Dynamic Methylation of an L1 Transduction Family during Reprogramming and Neurodifferentiation

Carmen Salvador-Palomeque,a Francisco J. Sanchez-Luque,a,b Patrick R. J. Fortuna,c Adam D. Ewing,a Ernst J. Wolvetang,c Sandra R. Richardson,a Geoffrey J. Faulknera,d,e

a Mater Research Institute, University of Queensland, Woolloongabba, Queensland, Australia
b Pfizer-University of Granada-Andalusian Government Center for Genomics and Oncological Research, Granada, Spain
c Australian Institute for Bioengineering and Nanotechnology, University of Queensland, St. Lucia, Queensland, Australia
d Queensland Brain Institute, University of Queensland, Brisbane, Queensland, Australia
e School of Biomedical Sciences, University of Queensland, Brisbane, Queensland, Australia

ABSTRACT The retrotransposon LINE-1 (L1) is a significant source of endogenous mutagenesis in humans. In each individual genome, a few retrotransposition-competent L1s (RC-L1s) can generate new heritable L1 insertions in the early embryo, primordial germ line, and germ cells. L1 retrotransposition can also occur in the neuronal lineage and cause somatic mosaicism. Although DNA methylation mediates L1 promoter repression, the temporal pattern of methylation applied to individual RC-L1s during neurogenesis is unclear. Here, we identified a de novo L1 insertion in a human induced pluripotent stem cell (hiPSC) line via retrotransposon capture sequencing (RC-seq). The L1 insertion was full-length and carried 5′ and 3′ transductions. The corresponding donor RC-L1 was part of a large and recently active L1 transduction family and was highly mobile in a cultured-cell L1 retrotransposition reporter assay. Notably, we observed distinct and dynamic DNA methylation profiles for the de novo L1 and members of its extended transduction family during neuronal differentiation. These experiments reveal how a de novo L1 insertion in a pluripotent stem cell is rapidly recognized and repressed, albeit incompletely, by the host genome during neurodifferentiation, while retaining potential for further retrotransposition.

KEYWORDS L1, LINE-1, methylation, neurogenesis, retrotransposon

LINE-1 (L1) retrotransposons are mobile genetic elements that occupy nearly 20% of the human genome (1) and are an endogenous source of mutagenesis (2). Approximately 100 retrotransposition-competent L1s (RC-L1s) are found in each individual, while the remaining ~500,000 L1 copies are immobile due to 5′ truncations, inversions, deletions, and other mutations (3, 4). Almost all RC-L1s belong to the L1-Ta subfamily (5). L1 retrotransposition is a copy-and-paste process involving an RNA intermediate and an L1-encoded protein machinery (6–8) that orchestrates L1 integration via a molecular process termed target-primed reverse transcription (TPRT) (9). An RC-L1 is 6 kb in length and contains a 5′ untranslated region (UTR), an antisense promoter and open reading frame (ORF0) (10, 11), two nonoverlapping sense open reading frames (ORF1 and ORF2), and a 3′ UTR that is punctuated by a poly(A) tract (12–14). Critically, ORF2p possesses endonuclease (EN) and reverse transcriptase (RT) activities required for L1 mobility (8, 15, 16), while new L1 insertions usually integrate at a degenerate L1 EN recognition motif (5′-TT/AAAA, where “/” represents the position cut by the L1 EN) (17) and are flanked by variable-length target site duplications (TSDs) (18, 19), which are hallmarks of TPRT. L1 is the only active autonomous human retrotransposon (5, 15) although other polyadenylated RNAs, including mRNAs and those of the Alu and...
SINE-VNTR-Alu (SVA) retrotransposon families, can be mobilized in trans by the L1 machinery (13, 20–24). The L1 5′ UTR has an internal RNA polymerase II promoter that directs L1 mRNA transcription (25) and is regulated by DNA methylation of a CpG island located nearby in the 5′ UTR (26–31). The host genome also restricts L1 activity through mechanisms limiting L1 mRNA production or otherwise hindering retrotransposition (32–34).

A minor fraction of RC-L1s in the human population are thought to generate the majority of new germ line L1 insertions and are highly mobile, or “hot,” when tested in cultured cell L1 retrotransposition assays (3, 8, 35–37). These experiments largely measure the enzymatic efficiency of L1s introduced in episomal vectors, and, importantly, a particular L1 locus may present multiple alleles with different retrotransposition efficiencies (38, 39). The endogenous regulation of a given RC-L1 may therefore be most clearly resolved in the spatiotemporal contexts where it produces new L1 insertions. An RC-L1 can be identified as the donor element for an L1 insertion through shared unique internal single nucleotide variants, or transductions (37, 40). 5′ transductions are thought to accompany <0.1% of L1-Ta insertions and likely arise when the L1 promoter, or another nearby promoter, initiates L1 mRNA transcription upstream of the canonical L1 transcription start site (1, 32, 41–43). In contrast, 3′ transductions are found alongside ~20% of new germ line L1 insertions and occur when L1 mRNA transcription bypasses the canonical L1 polyadenylation signal and terminates at an alternative downstream polyadenylation signal (1, 13, 44–48). Transductions have been used to trace RC-L1s responsible for pathogenic L1 insertions (32, 35, 49–53) and to reconstruct closely related RC-L1 lineages, or transduction families, in human populations (35, 44, 54).

Early embryogenesis provides a major developmental niche for heritable L1 retrotransposition events in mammals (55–57). Cultivated human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs) resembling the cells of the embryonic inner cell mass also express L1 mRNAs and support engineered and endogenous L1 retrotransposition (58–63). De novo L1 insertions arising during embryogenesis or later development can cause somatic mosaicism (55, 64–66). In particular, somatic L1 insertions have been reported in brain tissue (65, 67–74), while engineered L1 reporter genes mobilize during neurogenesis and in postmitotic neurons (63, 67, 75). Importantly, the L1-Ta subfamily is hypomethylated in hESCs and hiPSCs compared to methylation of neurons and other differentiated cells, suggesting genome-wide L1 promoter methylation is enforced during development (58, 61, 63, 67). However, the likely related temporal profiles of DNA methylation and somatic retrotransposition for individual RC-L1s that mobilize during neurogenesis are unresolved.

Here, we identified a reprogramming-associated de novo L1 insertion in a cultivated hiPSC line. This insertion was traced to a hot donor RC-L1 that was part of an extended and recently active transduction family. We then measured locus-specific DNA methylation among de novo, donor, and transduction family L1 promoters, as well as the L1-Ta subfamily genome-wide, at multiple points of neurodifferentiation. These experiments significantly elucidate the dynamic temporal profile of epigenetic L1 repression applied to new and extant L1 insertions during neurogenesis.

**RESULTS**

A de novo L1 insertion arising during reprogramming. To study endogenous retrotransposition during neurogenesis, we obtained two hiPSC lines (hiPSC-CRL1502 and hiPSC-CRL2429) generated via delivery of defined reprogramming factors to healthy human dermal fibroblasts (58, 76). We then differentiated each hiPSC line toward a neuronal phenotype for 156 days in culture (Fig. 1A) and applied retrotransposon capture sequencing (RC-seq) (58, 69, 77) to genomic DNA sampled from the parental fibroblasts (time point 0 [T₀]), hiPSCs (T₁), and several time points of differentiation (T₂ to T₇) (Table 1). Two earlier passages of each hiPSC line were also analyzed by RC-seq to better distinguish L1 insertions arising during reprogramming or cell cultivation (Table 1). Cells from each point of neurodifferentiation were characterized
by immunocytochemistry (Fig. 1A) and included neural epithelium (T₂), neural rosettes denoting immature neurons (T₃) and three stages of prolonged neuronal maturation (T₄ to T₆). Endogenous L1 insertions detected by RC-seq and absent from the reference genome were annotated as either polymorphic (previously published or present at T₀) or de novo (only present at T₁ or later in one time course). Two potential de novo L1 insertions were identified (see Table S1 in the supplemental material). We then performed insertion site-specific PCR validation for each event (Fig. 1B and Table 2) and found that one insertion, on chromosome 1 (Chr1), was de novo in hiPSC-CRL2429 cells at time point T₁, was carried through neurodifferentiation (Fig. 1C), and was absent from hiPSC-CRL1502 (Fig. 1C). PCR indicated that the other putative de novo event was polymorphic because it was found in the matched parental fibroblast population (Table 2 and Table S1).

We then cloned and capillary sequenced the entire de novo L1 insertion (Fig. 1D) and manually inspected the integration site for hallmarks of TPRT (8, 9, 16, 17). The L1 was full length, belonged to the L1-Ta subfamily, carried 5’= and 3’= transductions, was flanked by 16-nucleotide (nt) TSDs, inserted at a degenerate L1 endonuclease motif (5’=-TT/AAAG), and terminated with a 33-nt poly(A) tract. The 5’= and 3’= transductions were 10 nt and 44 nt in length, respectively, and the 3’= transduction was preceded by an internal 17-nt poly(A) tract (Fig. 1D). These features were consistent with endoge-
nous retrotransposition mediated via TPRT and, as confirmed by insertion site-specific PCR, showed that the de novo L1 insertion represented a bona fide retrotransposition event occurring during reprogramming, or very early in hiPSC-CRL2429 cultivation.

An extended human RC-L1 transduction family. The de novo L1 insertion was the first such example to be found in hiPSCs of an endogenous L1 insertion carrying both 5' and 3' transductions. These transductions uniquely indicated a donor L1 sequence on chromosome 3 that was heterozygous in the hiPSC-CRL2429 parental fibroblast population (Fig. 1E). The donor L1 was absent from the reference genome and was polymorphic in humans; it was previously shown to mobilize efficiently in vitro (35). To identify any other germ line L1 insertions closely related to the donor L1, we aligned the 3' transduced sequence to the reference genome and to the annotated 3' L1-genome junction sequences of polymorphic L1s carried by hiPSC-CRL2429 or hiPSC-CRL1502 (Table S1) or those annotated by previous studies (52, 58, 69, 70, 77, 78, 79–87). From this analysis, we reconstructed an extended L1 transduction family comprising 14 members (Table 3), including a plausible founder, or lineage progenitor (44), element for the family, which was homozygous in hiPSC-CRL2429 and located on chromosome 11 (Fig. 1E).

