ARTICLES

Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition

Alysson R. Muotri[1]*, Vi T. Chu[1]*†, Maria C. N. Marchetto[1], Wei Deng[1], John V. Moran[2] & Fred H. Gage[1]

Revealing the mechanisms for neuronal somatic diversification remains a central challenge for understanding individual differences in brain organization and function. Here we show that an engineered human LINE-1 (for long interspersed nuclear element-1; also known as L1) element can retrotranspose in neuronal precursors derived from rat hippocampus neural stem cells. The resulting retrotransposition events can alter the expression of neuronal genes, which, in turn, can influence neuronal cell fate *in vitro*. We further show that retrotransposition of a human L1 in transgenic mice results in neuronal somatic mosaicism. The molecular mechanism of action is probably mediated through Sox2, because a decrease in Sox2 expression during the early stages of neuronal differentiation is correlated with increases in both L1 transcription and retrotransposition. Our data therefore indicate that neuronal genomes might not be static, but some might be mosaic because of *de novo* L1 retrotransposition events.

Neural stem cells give rise to three main lineages: astrocytes, oligodendrocytes and neurons. Most of the cells from these lineages are generated during development, although many continue to be born throughout life in the adult mammalian brain[1–3]. In addition to the lineage decisions made by each cell type, many different types of cell are generated within each lineage. For example, within the neuronal lineage there are Purkinje, granule, pyramidal and basket cells, among others. Diversity also exists between individual cells within a cell type. Although incompletely understood, neuronal diversity is assumed to be governed by a combination of genetic and environmental factors[4,5]. By studying neurogenesis, one can examine early fate choices between lineages, type differences within a lineage, and diversity between individual neurons.

The glycosylated form of cystatin C (CCg) allows the propagation of multipotent neural progenitor cells (NPCs) from either single cells or polyclonal cell populations plated at low density in the presence of fibroblast growth factor-2 (FGF-2). By examining the transcription profiles generated from CCg-responsive NPCs, we discovered a class of overexpressed transcripts corresponding to L1s.

L1s are abundant retrotransposons that comprise about 20% of mammalian genomes[6–8]. Most L1s are retrotransposition-defective because they are 5′ truncated, contain internal rearrangements or harbour mutations within their open reading frames[9,10]. However, the average human genome is estimated to contain 80–100 retrotransposition-competent L1s (RC-L1s), and about 10% of these elements are classified as highly active or 'hot'[11]. By comparison, the mouse genome is estimated to contain at least 3,000 active L1s[12,13].

Here we show that an engineered human L1 harbouring a retrotransposition indicator cassette can retrotranspose in adult rat NPCs *in vitro*, and in the brains of transgenic mice *in vivo*. Retrotransposition events in rat NPCs can influence both neuronal gene expression and differentiation *in vitro*. Retrotransposition events were detected in both neurogenic and non-neurogenic areas of the adult mouse brain, indicating that retrotransposition is not peculiar to mature NPCs but also occurs during neuronal development. Expression analyses indicate that Sox2 might act to repress L1 transcription in rat adult hippocampus neural stem (HCN) cells. However, during neuronal differentiation, a decrease in Sox2 expression is correlated with a derepression of L1 transcription and an increase in L1 retrotransposition. These findings indicate that some neuronal genomes might not be static and might be influenced by *de novo* L1 retrotransposition events.

L1 transcripts are upregulated in CCg-responsive cells

CCg-responsive cells were isolated from heterogeneous HCN cells by plating them at low density in conditioned medium containing CCg (see Methods and Supplementary Methods for details). The resultant cell lines were then subjected to a microarray analysis to identify genes whose expression changed in response to culturing in CCg (Supplementary Fig. S1a and Supplementary Table S1). Surprisingly, a 1.5–2-fold enrichment of L1 transcripts in CCg-responsive cells was observed (Supplementary Fig. S1b and Supplementary Table S2). Semiquantitative reverse transcriptase polymerase chain reaction (RT–PCR) verified that L1 transcripts are enriched in CCg-responsive cells (Supplementary Fig. S1c).

Sox2 expression is inversely correlated with L1 promoter

We tested whether the promoter (that is, the 5′ untranslated region (UTR)) of a retrotransposition-competent human L1 could function in rat NPCs. The human L1 5′ UTR contains a YY1-binding site that is required for proper transcriptional initiation[14,15] as well as two Sox-binding sites[16] and a runt-domain transcription factor 3 (RUNX3)-binding site[17]. Sox proteins are expressed in a variety of tissues, including brain and testis[18], and Sox2 is expressed in both embryonic stem cells and neural progenitor populations[19]. We therefore tested the effects of Sox2 on the human L1 5′ UTR in HCN cells. Although Sox2 did not induce the expression of the reporter gene driven by the wild-type L1 5′ UTR, Sox2 overexpression stimulated the expression of the reporter gene driven by a promoter that lacks the first 100 base pairs (bp) of the L1 5′ UTR. The same trend was observed when the L1 promoter was cloned in an

[1]Laboratory of Genetics, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. [2]Department of Human Genetics and Internal Medicine, 1241 E. Catherine Street, University of Michigan Medical School, Ann Arbor, Michigan 48109-0618, USA. †Present address: Department of Cell Biology, Chemicon International, Inc., 28820 Single Oak Drive, Temecula, California 92590, USA.
*These authors contributed equally to this work.

antisense orientation with reference to the gene encoding luciferase (Supplementary Fig. S2a). Taken together, these findings indicate that factors that interact with the first 100 bp of the human L1 5′ UTR might be influenced by the Sox2 protein.

We next investigated the dynamics of L1 expression in HCN cells during neuronal differentiation. Sox2 expression was decreased during the first 24 h after the induction of differentiation and remained low for the duration of the 4-day assay (Fig. 1a). In contrast, L1 expression was stimulated up to tenfold during the first 24 h after the induction of differentiation. It then decreased steadily from day 2 to day 4, but its expression remained higher than that observed in undifferentiated cells (day 0; Fig. 1b). As expected, the neuron-specific synapsin promoter[20] was induced gradually during neuronal differentiation (Supplementary Fig. S2b). Additional experiments showed that a short interfering RNA (siRNA)-mediated 75% decrease in Sox2 protein in HCN cells led to a sixfold increase in L1 expression (Fig. 1c), whereas mutation of the Sox-binding sites in the L1 5′ UTR abolished the observed transcriptional repression by Sox2 protein (Fig. 1d). Last, chromatin immunoprecipitation (ChIP) revealed that Sox2 is associated with the endogenous rat L1 promoter in undifferentiated HCN cells (cultured with FGF-2) but is no longer associated with the rat L1 promoter 4 days after the induction of neuronal differentiation by FSK/RA (Fig. 1e). Consistently, histone deacetylase 1 (HDAC1) and methylation of histone H3 at Lys 9 (K9) (both associated with transcriptional silencing)[21] were also present in undifferentiated HCN cells, whereas acetylation of H3 at K9 and its methylation at Lys 4 (K4) (associated with transcriptional activation) were present only in differentiated cells (Fig. 1e). Together, these data indicate that Sox2 represses L1 transcription in HCN cells and that a decrease in Sox2 expression during neuronal differentiation is correlated with chromatin remodelling, allowing a transient stimulation of L1 transcription.

