs in silencing of retrotransposons in the mammalian brain. Proc Natl Acad Sci U S A. doi:10.1073/pnas.1609287113

Nasir, J., Floresco, S.B., O'Kusky, J.R., Diewert, V.M., Richman, J.M., Zeisler, J., Borowski, A., Marth, J.D., Phillips, A.G., Hayden, M.R., 1995. Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. Cell 81, 811–823.

Neto, J.L., Lee, J.-M., Afridi, A., Gillis, T., Guide, J.R., Dempsey, S., Lager, B., Alonso, I., Wheeler, V.C., Mouro Pinto, R., 2016. Genetic Contributors to Intergenerational CAG Repeat Instability in Huntington's Disease Knock-In Mice. Genetics. doi:10.1534/genetics.116.195578

Ng, C.W., Yildirim, F., Yap, Y.S., Dalin, S., Matthews, B.J., Velez, P.J., Labadorf, A., Housman, D.E., Fraenkel, E., 2013. Extensive changes in DNA methylation are

associated with expression of mutant huntingtin. Proc Natl Acad Sci U S A 110, 2354–2359. doi:10.1073/pnas.1221292110

Ngamphiw, C., Tongsima, S., Mutirangura, A., 2014. Roles of intragenic and intergenic L1s in mouse and human. PLoS ONE 9, e113434. doi:10.1371/journal.pone.0113434

Nguyen, G.D., Gokhan, S., Molero, A.E., Mehler, M.F., 2013. Selective roles of normal and mutant huntingtin in neural induction and early neurogenesis. PLoS ONE 8, e64368. doi:10.1371/journal.pone.0064368

O'Kusky, J.R., Nasir, J., Cicchetti, F., Parent, A., Hayden, M.R., 1999. Neuronal degeneration in the basal ganglia and loss of pallido-subthalamic synapses in mice with targeted disruption of the Huntington's disease gene. Brain Res 818, 468–479.

Orecchini, E., Doria, M., Antonioni, A., Galardi, S., Ciafrè, S.A., Frassinelli, L., Mancone, C., Montaldo, C., Tripodi, M., Michienzi, A., 2016. ADAR1 restricts LINE-1 retrotransposition. Nucleic Acids Res. doi:10.1093/nar/gkw834

Ostertag, E.M., DeBerardinis, R.J., Goodier, J.L., Zhang, Y., Yang, N., Gerton, G.L., Kazazian, H.H., 2002. A mouse model of human L1 retrotransposition. Nat Genet 32, 655–660. doi:10.1038/ng1022

Ostertag, E.M., Kazazian, H.H., 2001. Biology of mammalian L1 retrotransposons. Annu Rev Genet 35, 501–538. doi:10.1146/annurev.genet.35.102401.091032

Pace, J.K., Feschotte, C., 2007. The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. Genome Res 17, 422–432. doi:10.1101/gr.5826307

Packer, A.N., Xing, Y., Harper, S.Q., Jones, L., Davidson, B.L., 2008. The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. J Neurosci 28, 14341–14346. doi:10.1523/JNEUROSCI.2390-08.2008

Peaston, A.E., Evsikov, A.V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D., Knowles, B.B., 2004. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. Dev Cell 7, 597–606. doi:10.1016/j.devcel.2004.09.004

Peddigari, S., Li, P.W.-L., Rabe, J.L., Martin, S.L., 2013. hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. Nucleic Acids Res 41, 575–585. doi:10.1093/nar/gks1075

Perepelitsa-Belancio, V., Deininger, P., 2003. RNA truncation by premature polyadenylation attenuates human mobile element activity. Nat Genet 35, 363–366. doi:10.1038/ng1269

Perrat, P.N., DasGupta, S., Wang, J., Theurkauf, W., Weng, Z., Rosbash, M., Waddell, S., 2013. Transposition-driven genomic heterogeneity in the Drosophila brain. Science 340, 91–95. doi:10.1126/science.1231965

Petersén, A., Björkqvist, M., 2006. Hypothalamic-endocrine aspects in Huntington's disease. Eur J Neurosci 24, 961–967. doi:10.1111/j.1460-9568.2006.04985.x

Pezic, D., Manakov, S.A., Sachidanandam, R., Aravin, A.A., 2014. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. Genes Dev 28, 1410–1428. doi:10.1101/gad.240895.114

Pillai, R.S., Chuma, S., 2012. piRNAs and their involvement in male germline development in mice. Dev Growth Differ 54, 78–92. doi:10.1111/j.1440-169X.2011.01320.x

Pinto, R.M., Dragileva, E., Kirby, A., Lloret, A., Lopez, E., St Claire, J., Panigrahi, G.B., Hou, C., Holloway, K., Gillis, T., Guide, J.R., Cohen, P.E., Li, G.-M., Pearson, C.E., Daly, M.J., Wheeler, V.C., 2013. Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS Genet 9, e1003930. doi:10.1371/journal.pgen.1003930

Pittoggi, C., Sciamanna, I., Mattei, E., Beraldi, R., Lobascio, A.M., Mai, A., Quaglia, M.G., Lorenzini, R., Spadafora, C., 2003. Role of endogenous reverse transcriptase in murine early embryo development. Mol Reprod Dev 66, 225–236. doi:10.1002/mrd.10349