To further characterize the transduction family, we analyzed the complete internal sequence of eight of its members found in either hiPSC-CRL1502 or hiPSC-CRL2429, including the de novo L1 insertion. A consensus sequence was obtained for the lineage progenitor, donor, and de novo L1s, as well as for another L1 nonreference (Non-ref) element, named Non-ref_Chr3_p24.3, via capillary sequencing of multiple full-length amplicons derived from independent PCRs (Fig. 2). Internal and flanking sequences for four additional reference (Ref) elements (Ref_Chr7_q21.3, Ref_Chr1_p31.1a, Ref_Chr1_p31.1b, and Ref_Chr9_p23) were obtained from the reference genome assembly. The 5' and 3' L1-genome junctions of the remaining six nonreference elements (Non-ref_Chr3_p12.2-a, Non-ref_Chr3_p12.2-b, Non-ref_ChrX_p11.4, Non-ref_Chr17_q12, Non-ref_Chr1_p22.2, and Non-ref_Chr4_q12) were provided by previous studies (Table 3).

### Table 1 RC-seq library information

| hiPSC line and library DNA input | Time point | RC-seq readsb | RC-seq data source |
|----------------------------------|------------|---------------|--------------------|
| CRL1502                          | T0         | 44,033,582    | 99.95              | Klawitter et al. |
| Fibroblasts                      | T1         | 42,151,994    | 99.74              | This study       |
| hiPSCs p76                       | T2         | 33,972,001    | 99.77              | This study       |
| Neural epithelium                | T3         | 39,766,940    | 99.77              | This study       |
| Immature neurons                 | T4         | 47,155,514    | 99.78              | This study       |
| Neurons I                        | T5         | 44,381,111    | 97.91              | This study       |
| Neurons II                       | T6         | 36,222,610    | 99.77              | This study       |
| hiPSCs p15                       | Earlier hiPSC passage | 24,385,022 | 99.88 | Klawitter et al. |
| hiPSCs p40                       | Earlier hiPSC passage | 63,130,772 | 99.88 | Klawitter et al. |
| CRL2429                          | T0         | 24,386,590    | 99.41              | Klawitter et al. |
| Fibroblasts                      | T1         | 38,460,241    | 99.63              | This study       |
| hiPSCs p70                       | T2         | 40,174,554    | 99.79              | This study       |
| Neural epithelium                | T3         | 46,646,999    | 99.78              | This study       |
| Immature neurons                 | T4         | 27,279,492    | 99.79              | This study       |
| Neurons I                        | T5         | 46,018,310    | 99.77              | This study       |
| Neurons II                       | T6         | 36,033,944    | 99.54              | This study       |
| hiPSCs p11                       | Earlier hiPSC passage | 64,534,189 | 99.40 | Klawitter et al. |
| hiPSCs p40                       | Earlier hiPSC passage | 27,447,967 | 99.39 | Klawitter et al. |

| hiPSC line and library DNA input | Time point | RC-seq readsb | RC-seq data source |
|----------------------------------|------------|---------------|--------------------|
| hiPSCs p15 Earlier hiPSC passage |            | 24,385,022    | 99.88              | Klawitter et al. |
| hiPSCs p40 Earlier hiPSC passage |            | 63,130,772    | 99.88              | Klawitter et al. |

*p, passage.
bData from 2- by 150-mer reads.
cNeurons I, II, and III were harvested after 72, 112, and 156 days of differentiation in vitro, respectively.
TABLE 2 PCR primers used for validation and bisulfite sequencing

| Primer function and name | Sequence |
|--------------------------|----------|
| Genomic primers for empty/filled L1 validation reactions | |
| LineageProgenitor_Chr11_fwd | AGGAAACAGTGAGGGGAAGC |
| LineageProgenitor_Chr11_rev | TGAGGCCAGGATTCTATAC |
| Donor_Chr3_fwd | TGATGACAGTAAATAATGGGTAGATGA |
| Donor_Chr3_rev | CGGCTTCCTGACGTATT |
| DeNovo_Chr1_fwd | CGTGAACCCAGAGATGAG |
| DeNovo_Chr1_rev | ATCTGGCCTCAGGAACTTA |
| Non-ref_Chr3_fwd | TTGTGGGAGGCAAAATGAT |
| Non-ref_Chr3_rev | TATTTACTCAACCAAGG |

L1-specific primers for validation of 5’ and 3’ L1-genome junctions

| Primer function and name | Sequence |
|--------------------------|----------|
| hL1_273_rev | ACCCGATTTTCCAGTGTCGT |
| hL1_ACshort_fwd | AGATATACTAATGCTAGTATGAC |

Not/I-L1-genome junction-spanning primers for cloning full-length L1s

| Primer function and name | Sequence |
|--------------------------|----------|
| LineageProgenitor_Chr11_NotI_fwd | CAAGCGGCCGCTTACATTTTTAAAGAATTGTAGGG |
| Donor_Chr3_NotI_fwd | TTACGAGCAGATTCTGAGAG |
| DeNovo_Chr1_NotI_fwd | AGTCTTCAGGACAGGAA |
| Non-ref_Chr3_NotI_fwd | GAATCAGCACAACTGAAAG |

L1-specific primers for sequencing full-length L1s

| Primer function and name | Sequence |
|--------------------------|----------|
| L1_452_fwd | GCCCAGGCTTGCTTAGGTA |
| L1_1020_fwd | TGGTTGACCGAGCTGAGAGA |
| L1_1532_fwd | CCGTGAAGAAGGAACACTCA |
| L1_1966_fwd | GCAAAATCCACGCTAAATCA |
| L1_2494_fwd | AAATTCGAGTCAAGCAG |
| L1_3014_fwd | AACGACAGCAAGCTCAAGG |
| L1_3502_fwd | GAGGGCGCAGTCTGAGAG |
| L1_4022_fwd | CAATCAGGCAGGAGAG |
| L1_4472_fwd | CAGGCGAGGAGGAAG |
| L1_4973_fwd | TCCCGATACGCTACAAATG |
| L1_5492_fwd | TACCCAGGCACTAAGGG |

Primers for amplification of L1 promoters from bisulfite converted DNA

| Primer function and name | Sequence |
|--------------------------|----------|
| L1_Bis-LP | GATTTGTTTTTGGATTGTAAAATGGTT |
| L1_Bis-Donor | TGGTTGAGCAGAGTCAGAGAA |
| L1_BIS-DN | GTATTTGATAGATTTTGAAGG |
| L1_Bis-F | TAGGGAGGTGATGAGTGG |
| L1_Bis-R | ACTATAATGAAATCACCAC |

TABLE 3 Transduction family members

| Element | Genomic coordinate (hg19) | TSD | Full-length* | Identification source and/or reference(s) |
|---------|--------------------------|-----|-------------|------------------------------------------|
| Lineage progenitor L1 | Chr11: 95169381 AAAGAATTGTA | Y | Reference genome; 3 |
| Donor L1 | Chr3: 38626082 AAGATGAGTAAATAATG | Y | 35, 49, 71, 77, 79, 82–87 |
| De novo L1 | Chr1: 23179316 AAAGAATGGACATCG | Y | This study |
| Ref_Chr7_p21.3_a | Chr7: 96475963 GAAATGTCAGTTG | Y | Reference genome |
| Non-ref_Chr3_p24.3 | Chr3: 20749004 TAAAGACAC | Y | 35, 49, 71, 77, 79, 82, 83, 87 |
| Non-ref_Chr1_p31.1_a | Chr1: 84518060 AGAAAACAAAATCA | Y | Reference genome |
| Non-ref_Chr1_p31.1_b | Chr1: 83125969 AAAAAATGGTGTCAGTC | N | Reference genome |
| Ref_Chr9_p23 | Chr9: 12556931 GAAAATGGG | N | Reference genome |
| Non-ref_Chr3_p12.2_a | Chr3: 80590176 GAAAATGGG | Y | 35, 37, 49, 77, 79, 82, 83 |
| Non-ref_Chr3_p12.2_b | Chr3: 82144869 AGAAATGACATCG | Y | 49, 71, 77, 79, 83, 85, 86 |
| Non-ref_ChrX_p11.4 | ChrX: 38097551 AAAGGCAATG | Y | 49, 86 |
| Non-ref_Chr17_q12 | Chr17: 32813609 AAGAAGTACAGTAGT | N | 71, 77, 79, 82–84, 87 |
| Non-ref_Chr1_p22.2 | Chr1: 90914512 AAAAGCTCTTCCAG | N | 49, 71, 77, 79, 85, 86 |
| Non-ref_Chr4_q12 | Chr4: 53628490 TAAAGATGCCAT | N | 49, 71, 77, 79, 82, 83, 86 |

