

L1 Retrotransposons Are Transcriptionally Active in Hippocampus of Rat Brain

Somnath Mukherjee¹, Deepak Sharma¹, Kailash C. Upadhyaya^{1,2}

¹School of Life Sciences, Jawaharlal Nehru University, New Delhi, India;

²Amity Institute of Molecular Biology and Genomics, Amity University, Noida, India

Received August 2, 2015; Accepted February 24, 2016.

Key words: LINEs (Long Interspersed Nuclear Elements) – RT-PCR – Hippocampus – Retrotransposon – L1 activation

Abstract: LINE1 (L1) is an autonomous, non-LTR retrotransposon and the L1 family of retrotransposons constitute around 17%, 20% and 23% in the human, mouse and rat genomes respectively. Under normal physiological conditions, the retroelements remain by and large transcriptionally silent but are activated in response to biotic and abiotic stress conditions and during perturbation in cellular metabolism. They have also been shown to be transiently activated under certain developmental programs. Using RT-PCR, we show that the L1 elements are transcriptionally active in the hippocampus region of the brain of four-month-old rat under normal conditions without any apparent stress. Twenty non-redundant LINE1-specific reverse transcriptase (RTase) sequences from ORF2 region were isolated, cloned and sequenced. Full length L1 element sequences complementary to the isolated sequences were retrieved from the L1 database. *In silico* analysis was used to determine the presence of these retroelements proximal (up to 10 kb) to the genes transcriptionally active in the hippocampus. Many important genes were found to be in close proximity of the transcriptionally active L1 elements. Transcriptional activation of the elements possibly affects the expression of the neighbouring genes.

This study was supported from the University Grants Commission, New Delhi, India.

Mailing Address: Prof. Kailash C. Upadhyaya, PhD., Amity Institute of Molecular Biology and Genomics, Amity University, Noida, 201313, India; e-mail: kcupadhyaya@gmail.com

Introduction

Transposable elements make up sizeable components of all eukaryotic genomes, varying from 14% to over 80%. On the basis of their mechanism of transposition, they are categorized into two classes: (i) class I elements, called retrotransposons which move via an RNA intermediate by “copy-and-paste” mechanism; and (ii) class II elements, called transposons move via a DNA intermediate by “cut-and-paste” mechanism. The “copy-and-paste” mechanism implies that transposition of the element leads to increase in copy number whereas transposition by the “cut-and-paste” mechanism normally does not lead to increase in the copy number. Retroelements constitute a predominant class of elements in eukaryotic genomes and subdivided into two categories: LTR elements and non LTR elements. The mammalian genomes contain a preponderance of non-LTR retroelements. One of the transpositionally active groups of non-LTR elements in mammalian genomes is the L1 elements, LINE1 (Long Interspersed Nuclear Element). The L1 elements constitute approximately 17%, 20% and 23% in the human, mouse and rat genomes respectively (Lander et al., 2001; Waterston et al., 2002; Gibbs et al., 2004). They remain quiescent under normal physiological and developmental conditions (Furano, 2000; Ostertag and Kazazian, 2001; Moran and Gilbert, 2002); however, they could become active under environmental stress conditions (Muotri et al., 2009; Hunter et al., 2012, 2013; Terasaki et al., 2013) and certain pathological conditions (Beck et al., 2011; Kaer and Speek, 2013). The elements have shown to be active during the processes of differentiation (Muotri et al., 2005; Coufal et al., 2009), neurogenesis (Singer et al., 2010; Thomas et al., 2012), and embryogenesis (Garcia-Perez et al., 2007; van den Hurk et al., 2007; Castro-Diaz et al., 2014). Increased rate of L1 retrotransposition has also been reported in the neuronal tissues in schizophrenia and bipolar disorders (Coufal et al., 2011; Poduri et al., 2013; Bundo et al., 2014). Transpositional events in somatic cells have been shown to lead to mosaicism during neuronal differentiation (Muotri et al., 2005). Integration of the L1 elements into different locations within the genome may possibly affect the expression pattern of neighbouring genes. Activation of the elements in pathological conditions could be explained by certain degree of metabolic perturbation in the affected cells. As per the L1 database, over 11,600 full length L1 elements have been detected in human, over 6,500 in rat and over 2,000 in mouse genomes (<http://line1.bioapps.biozentrum.uniwuerzburg.de/l1base.php>).