Ponicsan, S.L., Kugel, J.F., Goodrich, J.A., 2010. Genomic gems: SINE RNAs regulate mRNA production. Curr Opin Genet Dev 20, 149–155. doi:10.1016/j.gde.2010.01.004

Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., Kandel, E.R., 2012. A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. Cell 149, 693–707. doi:10.1016/j.cell.2012.02.057

Rangasamy, D., 2013. Distinctive patterns of epigenetic marks are associated with promoter regions of mouse LINE-1 and LTR retrotransposons. Mob DNA 4, 27. doi:10.1186/1759-8753-4-27

Ratovitski, T., Chighladze, E., Arbez, N., Boronina, T., Herbrich, S., Cole, R.N., Ross, C.A., 2012. Huntingtin protein interactions altered by polyglutamine expansion as determined by quantitative proteomic analysis. Cell Cycle 11, 2006–2021. doi:10.4161/cc.20423

Reiner, A., Albin, R.L., Anderson, K.D., D'Amato, C.J., Penney, J.B., Young, A.B., 1988. Differential loss of striatal projection neurons in Huntington disease. Proc Natl Acad Sci U S A 85, 5733–5737.

Richardson, S.R., Doucet, A.J., Kopera, H.C., Moldovan, J.B., Garcia-Perez, J.L., Moran, J.V., 2015. The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. Microbiology spectrum 3, MDNA3–0061. doi:10.1128/microbiolspec.MDNA3-0061-2014

Rigamonti, D., Bauer, J.H., De-Fraja, C., Conti, L., Sipione, S., Sciorati, C., Clementi, E., Hackam, A., Hayden, M.R., Li, Y., Cooper, J.K., Ross, C.A., Govoni, S., Vincenz, C., Cattaneo, E., 2000. Wild-type huntingtin protects from apoptosis upstream of caspase-3. J Neurosci 20, 3705–3713.

Rosas, H.D., Liu, A.K., Hersch, S., Glessner, M., Ferrante, R.J., Salat, D.H., van der Kouwe, A., Jenkins, B.G., Dale, A.M., Fischl, B., 2002. Regional and progressive thinning of the cortical ribbon in Huntington's disease. Neurology 58, 695–701.

Rosenblatt, A., 2007. Neuropsychiatry of Huntington's disease. Dialogues Clin Neurosci 9, 191–197.

Rosenblatt, A., Brinkman, R.R., Liang, K.Y., Almqvist, E.W., Margolis, R.L., Huang, C.Y., Sherr, M., Franz, M.L., Abbott, M.H., Hayden, M.R., Ross, C.A., 2001. Familial influence on age of onset among siblings with Huntington disease. Am J Med Genet 105, 399–403.

Ross, C.A., Pantelyat, A., Kogan, J., Brandt, J., 2014. Determinants of functional disability in Huntington's disease: role of cognitive and motor dysfunction. Mov Disord 29, 1351–1358. doi:10.1002/mds.26012

Ryu, H., Lee, J., Hagerty, S.W., Soh, B.Y., McAlpin, S.E., Cormier, K.A., Smith, K.M., Ferrante, R.J., 2006. ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington's disease. Proc Natl Acad Sci U S A 103, 19176–19181. doi:10.1073/pnas.0606373103

Sapp, E., Kegel, K.B., Aronin, N., Hashikawa, T., Uchiyama, Y., Tohyama, K., Bhide, P.G., Vonsattel, J.P., DiFiglia, M., 2001. Early and progressive accumulation of

reactive microglia in the Huntington disease brain. J Neuropathol Exp Neurol 60, 161–172.

Savas, J.N., Ma, B., Deinhardt, K., Culver, B.P., Restituito, S., Wu, L., Belasco, J.G., Chao, M.V., Tanese, N., 2010. A role for huntington disease protein in dendritic RNA granules. J Biol Chem 285, 13142–13153. doi:10.1074/jbc.M110.114561

Savas, J.N., Makusky, A., Ottosen, S., Baillat, D., Then, F., Krainc, D., Shiekhattar, R., Markey, S.P., Tanese, N., 2008. Huntington's disease protein contributes to RNA-mediated gene silencing through association with Argonaute and P bodies. Proc Natl Acad Sci U S A 105, 10820–10825. doi:10.1073/pnas.0800658105

Sen, S.K., Huang, C.T., Han, K., Batzer, M.A., 2007. Endonuclease-independent insertion provides an alternative pathway for L1 retrotransposition in the human genome. Nucleic Acids Res 35, 3741–3751. doi:10.1093/nar/gkm317

Seong, I.S., Woda, J.M., Song, J.-J., Lloret, A., Abeyrathne, P.D., Woo, C.J., Gregory, G., Lee, J.-M., Wheeler, V.C., Walz, T., Kingston, R.E., Gusella, J.F., Conlon, R.A., MacDonald, M.E., 2010. Huntingtin facilitates polycomb repressive complex 2. Hum Mol Genet 19, 573–583. doi:10.1093/hmg/ddp524

Seredenina, T., Luthi-Carter, R., 2012. What have we learned from gene expression profiles in Huntington's disease? Neurobiol Dis 45, 83–98. doi:10.1016/j.nbd.2011.07.001

Severynse, D.M., Hutchison, C.A., Edgell, M.H., 1992. Identification of transcriptional regulatory activity within the 5' A-type monomer sequence of the mouse LINE-1 retroposon. Mamm Genome 2, 41–50.