*Y, yes; N, no.
FIG 2 The reprogramming-associated de novo L1 insertion belonged to an extended L1 transduction family. The diagram shows 14 members of this family, including the de novo L1 insertion. Two alleles of the lineage progenitor L1 were characterized. TSDs flanking each L1 are represented by blue arrows. The 5' UTR and 3' UTR sequences are shown in dark gray, while ORFs with known and unknown sequences are shown in white and light gray, respectively. Transduction colors match their source L1 locus: donor L1 -> de novo L1 (orange), Ref_Chr7_q21.3 -> Non-ref_Chr3_p24.3 (pink), lineage progenitor L1 -> all other family members (purple). Letter and number combinations within L1s correspond to L1.3 nucleotide (lowercase) and ORF1 and ORF2 amino acid (purple uppercase) positions (88). Nucleotide changes versus L1.3 and present in all, some, or one of the sequenced members of the transduction family are shown in gray, black, and blue, respectively. Nucleotide changes unique to the two alleles of the lineage progenitor are shown in green, and nucleotide changes unique to the donor L1 and de novo L1 are shown in pink.
common to both progenitor element alleles, in addition to shared 3’ transduced sequences (Fig. 2). The de novo and donor elements were identical in their L1 sequences, and the 5’ transduced sequence carried by the de novo insertion exactly matched the 10 nt directly upstream of the donor element. Surprisingly, in addition to the de novo L1 insertion, two other elements, Ref_Chr1_p31.1_a and Non-ref_ChrX_p11.4, each carried both 5’ and 3’ transductions, enabling us to unambiguously identify their respective donor L1 sequences (the lineage progenitor and Ref_Chr7_q21.3, respectively), which were also members of the transduction family (Fig. 2). Interestingly, the 539-nt 5’ transduction carried by Ref_Chr1_p31.1a was preceded by a single untemplated guanine, suggesting that the template mRNA was capped (18, 89), and utilized a transcription start site in the 5’ long terminal repeat (LTR7Y) sequence of a human endogenous retrovirus type H (HERV-H) provirus integrated 126 kb upstream of the lineage progenitor L1 (Fig. 2). This mRNA template incorporated two exons upstream of the lineage progenitor L1, which were spliced together and to the L1 via sites strongly resembling consensus mammalian splice donor and acceptor sequences (Fig. 3). Another element, Non-ref_Chr3_p24.3, incorporated a nonsense mutation predicted to truncate ORF2 prior to its RT domain. In sum, these experiments characterized relationships among members of a transduction family, which, in many cases, remain potentially capable of retrotransposition in the germ line, in tumors (37, 49), and, as shown here, in hiPSCs.

Transduction family mobilization in vitro. To assess the retrotransposition competence of several members of the transduction family, we employed a cultured-cell-engineered L1 retrotransposition reporter assay (8) in HeLa cells. Briefly, in this assay, an L1 sequence is cloned into a vector containing an antibiotic resistance cassette oriented antisense to the L1 copy, where the resistance gene contains an intron oriented in sense to the L1, meaning antibiotic resistance occurs only after splicing and retrotransposition of the reporter cassette (8, 90) (Fig. 4A). Through this approach, we tested the following elements: a known hot RC-L1 (L1.3) as a positive control (88, 91), an RT mutant L1 (L1.3 RT\(-\)) as a negative control (6), both detected alleles of the lineage progenitor L1, the donor L1 (identical in sequence to the de novo L1), and Non-ref_Chr3_p24.3, which contained an ORF2 stop codon in its RT domain (Fig. 2). Each
Among the tested elements, the lineage progenitor L1 allele 2 exhibited the highest retrotransposition frequency activity, at 135% of L1.3 (Fig. 4B). Consistent with the progenitor L1 allele 1 carrying two nonsynonymous mutations in ORF2 not found in allele 2, resulting in Q159H and D523H amino acid changes (Fig. 2), we found allele 1 retrotransposed at ~74% of the efficiency observed for allele 2 and at a similar efficiency as seen for L1.3 (Fig. 4B). Each progenitor L1 allele jumped at >10% of the efficiency of L1.3 and therefore met the definition of a hot RC-L1 (35). Notably, an allele of the progenitor L1 had previously been tested, albeit in an osteosarcoma cell line and with a different reporter system, and was found to present much more limited mobilization potential in vitro (3). The most likely explanation for this difference is that
the prior study tested an allele of the progenitor L1 not assayed here. This result further highlights the impact of allelic variation upon the retrotransposition efficiency of a given genomic RC-L1 copy (38, 39).

The donor L1 was sequenced from a line (hiPSC-CRL2429) established from a Caucasian individual. Apart from a single nucleotide mutation in its 3’ UTR, this L1 was identical to one identified in a Japanese individual by a previous study, which reported its retrotransposition efficiency as 101% of L1.3 in the same reporter assay (35). Here, the donor L1 jumped at 117% of L1.3, corroborating the prior experimental results and confirming that retrotransposition-competent alleles of this L1 exist in multiple human populations. Finally, L1.3 RT− and Non-ref_Chr3_p24.3 did not retrotranspose, consistent with disabled ORF2 RT activity in each case (Fig. 4B). Overall, these results demonstrate that the de novo L1, its donor sequence, and the progenitor element of the transduction family were all hot RC-L1s in vitro.

**L1 promoter methylation is dynamic during neurodifferentiation.** Full-length L1 mRNA transcription is a prerequisite for L1 retrotransposition in cis and is directed by an internal promoter located in the L1 5’ UTR (25). DNA methylation of an adjacent CpG island mediates repression of the L1 promoter (26, 31). Genome-wide, the L1-Ta subfamily is thought to be broadly hypomethylated in pluripotent cells and then methylated during differentiation, including in mature neurons (40, 49, 58, 61, 63, 67). However, the temporal methylation patterns for the L1-Ta subfamily and individual L1-Ta promoters during the various stages of neurodifferentiation to date have not been resolved. It is also unknown how quickly methylation is established upon new L1 insertions that arise in pluripotent cells. To address these questions, we applied a multiplexed L1 locus-specific bisulfite sequencing approach (52, 78) (Fig. 5A and Table 2) to assess DNA methylation among the de novo, donor, and progenitor L1 5’ UTR sequences, as well as the L1-Ta subfamily genome wide. This analysis was performed for both hiPSC lines and their parental fibroblasts and derivative neuronal cell populations, as surveyed by RC-seq, with the exception of the de novo L1, which was present only in hiPSC-CRL2429 (Fig. 5B and 6).

Considering general trends observed in both hiPSC lines, the L1-Ta subfamily and individual L1 promoters were most methylated in fibroblasts and differentiatated neurons and least methylated in hiPSCs and the earliest stages of neurodifferentiation (Fig. 5B and 6A). For example, 66.6%, 31.1%, and 61.0% of CpG dinucleotides surveyed in the donor L1 were methylated, on average, in hiPSC-CRL2429 fibroblasts, hiPSCs, and mature neurons, respectively. Among the two hiPSC lines, the highly significant (P < 0.0001, paired t test with Bonferroni correction) reductions in methylation observed for the donor L1 during hiPSC derivation (25.0% on average) far exceeded that seen for the lineage progenitor (12.5%) and L1-Ta subfamily (2.9%) (Fig. 5C and 6B). The lineage progenitor L1 was significantly (P < 0.001, paired t test) more methylated than the donor L1 at all time points in each hiPSC line, with the L1-Ta subfamily being methylated to a level between that of the lineage progenitor L1 and donor L1 at most time points (Fig. 5C and 6B). Notably, we observed a significant (P < 0.001, paired t test with Bonferroni correction) reduction in methylation (23.1% average decrease) for all amplicons at T5 in hiPSC-CRL2429, followed by a significant (P < 0.01) increase in methylation at T6 (20.1% average increase) (Fig. 5C). This trend was also observed at T5 for hiPSC-CRL1502, except for the donor L1 (Fig. 6B). The reasons for this pattern are presently unclear (see Discussion). Overall, these results demonstrate that DNA methylation is far more dynamic during reprogramming and differentiation for a donor L1 that can mobilize during or shortly after reprogramming than is seen for the vast majority of L1-Ta subfamily elements.

The de novo L1, which arose in hiPSC-CRL2429, could be detected at its 5’ L1-genome junction by site-specific PCR at time points T5 through T6 (Fig. 1C). However, as assessed by the number of unique sequencing reads generated, the PCR amplicon pool for the de novo L1 was very low in complexity at T6, perhaps due to a low percentage of cells carrying the mutation, and we therefore excluded T6 from further
L1 promoter DNA methylation is dynamic during hiPSC-CRL2429 reprogramming and neurodifferentiation. (A) L1 bisulfite sequencing analysis design. CpG dinucleotides are indicated by circles above the L1 5' UTR, and their nucleotide positions are provided below. A common reverse primer (black) is combined with either an L1-Ta subfamily forward primer (purple) or an L1 locus-specific forward primer (pink) to generate PCR amplicons for multiplexed paired-end Illumina 2-by-300-mer sequencing, resolving each amplicon in full. (B) L1 CpG methylation patterns in hiPSC-CRL2429 fibroblasts, hiPSCs, and neural cells derived in vitro. Each cartoon panel corresponds to an amplicon (L1-Ta subfamily or specific L1 locus) and displays 50 random, nonidentical sequences (black circle).

(Continued on next page)
The de novo L1 was nonetheless consistently less methylated than its donor L1 in hiPSC-CRL2429 time points T2 through T6, with average values across these stages of 41.6% and 53.8%, respectively (Fig. 5B). Methylation ultimately increased upon the de novo L1 during neurodifferentiation, but even in neurons we observed a significant number of cells in which the de novo L1 promoter was fully demethylated. For the donor L1 and the L1-Ta subfamily, we also observed instances of cells in which these promoters were fully demethylated at various points of neuronal differentiation.

FIG 6 L1 CpG methylation patterns in hiPSC-CRL1502 fibroblasts, hiPSCs, and neural cells derived in vitro. (A) Each cartoon panel corresponds to an amplicon (L1-Ta subfamily or specific L1 locus) and displays 50 random, nonidentical sequences (black circle, methylated CpG; white circle, unmethylated CpG; ×, mutated CpG). The percentage of methylated CpG is indicated in the lower right corner of each cartoon. (B) L1 promoter CpG methylation levels for the hiPSC-CRL1502 neurodifferentiation time course. Values represent the means ± standard deviations of CpG methylation of the corresponding 50 reads for each amplicon, as presented in panel A. Statistical analyses involved paired t tests, with a Bonferroni multiple-testing correction where appropriate. *, P < 0.01; **, P < 0.001; ***, P < 0.0001.