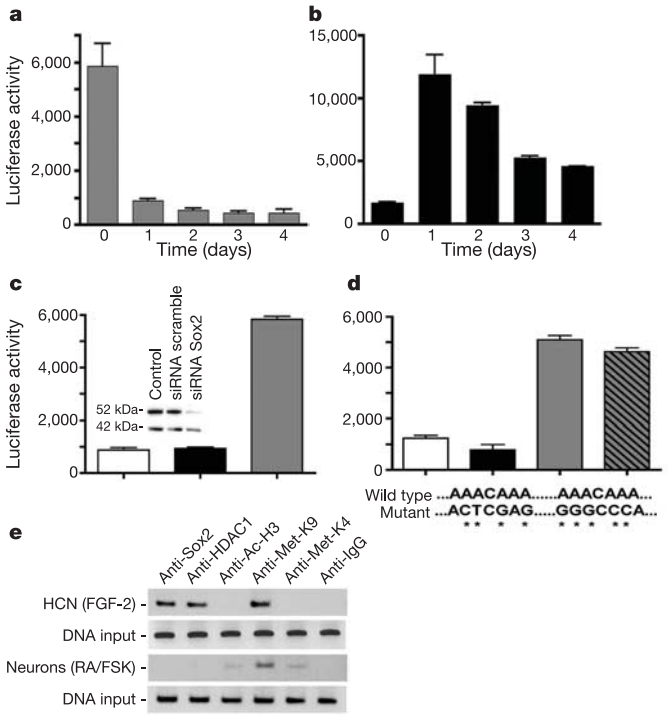

Figure 1 | L1 expression correlates with decreased *Sox2* expression in HCN cells. **a**, *Sox2* promoter is downregulated during neuronal differentiation. HCN cells were transfected with a luciferase reporter gene driven by the *Sox2* promoter (day 0). The cells were then grown in N2 medium in the absence of FGF-2 but with RA/FSK to stimulate neuronal differentiation (days 1–4). **b**, HCN cells were treated as described in **a** and the expression of a luciferase reporter gene driven by the L1 5′ UTR was followed for 4 days. **c**, SiRNA inactivation of *Sox2* transcripts correlates with increased activity in the L1 5′ UTR promoter region. HCN cells were transfected with a luciferase reporter gene driven by a wild-type L1 promoter (control; white bar) in the presence of irrelevant siRNAs (siRNA scramble; black bar) or siRNA against the *Sox2* mRNA (grey bar). Transfection of HCN cells with siRNAs against *Sox2* mRNA decreased Sox2 protein levels by about 75% (inset). **d**, Mutations in the binding sites of Sox2 protein in the L1 5′ UTR promoter region (grey bar) increased the luciferase activity in comparison with the intact L1 5′ UTR (white bar) in HCN cells. Sox2 overexpression cannot induce the wild-type L1 5′ UTR luciferase activity (black bar) but can activate the mutant L1 5′ UTR (hatched bar). **e**, Recruitment of Sox2, HDAC1 and histone H3 modification on the endogenous rat L1 promoter region during neuronal differentiation by ChIP. Extracts of formaldehyde-fixed HCN cells were precipitated with specific antibodies, either in undifferentiated cells (FGF-2) or after induction to neuronal differentiation (RA/FSK), and then analysed by PCR with primers for the L1 5′ UTR. Sox2 and HDAC1 were associated with the rat L1 promoter in undifferentiated cells only. Modifications of histone H3 associated with the L1 promoter indicate a dynamic chromatin structure, from a transcriptional silencing status (undifferentiated cells) to a transcriptional activation (neurons). Antibodies specific for methylated K9 in H3 (Met-K9), but not antibodies specific for methylated K4 (Met-K4) or acetylated K9 (Ac-H3), highly precipitated the L1 5′ UTR sequence from HCN cell extracts. In contrast, antibodies specific for either Met-K4 or Ac-H3 precipitated the L1 5′ UTR from neuronal extracts only. Error bars show s.e.m.

Rat NPCs can support human L1 retrotransposition

HCN cells are a heterogeneous population composed of Sox2-positive neural stem cells (NSCs) and Sox2-negative NPCs. To test whether L1 can retrotranspose in HCN cells, we obtained L1 expression constructs harbouring a retrotransposition indicator cassette in their respective 3′ UTRs. The indicator cassette consists of the gene encoding enhanced green fluorescence protein (EGFP) in the opposite orientation to the L1 transcript, a heterologous promoter (pCMV) and a polyadenylation signal (pA) (Fig. 2a). The EGFP gene is interrupted by an intron (IVS2 of the γ-globin gene) in the same transcriptional orientation as the L1 transcript. This arrangement ensures that EGFP-positive cells arise only when a transcript initiated from the promoter driving L1 expression is spliced, reverse-transcribed and integrated into chromosomal DNA, thereby allowing expression of the retrotransposed EGFP gene from the pCMV promoter[22,23].

Primary cultures of fresh rat adult hippocampus-derived neural progenitor (AHNP) cells, previously established HCN cells, rat primary neurons and astrocytes derived from the hippocampus, rat mesenchymal stem cells, rat fibroblasts, human $CD34^+$ lymphocytes and human HeLa cells were electroporated with either an active L1 element (LRE3-*EGFP*) or a retrotransposition-defective L1 (JM111-*EGFP*) that contained two missense mutations in open reading frame 1 (ORF1)[22–24]. Cells harbouring the L1 expression constructs were selected by the addition of puromycin to the culture medium, and puromycin-resistant cells were screened for EGFP expression by flow cytometry. After 3 days, EGFP expression was detected only in HeLa cells (1.50 ± 0.25%; data not shown). After 7 days, EGFP expression was detected in HCN cells (0.75 ± 0.07%) and AHNP cells (1.50 ± 0.50%). EGFP-positive cells were not detected in any of the other cell lines, even after several passages (Supplementary Fig. S3). PCR confirmed the presence of the retrotransposed (that is, spliced) EGFP gene in HCN-positive cells; sequencing of the PCR products confirmed the precise splicing of the intron (Fig. 2b; data not shown). Thus we conclude that a subset of cells present in the HCN and AHNP populations can support L1 retrotransposition.

To discriminate which kind of NPCs in the AHNP and HCN cell populations (that is, neuronal or glial progenitors) can support retrotransposition, individual EGFP-positive and EGFP-negative puromycin-resistant cells were collected by fluorescence-activated cell sorting (FACS) 7 days after electroporation. EGFP-positive cells

remained green during their first week in culture and often displayed a differentiated morphology when compared with cells in the initial HCN population (Fig. 2c, insets). However, EGFP expression was reduced over time and only a few cells remained EGFP-positive after 2 weeks in culture (Fig. 2c). Unexpectedly, all puromycin-resistant EGFP-negative clones also yielded a PCR product that corresponded in size to the retrotransposed EGFP gene (Fig. 2d). This observation might represent an event after FACS; a truncation of the 5′ end or the retrotransposed EGFP gene might undergo epigenetic silencing either during or soon after L1 retrotransposition (see below).