Sipione, S., Rigamonti, D., Valenza, M., Zuccato, C., Conti, L., Pritchard, J., Kooperberg, C., Olson, J.M., Cattaneo, E., 2002. Early transcriptional profiles in huntingtin-inducible striatal cells by microarray analyses. Hum Mol Genet 11, 1953–1965.

Smallwood, S.A., Kelsey, G., 2012. De novo DNA methylation: a germ cell perspective. Trends Genet 28, 33–42. doi:10.1016/j.tig.2011.09.004

Sookdeo, A., Hepp, C.M., McClure, M.A., Boissinot, S., 2013. Revisiting the evolution of mouse LINE-1 in the genomic era. Mob DNA 4, 3. doi:10.1186/1759-8753-4-3

Specchia, V., Piacentini, L., Tritto, P., Fanti, L., D'Alessandro, R., Palumbo, G., Pimpinelli, S., Bozzetti, M.P., 2010. Hsp90 prevents phenotypic variation by

suppressing the mutagenic activity of transposons. Nature 463, 662–665. doi:10.1038/nature08739

Speek, M., 2001. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol Cell Biol 21, 1973–1985. doi:10.1128/MCB.21.6.1973-1985.2001

Steffan, J.S., Bodai, L., Pallos, J., Poelman, M., McCampbell, A., Apostol, B.L., Kazantsev, A., Schmidt, E., Zhu, Y.Z., Greenwald, M., Kurokawa, R., Housman, D.E., Jackson, G.R., Marsh, J.L., Thompson, L.M., 2001. Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. Nature 413, 739–743. doi:10.1038/35099568

Steffan, J.S., Kazantsev, A., Spasic-Boskovic, O., Greenwald, M., Zhu, Y.Z., Gohler, H., Wanker, E.E., Bates, G.P., Housman, D.E., Thompson, L.M., 2000. The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. Proc Natl Acad Sci U S A 97, 6763–6768. doi:10.1073/pnas.100110097

Stetson, D.B., Ko, J.S., Heidmann, T., Medzhitov, R., 2008. Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell 134, 587–598. doi:10.1016/j.cell.2008.06.032

Swami, M., Hendricks, A.E., Gillis, T., Massood, T., Mysore, J., Myers, R.H., Wheeler, V.C., 2009. Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. Hum Mol Genet 18, 3039–3047. doi:10.1093/hmg/ddp242

Symer, D.E., Connelly, C., Szak, S.T., Caputo, E.M., Cost, G.J., Parmigiani, G., Boeke, J.D., 2002. Human l1 retrotransposition is associated with genetic instability in vivo. Cell 110, 327–338.

Takano, H., Gusella, J.F., 2002. The predominantly HEAT-like motif structure of huntingtin and its association and coincident nuclear entry with dorsal, an NF-kB/Rel/dorsal family transcription factor. BMC Neurosci 3, 15.

Taylor, M.S., LaCava, J., Mita, P., Molloy, K.R., Huang, C.R.L., Li, D., Adney, E.M., Jiang, H., Burns, K.H., Chait, B.T., Rout, M.P., Boeke, J.D., Dai, L., 2013. Affinity proteomics reveals human host factors implicated in discrete stages of LINE-1 retrotransposition. Cell 155, 1034–1048. doi:10.1016/j.cell.2013.10.021

Telenius, H., Kremer, B., Goldberg, Y.P., Theilmann, J., Andrew, S.E., Zeisler, J., Adam, S., Greenberg, C., Ives, E.J., Clarke, L.A., 1994. Somatic and gonadal

mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat Genet 6, 409–414. doi:10.1038/ng0494-409

Tomé, S., Manley, K., Simard, J.P., Clark, G.W., Slean, M.M., Swami, M., Shelbourne, P.F., Tillier, E.R.M., Monckton, D.G., Messer, A., Pearson, C.E., 2013. MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington's disease mice. PLoS Genet 9, e1003280. doi:10.1371/journal.pgen.1003280

Trelogan, S.A., Martin, S.L., 1995. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. Proc Natl Acad Sci U S A 92, 1520–1524.

Trottier, Y., Biancalana, V., Mandel, J.L., 1994. Instability of CAG repeats in Huntington's disease: relation to parental transmission and age of onset. J Med Genet 31, 377–382.