FIG 5 Legend (Continued)
methylated CpG; white circle, unmethylated CpG; ×, mutated CpG). The percentage of methylated CpG is indicated in the lower right corner of each cartoon. (C) L1 promoter CpG methylation levels for the hiPSC-CRL2429 neurodifferentiation time course. Values represent the means ± standard deviations of CpG methylation of the corresponding 50 reads for each amplicon, as presented in panel B. Statistical analyses involved paired t tests, with a Bonferroni multiple-testing correction where appropriate. *, P < 0.01; **, P < 0.001; ***, P < 0.0001.

April 2019 Volume 39 Issue 7 e00499-18 mcb.asm.org
Given the disparate methylation levels observed for the de novo and donor L1 promoter regions compared to the level of the lineage progenitor L1, we examined predicted DNA-binding protein motifs (92) affected by sequence variation among these elements (Fig. 2). The 10-nt 5’ transduction carried by the de novo L1 insertion incorporated a perfect FOX (forkhead box) protein binding motif (93). Members of the FOX protein family can act as “pioneer” factors in the developmental activation of promoters located in heterochromatin (94). In addition, the T708C nucleotide mutation present in the de novo and donor L1 copies greatly increased the predicted binding affinity for retinoid X receptor (RXR) proteins to this site. RXR proteins are known to respond to vitamin A (95), which is a component of the B-27 medium used here for neurodifferentiation. Conversely, the C581A nucleotide mutation carried by the lineage progenitor L1, and not by the de novo or donor L1 sequences or any other member of the transduction family, removed a key nucleotide mismatch from the core of a predicted PU.1 binding motif. PU.1 is established to recruit DNA methyltransferases to genomic loci and to form a repressor complex with MeCP2, which is a key mediator of L1 silencing (96–98). These in silico analyses suggested that differential DNA-binding protein activity as a result of sequence variation may impact the methylation and transcriptional state of members of the transduction family.

DISCUSSION

The L1 transduction family identified here is the largest found to date and adds to other such families characterized by previous studies (35, 44, 54). Although the extent of the transduction family is revealed here, it is likely that additional members will be identified in the future. It should also be noted that each transduction family member, aside from the de novo L1, was either present in the reference genome or identified by earlier works (Table 3). Unusually, in addition to 3’ transduced sequences, 3 of the 14 family members carried 5’ transductions. This 5’ transduction frequency (21.4%) is exceptionally high, given how rarely such events are found in the human germ line (1). Two of the 5’ transductions were relatively short (10 nt, de novo L1; 18 nt, Non-ref ChrX_p11.4) and likely resulted from the L1 promoter directing mRNA transcriptional initiation upstream of L1 position +1. The third 5’ transduction identified was significantly longer (539 nt, Ref Chr1_p31.1_a) and resulted from transcription initiated by the 5’ LTR of an upstream HERV-H proviral sequence, followed by splicing of this mRNA into a site adjacent to the donor L1. The inclusion of both LTR and internal HERV-H sequences in an L1 5’ transduction was an intriguing result as most heritable L1 insertions appear to arise early in mammalian embryogenesis (55, 56), and HERV-H elements are highly expressed in pluripotent cells (99–103). To speculate, this example demonstrates how HERV-H activation in the early embryo could lead to L1 mobilization. Nonetheless, it remains unclear why 5’ transductions are generally so frequent in this family and not in other transduction families (35, 44, 54). One possibility, an ORF2p amino acid change supporting elevated RT processivity and therefore increased average L1 insertion length, was excluded by an inspection of nonsynonymous sequence variants in this region (Fig. 2). Also excluded was the more likely possibility of mutations in known YY1, RUNX3, or SOX transcription factor binding sites (41, 104, 105) in the lineage progenitor L1 5’UTR or in alternative predicted sites located in the immediate 100 nt of its 5’ genomic flank, which may alter the accuracy of RNA polymerase II transcriptional initiation (Fig. 2). Otherwise, the family exhibited extensive variation in 3’ transduction and poly(A) tail length, as reported elsewhere for L1 insertions arising from a common donor L1 in the human population and cancer genomes (32, 37, 44, 49, 52, 78).

The discovery of a de novo L1 insertion in hiPSC-CRL2429 corroborates previous reports of endogenous and engineered L1 retrotransposition associated with reprogramming and hiPSC cultivation (58, 61). L1-mediated mutagenesis is potentially an important consideration for the use of hiPSCs in biomedical applications and as models of disease because the phenotypic properties of hiPSCs and their cellular derivatives...
could be compromised as a result of de novo L1 insertions \((58, 106)\). We demonstrate here that an endogenous L1 insertion arising in an hiPSC line is maintained during neurodifferentiation, indicating that such events can be present in differentiated cell lines derived from hiPSCs. In this case, the L1 was intergenic, and the accompanying transductions did not include protein-coding exons or regulatory elements \((47)\), lessening the probability of a functional impact in neurons carrying the L1 insertion. Although endogenous L1 retrotransposition is established to occur in the neuronal lineage \((65)\), we did not identify any additional de novo L1 insertions that were restricted to neural cells. These events were likely to each be carried by very few cells, meaning that they may not accrue sufficient RC-seq read depth to meet the detection thresholds used here. Nonetheless, it is plausible that de novo L1 insertions that impact the phenotype of hiPSC-derived cells will be identified in the future, especially as gene expression changes have been observed coincident with intronic L1 insertions arising during hiPSC generation \((58)\).

DNA methylation is thought to be established on L1 sequences very early in mammalian embryogenesis \((27, 28, 58, 61, 63, 67)\) and maintained in mature neurons. To our knowledge, L1 promoter methylation has not been explored for the various multipotent and immature neuronal cell types that arise during neurogenesis. Using in vitro hiPSC neurodifferentiation to represent neuronal development and maturation in vivo, we found that L1 promoter methylation was highly dynamic and increased as neurons matured. In each hiPSC line studied, we observed cells at multiple stages of neurodifferentiation, including mature neurons, where the donor L1 and other L1-Ta promoters were fully demethylated. Although the donor L1 was demethylated in hiPSCs compared to the methylation level of the matching parental fibroblasts, the absolute magnitudes of this change were dissimilar in the two lines \((35.5\% \text{ and } 14.4\% \text{ for hiPSC-CRL2429 and hiPSC-CRL1502, respectively})\). This perhaps reflected natural variation in the cohort of RC-L1s hypomethylated in each individual, before and after reprogramming. At time point \(T_5\), which follows a gliogenic switch \((107–109)\) during neural differentiation, we also observed a consistent reduction in L1 promoter methylation. This phenomenon could reflect a genome-wide reduction in DNA methylation specific to this stage of neurodifferentiation, perhaps due to a shift in the proportion of glial and neuronal cells present in culture, and warrants further study.

The de novo L1 insertion appeared to be rapidly targeted for repression by the host genome. During neurodifferentiation, similar transitions in methylation were observed for the de novo, donor and lineage progenitor L1s, and the L1-Ta subfamily even if the absolute methylation levels were very different among these elements. This result was consistent with epigenomic remodeling during reprogramming and neurodifferentiation \((110, 111)\) impacting the ground state of L1 methylation genome-wide. It also suggested that the de novo L1 insertion was quickly identified and regulated by the same pathways acting upon extant L1 copies on the genome even if the degree of methylation upon the de novo L1 was significantly lower than that applied to the transduction family and its ancestral L1-Ta subfamily. L1 5' UTR sequence variants, for example the C581A nucleotide mutation carried by the lineage progenitor L1 and predicted to increase DNA methylation mediated by PU.1, could contribute to differential methylation patterns among members of the transduction family. It is also notable that the de novo L1 remained retrotransposition competent, as do many other L1 insertions occurring in hiPSCs or arising during human embryogenesis \((57, 58)\). To speculate, if hiPSCs are taken as a model of very early development, a milieu where most heritable L1 insertions arise \((55)\), it is plausible that RC-L1 insertions arising de novo in this context will be incompletely methylated during later development and therefore possess a disproportionate capacity for further mobilization in the soma. Ultimately, hiPSCs and hESCs present accessible models to predict how L1 subfamilies and individual L1 loci are regulated. Additional work is required to test whether these patterns are observed during mammalian development in vivo.
MATERIALS AND METHODS

hiPSC generation and neuronal differentiation. Human induced pluripotent stem cell lines were epistemally derived as previously described (76). Neuronal differentiation was performed as described previously (112) with slight modifications. Prior to neuronal differentiation, feeder-free hiPSCs were cultured in murine embryonic fibroblast (MEF)-conditioned KOSR medium supplemented with 100 ng/ml basic fibroblast growth factor [b-FGF]. Initiation of neuronal differentiation occurred with the supplementation of dual SMAD inhibitors SB431542 (10 μM) and dorsomorphin (1 μM) into knockout serum replacement (KOSR) medium, which was gradually exchanged for 3 N medium (1:1 medium mix of N-2- and B-27-containing medium comprised of 1:1 neurobasal/Dulbecco’s modified Eagle’s medium [DMEM]–F-12 supplemented with 2% B-27, 1% N-2, 2 mM GlutaMax, 2.5 μg/ml insulin, 0.05 mM nonesential amino acids [NEAA], 0.05 mM beta-mercaptoethanol [all from Life Technologies]) in 25% incremental steps on days 4, 6, 8, and 10. Neural rosettes were selectively harvested and plated on Matrigel-coated TC dishes and expanded in 3 N medium supplemented with 20 ng/ml b-FGF. Around day 30 early neuronal progenitors were harvested with Accutase and seeded onto poly-L-ornithine/laminin-coated dishes (0.01% weight/volume and 20 μg/ml, respectively), and maintained in 3 N medium for the remainder of neurodifferentiation.