Expression of EGFP in neuronal differentiation

To test whether clones harbouring L1 retrotransposition events remained multipotent, we stimulated their differentiation into the three main neural cell types: neurons, astrocytes and oligodendrocytes (see Methods for differentiation conditions). Despite containing a retrotransposed EGFP gene, all of the clones tested at the onset of this experiment were EGFP-negative, either because they never expressed the retrotransposed EGFP gene ('negative clones') or because the retrotransposed EGFP gene was silenced during culturing in the presence of FGF-2 ('positive clones'). Under strong induction conditions, each clone tested was capable of undergoing differentiation into the three neural cells types at a similar rate to that of HCN cells (Supplementary Fig. S4a). Reactivation of the retrotransposed EGFP gene could be detected only during neuronal differentiation (Supplementary Fig. S4b). Time-lapse imaging revealed that EGFP expression began as early as 2 h after the induction of neuronal differentiation, indicating that it might have resulted from alterations in chromatin structure rather than new L1 retrotransposition events (Supplementary movie). EGFP-positive cells were rarely detected after the induction of astrocyte or oligodendrocyte differentiation, and they never co-localized with astrocyte (Gfap) or oligodendrocyte (Rip) markers, unlike the cells expressing the control CMV-driven EGFP cassette (Supplementary Fig. S4c), indicating that expression of the retrotransposed EGFP gene during neuronal differentiation in the analysed clones might be due to epigenetic modifications rather than to

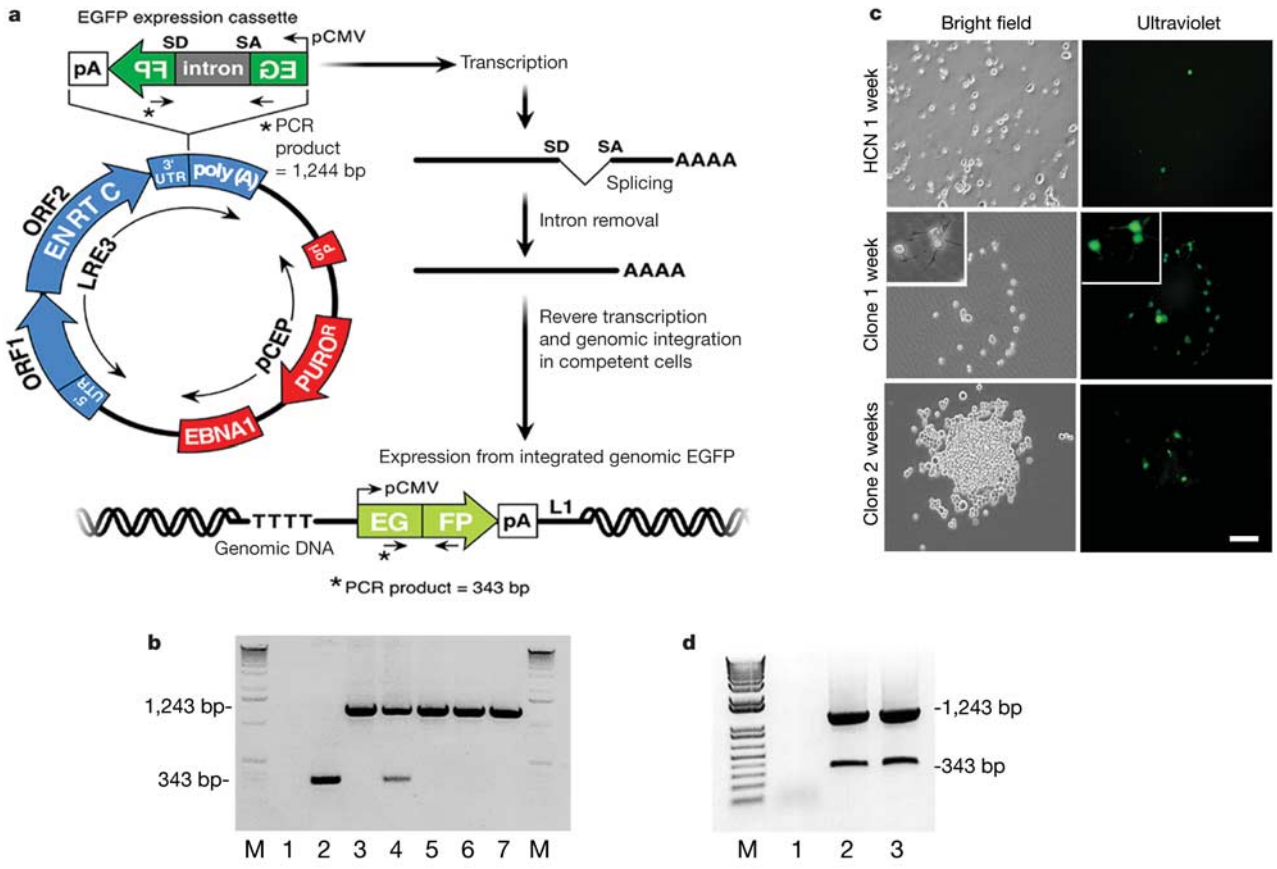

Figure 2 | NPCs can support L1 retrotransposition. **a**, Diagram of the L1-*EGFP* retrotransposition indicator construct. Retrotransposition-competent human L1s contain a 5′ UTR harbouring an internal promoter[14,15], two open reading frames (ORF1 and ORF2; blue arrows), and a 3′ UTR that ends in a poly(A) tail[6,9]. The EGFP retrotransposition indicator cassette consists of a backward copy of an EGFP gene whose expression is controlled by the CMV minimal promoter (pCMV) and the thymidine kinase polyadenylation sequence (pA). The EGFP gene is also interrupted by an intron that is in the same transcriptional orientation as the L1. This arrangement ensures that EGFP expression will become activated only on L1 retrotransposition (see the text for details)[22,23]. PCR primers flanking the intron in the EGFP gene are indicated at the bottom. The splice donor (SD) and splice acceptor (SA) sites of the intron are indicated. **b**, PCR analysis of genomic DNA isolated from different cell populations transfected with LRE3-*EGFP*. PCR was conducted with the primers shown in **a**. The 1,243-bp PCR product corresponds to the original L1 vector harbouring the intron-containing EGFP indicator cassette. The 343-bp PCR product, diagnostic for the loss of the intron, indicates a retrotransposition event (lane 4). Lane M, molecular mass standards; lane 1, water; lane 2, positive control for the EGFP gene lacking an intron; lane 3, positive control for the EGFP gene containing the intron; lane 4, NPCs, 7 days after transfection; lane 5, NPCs, 3 days after transfection; lane 6, mesenchymal stem cells, 7 days after transfection; lane 7, fibroblast cells, 7 days after transfection. **c**, Despite containing a retrotransposed EGFP gene, cells that were initially EGFP-positive can no longer express EGFP. **d**, Puromycin-resistant clones that never expressed EGFP harbour L1 retrotransposition events, as indicated by the presence of the 343-bp PCR product. Lane M, molecular mass standards; lane 1, water; lane 2, $EGFP^+$ clone; lane 3, $EGFP^-$ clone.

neural cell-type-specific differences in pCMV promoter specificity.

To gain insights into the kinds of chromatin alterations that could be involved in EGFP gene reactivation during neuronal differentiation, we treated the clones with either 500 ng ml^{-1} 5-azacytidine (5-Aza; a demethylating agent) for 4 days or 1 μM trichostatin A (TsA; an inhibitor of histone deacetylation activity) for 1 day in N2 medium. Both treatments caused an increase in EGFP expression in the different clones assayed. Moreover, immunostaining revealed that most of the EGFP-positive cells (63% in the TsA treatment and 30% in the 5-Aza treatment) co-localized with Map2(a + b) but not with Gfap (Supplementary Fig. S4d–f), which is consistent with recent reports showing that epigenetic modifications accompany neuronal differentiation of NSCs[21,25].