Ueno, M., Okamura, T., Mishina, M., Ishizaka, Y., 2016. Modulation of long interspersed nuclear element-1 in the mouse hippocampus during maturation. Mob Genet Elements 6, e1211980. doi:10.1080/2159256X.2016.1211980

Upton, K.R., Gerhardt, D.J., Jesuadian, J.S., Richardson, S.R., Sánchez-Luque, F.J., Bodea, G.O., Ewing, A.D., Salvador-Palomeque, C., van der Knaap, M.S., Brennan, P.M., Vanderver, A., Faulkner, G.J., 2015. Ubiquitous L1 mosaicism in hippocampal neurons. Cell 161, 228–239. doi:10.1016/j.cell.2015.03.026

Van den Hurk, J.A.J.M., Meij, I.C., Seleme, M. del C., Kano, H., Nikopoulos, K., Hoefsloot, L.H., Sistermans, E.A., de Wijs, I.J., Mukhopadhyay, A., Plomp, A.S., de Jong, P.T.V.M., Kazazian, H.H., Cremers, F.P.M., 2007. L1 retrotransposition can occur early in human embryonic development. Hum Mol Genet 16, 1587–1592. doi:10.1093/hmg/ddm108

Van Meter, M., Kashyap, M., Rezazadeh, S., Geneva, A.J., Morello, T.D., Seluanov, A., Gorbunova, V., 2014. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. Nat Commun 5, 5011. doi:10.1038/ncomms6011

Vashishtha, M., Ng, C.W., Yildirim, F., Gipson, T.A., Kratter, I.H., Bodai, L., Song, W., Lau, A., Labadorf, A., Vogel-Ciernia, A., Troncosco, J., Ross, C.A., Bates, G.P., Krainc, D., Sadri-Vakili, G., Finkbeiner, S., Marsh, J.L., Housman, D.E., Fraenkel, E., Thompson, L.M., 2013. Targeting H3K4 trimethylation in Huntington disease. Proc Natl Acad Sci U S A 110, E3027–36. doi:10.1073/pnas.1311323110

Viollet, S., Monot, C., Cristofari, G., 2014. L1 retrotransposition: The snap-velcro model and its consequences. Mob Genet Elements 4, e28907. doi:10.4161/mge.28907

Vitullo, P., Sciamanna, I., Baiocchi, M., Sinibaldi-Vallebona, P., Spadafora, C., 2012. LINE-1 retrotransposon copies are amplified during murine early embryo development. Mol Reprod Dev 79, 118–127. doi:10.1002/mrd.22003

Von Schimmelmann, M., Feinberg, P.A., Sullivan, J.M., Ku, S.M., Badimon, A., Duff, M.K., Wang, Z., Lachmann, A., Dewell, S., Ma'ayan, A., Han, M.-H., Tarakhovsky, A., Schaefer, A., 2016. Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration. Nat Neurosci 19, 1321–1330. doi:10.1038/nn.4360

Wallace, N.A., Belancio, V.P., Deininger, P.L., 2008. L1 mobile element expression causes multiple types of toxicity. Gene 419, 75–81. doi:10.1016/j.gene.2008.04.013

Warrick, J.M., Chan, H.Y., Gray-Board, G.L., Chai, Y., Paulson, H.L., Bonini, N.M., 1999. Suppression of polyglutamine-mediated neurodegeneration in Drosophila by the molecular chaperone HSP70. Nat Genet 23, 425–428. doi:10.1038/70532

Weber, M.J., 2006. Mammalian small nucleolar RNAs are mobile genetic elements. PLoS Genet 2, e205. doi:10.1371/journal.pgen.0020205

Wellington, C.L., Ellerby, L.M., Hackam, A.S., Margolis, R.L., Trifiro, M.A., Singaraja, R., McCutcheon, K., Salvesen, G.S., Propp, S.S., Bromm, M., Rowland, K.J., Zhang, T., Rasper, D., Roy, S., Thornberry, N., Pinsky, L., Kakizuka, A., Ross, C.A., Nicholson, D.W., Bredesen, D.E., Hayden, M.R., 1998. Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract. J Biol Chem 273, 9158–9167.

Wheeler, V.C., Auerbach, W., White, J.K., Srinidhi, J., Auerbach, A., Ryan, A., Duyao, M.P., Vrbanac, V., Weaver, M., Gusella, J.F., Joyner, A.L., MacDonald, M.E., 1999. Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. Hum Mol Genet 8, 115–122.

Wheeler, V.C., Gutekunst, C.-A., Vrbanac, V., Lebel, L.-A., Schilling, G., Hersch, S., Friedlander, R.M., Gusella, J.F., Vonsattel, J.-P., Borchelt, D.R., MacDonald, M.E., 2002. Early phenotypes that presage late-onset neurodegenerative disease

allow testing of modifiers in Hdh CAG knock-in mice. Hum Mol Genet 11, 633–640.

Wheeler, V.C., Lebel, L.-A., Vrbanac, V., Teed, A., te Riele, H., MacDonald, M.E., 2003. Mismatch repair gene Msh2 modifies the timing of early disease in Hdh(Q111) striatum. Hum Mol Genet 12, 273–281.

White, J.K., Auerbach, W., Duyao, M.P., Vonsattel, J.P., Gusella, J.F., Joyner, A.L., MacDonald, M.E., 1997. Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. Nat Genet 17, 404–410. doi:10.1038/ng1297-404

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., Schulman, A.H., 2007. A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8, 973–982. doi:10.1038/nrg2165

Wilda, M., Bächner, D., Zechner, U., Kehrer-Sawatzki, H., Vogel, W., Hameister, H., 2000. Do the constraints of human speciation cause expression of the same set of genes in brain, testis, and placenta? Cytogenet Cell Genet 91, 300–302. doi:56861