Immunocytochemistry. Neural cultures were grown on Matrigel-coated plastic coverslips in 3 N medium and were fixed in 4% paraformaldehyde (Sigma) in phosphate-buffered saline (PBS) for 15 min at room temperature and permeabilized in 0.01% Triton X-100 (Ajax Finechem) in PBS for 15 min at room temperature. All cells were blocked for 1 h with 10% goat serum (InVitrogen) in PBS. Primary antibodies used were OCT4 (1:10; Millipore), NANOG (1:10; Millipore), CUX1 (1:100; Abcam), glial fibrillary acidic protein (GFAP) (1:200; Dako), TUBB3/TUJ1 (1:1,000; Covance), BRN2 (1:100; Abcam), PAX6 (1:1,000; Developmental Studies Hybridoma Bank [DSHB]), anti-phospho-histone H3 (Ser10) (1:200; Cell Signaling Technology) and were applied for 3 to 4 h at room temperature or overnight at 4°C. Isotype- and species-matched Alexa Fluor-conjugated secondary antibodies (1:1,000; Invitrogen) were applied for 1 h at room temperature. Cells were washed in PBS and mounted on glass slides with ProLong Gold antifade containing 4′,6′-diamidino-2-phenylindole (DAPI; Invitrogen) and imaged using an Olympus IX51 (Olympus) fluorescence microscope equipped with a MicroPublisher, version 3.3, real-time viewing (RTV) charge-coupled-device (CCD) camera (Qimaging) using Q-Capture Pro, version 6.0, software.

Nucleic acid extraction. A total of approximately 500,000 cells per time point were pelleted (1,000 rpm for 5 min) and then washed with Dulbecco’s phosphate-buffered saline (DPBS) (14190144; Gibco) and pelleted again (1,000 rpm for 5 min) and resuspended in 100 μl of UltraPure DNase/RNase-free distilled water (10977023; Gibco). Cells were lysed in 10 mM Tris, pH 9.0, and 1 mM EDTA, with 2% SDS and 100 μg/ml proteinase K at 65°C. A final concentration of 10 μg/ml RNase A was added to each sample and incubated at 37°C for 30 min. DNA was extracted using phenol-chloroform-isooamyl alcohol (25:24:1) and chloroform-isooamyl alcohol (24:1). DNA was precipitated with 0.1 volume of 3 M sodium acetate and 2.5 volumes of 100% isopropanol. Precipitated DNA was washed in 0.8 ml of 75% ethanol (EtOH), slightly air dried, and resuspended in 50 μl of UltraPure DNase/RNase-free distilled water (10977023; Gibco). The quality and quantity of DNA were assessed by NanoDrop (Thermo Fisher Scientific).

RC-seq. Genomic DNA from time points T1 to T6 for each hiPSC line was analyzed by retrotransposon capture sequencing (RC-seq), as described previously (69). Each library was constructed from 2 μg of input genomic DNA (gDNA) and sequenced in multiplex on an Illumina HiSeq 2500 instrument (Macrogen, South Korea). Fibroblast samples (time point T0) were previously analyzed by RC-seq (58). A total of 726,181,832 paired-end 2- by 150-mer reads were generated across 18 libraries (Table 1). RC-seq data were analyzed on the http://broadinstitute.github.io/pcard website (https://github.com/adamewing/tebreak). Reads were aligned to the hg19 reference genome sequence using Burrows-Wheeler Aligner maximal exact match (BWAMEM) (113) with parameters -Y and -M. Duplicate reads were marked with Picard MarkDuplicates (http://broadinstitute.github.io/picard). Candidate nonreference genome L1 insertions that were (i) detected in only one of the two hiPSC lines analyzed, (ii) absent from the matching parental fibroblasts, and (iii) did not correspond to a known nonreference germ line transposable element insertions (35, 49, 77, 79–87, 114–116) were annotated as putatively de novo (see Table S1 in the supplemental material). The remaining nonreference L1 insertions were annotated as polymorphic.

PCR validation of L1 insertions. RC-seq reads indicating putative de novo L1 insertions were manually inspected, and primers (Table 2) were designed to PCR amplify integration sites and identify the hallmark L1 integrase (int) consensus L1 retrotransposition events (117). Empty/filled-site, 5’ L1-genome junction, and 3’ L1-genome junction PCR amplifications were performed. Primers were situated within flanking genomic DNA sequences for empty/filled-site PCRs. The same flanking primers were paired with appropriate L1-specific primers for L1-genome junction assays. Expand long-range enzyme was used for empty/filled-site PCRs using 1.75 U of Expand Long Template enzyme (04829069001; Roche), 5 μl of 5× buffer with 12.5 mM MgCl2, 1.25 μl of 100% dimethyl sulfoxide (DMSO), 1.25 μl of 10 mM deoxyxynucleoside triphosphates (dNTPs), 1 μl of primer mix (25 μM each primer), 4 ng of genomic DNA template, and molecular-grade water in a final volume of 25 μl under the following PCR conditions: 92°C for 2 min, followed first by 10 cycles at 92°C for 10 s, 59°C for 15 s, and 68°C for 6.5 min and then by 30 cycles at 92°C for 2 min, 59°C for 15 s, and 68°C for 6.5 min plus 20 s of extension time per cycle, with a single extension step at 68°C for 10 min. The 5’ and 3’ L1-genome junction PCRs were performed using 2 U of MyTaq hot-start DNA polymerase (Bio-21112; Bioline), 1× PCR buffer, 1 μM each primer, 5 ng of genomic DNA template, and molecular-grade water in a final volume of 25 μl. Cycling conditions were as follows: 95°C for 2 min, followed by 35 cycles at 95°C for 30 s, 58°C for 30 s, and 72°C for 3 min, with a single extension step of 72°C for 5 min. Amplified fragments were resolved on 1% and 2% agarose gels (1× Tris-acetate-EDTA buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE) buffer with 12.5 mM Tris-acetate-EDTA (TAE)
[TAE buffer] stained with SybrSafe (Life Technologies) for empty/filled-site and 5’ and 3’ junction PCR assays, respectively, and imaged using a Typhoon FLA 9500 (GE Healthcare Life Sciences, USA). Amplicons of the expected size were excised from the gels, and DNA was extracted using a QIAquick gel extraction kit (28704; Qiagen), followed by capillary sequencing to confirm and characterize L1 insertion structural features.

**L1 genotyping and cloning.** To facilitate cloning of full-length L1 insertions, a NotI restriction enzyme sequence (5’-GC/GGCGC) was introduced at the 5’ end of each forward primer close to the L1-genome junction. Purified PCR products (500 ng) approximately 6 kbp in size were digested with NotI and Bst217I (R3138; New England Biolabs) in 1× CutSmart buffer at 37°C for 1 h. Digestion reactions were run in 2% agarose gels (1× TAE buffer), purified by phenol-chloroform extraction, and cloned into the vector TOPO-XL PCR cloning kit (K4700-20; Life Technologies) according to the manufacturer’s instructions. Five microliters of the ligation product was used to transform One Shot TOP10 electrocompetent bacteria as per the manufacturer’s instructions. LB agar containing 0.5 μg/ml of kanamycin was used to plate bacteria, which were incubated at 37°C overnight. Single colonies were picked and transferred to 5 ml of LB liquid containing 0.5 μg/ml of kanamycin for Miniprep plasmid purification (12143; Qiagen).

To filter induced PCR mutations and distinguish possible allelic variants, at least four independent PCR products, and clones from each L1 transduction family member were capillary sequenced using 12 overlapping primer pairs (Table 2) distributed at ~500-bp intervals covering the entire L1 sequence. Each independent clone sequence was then manually assembled and aligned with the other clones of the same element using Clustal Omega ([https://www.ebi.ac.uk/Tools/msa/clustalo/](https://www.ebi.ac.uk/Tools/msa/clustalo/)). For each L1, a consensus sequence was obtained, and a mutation-free construct was reconstructed by performing multiple restriction enzyme digestions. The desired fragments were resolved in a 2% agarose gel (1× TAE buffer), purified, and ligated into a pCEP4 vector using T4 ligase in a 5:1 (insert/vector) ratio. Five microliters of the ligation product was used to transform One Shot TOP10 chemically competent bacteria (C404010; Invitrogen) as per the manufacturer’s instructions. LB agar containing 1 μg/ml of ampicillin was used to plate the bacteria, and these were incubated at 37°C overnight. Single colonies were picked and transferred to 5 ml of LB liquid containing ampicillin for Miniprep plasmid purification. To verify the fidelity of the resultant clones, these were capillary sequenced, as described above, using 12 different primers covering the entire L1 sequence.

Retrotransposition indicator plasmids termed L1.3 and L1.3 RT− were generated through modification of the pCEP4 backbone of pJM101/L1.3 (14, 91) and pJM105/L1.3 (118) by removing a BglII fragment containing the cytomegalovirus (CMV) promoter. The full L1 3’ UTR, except for a point mutation disrupting the native L1 polyadenylation signal, was reintroduced, and a PacI site was incorporated between the L1.3 3’ UTR and the Neo cassette (F. J. Sanchez-Luque and G. J. Faulkner, unpublished data). The mutation-free full-length transduction family members described above were then introduced into this retrotransposition indicator backbone.

DNA-binding protein motif analyses of the lineage progenitor, donor, and de novo L1 sequences were performed using the Catalog of Inferred Sequence Binding Preferences (CIS-BP) database (92).