L1s can insert into neuronally expressed genes

We next used inverse PCR[26] to characterize the post-integration sites from 17 clones (independent of the initial EGFP status to avoid expression bias) harbouring detectable L1 retrotransposition events by PCR. Interestingly, some of the insertions (clones B2, 22, 2, 10, 6 and 5; Table 1) occurred within neuronally expressed genes (see Supplementary Information). Comparison of the post-integration sequence with the pre-integration site present in the University of California Santa Cruz genome database revealed that the characterized retrotransposition events occurred by conventional endonuclease-dependent retrotransposition[26] (Supplementary Information). Curiously, clones 2 and 5 contained independent L1 retrotransposition events into the *Slc6a6* gene (Table 1). Although the number of characterized insertions is relatively small, our data indicate that L1 retrotransposition events can occur in neuronally expressed genes. The ability of L1 to retrotranspose into genes is consistent with results from previous studies performed in transformed cultured cell lines[27–29].

Retrotransposition events in NPCs can affect cell fate

In a mixed differentiation protocol (Supplementary Fig. S4a), most HCN clones harbouring detectable L1 retrotransposition events had a tendency to differentiate into neurons rather than glial cells (Table 1). To test whether new L1 retrotransposition events could influence the phenotype of the NPCs *in vitro*, we fully characterized the insertion site present in clone 22 (Cl 22). The L1 was inserted, in the antisense orientation, into the 5′ UTR of the rat neuronal *chapsyn110*/*Psd-93* gene (Fig. 3a). RT–PCR with primers specific for *Psd-93* revealed that RNA levels were about tenfold higher in Cl 22 cells than in the original HCN population. Moreover, *Psd-93* was not overexpressed in Cl 28, which contains two different L1 insertions (Fig. 3b). The Psd-93 protein level was also 3–5-fold higher in Cl 22, and the increase in Psd-93 protein level was correlated with a decrease in Sox2 expression (Fig. 3c).

Table 1 | Analyses of L1 insertions in HCN cell-derived clones

Clone	Initial EGFP expression	Phenotype	Locus	L1 insertional target site description
HCN	Negative		–	–
B2	Negative		1q43	Inside olfactory receptor *Olr346*
3	Positive		2q12	Inside *Dhfr-1* (dehidrofolate reductase-1)
28	Positive		6q24 Xq22	Inside pseudogene similar to human glyceraldehydes-3-phosphate dehydrogenase (*G3PDH*) Inside Loc317538, similar to dystrophin major muscle isoform
B9	Positive		3q34	150 kb distant from gene similar to brain type ryanodine receptor 3 (*RyR3*), a Ca^{2+} release channel, and 55 kb distant from gene similar to the Notchless WD repeat
22	Positive		1q32	Inside *Chapsyn110*, postsynaptic protein 93 (*Psd-93*)
7	Positive		18q12.3	25 kb distant from similar to human *G3PDH*
2	Positive		4q34	Inside Slc6a6 (solute carrier family 6), neurotransmitter transporter, taurine
F10	Negative	n.d.	1q31	2 kb distant from a number of predicted genes with unknown function
9	Negative	n.d.	19q11	Inside Loc364967, predicted protein similar to spermatogenesis associated glutamate-rich protein 4f from *Mus musculus*
6	Negative		6q16	Inside *NVP-3*, neural visinin-like Ca^{2+}-binding protein type 3
5	Positive		4q34	Inside Slc6a6 (solute carrier family 6), neurotransmitter transporter, taurine
26	Positive		2q12	60 kb distant from gene similar to *Xrcc4* DSB repair
13	Positive		3p13	20 kb distant from *Nidogen 2*, a base membrane protein involved in cell adhesion
21	Positive		2q32	50 kb distant from hypothetical kazal type serine protease inhibitor domain, with EF-hand Ca^{2+} binding and N-CAM L1 motif
B5	Negative	n.d.	19q11	Inside hypothetical protein FLJ10846-like, with unknown function, and 30 kb distant from CG9882-PA with TPR and ribosomal protein L11 methyltransferase domains
10	Positive		18p11	Inside *rCNR* gene, a cadherin-related neuronal receptor
1	Positive		7q33	Inside *LRTP*, similar to mouse testis specific protein with leucine-rich repeat domain and involved in regulation of protein phosphatases

Phenotypes were determined by exposing the clones to mixed differentiation (RA/FBS) medium, and L1 target sites were determined by inverse PCR (see Methods). Percentages of oligodendrocytes in these clones were too low to display or nonexistent (data not shown). White, undifferentiated; black, neurons (Map2(a + b)); grey, astrocytes (Gfap). n.d., not determined.

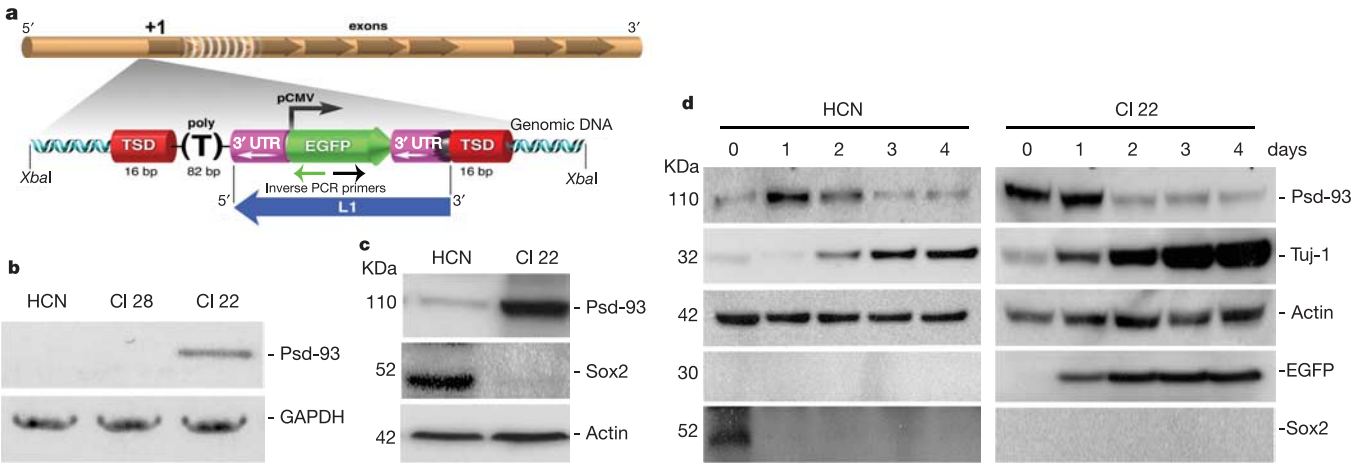

Figure 3 | L1 retrotransposition events can modify neuronal gene expression. **a**, Schematic representation of an L1 insertion into the *Chapsyn*110/*Psd-93* gene. The primers used in the inversed PCR experiments are indicated by the green and black arrows. **b**, RT–PCR showed that *Psd-93* expression is higher in the clone harbouring the L1 insertion in the *Psd-93* gene (Cl 22) than in naive HCN cells, or another cell line harbouring L1 insertions at different loci (Cl 28). GAPDH, glyceraldehydes-3-phosphate dehydrogenase. **c**, Western blot analysis revealed a higher expression of Psd-93 protein but a lower expression of Sox2 protein in Cl 22 cells than in naive HCN cells. **d**, Kinetic analysis of *Psd-93* expression during neuronal differentiation. In HCN cells, *Psd-93* was upregulated during early stages of neuronal differentiation (day 1) and then decreased over the course of the 4-day assay. By comparison, in Cl 22, *Psd-93* was expressed initially at high levels (day 0) and was upregulated to a smaller extent on neuronal differentiation (day 1). However, the decrease in expression mirrored that seen in naive HCN cells (days 2–4). As controls, we monitored the expression of actin, EGFP, Tuj-1 and Sox2 protein levels in both cell lines.