L1 encodes two proteins: ORF1, an RNA binding protein and ORF2, an endonuclease and reverse transcriptase. Both the proteins are required for L1 transposition. L1 elements propagate in the host genome via an mRNA. Moran and Gilbert (2002) showed that the integrated full length L1 element is generally flanked by variably sized target site duplications. The L1 elements are thought to be transcribed by RNA polymerase II. Transcription into full-length RNA is the first step in the L1 retrotransposition that provides a template for the synthesis of

both L1-encoded proteins and a DNA copy (Furano, 2000; Ostertag and Kazazian, 2001; Moran and Gilbert, 2002). Then the reverse transcribed DNA copy may be integrated into the host genome. The increased activation of L1 retrotransposition in the neuronal genome has also been observed in mental disorders including schizophrenia and bipolar disorders (Coufal et al., 2011; Poduri et al., 2013; Bundo et al., 2014). It is conjectured that increased retrotranspositional activity of the L1 elements in schizophrenia may be triggered by genetic and/or environmental factors in the early neural development and that could be the contributing factor for the susceptibility and pathophysiology of the disease (Bundo et al., 2014).

We analysed L1 expression in different regions of brain of four-month-old rats by RT-PCR using primers designed for the L1-specific reverse transcriptase domain. We observed the presence of L1-specific reverse transcriptase (RTase) transcripts in the hippocampus under normal physiological conditions. Analysis of diversity of RTase domain of different L1 elements indicates the presence of transcript from at least 19 different L1 elements. *In silico* analysis reveals that the L1 elements active in the hippocampus are located within introns and in the close vicinity of upstream or downstream of the genes involved in neural related functions.

Material and Methods

Animals

Four-month-old male Wistar albino rats (6 for each age group) were used for all the experiments. Animals were maintained in the animal house facility of Jawaharlal Nehru University (JNU), New Delhi, at a constant temperature of 25 °C, humidity of 55% and 12 h dark and 12 h light cycle (light from 06:00 to 18:00 h). The animals were fed standard chow rat feed (Hindustan Lever Ltd., India) and water *ad libitum*. All animal experiments were approved by the JNU-Institutional Animal Ethics Committee; and all the institutional guidelines were adhered to in the care and treatment of the animals. Rats were sacrificed by cervical dislocation. Different regions of brain namely cerebral cortex, hippocampus, cerebellum, medulla and pons were excised, rapidly frozen in liquid nitrogen and stored at –80 °C.

RNA isolation and RT-PCR

Total RNA from the 4-month-old normal rat brain regions was extracted using TriZol (Invitrogen, USA) according to the manufacturer's instructions. To eliminate the possibility of genomic DNA contamination, the RNA samples were treated with the RNase free DNase I (Fermentas) according to the manufacturer's instructions. The concentration of RNA was determined and integrity of RNA in the samples was confirmed by 1.2% (w/v) MOPS agarose gel electrophoresis.

Isolation of reverse transcriptase sequences

Forward and reverse primers for PCR amplification were designed from the conserved domain within the ORF2 of the L1 element (Dobigny et al., 2004).

RT-PCR was carried out using Promega's Access RT-PCR system with slight modifications: 1 µg RNA template, 1.5 mM MgCl₂, and 200 nM of each of the primers (5'-ATT CTR TTC CAT TGG TCT A-3' and 5'-CCA TGC TCA TSG ATT GG-3'). The amplification was performed in a MJ Research thermocycler with the following parameters: 94 °C for 2 min, followed by 30 cycles of 94 °C for 30 s, 52 °C for 30 s and 72 °C for 30 s, followed by a final extension step at 72 °C for 10 min. The amplified products were eluted from 1.5% agarose gel and cloned into a pGEMTeasy vector (Promega). Cloned amplicons were sequenced by Scigenome, India.

Similarity among the amplified sequences

Thirty seven different nucleotide sequences obtained were checked for non-redundant sequences and the sequences over 90% similarities were excluded. We used the Mview tool for this analysis (<http://www.ebi.ac.uk/tools/msa>). Mview transforms the Sequence Similarity Search results into Multiple Sequence Alignment (Brown et al., 1998). Of all the sequences, 20 non-redundant sequences were used for further analysis.

Distribution of L1 elements on different chromosomes

The sequence of each chromosome was analysed using Ensembl BLAST/BLAT (Hubbard et al., 2002) online program (<http://asia.ensembl.org/Multi/blastview>). BLAT analysis was performed with “exact matches” search sensitivity. Top 1,000 results for each sequence were sorted by raw score. Out of all results only those alignments were selected which showed approximately 95% sequence similarity and 90% length match with query sequences. Total number of alignments per chromosome were calculated and converted to percentages for each query sequence. To show the relationship between L1 density and chromosome size, Pearson's correlation test was performed. The statistical analyses and graphs were built using Prism GraphPad v5.