**Retrotransposition assay.** HeLa-JVM cells grown in a humidified, 5% CO₂ incubator at 37°C in high-glucose Dulbecco’s modified Eagle’s medium (DMEM) without pyruvate (11965-092; Gibco), supplemented with 10% fetal bovine serum (26400-044; Gibco), 2 mM L-glutamine, 100 U/ml penicillin, and 100 μg/ml streptomycin (10378-016; Gibco) (DMEM complete). Plasmid DNA was purified using a Midi kit (13343; Qiagen) and diluted in sterile water to 0.5 μg/μl. Cells were transfected and seeded at 5×10⁵ cells/well in six-well plates using FuGENE HD transfection reagent (Promega) at a ratio of 4 μl to 1 μg of plasmid DNA. Selection with G418 began 72 h after transfection and continued every 48 h for 14 days (6). Transfection efficiency assays were performed in parallel by cotransfection of pCAG-enhanced green fluorescent protein (EGFP) with L1 reporter plasmids, as described above, with 0.5 μg of each construct and 0.5 μg of pCAG-EGFP. Cells were analyzed by flow cytometry 48 h posttransfection on a Cytoflex flow cytometer (Beckman-Coulter) at the Translational Research Institute Flow Cytometry Core. The results were used to normalize the G418-resistant colony counts with the percentage of EGFP-positive cells for each L1 reporter construct obtained in the retrotransposition assay, as performed previously (118).

**L1 CpG methylation analyses.** L1-Ta subfamily-wide and L1 locus-specific bisulfite sequencing for each time point in hiPSC-CRL1502 and hiPSC-CRL2429 was performed as described previously (52). Briefly, 500 ng of gDNA was bisulfite treated using an EZ DNA Methylation Lightning kit (Zymo Research), allowing 20 min desulfonation time and eluting in a 25-μl volume. Primers L1_Bis-F and L1_Bis-R were used to amplify the L1-Ta 5’ UTR region containing a CpG island (Table 2), while for the L1 locus-specific reactions, L1_Bis-R was combined with one of three forward primers placed in the genomic flank of the lineage progenitor, donor, and de novo L1 insertions (L1_Bis-PL, L1_Bis-Donor, and L1_Bis-DN, respectively). PCRs incorporated 1 U of MyTaq hot-start DNA polymerase (BIO-21112; Bioline), 2 μl of bisulfite-treated gDNA from each sample, 1× reaction buffer, and 2 μl of each primer, in a 20-μl final volume. PCR cycling conditions were as follows: 95°C for 2 min, followed by 40 cycles of 95°C for 30 s, 54°C for 30 s, and 72°C for 30 s, with a single extension step at 72°C for 5 min. Barcoded libraries were prepared from amplicons pooled by time point and sample using a TruSeq DNA PCR-free library preparation kit (FC-121-3001/2; Illumina) and subjected to multiplexed paired-end 2- by 300-mer sequencing using an Illumina MiSeq platform. Data were processed as described previously (52) and visualized using QUMA (119) with default parameters.

**Accession number(s).** RC-seq FASTQ files were deposited in the European Nucleotide Archive under accession number PRJEB27103.
SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/MCB.00499-18.

SUPPLEMENTAL FILE 1, XLSX file, 0.2 MB.

ACKNOWLEDGMENTS

C.-S.-P., F.J.S.-L., P.R.J.F., and S.R.R. performed experiments. E.J.W. and G.J.F. provided resources. C.-S.-P., A.D.E., and G.J.F. performed bioinformatic analyses. C.-S.-P., S.R.R., F.J.S.-L., and G.J.F. designed experiments. C.-S.-P., S.R.R., and G.J.F. prepared figures, conceived the study, and wrote the manuscript. All authors commented on the final manuscript.

We thank John V. Moran for pJM101/L1.3 and pJM105/L1.3 plasmids and the HeLa-JVM cell line. We thank Seth Cheetham for helpful discussions.

This work was supported by a University of Queensland International Scholarship (to C.-S.-P.), the People Program (Marie Curie Actions) of the European Union’s Seventh Framework Program (FP7/2007-2013) under REA grant agreement PIOF-GA-2013-623324 awarded to F.J.S.-L., and Australian Research Council Discovery Early Career Researcher Award (DE150101117) and Discovery Project (DP170101198) grants awarded to A.D.E. G.J.F. acknowledges the support of the Mater Foundation, a CSL Researcher Award (DE150101117) and Discovery Project (DP170101198) grants awarded to A.D.E. G.J.F. acknowledges the support of the Mater Foundation, a CSL Researcher Award (DE150101117) and Discovery Project (DP170101198) grants awarded to A.D.E.

We declare that we have no competing interests.

REFERENCES

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham D, Durbin R, French L, Graham D, et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860–921. https://doi.org/10.1038/35073626

2. Kazazian HH, Jr, Moran JV. 2017. Mobile DNA in health and disease. N Engl J Med 377:361–370. https://doi.org/10.1056/NEJMra1510092

3. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Boeke JD, Kazazian HH, Jr, Komma D. 2014. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol Cell Biol 24:1805–1808. https://doi.org/10.1128/MCB.24.6.1973-1985.2004

4. Myers JS, Vincent BJ, Uddal H, Watkins WS, Morrish TA, Kilroy GE, Svergold GD, Henke J, Henke L, Moran JV, Jorde LB, Batzer MA. 2002. A comprehensive analysis of recently integrated human Ta L1 elements. Am J Hum Genet 71:312–326. https://doi.org/10.1086/341718

5. Mills RE, Bennett EA, Iukow RC, Devine SE. 2007. Which transposable elements are active in the human genome? Trends Genet 23:183–190. https://doi.org/10.1016/j.tig.2007.02.006

6. Weil W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JD, Moran JV. 2001. Human L1 retrotransposition: cis preference versus trans complementation. Mol Cell Biol 21:1429–1439. https://doi.org/10.1128/MCB.21.4.1429-1439.2001

7. Boeke JD, Garfinkel DJ, Styles CA, Fink GR. 1985. Ty elements transpose through an RNA intermediate. Cell 40:491–500. https://doi.org/10.1016/0092-8674(85)90197-7

8. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH, Jr. 1996. High frequency retrotransposition in cultured mammalian cells. Cell 87:917–927. https://doi.org/10.1016/s0092-8674(96)00198-4

9. Luan DD, Korman MH, Jakubczak JL, Eickbush TH. 1993. Reverse transcription of 28RNA is primed by a nick at the chromosomal target site; a mechanism for non-LTR retrotransposition. Cell 72:595–605. https://doi.org/10.1016/0092-8674(93)90078-5

10. Denil AM, Narvaiza J, Kerman BE, Penin M, Benner C, Marchetto MC, Diedrich JK, Aslanian A, Ma J, MoreSCO JJ, Moore L, Hunter T, Saghatelian A, Gage FH. 2015. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell 163:583–593. https://doi.org/10.1016/j.cell.2015.09.025

11. Hilsenbeck S, M. 2001. Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol Cell Biol 21:1973–1985. https://doi.org/10.1128/MCB.21.9.1973-1985.2001

12. Scott AF, Schmeckpeper BJ, Abdelrazik M, Coney CT, O’Hara B, Rossiter JP, Cooley T, Heath P, Smith KD, Margolet L. 1987. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1:113–125. https://doi.org/10.1016/0888-7543(87)90003-6

13. Doucet AJ, Wilusz JE, Miyoshi T, Liu Y, Moran JV. 2015. A 3' poly(A) tract is required for LINE-1 retrotransposition. Mol Cell 60:728–741. https://doi.org/10.1016/j.molcel.2015.10.012

14. Dombroski BA, Mathias SL, Nanthakumar E, Scott AF, Kazazian HH, Jr, Moran JV. 1997. Isolation of an active human transposable element. Science 275:754–755. https://doi.org/10.1126/science.1722352

15. Feng Q, Moran JV, Kazazian HH, Jr, Boeke JD. 1996. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905–916. https://doi.org/10.1016/S0092-8674(00)81997-2

16. Jurka J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retrotransposons. Proc Natl Acad Sci U S A 94:1872–1877. https://doi.org/10.1073/pnas.94.5.1872

17. Scott AF, Kazazian HH, Jr, Moran JV. 2001. Human L1 retrotransposition: cis preference versus trans complementation. Mol Cell Biol 21:1429–1439. https://doi.org/10.1128/MCB.21.4.1429-1439.2001

18. Gilbert N, Lutz S, Morrish TA, Moran JV. 2005. Multiple fates of L1 retrotransposition intermediates in cultured human cells. Mol Cell Biol 25:7780–7795. https://doi.org/10.1128/MCB.25.17.7780-7795.2005

19. Grimaldi G, Skowronski J, Singer MF. 1984. Defining the beginning and end of KpnI family segments. EMBO J 3:1753–1759. https://doi.org/10.1002/j.1460-2075.1984.tb02042.x

20. Ahl V, Keller H, Schmidt S, Weichenrieder O. 2015. Retrotransposition and crystal structure of an Alu RNP in the ribosome-stalling conformation. Mol Cell 60:715–727. https://doi.org/10.1016/j.molcel.2015.10.003

21. Ennault C, Maestre J, Heidmann T. 2000. Human LINE retrotransposons generate processed pseudogenes. Nat Genet 24:363–367. https://doi.org/10.1038/74184

22. Demir M, Ennault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet 35:41–48. https://doi.org/10.1038/ng1223

23. Raiz J, Damert A, Chira S, Held U, Klawitter S, Hamdorf M, Lower J,
L1 Methylation during hiPSC Differentiation

Molecular and Cellular Biology

41. Athanikar JN, Badge RM, Moran JV. 2004. A YY1-binding site is required for active human LINE-1 transcription initiation. Nucleic Acids Res 32:3846–3855.

42. Symer DE, Connelly C, Szak ST, Caputo EM, Cost CJ, Parmigiani G, Boeke JD. 2002. Human L1 retrotransposition is associated with genetic instability in vivo. Cell 110:327–338.

43. Larson PA, Moldovan JB, Jasti N, Kidd JM, Beck CR, Moran JV. 2018. Spliced integrated retrotransposed element (SIRE) formation in the human genome. PLoS Biol 16:e2003067.

44. Macfarlane CM, Collier P, Rahbari R, Beck CR, Wagstaff JF, Igoe S, Moran JV, Badge RM. 2013. Transcription-specific ATLAS reveals a cohort of highly active L1 retrotransposons in human populations. Hum Mutat 34:974–985. https://doi.org/10.1002/humu.22327.