We next examined the expression of *Psd-93* during neuronal differentiation. In HCN cells, Psd-93 protein was induced during the early stages (24 h) of neuronal differentiation, at a similar time as the retrotransposed EGFP gene was being reactivated and Sox2 was downregulated. Psd-93 protein levels then decreased during neuronal maturation until they reached a baseline expression level on day 4 of the assay (Fig. 3d). By comparison, in Cl 22 cells, the initial level of Psd-93 protein was higher than that observed in HCN cells and, after differentiation, was induced to a smaller extent than in HCN cells. However, the downregulation of the Psd-93 protein after 48 h was similar to that observed in HCN cells (Fig. 3d). Thus, despite its initial overexpression, further regulation of the *Psd-93* gene activity in Cl 22 remained unaffected, indicating that the entire locus might have been subject to higher-order regulation.

To verify whether the initial overexpression of *Psd-93* in Cl 22 affected neuronal differentiation, we used specific siRNAs to

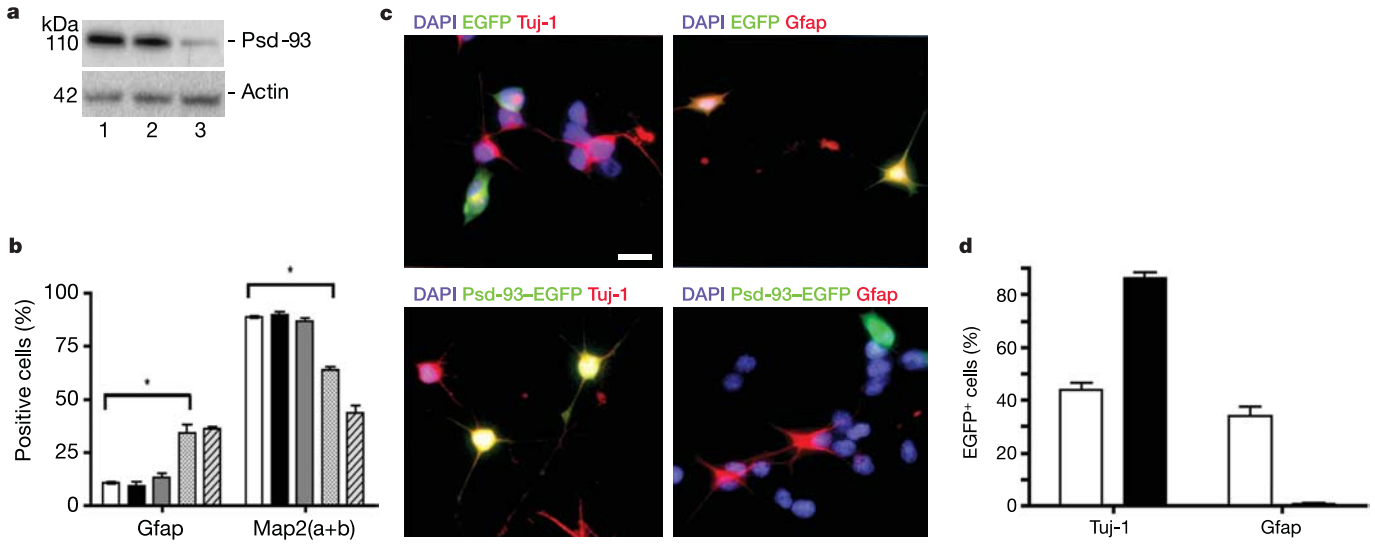

Figure 4 | An L1 retrotransposition event can drive neuronal maturation through *Psd-93* overexpression. **a**, siRNA against *Psd-93* transcripts (lane 3) results in a 70% decrease in Psd-93 protein levels in Cl 22 cells (lane 2, without siRNA). Control cells (lane 1) were treated with non-specific siRNAs. **b**, Both Cl 22 cells (white bars) and Cl 22 cells treated with irrelevant siRNAs (black and grey bars) have a tendency to differentiate into neurons when plated in mixed (RA/FBS) differentiation medium. By comparison, lower levels of *Psd-93* transcripts (stippled bars) attenuated the strong neuronal bias of Cl 22 cells to that observed in naive HCN cells (hatched bars) (*$P = 0.0056$). **c**, *Psd-93* overexpression induced neuronal differentiation/maturation of HCN cells plated in RA/FBS. The Psd-93–EGFP fusion protein co-localizes with a neuronal marker (Tuj-1) but not with an astrocyte marker (Gfap; shown in red). EGFP expression is shown in green. Co-localization of Psd-93–Egfp and Tuj-1 is shown in the lower left panel. DAPI, 4,6-diamidino-2-phenylindole. Scale bar, 10 μm. **d**, Quantification of panel **c**. EGFP-positive (fused (black bars) or not (white bars) with *Psd-93*) cells were scored for their co-localization with Tuj-1 or Gfap markers. The percentage of each cell type is indicated on the *y* axis. Error bars in all panels show s.e.m.

transiently downregulate *Psd-93* mRNA transcripts (Fig. 4a). Remarkably, 4 days after electroporation, the resultant Cl 22 cells seemed less differentiated and showed a decrease in cellular processes (data not shown), and protein levels more closely resembled those of naive HCN cells (Fig. 4b). Consistent with this notion was our observation that overexpression of a *Psd-93–egfp* fusion protein in HCN cells was able to induce neuronal differentiation under the mixed condition RA/FBS (Fig. 4c, d). Together, these observations indicate that retrotransposition of an engineered human L1 into this gene can cause an alteration in gene expression, which results in a phenotypic change in cell behaviour.

L1 somatic retrotransposition in mouse brain

To determine whether L1 retrotransposition can occur *in vivo*, we generated a transgenic animal harbouring a retrotransposition-competent human L1 element ($L1_{RP}$) under the control of its endogenous promoter[23,30]. The EGFP reporter gene is under control of the CMV promoter; thus, the retrotransposed EGFP gene has the potential to be expressed ubiquitously in mice. We obtained two independent founders (Fo4 and Fo6) containing the $L1_{RP}$/EGFP transgene. They were bred with wild-type C57BL/6J mice, and the resultant progeny were screened for L1 retrotransposition events by genotyping tail DNA with PCR primers flanking the intron present in the EGFP retrotransposition indicator cassette[23]. Both founders were positive for the transgene. Figure 5a is a diagram of the progeny of Fo4. Animal 1 lacked both the $L1_{RP}$/EGFP transgene and a retrotransposed EGFP gene. Animals 2 and 3 contained only the $L1_{RP}$/EGFP transgene. Animal 4 lacked the $L1_{RP}$/EGFP transgene but contained a retrotransposed EGFP gene. Animal 5 contained both the $L1_{RP}$/EGFP transgene and the retrotransposed EGFP gene.