Location of genes in the vicinity of intact L1 elements

Each chromosomal sequence was submitted to the L1 database (<http://line1.bioapps.biozentrum.uni-wuerzburg.de/l1base.php>) to know intact L1 elements-associated sequences within the rat genome (using RGSC v3.4-Ens30.34; 6108 Entries, Last Update: 2007-10-14 21:07:58). The first top 100 were sorted by intactness score, which also provided information about genes associated with each resultant L1 element. Each L1 UID was selected and checked for the vicinity of genes belonging to a specified probable chromosomal location using Ensembl Rat Genome Database RGSC v3.4. The chromosomal location was ascertained in “Ensembl Genome Browser” (<http://asia.ensembl.org/index.html>) in its required input format (e.g. Chromosomal Position-7:105925414-105916554). Only those genes were selected for further analysis, which were in the direct vicinity i.e. located within the gene

or located up to 10 kb from the upstream or downstream to an annotated protein coding gene. Thus, data were rearranged manually to show the presence of intact L1 element in the vicinity of a gene (upstream or downstream or intragenic). The results of Ensembl Genome Browser for the genes associated with intact L1 elements were recorded and tabulated. The expression of each gene in rat hippocampus was assessed using “Atlas Gene Expression Database” (<http://www.ebi.ac.uk/gxa/>).

Results

Presence of L1 specific transcripts in the hippocampus

The L1 retrotransposons are present in multiple copies in the rat genome and only a few copies are known to be transcriptionally active under stress conditions. We

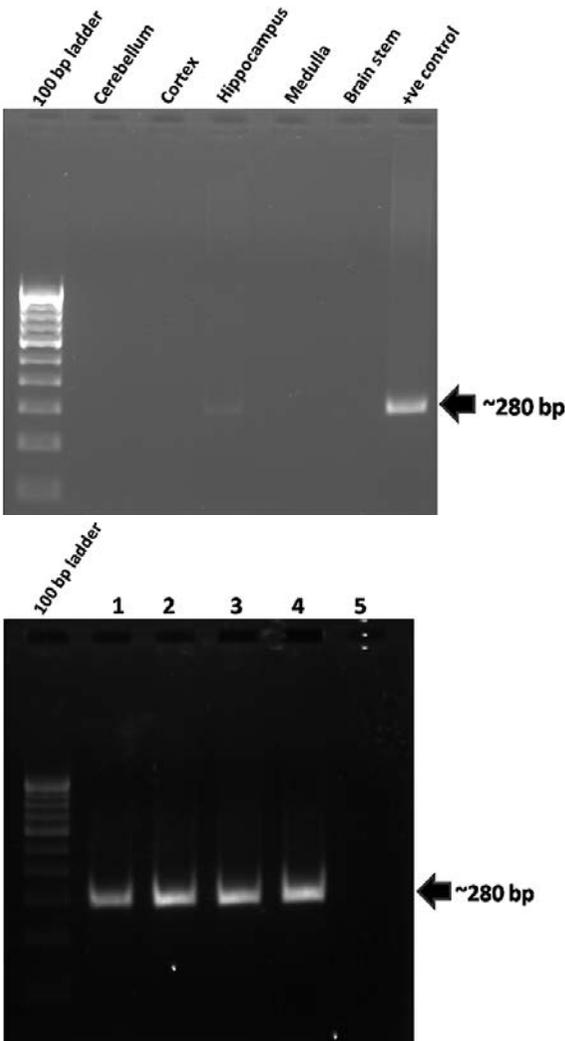


Figure 1 – RT-PCR profile of the L1 reverse transcriptase domain amplified using RNA isolated from different regions of brain. (A) RT-PCR amplification profiles of the L1 reverse transcriptase domain of different regions of brain: cerebellum, cortex, hippocampus, medulla oblongata, and pons. Approximately 280-bp amplicon (positive control) is expected using the specific primers. Amplification product is visible only in the hippocampus. (B) Lanes 1–4 – RT-PCR amplification of the L1 reverse transcriptase domain from the hippocampus of different rats. Lane 5 – negative control.

analysed the presence of L1 specific transcripts in various regions of rat brain under normal growth conditions. The degenerate primers corresponding to the L1-specific RTase domain are expected to produce an amplicon of ~280 bp. RNA isolated from five different regions (cerebral cortex, cerebellum, hippocampus, medulla oblongata and pons) of four-month-old rat brain were used for RT-PCR. The RT-PCR profiles show the presence of the L1-specific transcript only in the RNA isolated from the hippocampus region (Figure 1A). The results indicate that the L1 retrotransposons are transcriptionally active in the hippocampus region of rat brain. No L1-specific transcripts were observed in any other region of the brain. The experiments were repeated several times confirming that detectable transcriptional activity of L1 elements is only present in the hippocampus (Figure 1B).