Williams, J.K., Hamilton, R., Nehl, C., McGonigal-Kenney, M., Schutte, D.L., Sparbel, K., Birrer, E., Tripp-Reimer, T., Friedrich, R., Penziner, E., Jarmon, L., Paulsen, J., 2007. No one else sees the difference: "family members" perceptions of changes in persons with preclinical Huntington disease. Am J Med Genet B Neuropsychiatr Genet 144B, 636–641. doi:10.1002/ajmg.b.30479

Wissing, S., Montano, M., Garcia-Perez, J.L., Moran, J.V., Greene, W.C., 2011. Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. J Biol Chem 286, 36427–36437. doi:10.1074/jbc.M111.251058

Wissing, S., Muñoz-Lopez, M., Macia, A., Yang, Z., Montano, M., Collins, W., Garcia-Perez, J.L., Moran, J.V., Greene, W.C., 2012. Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. Hum Mol Genet 21, 208–218. doi:10.1093/hmg/ddr455

Yang, N., Kazazian, H.H., 2006. L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol 13, 763–771. doi:10.1038/nsmb1141

Yoder, J.A., Walsh, C.P., Bestor, T.H., 1997. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13, 335–340.

Yu, F., Zingler, N., Schumann, G., Strätling, W.H., 2001. Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res 29, 4493–4501.

Zeitlin, S., Liu, J.P., Chapman, D.L., Papaioannou, V.E., Efstratiadis, A., 1995. Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. Nat Genet 11, 155–163. doi:10.1038/ng1095-155

Zemojtel, T., Penzkofer, T., Schultz, J., Dandekar, T., Badge, R., Vingron, M., 2007. Exonization of active mouse L1s: a driver of transcriptome evolution? BMC Genomics 8, 392. doi:10.1186/1471-2164-8-392

Zhang, A., Dong, B., Doucet, A.J., Moldovan, J.B., Moran, J.V., Silverman, R.H., 2014. RNase L restricts the mobility of engineered retrotransposons in cultured human cells. Nucleic Acids Res 42, 3803–3820. doi:10.1093/nar/gkt1308

Zhao, K., Du, J., Han, X., Goodier, J.L., Li, P., Zhou, X., Wei, W., Evans, S.L., Li, L., Zhang, W., Cheung, L.E., Wang, G., Kazazian, H.H., Yu, X.-F., 2013. Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutières syndrome-related SAMHD1. Cell Rep 4, 1108–1115. doi:10.1016/j.celrep.2013.08.019

Zuccato, C., Ciammola, A., Rigamonti, D., Leavitt, B.R., Goffredo, D., Conti, L., MacDonald, M.E., Friedlander, R.M., Silani, V., Hayden, M.R., Timmusk, T., Sipione, S., Cattaneo, E., 2001. Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. Science 293, 493–498. doi:10.1126/science.1059581

Zuccato, C., Tartari, M., Crotti, A., Goffredo, D., Valenza, M., Conti, L., Cataudella, T., Leavitt, B.R., Hayden, M.R., Timmusk, T., Rigamonti, D., Cattaneo, E., 2003. Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. Nat Genet 35, 76–83. doi:10.1038/ng1219