The progeny (a total of seven) containing only the $L1_{RP}$/EGFP transgene from both founders (for example animals 2 and 3; Fig. 5a) were selected for further analysis. We killed adult animals and used anti-EGFP antibodies to detect L1 retrotransposition events by immunofluorescence microscopy in tissues from different organs. EGFP-positive cells were detected in germ cells (ovary and testes, previously shown to express L1 ORF1[31]) but not in other somatic tissues (Supplementary Figs S5 and S6). EGFP-positive cells also were found in the brains of both male and female transgenic animals (for example striatum, cortex, hypothalamus, hilus, cerebellum, ventricles, amygdala and hippocampus; Fig. 5c–k). EGFP-positive cells co-localized only with a neuronal marker (NeuN) and not with oligodendrocyte (glutathione *S*-transferase π, GSTπ) or astrocyte (S100-β) markers, indicating that L1 retrotransposition might have occurred in neuronal precursor cells rather than glial precursor cells or a common precursor cell early during embryogenesis (Fig. 5i–k). Because the migration and maturation of the different brain EGFP-positive cell types occurred in distinct regions in the wild-type brain, L1 retrotransposition most probably occurred in several distinct NPCs at different times during brain development.

Analysis of embryos at embryonic day 8.5 (E8.5) showed no EGFP-positive cells. At this stage the neural tube is already defined and separated from tissues with the same embryonic origin (such as the skin). By comparison, EGFP-positive cells were detected in the

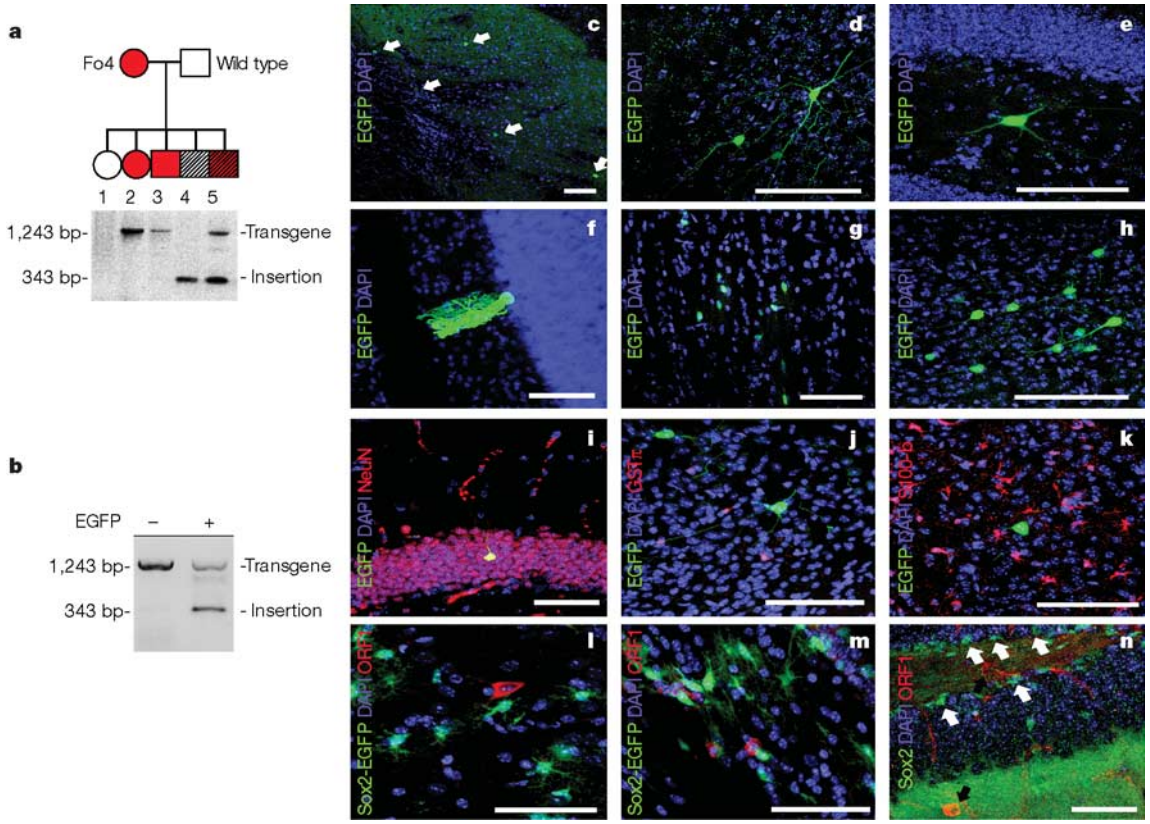

Figure 5 | L1 retrotransposition detection in the brains of transgenic mice. a, Genotyping results on tail DNA from offspring of founder Fo4 (animals 1–5). The PCR primers detected the two possible EGFP configurations[33]. Red circles (female) and red squares (male) indicate animals containing the $L1_{RP}$/EGFP transgene; shading indicates animals containing the retrotransposed EGFP gene. **b**, PCR on genomic DNAs isolated from laser-captured cells. About 50 cells were collected in each population (EGFP-positive and EGFP-negative) from the brains of two animals harbouring the $L1_{RP}$/EGFP transgene. EGFP-positive cells were detected in different regions of the mouse brain, such as striatum (**c**, **l**) (white arrows in **c** indicate EGFP-positive cells), cortex (**d**), hypothalamus (**k**), hilus (**e**), cerebellum (**f**), ventricles (**g**, **l**, **m**), amygdala (**h**) and hippocampus (**i**, **n**). ORF1-positive cells were also found in different regions of the brain, such as the ventricular zone (**l**, **m**) and the dentate gyrus of the hippocampus (**n**; white arrows indicate Sox2-positive cells, black arrows indicate ORF1-positive cells). Scale bar, 10 μm.

cephalic neural tube but not in other regions or organs of the E10.5 embryo. Taken together, these results indicate that L1 retrotransposition might take place during both embryonic and adult neurogenesis (Supplementary Fig. S5).

To confirm that the brain cells contained L1 retrotransposition events, we next isolated EGFP-positive and EGFP-negative cells by laser-capture microscopy and subjected genomic DNAs extracted from those cells to PCR analysis (Fig. 5b). EGFP-positive cells contained both the $L1_{RP}$/EGFP transgene and the retrotransposed EGFP gene. Sequence analysis confirmed that the 343-bp PCR product is the precisely spliced EGFP gene (data not shown). EGFP-negative cells contained the $L1_{RP}$/EGFP transgene and a very faint PCR product corresponding to the spliced EGFP gene. This faint amplicon might represent silenced EGFP-gene insertions from mature neurons.

Finally, consistent with the hypothesis that Sox2 acts to repress L1 expression in NSCs and that a decrease in Sox2 expression in NPCs is correlated with an increase in L1 expression and retrotransposition, our initial experiments show that anti-Sox2 antibodies do not co-localize with cells that stain positively with a mouse anti-ORF1p antibody in Sox2-EGFP transgenic animals[32] (Fig. 5l, m) or wild-type C57BL/6J brain sections (see Fig. 5n, for example).

Discussion

Previous studies have indicated that L1 retrotransposition can occur in germ cells or in early embryogenesis, before the germ line becomes a distinct lineage[33,34], whereas a cultured cell retrotransposition assay has revealed that human and mouse L1 elements can retrotranspose in a variety of transformed or immortalized cultured cell lines[22,26,35]. There is only one documented example of a genuine somatic retrotransposition event *in vivo*, and it occurred into the adenomatous polyposis coli gene of a colorectal tumour[36]. Our data show that rat NPCs can support retrotransposition of an engineered human L1 and that a human L1 element can undergo somatic retrotransposition in the mouse brain. We were unable to detect L1 insertions in other somatic tissues, indicating that the frequency of L1 retrotransposition might be lower in those tissues than in NPCs.