Existence of different families of L1 transcripts

The reverse transcriptase domain of L1 retrotransposons is highly conserved at the amino acid level but not at the nucleotide level. Since a large number of different L1 retrotransposons are present in the genome, the kinetic complexity of the amplified transcript was analysed in order to identify transcriptionally active elements. Therefore, the amplified products were cloned and sequenced. Of the 37 clones sequenced, 20 were non-redundant sequences for L1-specific RTase domain suggesting that they possibly have originated from different L1 elements. All these sequences originating from transcriptionally active elements

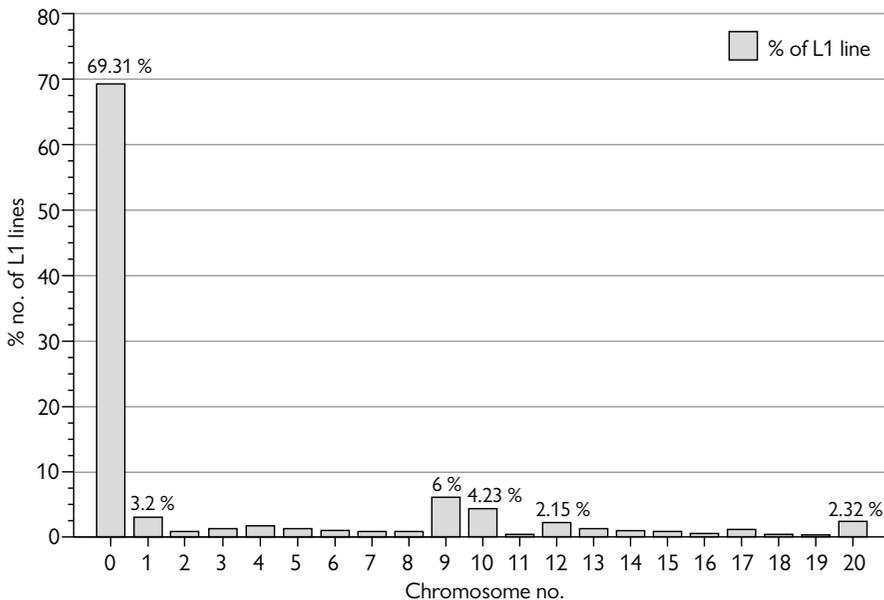


Figure 2 – Distribution of full length L1 elements in the individual chromosomes.

were analysed *in silico* to assess the physical distribution of the hippocampus active L1 elements in the rat genome. Since the truncated elements could not be functional, only full length elements were taken into consideration. Analysis indicates a skewed distribution of the full length L1 elements on different rat chromosomes. Chromosome 1 contains over 69% of the full length elements and chromosomes 19 and 20 are almost devoid of any full length L1 elements (Figure 2). The chromosome 1 contains majority of full length elements but only a tiny fraction of these active in the hippocampus. The elements were considered active on the basis of sequence homology to the RT-transcript. And out of these

A 1. Upstream L1 localization to a gene:



2. Downstream L1 localization to a gene:



3. Intragenic L1 localization to a gene:

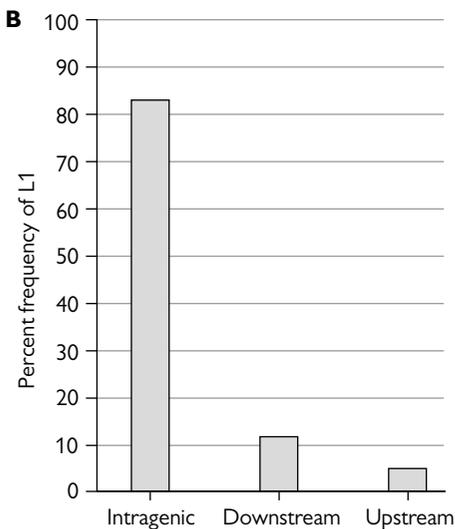
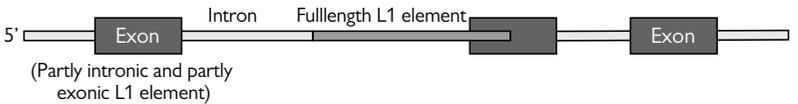


Figure 3A – Diagrammatic representation showing localization of full length L1 elements within or in the vicinity of the genes.

Figure 3B – Relative proportion of location of the L1 elements within or in the proximity to the genes.

a very few of the full length elements are located in the vicinity of the protein coding genes. Chromosomes 2, 4, 9, 13 and 14 contain 13, 9, 10, 10 and 7% of the transcriptionally active L1 elements respectively.

L1 elements are associated in close vicinity with genes

On searching of the L1 Rn database, we found 77 full length elements in the vicinity of genes. Of these full length elements, 63 are localized in the intragenic regions and the remaining ones were located upstream or downstream (within 10 kb) of the genes (Figure 3A and B). Figure 3B shows a relative distribution of the elements within or in the vicinity of the genes. At least 20 different members of the L1 family appear to be transcriptionally active in the hippocampus. The integration of a large proportion of these elements is within genes, either within an intron or in some cases also includes part of an exon. Functionality of the gene containing element within an intron may not be affected provided the integration does not affect the splicing.