APPENDIX

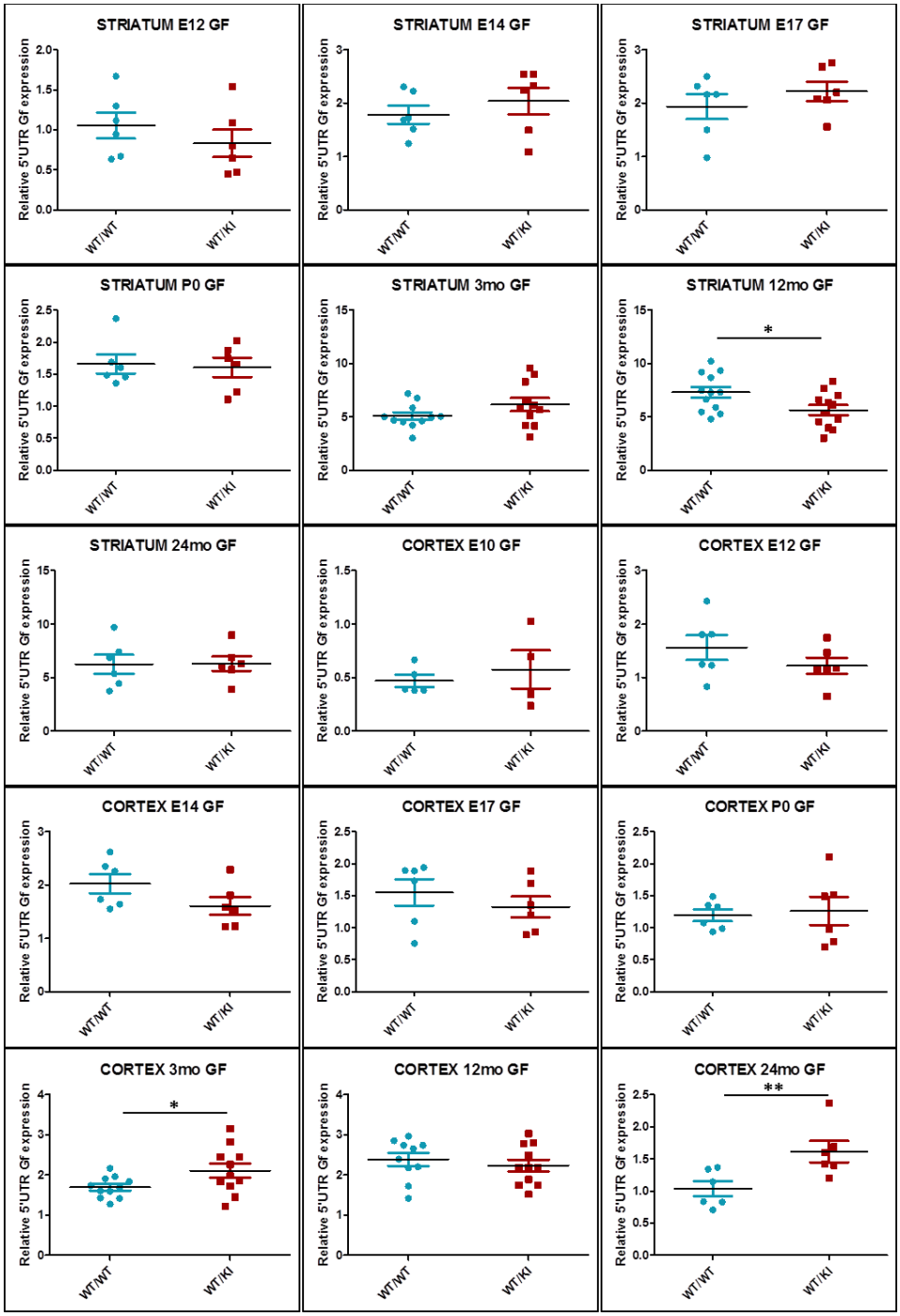


Figure 32. Expression analysis of L1 5'UTR Gf elements in striatum and cortex of WT/WT and WT/KI mice. Scatter plots display interindividual variability between biological replica at each developmental stage. E10, E12, E14, E17 means embryonal day 10, 12, 14, 17, respectively; P0, date of birth; 3mo, 3 months of age; 12mo,12months of age; 24mo, 24 months of age. Relative 5'UTR L1 mRNA levels are normalized with the housekeeping gene UbC gene. Each dot of the scatter plot represent the average of at least 3 independent qPCR replica. The average value between biological replica is shown with black horizontal line. Error bars indicate standard deviation. *P<0.05, **P<0.01 resulting from Mann Whitney unpaired test.