Two previous reports documented the retrotransposition of an engineered human L1/EGFP transgene in mice. Ostertag *et al.* demonstrated that a human L1 element could retrotranspose in the male germline before the onset of meiosis II[33]. However, because the sperm-specific acrosin promoter drove the expression of the retrotransposed EGFP gene, it is unlikely that these authors would have been able to detect the expression of the EGFP protein in the brain. Prak *et al.*[34] used the same $L1_{RP}$/EGFP transgene described in this study, and the authors detected EGFP-positive cells in the testes but not in the brain or other tissues. However, despite detecting a low level of *EGFP* mRNA expression in a variety of different tissues by RT–PCR[34], the authors only analysed EGFP expression directly; they did not use anti-EGFP antibodies to detect protein expression.

Our expression analyses lead us to propose that L1 retrotransposons are silenced in NSCs growing in FGF-2 owing to *Sox2* repression. Downregulation of *Sox2* triggered by CCg might lead to chromatin modifications. Indeed, our inability to find overexpressed neuronal genes in NPCs, coupled with the finding that all the clones selected for retrotransposition remain multipotent, indicates that L1 can retrotranspose during early stages of neuronal differentiation.

Our results are consistent with previous analyses that retrotransposition events generated from an engineered human L1 often insert into genes[27–29]. Indeed, we found that some L1 retrotransposition events in NPCs integrated into genes that are expressed in neurons, whereas other insertions were located in gene 'deserts' and/or repetitive sequences. In one instance, we have shown that retrotransposition of an engineered human L1 into the *Psd-93* gene can lead to its overexpression, which influences the differentiation pattern of the NPCs. These data provide proof in principle that new L1 retrotransposition events can affect the expression of neuronal genes *in vitro*. As indicated from previous analyses[35,37], we predict that L1 insertions into genes may also repress their expression or may lead to alternative splicing patterns[38].

Thus, our findings indicate that an engineered human L1 can retrotranspose in rat NPCs and indicate that individual neurons might be mosaic with respect to L1 content. Future experiments will focus on whether endogenous L1s naturally retrotranspose in NPCs and whether this process has any developmental significance. However, with those caveats being clearly stated, it is tempting to speculate that some of the genomic changes necessary for the uniqueness of individuals within a population, as defined by their neural circuitry, might be driven, in part, by the activities of mobile elements.

METHODS

Cell culture and transfection. HCN-A94 and freshly isolated AHNPCs cells were prepared and cultured as described[39]. Plasmid transfections were performed with Fugene6 (Roche) for luciferase assay and with a rat NSC nucleofector electroporation kit in all other experiments in accordance with the manufacturer's instructions (Amaxa Biosystem).

Luciferase assay and siRNA sequences. Luciferase activity was measured with the Dual-Luciferase reporter assay system (Promega) in accordance with the manufacturer's protocol. The L1 5′ UTR-Sox2 mutant plasmid was a gift from J. Athanikar. The *Sox2* promoter and complementary DNA were gifts from A. Rizzino. The Synapsin-1 promoter region was a gift from G. Thiel. The Psd-93–*egfp* plasmid was a gift from D. Bredt.

ChIP. The ChIP assay was performed essentially in accordance with the manufacturer's protocol with a ChIP assay kit (Upstate) and primers for the rat L1 5′ UTR promoter region.

Immunofluorescence and immunoblotting. Immunofluorescence was performed as described previously[40]. The mouse polyclonal ORF1 antibody was a gift from S. L. Martin. Western blotting was performed with standard protocols.

Expression profile analyses. To normalize for processing errors, each replicate consisted of cells that were plated at the same time from the same starting bulk population. The following cells were profiled: CCg-responsive cells, neurons and astrocytes (twenty 96-well plates). The resulting cRNA from each sample was hybridized to DNA microarrays (Affymetrix Rat Genome RG-U34A). Data analysis was performed with Affymetrix MAS4.0 software to search for transcripts differentially expressed in CCg-responsive NPCs in comparison with proliferating bulk progenitors and differentiated cells. The results were analysed with three different data mining tools: Bullfrog, an empirically based filtering algorithm[41], dChip software for model-based expression analysis[42], and Felix Naef's statistically based perfect-match-only algorithm[43].

RT–PCR analysis. Total RNA was prepared with an RNAeasy kit (Qiagen). First-strand cDNA synthesis was performed with the Superscript II kit (Invitrogen). PCR was performed with *Taq* polymerase (BMH).

Retrotransposition assays. Antibiotic selection (puromycin, $1\,\mu g\,ml^{-1}$) was begun 48 h after electroporation. After 7 days, transfected puromycin-resistant cells were analysed with a Becton Dickenson FACStar Plus containing a blue argon laser (488 nm) and fluorescein filter sets (530/30 nm bandpass). The EGFP PCR primers used here were as described previously[23]. The L1 retrotransposition cassettes were gifts from H. H. Kazazian Jr.

Inverse PCR. Genomic DNAs were digested with *Ssp*I or *Xba*I, extracted with phenol and then chloroform, and subjected to overnight ligation. The products then were re-extracted, ethanol-precipitated and subjected to first-round PCR amplification with the primers for the EGFP expression cassette. The L1 pre-integration sequence was identified with Blast (http://www.ncbi.nlm.nih.gov/BLAST/) and the Celera database (http://www.celeradiscoverysystem.com/).

L1 transgenic animals. Transgenic mice were generated with the standard pronuclear injection protocol[44]. The potential founders were screened by PCR by using primers described previously[23]. The $L1_{RP}$/EGFP construct was a gift from H. H. Kazazian Jr.

Tissue preparation. Animals were killed with an overdose of anaesthetics and perfused transcardially with 4% paraformaldehyde in phosphate buffer (0.1 M, pH 7.4). Brains were cut coronally (40 μm) on a sliding microtome from a solid-CO_2-cooled block before use.

Other methods. More details can be found in the Supplementary Methods.

Received 7 March; accepted 20 April 2005.

1. Lie, D. C., Song, H., Colamarino, S. A., Ming, G. L. & Gage, F. H. Neurogenesis in the adult brain: new strategies for central nervous system diseases. *Annu. Rev. Pharmacol. Toxicol.* **44**, 399–421 (2004).