Discussion

The rat genome contains 23% of the L1 family of retroelements and a proportion of the elements may potentially be mobile. They are known to be under tight regulatory controls and remain transcriptionally or transpositionally silent under normal conditions. They are known to be activated under environmental stress or metabolic perturbation (Muotri et al., 2009; Hunter et al., 2013; Kaer and Speek, 2013). We show that the retroelements that normally are transcriptionally/transpositionally activated in response to stress or metabolic perturbation could be active under normal physiological conditions in the hippocampus. Since the transcription of the elements is a prerequisite for retrotransposition, it is very likely that at least a part of L1-derived transcript would lead to retrotransposition. Besides retrotransposition activity has shown to be increased during embryogenesis and neurogenesis. Since the hippocampus is the primary site for neurogenesis it is possible that the enhanced transcriptional activity of L1 elements in the hippocampus may directly be related to neurogenesis.

Enhanced transpositional activation of L1 elements has been associated with high neuronal turn over leading to neuronal diversity and plasticity. In adult humans, 700 new neurons are added in the hippocampus every day. Substantial neurogenesis takes place throughout life in the human hippocampus except a modest decline during aging (Spalding et al., 2013). New neurons are continuously added and this high degree of neurogenesis continuously remodels hippocampal circuits (Akers et al., 2014). This remodelling possibly is a consequence of synthesis of new neurons replenishing the old ones. It is tempting to conjecture that the increased retrotransposition during neurogenesis may be related to the degradation of old neurons. However, there are contrasting observations with respect to increase in copy number of L1 elements in the hippocampus. Coufal et al. (2009) observed

an increased in copy number of L1 elements in the hippocampus and other regions of adult human brain as compared to the copy number of L1 elements in heart or liver. On the other hand, Evrony et al. (2012) analysed the genomic variability among the individual neurons of the brain by single neuron sequencing and concluded that no detectable level of L1 induced integration in neurons. They suggested that L1 was not a major generator of neuronal diversity.

The retroelements have mostly been considered evolutionary byproducts even though mammalian genomes contain a significant proportion of potentially mobile retrotransposons (Speek, 2001; Baillie et al., 2011). The transcriptional control on the elements is being exercised either through epigenetic modulations or by specific miRNAs (Macia et al., 2011). The regulatory controls are known to be relaxed during pathological conditions especially in cancer. Extensive transduction of non-repetitive DNA mediated by L1 retrotransposition in cancer genomes has been observed (Tubio et al., 2014). A significant proportion of protein coding genes and regulatory domains of the genes have been shown to be evolved from retrotransposons (Kokošar and Kordiš, 2013). The proximity of the active retroelements to protein coding genes may affect their expression. The L1 elements could be integrated within the gene in either orientation. The human and mouse L1 elements may have both upstream and downstream transcriptional capabilities due to presence of sense and anti-sense promoters at their 5' UTRs. Expression of genes such as *Nkain3*, *Mtif2*, *Atf6*, *Rabgap11*, *Dock1*, *Lin7a*, *Cyp20a1* etc. located in close proximity to the L1 elements may possibly be modulated by the active element in the hippocampus (Speek, 2001; Faulkner et al., 2009). Therefore, it is conjectured that limited transcriptional activation of elements may have specific role in neuronal differentiation.

References

- Akers, K. G., Martinez-Canabal, A., Restivo, L., Yiu, A. P., De Cristofaro, A., Hsiang, H. L., Wheeler, A. L., Guskjolen, A., Niibori, Y., Shoji, H., Ohira, K., Richards, B. A., Miyakawa, T., Josselyn, S. A., Frankland, P.W. (2014) Hippocampal neurogenesis regulates forgetting during adulthood and infancy. *Science* **344**, 598–602.
- 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.
- 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.
- Brown, N. P., Leroy, C., Sander, C. (1998) MView: a web-compatible database search or multiple alignment viewer. *Bioinformatics* **14**, 380–381.
- 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.