45. Goodier JL, Ostertag EM, Kazazian HH, Jr. 2000. Transduction of 3'-flanking sequences is common in L1 retrotransposition. Hum Mol Genet 9:653–657. https://doi.org/10.1093/hmg/9.4.653.

46. Moran JV. 1999. Human L1 retrotransposition: insights and peculiarities learned from a cultured cell retrotransposition assay. Genetica 107:39–51. https://doi.org/10.1023/A:1004035023354.

47. Moran JV, DeBerardinis RJ, Kazazian HH, Jr. 1999. Exxon shuffling by L1 retrotransposition. Science 283:1530–1534. https://doi.org/10.1126/science.283.5407.1530.

48. Pickeral OK, Makalowski W, Boguski MS, Boyle JD. 2000. Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. Genome Res 10:411–415. https://doi.org/10.1101/gr.241661.114.

49. Tubio JMC, Li Y, Ju YS, Marticorena I, Cooke SL, Tojo M, Gundem G, Pipinikas CP, Zamora J, Raine K, Meneses A, Roman-Garcia P, Fullam GM, Gerstung M, Shlien A, Trapery PS, Papamaniou E, Knappskog S, Van Loo P, Ramakrishna M, Davies HR, Marshall J, Wedge JW, Butler AP, Nik-Zainal S, Alexandrov L, Behjati S, Yates LR, Bolli N, Mudie L, Hardy C, Martin S, McLaren S, O'Meara S, Anderson E, Maddison M, Gamble S, Foster C, Warren AY, Whitaker H, Brewer D, Eeles R, Cooper C, Neal D, Lynch AG, Visakorpi T, Isaacs WB, Veer LV, Caldas C, et al. 2014. Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science 345:1251343. https://doi.org/10.1126/science.1251343.

50. Solyom S, Ewing AD, Hancks DC, Takehisa Y, Awano H, Matsuo M, Kazazian HH, Jr. 2012. Pathogenic orphan transduction created by a nonrepetitive LINE-1 retrotransposon. Hum Mutat 33:369–371. https://doi.org/10.1002/humu.21663.

51. Brokken BA, Meischl C, Ostertag EM, de Boer M, Zhang Y, Neijens H, Roos D, Kazazian HH, Jr. 2002. Evidence consistent with human L1 retrotransposition in maternal meiosis I. Am J Hum Genet 71:327–336. https://doi.org/10.1086/341722.

52. Nguyen THM, Carreira PE, Sanchez-Luque FJ, Schauer SN, Fagg AC, Richardson SR, Davies CM, Jesuadian JS, Kempen MHC, Troskie RL, James C, Beaver EA, Wallis TP, Coward JIG, Chetty NP, Crandon AJ, Van Dijk AJ, Dames PE, Perrin LC, Hoelder JP, Ewing AD, Upton KR, Faulkner GJ. 2018. L1 retrotransposition heterogeneity in ovarian tumor cell evolution. Cell Rep 23:3730–3740. https://doi.org/10.1016/j.celrep.2018.05.090.

53. Holmes SE, Dombroski BA, Krebs CM, Boehm CD, Kazazian HH, Jr. 1994. A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. Nat Genet 7:143–148. https://doi.org/10.1038/ng0494-143.

54. Szak ST, Pickeral OK, Landsman D, Boeke JD. 2003. Identifying related L1 retrotransposition events by analyzing 3’ transduced sequences. Genome Biol 4:R30. https://doi.org/10.1186/gb-2003-4-5-r30.

55. Richardson SR, Gerdes P, Gerhardt DJ, Sanchez-Luque FJ, Bodea GO, Munoz-Lopez M, Jesuadian JS, Kempen MHC, Carreira PE, Jendeleh JA, Garcia-Perez JL, Kazazian HH, Jr, Ewing AD, Faulkner GJ. 2017. Heritable L1 retrotransposition in the mouse primordial germline and early embryo. Genes 8:1395–1405. https://doi.org/10.3390/genes8050139.

56. Richardson SR, Faulkner GJ. 2018. Heritable L1 retrotransposition events during development: understanding their origins. Bioessays 40:e1700189. https://doi.org/10.1002/bies.201700189.

57. van den Hurk JA, Meij IC, Seleme MC, Kano H, Nikopoulos K, Hoefsloot H, Garcia-Perez JL, Kazazian HH, Jr, Ewing AD, Faulkner GJ. 2017. Heritable L1 retrotransposition created by a cultured cell retrotransposition assay. J Mol Biol 430:1251343. https://doi.org/10.1002/humu.22327.
stem cells. Hum Mol Genet 16:1569–1577. https://doi.org/10.1093/hmg/ddm105.

60. Wissing S, Montano M, Garcia-Perez JL, Moran JV, Greene WC. 2011. Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells. J Biol Chem 286:36427–36437. https://doi.org/10.1074/jbc.M111.251058.

61. Wissing S, Munoz-Lopez M, Macia A, Yang Z, Montano M, Collins W, Garcia-Perez JL, Moran JV, Greene WC. 2012. Reprogramming somatic cells into iPSCs activates LINE-1 retroelement mobility. Hum Mol Genet 21:208–218. https://doi.org/10.1093/hmg/ddr455.

62. Macia A, Widmann TJ, Heras SR, Ayllon V, Sanchez L, Benkaddour-Boumzaoud M, Munoz-Lopez M, Rubio A, Amador-Cubero S, Blanco-Jimenez E, Garcia-Castro J, Menendez P, Ng P, Muotri AR, Goodier JL, Garcia-Perez JL. 2017. Engineered LINE-1 retrotransposition in nondi- viding human neurons. Genome Res 27:335–348. https://doi.org/10 .1101/gr.208805.116.

63. An W, Han JS, Wheelan SJ, Luquette LJ, Lohr JG, Harris CC, Ding L, Wilson RK, Wheeler DA, Gibbs RA, Kucherlapati R, Lee C, Karchenkov PV, Park PJ, The Cancer Genome Atlas Research Network. 2012. Landscape of somatic retrotransposition in human cancers. Science 337:967–971. https://doi.org/10.1126/science.122077.

64. Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH, Jr. 2009. L1 retrotransposition occurs mainly in embryonic and early fetal human genomes. Dev 23:1303–1312. https://doi.org/10.1016/j.gast.2008.09.009.

65. Faulkner GJ, Garcia-Perez JL. 2017. L1 mosaicism in mammals: extent, effects, and evolution. Trends Genet 33:802–816. https://doi.org/10 .1016/j.tig.2017.07.004.

66. An W, Han JS, Wheelan SJ, Davis ES, Coombes CE, Ye P, Triplet C, Boeke JD. 2006. Active retrotransposition by a synthetic L1 element in mice. Proc Natl Acad Sci U S A 103:18662–18667. https://doi.org/10.1073/pnas.0605300103.

67. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O’Shea KS, Brennan PM, Rizzu P, Smith S, Fell M, Talbot RT, Gustincich S, Freeman TC, Mattick JS, Hume DA, Heutink P, Carninci P, Ciccarelli FD, Garcia-Perez JL, Faulkner GJ. 2013. Ubiquitous L1 mosaicism in human neural progenitor cells. Nature 460:1127–1131. https://doi.org/10.1038/nature108248.

68. Faulkner GJ, Billon V. 2018. L1 retrotransposition in the soma: a field jumping ahead. Mob DNA 9:22. https://doi.org/10.1186/s13100-018-0128-1.

69. Upton KR, Gerhardt DJ, Jasin SJ, Richardson SR, Sanchez-Luquey FJ, Bodea GO, Ewing AD, Salvador-Palomeque C, de la Vaca K, Brennan PM, Vandenver A, Faulkner GJ. 2015. Ubiquitous L1 mosaicism in hippocampal neurons. Cell 161:228–239. https://doi.org/10.1016/j.cell.2015.03.026.

70. Banerji JK, Barrett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sipio F, Brennan PM, Rizzi P, Smith S, Fell M, Talbot RT, Gustin, S, Freeman TC, Mattick JS, Hume DA, Deutink P, Carninci P, Jeddeloh JA, Faulkner GJ. 2011. Somatic retrotransposition alters the genetic landscape of the human brain. Nature 479:534–537. https://doi.org/10.1038/nature10531.

71. Evrony GD, Cai X, Lee E, Mills LB, Elhosary PC, Lehmann HS, Parker JJ, Ateshian TA, Alves FI, Butcher CR, Herdy JR, Sarkar A, Lasken RS, Muotri AR, Goodier JL, Wissenberg GJ, Billon V, Freeman TC, Garcia-Castro J, Menendez P, Ng P, Muotri AR, Goodier JL, Garcia-Perez JL. 2017. Engineered LINE-1 retrotransposition in nondi- viding human neurons. Genome Res 27:335–348. https://doi.org/10 .1101/gr.208805.116.

72. Evrony GD, Cai X, Lee E, Mills LB, Elhosary PC, Lehmann HS, Parker JJ, Ateshian TA, Alves FI, Butcher CR, Herdy JR, Sarkar A, Lasken RS, Muotri AR, Goodier JL, Garcia-Perez JL. 2017. Engineered LINE-1 retrotransposition in nondi- viding human neurons. Genome Res 27:335–348. https://doi.org/10 .1101/gr.208805.116.

73. Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ, Lohr JG, Harris CC, Ding L, Wilson RK, Wheeler DA, Gibbs RA, Kucherlapati R, Lee C, Karchenkov PV, Park PJ, The Cancer Genome Atlas Research Network. 2012. Landscape of somatic retrotransposition in human cancers. Science 337:967–971. https://doi.org/10.1126/science.122077.

74. Iskow RC, McCabe MT, Mills RE, Torene S, Pittard WS, Neufeld AW, Vall, Meir EG, Vertino PM, Devine SE. 2010. Natural mutagenesis of human genomes by endogenous retrotransposons. PLoS Genet 6:141253–1261. https://doi.org/10.1371/journal.pgen.10012236.