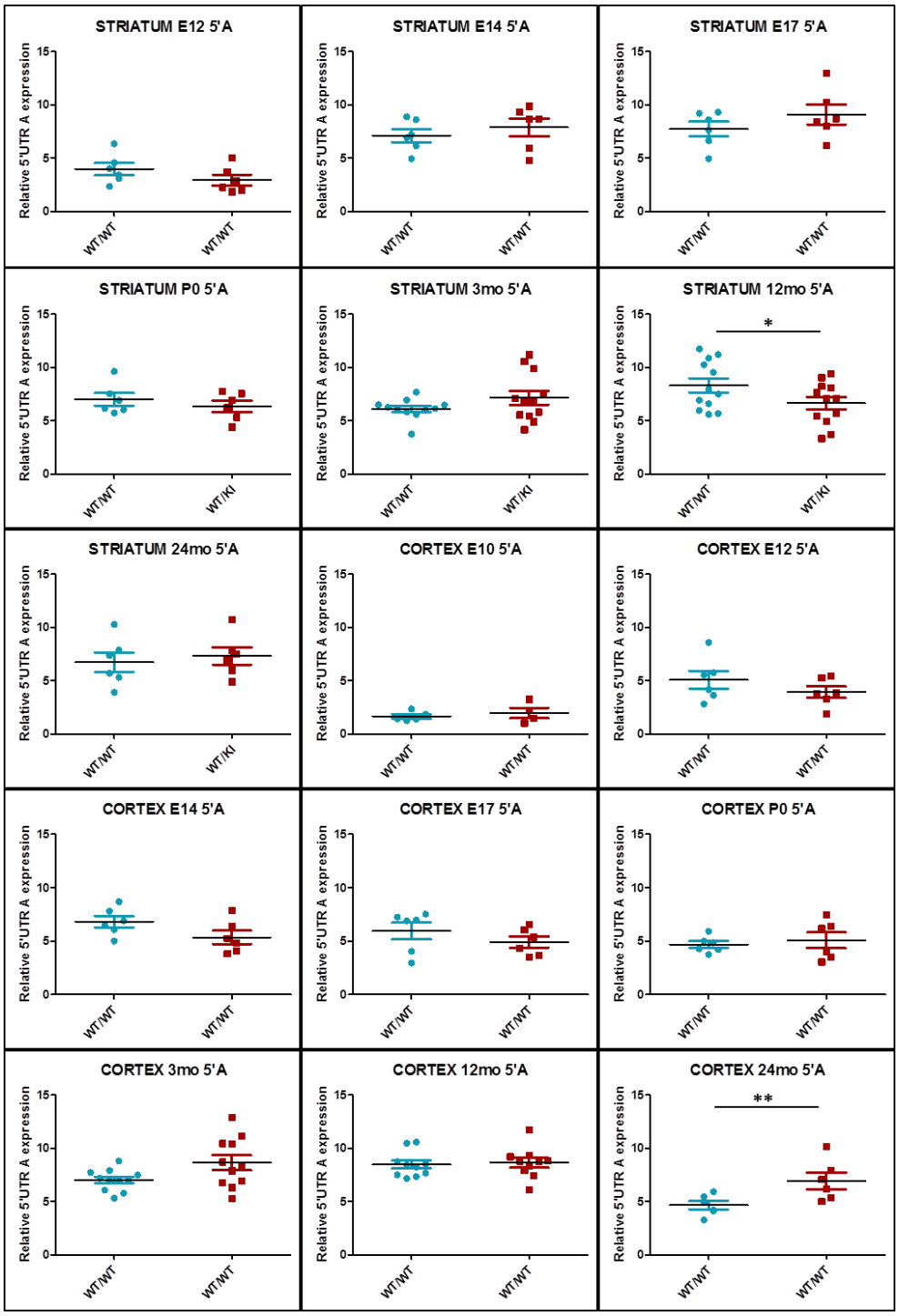


Figure 33. Expression analysis of L1 5'UTR A elements in striatum and cortex of WT/WT and WT/KI mice. Scatter plots display interindividual variability between biological replica at each developmental stage. E10, E12, E14, E17 means embryonal day 10, 12, 14, 17, respectively; P0, date of birth; 3mo, 3 months of age; 12mo,12months of age; 24mo, 24 months of age. Relative 5'UTR L1 mRNA levels are normalized with the housekeeping gene UbC. Each dot of the scatter plot represent the average of at least 3 independent qPCR replica. The average value between biological replica is shown with black horizontal line. Error bars indicate standard deviation. *P<0.05, **P<0.01 resulting from Mann Whitney unpaired test.