- 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.
- 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.
- Coufal, N. G., Garcia-Perez, J. L., Peng, G. E., Marchetto, M. C., 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**, 20382–20388.
- Dobigny, G., Ozouf-Costaz, C., Waters, P. D., Bonillo, C., Coutanceau, J. P., Volobouev, V. (2004) LINE-1 amplification accompanies explosive genome repatterning in rodents. *Chromosome Res.* **12**, 787–793.
- 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.
- 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., Forrester, A. R. R., Suzuki, H., Hayashizaki, Y., Hume, D. A., Orlando, V., Grimmond, S. M., Carninci, P. (2009) The regulated retrotransposons transcriptome of mammalian cells. *Nat. Genet.* **41**, 563–571.
- Furano, A. V. (2000) The biological properties and evolutionary dynamics of mammalian LINE-1 retrotransposons. *Prog. Nucleic Acid Res. Mol. Biol.* **64**, 255–294.
- Garcia-Perez, J. L., Marchetto, M. C., 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.
- Gibbs, R. A., Weinstock, G. M., Metzker, M. L., Muzny, D. M., Sodergren, E. J., Scherer, S., Scott, G., Steffen, D., Worley, K. C., Burch, P. E., Okwuonu, G., Hines, S., Lewis, L., DeRamo, C., Delgado, O., Dugan-Rocha, S., Miner, G., Morgan, M., Hawes, A., Gill, R., Celera, Holt, R. A., Adams, M. D., Amanatides, P. G., Baden-Tillson, H., Barnstead, M., Chin, S., Evans, C. A., Ferriera, S., Fosler, C., Glodek, A., Gu, Z., Jennings, D., Kraft, C. L., Nguyen, T., Pfannkoch, C. M., Sitter, C., Sutton, G. G., Venter, J. C., Woodage, T., Smith, D., Lee, H. M., Gustafson, E., Cahill, P., Kana, A., Doucette-Stamm, L., Weinstock, K., Fectel, K., Weiss, R. B., Dunn, D. M., Green, E. D., Blakesley, R. W., Bouffard, G. G., De Jong, P. J., Osoegawa, K., Zhu, B., Marra, M., Schein, J., Bosdet, I., Fjell, C., Jones, S., Krzywinski, M., Mathewson, C., Siddiqui, A., Wye, N., McPherson, J., Zhao, S., Fraser, C. M., Shetty, J., Shatsman, S., Geer, K., Chen, Y., Abramzon, S., Nierman, W. C., Havlak, P. H., Chen, R., Durbin, K. J., Egan, A., Ren, Y., Song, X. Z., Li, B., Liu, Y., Qin, X., Cawley, S., Worley, K. C., Cooney, A. J., D'Souza, L. M., Martin, K., Wu, J. Q., Gonzalez-Garay, M. L., Jackson, A. R., Kalafus, K. J., McLeod, M. P., Milosavljevic, A., Virk, D., Volkov, A., Wheeler, D. A., Zhang, Z., Bailey, J. A., Eichler, E. E., Tuzun, E., Birney, E., Mongin, E., Ureta-Vidal, A., Woodwark, C., Zdobnov, E., Bork, P., Suyama, M., Torrents, D., Alexandersson, M., Trask, B. J., Young, J. M., Huang, H., Wang, H., Xing, H., Daniels, S., Gietzen, D., Schmidt, J., Stevens, K., Vitt, U., Wingrove, J., Camara, F., Mar Albà, M., Abril, J. F., Guigo, R., Smit, A., Dubchak, I., Rubin, E. M., Couronne, O., Poliakov, A., Hübner, N., Ganten, D., Goesele, C., Hummel, O., Kreitler, T., Lee, Y. A., Monti, J., Schulz, H., Zimdahl, H., Himmelbauer, H., Lehrach, H., Jacob, H. J., Bromberg, S., Gullings-Handley, J., Jensen-Seaman, M. I., Kwitek, A. E., Lazar, J., Pasko, D., Tonellato, P. J., Twigger, S., Ponting, C. P., Duarte, J. M., Rice, S., Goodstadt, L., Beatson, S. A., Emes, R. D., Winter, E. E., Webber, C., Brandt, P., Nyakatura, G., Adetobi, M., Chiaromonte, F., Elnitski, L., Eswara, P., Hardison, R. C., Hou, M., Kolbe, D., Makova, K., Miller, W., Nekrutenko, A., Riemer, C., Schwartz, S., Taylor, J., Yang, S., Zhang, Y., Lindpaintner, K., Andrews, T. D., Caccamo, M., Clamp, M., Clarke, L., Curwen, V., Durbin, R., Eyas, E., Searle, S. M., Cooper, G. M., Batzoglou, S., Brudno, M., Sidow, A., Stone, E. A., Venter, J. C., Payseur, B. A., Bourque, G., López-Otín, C., Puente, X. S., Chakrabarti, K., Chatterji, S., Dewey, C., Pachter, L., Bray, N., Yap, V. B., Caspi, A., Tesler, G., Pevzner, P. A., Haussler, D.,