2. Taupin, P. & Gage, F. H. Adult neurogenesis and neural stem cells of the central nervous system in mammals. *J. Neurosci. Res.* **69,** 745–749 (2002).
3. Gage, F. H. Mammalian neural stem cells. *Science* **287,** 1433–1438 (2000).
4. Kempermann, G., Kuhn, H. G. & Gage, F. H. More hippocampal neurons in adult mice living in an enriched environment. *Nature* **386,** 493–495 (1997).
5. van Praag, H., Christie, B. R., Sejnowski, T. J. & Gage, F. H. Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proc. Natl Acad. Sci. USA* **96,** 13427–13431 (1999).
6. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).
7. Gibbs, R. A. *et al.* Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428,** 493–521 (2004).
8. Waterston, R. H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420,** 520–562 (2002).
9. Grimaldi, G., Skowronski, J. & Singer, M. F. Defining the beginning and end of KpnI family segments. *EMBO J.* **3,** 1753–1759 (1984).
10. Moran, J. V. & Gilbert, N. in *Mobile DNA II* (eds Craig, N., Craggie, R., Gellert, M. & Lambowitz, A.) 836–869 (ASM Press, Washington DC, 2002).
11. Brouha, B. *et al.* Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl Acad. Sci. USA* **100,** 5280–5285 (2003).
12. Goodier, J. L., Ostertag, E. M., Du, K. & Kazazian, H. H. Jr A novel active L1 retrotransposon subfamily in the mouse. *Genome Res.* **11,** 1677–1685 (2001).
13. DeBerardinis, R. J., Goodier, J. L., Ostertag, E. M. & Kazazian, H. H. Jr Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nature Genet.* **20,** 288–290 (1998).
14. Swergold, G. D. Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* **10,** 6718–6729 (1990).
15. Athanikar, J. N., Badge, R. M. & Moran, J. V. A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucleic Acids Res.* **32,** 3846–3855 (2004).
16. Tchenio, T., Casella, J. F. & Heidmann, T. Members of the SRY family regulate the human LINE retrotransposons. *Nucleic Acids Res.* **28,** 411–415 (2000).
17. Yang, N., Zhang, L., Zhang, Y. & Kazazian, H. H. Jr. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res.* **31,** 4929–4940 (2003).
18. Wegner, M. From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res.* **27,** 1409–1420 (1999).
19. Zappone, M. V. *et al.* Sox2 regulatory sequences direct expression of a β-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. *Development* **127,** 2367–2382 (2000).
20. Lietz, M., Hohl, M. & Thiel, G. RE-1 silencing transcription factor (REST) regulates human synaptophysin gene transcription through an intronic sequence-specific DNA-binding site. *Eur. J. Biochem.* **270,** 2–9 (2003).
21. Hsieh, J. & Gage, F. H. Epigenetic control of neural stem cell fate. *Curr. Opin. Genet. Dev.* **14,** 461–469 (2004).
22. Moran, J. V. *et al.* High frequency retrotransposition in cultured mammalian cells. *Cell* **87,** 917–927 (1996).
23. Ostertag, E. M., Prak, E. T., DeBerardinis, R. J., Moran, J. V. & Kazazian, H. H. Jr Determination of L1 retrotransposition kinetics in cultured cells. *Nucleic Acids Res.* **28,** 1418–1423 (2000).
24. Brouha, B. *et al.* Evidence consistent with human L1 retrotransposition in maternal meiosis I. *Am. J. Hum. Genet.* **71,** 327–336 (2002).
25. Hsieh, J., Nakashima, K., Kuwabara, T., Mejia, E. & Gage, F. H. Histone deacetylase inhibition-mediated neuronal differentiation of multipotent adult neural progenitor cells. *Proc. Natl Acad. Sci. USA* **101,** 16659–16664 (2004).
26. Morrish, T. A. *et al.* DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. *Nature Genet.* **31,** 159–165 (2002).
27. Moran, J. V., DeBerardinis, R. J. & Kazazian, H. H. Jr Exon shuffling by L1 retrotransposition. *Science* **283,** 1530–1534 (1999).
28. Gilbert, N., Lutz-Prigge, S. & Moran, J. V. Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110,** 315–325 (2002).
29. Symer, D. E. *et al.* Human l1 retrotransposition is associated with genetic instability *in vivo*. *Cell* **110,** 327–338 (2002).
30. Schwahn, U. *et al.* Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nature Genet.* **19,** 327–332 (1998).
31. Trelogan, S. A. & Martin, S. L. Tightly regulated, developmentally specific expression of the first open reading frame from LINE-1 during mouse embryogenesis. *Proc. Natl Acad. Sci. USA* **92,** 1520–1524 (1995).
32. D'Amour, K. A. & Gage, F. H. Genetic and functional differences between multipotent neural and pluripotent embryonic stem cells. *Proc. Natl Acad. Sci. USA* **100(** Suppl 1), 11866–11872 (2003).
33. Ostertag, E. M. *et al.* A mouse model of human L1 retrotransposition. *Nature Genet.* **32,** 655–660 (2002).
34. Prak, E. T., Dodson, A. W., Farkash, E. A. & Kazazian, H. H. Jr. Tracking an embryonic L1 retrotransposition event. *Proc. Natl Acad. Sci. USA* **100,** 1832–1837 (2003).
35. Han, J. S., Szak, S. T. & Boeke, J. D. Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. *Nature* **429,** 268–274 (2004).
36. Miki, Y. *et al.* Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* **52,** 643–645 (1992).
37. Perepelitsa-Belancio, V. & Deininger, P. RNA truncation by premature polyadenylation attenuates human mobile element activity. *Nature Genet.* **35,** 363–366 (2003).
38. Ostertag, E. M. & Kazazian, H. H. Jr. Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.* **35,** 501–538 (2001).
39. Gage, F. H., Ray, J. & Fisher, L. J. Isolation, characterization, and use of stem cells from the CNS. *Annu. Rev. Neurosci.* **18,** 159–192 (1995).
40. Gage, F. H. *et al.* Survival and differentiation of adult neuronal progenitor cells transplanted to the adult brain. *Proc. Natl Acad. Sci. USA* **92,** 11879–11883 (1995).
41. Zapala, M. A., Lockhart, D. J., Pankratz, D. G., Garcia, A. J. & Barlow, C. Software and methods for oligonucleotide and cDNA array data analysis. *Genome Biol.* **3,** SOFTWARE0001.1–0001.9 (2002).
42. Li, C. & Wong, W. H. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl Acad. Sci. USA* **98,** 31–36 (2001).
43. Naef, F., Hacker, C. R., Patil, N. & Magnasco, M. Empirical characterization of the expression ratio noise structure in high-density oligonucleotide arrays. *Genome Biol.* **3,** RESEARCH0018.1–0018.11 (2002).
44. Hogan, J., Beddington, R., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. L. Gage for editorial comments, H. Suh for Sox2-EGFP brain sections and embryo advice, P. Taupin for assistance during CCg experiments, and J. L. Garcia-Perez and R. Badge for critical comments on the manuscript. A.R.M. is a Pew Latin-America Fellow. V.T.C. was supported by grants from Lynn and Edward Streim and the Neuroplasticity of Aging Training Grant. J.V.M. was supported by grants from the National Institutes of Health and the W. M. Keck Foundation, and F.H.G. was supported by the Lookout Fund, The Christopher Reeve Paralysis Foundation, Max Planck Research Award Program, by the German Federal Ministry for Education, Science, Research and Technology and the National Institutes of Health: National Institute on Aging and National Institute of Neurological Disease and Stroke.

Author Contributions A.R.M. is the leading author. He contributed to the concept, designed, performed the experiments and analysed the data, and wrote the manuscript. V.T.C. designed and performed the microarrays experiments. M.C.N.M. designed, performed and analysed the inverse PCR data and some tissue culture experiments and revised the manuscript. W.D. performed the transgenic experiment. J.V.M. contributed reagents, and performed data analyses and manuscript revision. F.H.G. is the senior author. He contributed to the concept, analysed the data, revised the manuscript and provided financial support.

Author Information Microarray data have been deposited in the GEO archive under accession number GSE2499, and the Cl22 L1 insertion sequence has been deposited in GenBank under accession number AY995186. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.H.G. (gage@salk.edu).