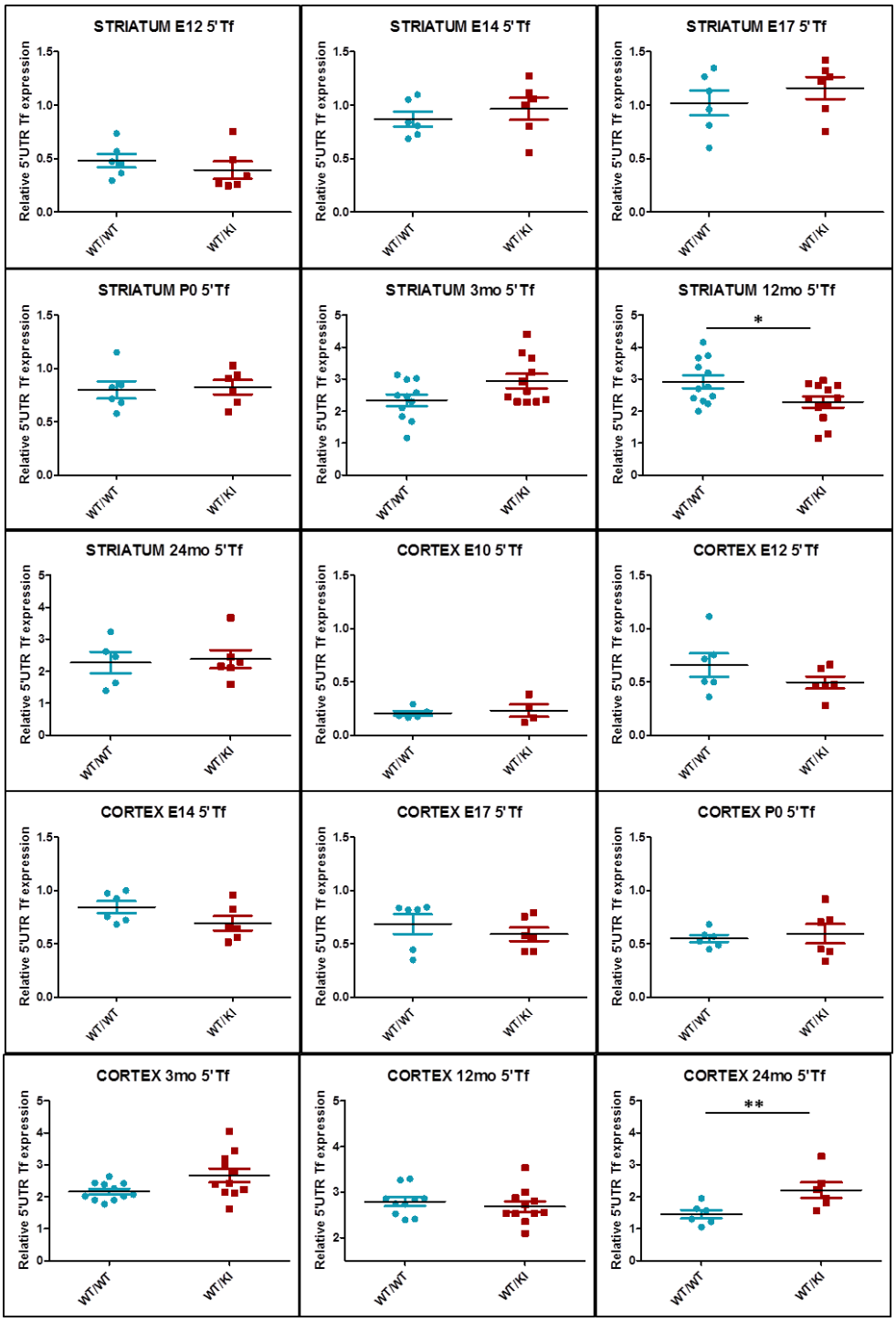


Figure 34. Expression analysis of L1 5'UTR Tf elements in striatum and cortex of WT/WT and WT/KI mice. Scatter plots display interindividual variability between biological replica at each developmental stage. E10, E12, E14, E17 means embryonal day 10, 12, 14, 17, respectively; P0, date of birth; 3mo, 3 months of age; 12mo,12months of age; 24mo, 24 months of age. Relative 5'UTR L1 mRNA levels are normalized with the housekeeping gene UbC. Each dot of the scatter plot represent the average of at least 3 independent qPCR replica. The average value between biological replica is shown with black horizontal line. Error bars indicate standard deviation. *P<0.05, **P<0.01 resulting from Mann Whitney unpaired test.

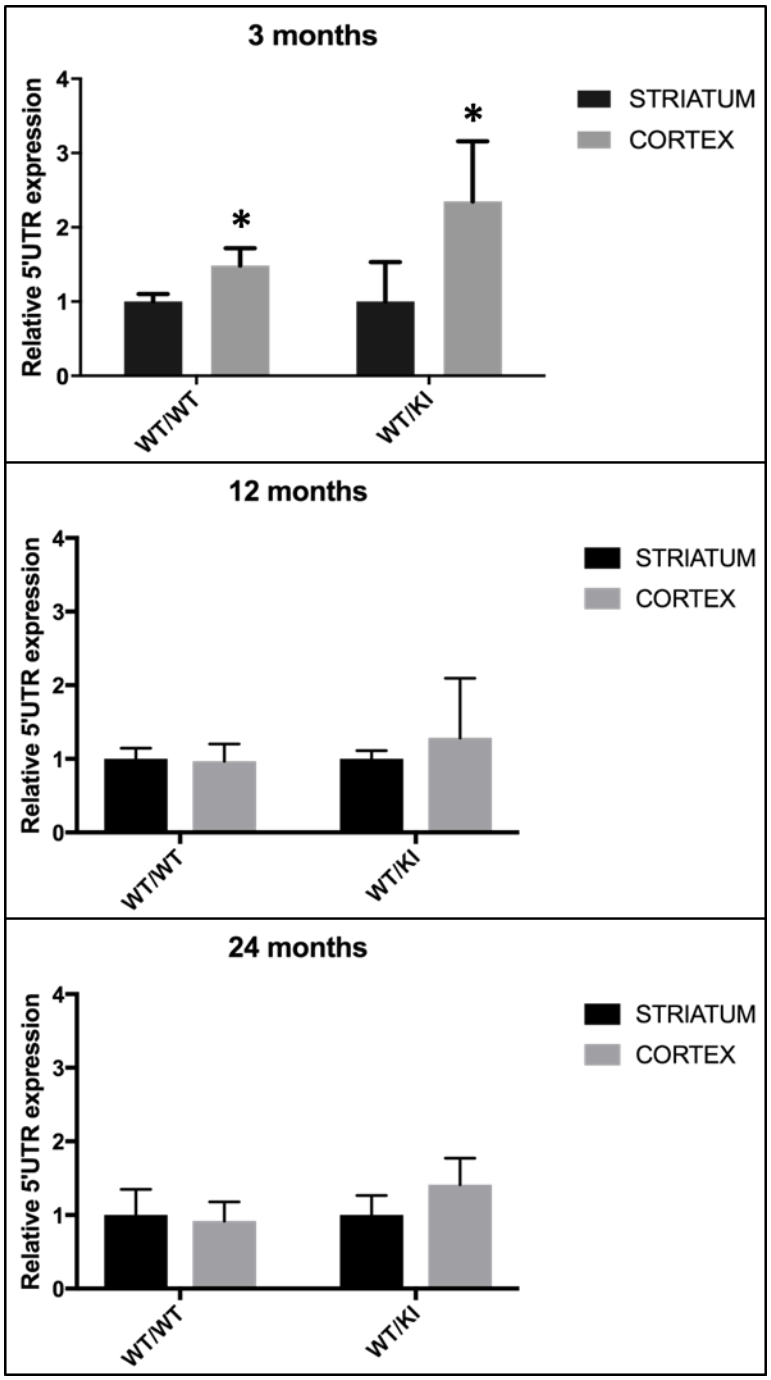


Figure 35. Comparison of full length L1 expression between striatum and cortex in adult mice. Relative quantification of 5'UTR Gf L1 mRNA transcript levels in striatum (black) and cortex (grey) of adult WT/WT and WT/KI mice. 3mo, 3 months of age; 12mo, 12 months of age; 24mo, 24 months of age. Relative 5'UTR L1 mRNA levels are normalized with the housekeeping gene UbC. At 3mo of age, L1 expression is significanty increased in the cortex as compared to the striatum. 3 independent qPCR replica were performed on each biological sample. Each bar represent average value of twelve independent biological replica (n=12). Error bars indicate standard error. *P<0.05 resulting from *Mann Whitney* unpaired test.