- Roskin, K. M., Baertsch, R., Clawson, H., Furey, T. S., Hinrichs, A. S., Karolchik, D., Kent, W. J., Rosenbloom, K. R., Trumbower, H., Weirauch, M., Cooper, D. N., Stenson, P. D., Ma, B., Brent, M., Arumugam, M., Shteynberg, D., Copley, R. R., Taylor, M. S., Riethman, H., Mudunuri, U., Peterson, J., Guyer, M., Felsenfeld, A., Old, S., Mockrin, S., Collins, F.; Rat Genome Sequencing Project Consortium (2004) Genome sequence of the brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493–521.
- Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T., Durbin, R., Eyras, E., Gilbert, J., Hammond, M., Huminecki, L., Kasprzyk, A., Lehvaslaiho, H., Lijnzaad, P., Melsopp, C., Mongin, E., Pettett, R., Pockock, M., Potter, S., Rust, A., Schmidt, E., Searle, S., Slater, G., Smith, J., Spooner, W., Stabenau, A., Stalker, J., Stupka, E., Ureta-Vidal, A., Vastrik, I., Clamp, M. (2002) The Ensembl genome database project. *Nucleic Acids Res.* **30**, 38–44.
- Hunter, R. G., Murakami, G., Dewell, S., Seligsohn, M., Baker, M. E., Datsun, N. A., McEwen, B. S., Pfaff, D. W. (2012) Acute stress and hippocampal histone H3 lysine 9 trimethylation, a retrotransposon silencing response. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17657–17662.
- Hunter, R. G., McEwen, B. S., Pfaff, D. W. (2013) Environmental stress and transposon transcription in the mammalian brain. *Mob. Genet. Elements* **3**, e24555.
- Kaer, K., Speek, M. (2013) Retroelements in human disease. *Gene* **518**, 231–241.
- Kokošar, J., Kordiš, D. (2013) Genesis and regulatory wiring of retroelement-derived domesticated genes: a phylogenomics perspective. *Mol. Biol. Evol.* **30**, 1015–1031.
- Lander, E. S., Linton, L. M., Birren, B. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.
- 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.
- Moran, J. V., Gilbert, N. (2002) Mammalian LINE-1 retrotransposons and related elements. In: *Mobile DNA II*. Craig, N., Craggie, R., Gellert, M., Lambowitz, A., pp. 836–869, ASM Press, Washington, D.C.
- Muotri, A. R., Chu, V. T., Marchetto, M. C., Deng, W., Moran, J. V., Gage, F. H. (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **435**, 903–910.
- Muotri, A. R., Zhao, C., Marchetto, M. C., Gage, F. H. (2009) Environmental influence on L1 retrotransposons in adult hippocampus. *Hippocampus* **19**, 1002–1007.
- Ostertag, E. M., Kazazian, H. H. Jr. (2001) Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* **35**, 501–538.
- Poduri, A., Evrony, G. D., Cai, X., Walsh, C. A. (2013) Somatic mutation, genomic variation and neurological disease. *Science* **341**, 1237758.
- Singer, T., McConnell, M. J., Marchetto, M. C., Coufal, N. G., Gage, F. H. (2010) LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci.* **33**, 345–354.
- Spalding, K. L., Bergmann, O., Alkass, K., Bernard, S., Salehpour, M., Huttner, H. B., Boström, E., Westerlund, I., Vial, C., Buchholz, B. A., Possnert, G., Mash, D. C., Druid, H., Frisén, J. (2013) Dynamics of hippocampal neurogenesis in adult humans. *Cell* **153**, 1219–1227.
- Speek, M. (2001) Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol. Cell. Biol.* **21**(6), 1973–1985.
- Terasaki, N., Goodier, J. L., Cheung, L. E., Wang, Y. J., Kajikawa, M., Kazazian, H. H. Jr., Okada, N. (2013) *In vitro* screening for compounds that enhance human L1 mobilization. *PLoS One* **8**, e74629.
- Thomas, C. A., Paquola, C., Muotri, A. R. (2012) LINE-1 retrotransposition in the nervous system. *Annu. Rev. Cell Dev. Biol.* **28**, 555–573.
- Tubio, J. M., Li, Y., Ju, Y. S., Martincorena, I., Cooke, S. L., Tojo, M., Gundem, G., Pipinikas, C. P., Zamora, J., Raine, K.,