75. Helman E, Lawrence MS, Stewart C, Sougnez C, Getz G, Meyerson M. 2014. Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. Genome Res 24:1053–1063. https://doi.org/10.1101/gr.2745804.

76. Kuhn A, Ong YM, Cheng CY, Wong TY, Quake SR, Burkholder WF. 2014. Landscape dissequilibrations and signatures of positive selection around LINE-1 retrotransposons in the human genome. Proc Natl Acad Sci U S A 111:8131–8136. https://doi.org/10.1073/pnas.1401532111.

77. Stewart C, Kural D, Stromberg MP, Walker JA, Konkel MK, Stutz AM, Urban AE, Grubert F, Lam HY, Lee WP, Busby Indar AP, Garrison E, Huff C, Jing X, Snyder MP, Jorde LB, Matzor MA, Korbel JD, Mart N, Genome Project. 2011. A comprehensive map of mobile element insertion polymorphisms in humans. PLoS Genet 7:e1002236. https://doi.org/10.1371/journal.pgen.1002236.

78. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, Fritz MH-Y, Konkel MK, Malhotra A, Stutz AM, Shi X, Casale FP, Chen J, Hormozdini F, Dayama G, Chen K, Malig M, Chaisson MJP, Walter K, Meiers S, Kashin S, Garrison E, Auton A, Lam HYK, Mu XJ, Alkan C, Antaki D, Bae T, Cerveira E, Chines P, Chong Z, Clarke L, Dal E, Ding L, Emery S, Fan X, Gujral M, Kahveci F, Kidd JM, Kong Y, Lamejier E-W, McCarthy S, Flicek P, Gibbs RA, Math R, Gysin CE, Menelaoa A, et al. 2015. An integrated map of structural variation in 2,504 human genomes. Nature 526:75–81. https://doi.org/10.1038/nature15394.

79. Wang J, Song L, Grover D, Atrakz S, Batzer MA, Liang P. 2006. dbRIP: a highly integrated database of retrotransposition insertion polymorphisms in humans. Hum Mutat 27:323–329. https://doi.org/10.1002/humu.20107.

80. Dombrowski BA, Scott AF, Kazazian HH, Jr. 1993. Two additional potential retrotransposons isolated from a human L1 subfamily that contains an active retrotransposable element. Proc Natl Acad Sci U S A 90:6513–6517. https://doi.org/10.1073/pnas.90.14.6513.

81. Lavié L, Maldener E, Brouha B, Meese EU, Mayer J. 2004. The human L1 promoter: variable transcription initiation sites and a major impact of upstream flanking sequence on promoter activity. Genome Res 14:2253–2260. https://doi.org/10.1101/gr.274508.

82. Kopera HC, Larson PA, Moldovan JB, Richardson SR, Liu Y, Moran JV. 2016. LINE-1 cultured cell retrotransposition assay. Methods Mol Biol 1400:139–156. https://doi.org/10.1007/978-1-4939-3372-3_10.

83. Sassaman DM, Dombrowski BA, Moran JV, Kimberland ML, Naas TP, DesMardeins RJ, Gabriel A, Swergold GD, Kazazian HH, Jr. 1997. Many human L1 elements are capable of retrotransposition. Nat Genet 16:37–43. https://doi.org/10.1038/ng0597-37.
92. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, Zheng H, Goity A, van Bakel H, Lozano JC, Galli M, Lewsey MG, Huang E, Mukherjee T, Chen X, Reece-Hoyes JS, Govindarajan S, Shaulsky G, Walhout AJM, Bouget FY, Ratsch G, Larrondo LF, Ecker JR, Hughes TR. 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158:1431–1443. https://doi.org/10.1016/j.cell.2014.08.009.

93. Pierrou S, Hellyquist M, Samuelsson E, Enerback S, Carlsson P. 1994. Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA binding. EMBO J 13:5002–5012. https://doi.org/10.1002/j.1460-2075.1994.tb06827.x.

94. Tchenio T, Casella JF, Heidmann T. 2000. Members of the SRY family regulate the human LINE retrotransposons. Nucleic Acids Res 28:411–415. https://doi.org/10.1093/nar/28.2.411.

95. Yang N, Zhang L, Zhang Y, Kazazian HH, Jr. 2003. An important role for RUNX3 in human L1 transcription and retrotransposition. Nucleic Acids Res 31:4929–4940. https://doi.org/10.1093/nar/gkg663.

96. Miller FD, Gauthier AS. 2007. Timing is everything: making neurons versus glia in the developing cortex. Neuron 54:357–369. https://doi.org/10.1016/j.neuron.2007.04.019.

97. Sun YE, Martinowich K, Ge W. 2003. Making and repairing the mammalian brain—signaling toward neurogenesis and gliogenesis. Semin Cell Dev Biol 14:161–166. https://doi.org/10.1016/S1041-8365(03)00007-7.

98. Lister R, Mukamel EA, Nery JR, Urlich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, WU JC, Rao A, Esteller M, He C, Haghihg FG, Sejnowski TJ, Behrens MM, Ecker JR. 2013. Global epigenomic reconfiguration during mammalian brain development. Science 341:1237905. https://doi.org/10.1126/science.1237905.

99. Papp B, Plath K. 2013. Epigenetics of reprogramming to induced pluripotency. Cell 152:1324–1343. https://doi.org/10.1016/j.cell.2013.02.043.

100. Shi Y, Kirwan P, Livesey FJ. 2012. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. Nat Protoc 7:1836–1846. https://doi.org/10.1038/nprot.2012.116.

101. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 1303.3997. [q-bio.GN]. https://arxiv.org/abs/1303.3997.

102. Hormodziar F, Alkan C, Ventura M, Hajirasouliha I, Malig M, Hach F, Yorukoglu D, Dao P, Bakhshi M, Sainhulp SC, Eichler EE. 2011. Alu repeat discovery and characterization within human genomes. Genome Res 21:840–849. https://doi.org/10.1101/gr.130398.110.

103. Witherspoon DJ, Xing J, Zhang Y, Watkins WS, Batzer MA, Jorde L. 2010. Mobile element scanning (ME-Scan) by targeted high-throughput sequencing. BMC Genomics 11:410. https://doi.org/10.1186/1471-2164-11-410.

104. Tsvetkov I, Casella JF, Heidmann T. 2000. Members of the SRY family regulate the human LINE retrotransposons. Nucleic Acids Res 28:411–415. https://doi.org/10.1093/nar/28.2.411.

105. Yang N, Zhang L, Zhang Y, Kazazian HH, Jr. 2003. An important role for RUNX3 in human L1 transcription and retrotransposition. Nucleic Acids Res 31:4929–4940. https://doi.org/10.1093/nar/gkg663.

106. Tapia N, Scholer HR. 2016. Molecular obstacles to clinical translation of iPSCs. Cell Stem Cell 19:298–309. https://doi.org/10.1016/j.stem.2016.06.017.

107. Rovitch DH, Kriegstein AR. 2010. Developmental genetics of vertebrate glial-cell specification. Nature 468:214–222. https://doi.org/10.1038/nature09611.

108. Miller FD, Gauthier AS. 2007. Timing is everything: making neurons versus glia in the developing cortex. Neuron 54:357–369. https://doi.org/10.1016/j.neuron.2007.04.019.

109. Sun YE, Martinowich K, Ge W. 2003. Making and repairing the mammalian brain—signaling toward neurogenesis and gliogenesis. Semin Cell Dev Biol 14:161–166. https://doi.org/10.1016/S1041-8365(03)00007-7.

110. Lister R, Mukamel EA, Nery JR, Urlich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, Yu M, Tonti-Filippini J, Heyn H, Hu S, Wu JC, Rao A, Esteller M, He C, Haghihg FG, Sejnowski TJ, Behrens MM, Ecker JR. 2013. Global epigenomic reconfiguration during mammalian brain development. Science 341:1237905. https://doi.org/10.1126/science.1237905.

111. Papp B, Plath K. 2013. Epigenetics of reprogramming to induced pluripotency. Cell 152:1324–1343. https://doi.org/10.1016/j.cell.2013.02.043.

112. Shi Y, Kirwan P, Livesey FJ. 2012. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. Nat Protoc 7:1836–1846. https://doi.org/10.1038/nprot.2012.116.

113. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 1303.3997. [q-bio.GN]. https://arxiv.org/abs/1303.3997.

114. Hormodziar F, Alkan C, Ventura M, Hajirasouliha I, Malig M, Hach F, Yorukoglu D, Dao P, Bakhshi M, Sainhulp SC, Eichler EE. 2011. Alu repeat discovery and characterization within human genomes. Genome Res 21:840–849. https://doi.org/10.1101/gr.130398.110.

115. Witherspoon DJ, Xing J, Zhang Y, Watkins WS, Batzer MA, Jorde L. 2010. Mobile element scanning (ME-Scan) by targeted high-throughput sequencing. BMC Genomics 11:410. https://doi.org/10.1186/1471-2164-11-410.

116. Witherspoon DJ, Xing J, Zhang Y, Watkins WS, Batzer MA, Jorde L. 2010. Mobile element scanning (ME-Scan) by targeted high-throughput sequencing. BMC Genomics 11:410. https://doi.org/10.1186/1471-2164-11-410.

117. Richardson SR, Morell S, Faulkner GJ. 2014. L1 retrotransposons and somatic mosaicism in the brain. Annu Rev Genet 48:1–27. https://doi.org/10.1146/annurev-genet-120213-092412.

118. Wei W, Morrish TA, Alisch RS, Moran IV. 2000. A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. Anal Biochem 284:435–438. https://doi.org/10.1006/abio.2000.4675.

119. Kumaki Y, Oda M, Okano M. 2008. QUMA: quantification tool for methylation analysis. Nucleic Acids Res 36:W170–W175. https://doi.org/10.1093/nar/gkn294.