- Menzies, A., Roman-Garcia, P., Fullam, A., Gerstung, M., Shlien, A., Tarpey, P. S., Papaemmanuil, E., Knappskog, S., Van Loo, P., Ramakrishna, M., Davies, H. R., Marshall, J., Wedge, D. C., Teague, J. W., Butler, A. P., Nik-Zainal, S., Alexandrov, L., Behjati, S., Yates, L. R., Bolli, N., Mudie, L., Hardy, C., Martin, S., McLaren, S., O'Meara, S., Anderson, E., Maddison, M., Gamble, S.; ICGC Breast Cancer Group; ICGC Bone Cancer Group; ICGC Prostate Cancer Group, Foster, C., Warren, A. Y., Whitaker, H., Brewer, D., Eeles, R., Cooper, C., Neal, D., Lynch, A. G., Visakorpi, T., Isaacs, W. B., van't Veer, L., Caldas, C., Desmedt, C., Sotiriou, C., Aparicio, S., Foekens, J. A., Eyfjörd, J. E., Lakhani, S. R., Thomas, G., Myklebost, O., Span, P. N., Børresen-Dale, A. L., Richardson, A. L., Van de Vijver, M., Vincent-Salomon, A., Van den Eynden, G. G., Flanagan, A. M., Futreal, P. A., Janes, S. M., Bova, G. S., Stratton, M. R., McDermott, U., Campbell, P. J. (2014) Extensive transduction of non-repetitions DNA mediated by L1 retrotransposition in cancer genomes. *Science* **345**, 1251343.
- van den Hurk, J. A., Meij, I. C., Seleme, M. C., Kano, H., Nikopoulos, K., Hoefsloot, L. H., Sistermans, E. A., de Wijs, I. J., Mukhopadhyay, A., Plomp, A. S., de Jong, P. T., Kazazian, H. H., Cremers, F. P. (2007) L1 retrotransposition can occur early in human embryonic development. *Hum. Mol. Genet.* **16**, 1587–1592.
- Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M. R., Brown, D. G., Brown, S. D., Bult, C., Burton, J., Butler, J., Campbell, R. D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A. T., Church, D. M., Clamp, M., Clee, C., Collins, F. S., Cook, L. L., Copley, R. R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K. D., Deri, J., Dermitzakis, E. T., Dewey, C., Dickens, N. J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D. M., Eddy, S. R., Elnitski, L., Emes, R. D., Esvara, P., Eyas, E., Felsenfeld, A., Fewell, G. A., Flicek, P., Foley, K., Frankel, W. N., Fulton, L. A., Fulton, R. S., Furey, T. S., Gage, D., Gibbs, R. A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Graffham, D., Graves, T. A., Green, E. D., Gregory, S., Guigó, R., Guyer, M., Hardison, R. C., Haussler, D., Hayashizaki, Y., Hillier, L. W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D. B., Johnson, L. S., Jones, M., Jones, T. A., Joy, A., Kamal, M., Karlsson, E. K., Karolchik, D., Kasprzyk, A., Kawai, J., Keibler, E., Kells, C., Kent, W. J., Kirby, A., Kolbe, D. L., Korf, I., Kucherlapati, R. S., Kulbokas, E. J., Kulp, D., Landers, T., Leger, J. P., Leonard, S., Letunic, I., Levine, R., Li, J., Li, M., Lloyd, C., Lucas, S., Ma, B., Maglott, D. R., Mardis, E. R., Matthews, L., Mauceli, E., Mayer, J. H., McCarthy, M., McCombie, W. R., McLaren, S., McLay, K., McPherson, J. D., Meldrim, J., Meredith, B., Mesirov, J. P., Miller, W., Miner, T. L., Mongin, E., Montgomery, K. T., Morgan, M., Mott, R., Mullikin, J. C., Muzny, D. M., Nash, W. E., Nelson, J. O., Nhan, M. N., Nicol, R., Ning, Z., Nusbaum, C., O'Connor, M. J., Okazaki, Y., Oliver, K., Overton-Larty, E., Pachter, L., Parra, G., Pepin, K. H., Peterson, J., Pevzner, P., Plumb, R., Pohl, C. S., Poliakov, A., Ponce, T. C., Ponting, C. P., Potter, S., Quail, M., Reymond, A., Roe, B. A., Roskin, K. M., Rubin, E. M., Rust, A. G., Santos, R., Sapojnikov, V., Schultz, B., Schultz, J., Schwartz, M. S., Schwartz, S., Scott, C., Seaman, S., Searle, S., Sharpe, T., Sheridan, A., Shownkeen, R., Sims, S., Singer, J. B., Slater, G., Smit, A., Smith, D. R., Spencer, B., Stabenau, A., Stange-Thomann, N., Sugnet, C., Suyama, M., Tesler, G., Thompson, J., Torrents, D., Trevaskis, E., Tromp, J., Ucla, C., Ureta-Vidal, A., Vinson, J. P., Von Niederhausern, A. C., Wade, C. M., Wall, M., Weber, R. J., Weiss, R. B., Wendl, M. C., West, A. P., Wetterstrand, K., Wheeler, R., Whelan, S., Wierzbowski, J., Willey, D., Williams, S., Wilson, R. K., Winter, E., Worley, K. C., Wyman, D., Yang, S., Yang, S. P., Zdobnov, E. M., Zody, M. C., Lander, E. S. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562.