L1 retrotransposition can occur early in human embryonic development

José A.J.M. van den Hurk¹, Iwan C. Meij², Maria del Carmen Seleme³, Hiroki Kano³, Konstantinos Nikopoulos¹,³, Lies H. Hoefsloot¹, Erik A. Sistermans¹, Ilse J. de Wijs¹, Arijit Mukhopadhyay¹,³, Astrid S. Plomp⁴,⁵, Paulus T.V.M. de Jong⁴,⁶,⁷, Haig H. Kazazian³ and Frans P.M. Cremers¹,³,⁸,*

¹Department of Human Genetics, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands, ²Max Delbrück Centre for Molecular Medicine, 13125 Berlin, Germany, ³Department of Genetics, University of Pennsylvania School of Medicine, Pennsylvania, PA 19104-6145, USA, ⁴Department of Neuromedical Genetics, The Netherlands Institute for Neuroscience, 1105 BA Amsterdam, The Netherlands, ⁵Department of Clinical Genetics, ⁶Department of Ophthalmology, Academic Medical Centre, 1105 AZ Amsterdam, The Netherlands, ⁷Department of Epidemiology and Biostatistics, Erasmus Medical Centre, 3015 GD Rotterdam, The Netherlands and ⁸Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen Medical Centre, 6525 GA Nijmegen, The Netherlands

Received April 13, 2007; Revised and Accepted April 19, 2007

L1 elements are autonomous retrotransposons that can cause hereditary diseases. We have previously identified a full-length L1 insertion in the CHM (choroideremia) gene of a patient with choroideremia, an X-linked progressive eye disease. Because this L1 element, designated L1CHM, contains two 3'-transductions, we were able to delineate a retrotransposition path in which a precursor L1 on chromosome 10p15 or 18p11 retrotransposed to chromosome 6p21 and subsequently to the CHM gene on chromosome Xq21. A cell culture retrotransposition assay showed that L1CHM is one of the most active L1 elements in the human genome. Most importantly, analysis of genomic DNA from the CHM patient’s relatives indicated somatic and germ-line mosaicism for the L1 insertion in his mother. These findings provide evidence that L1 retrotransposition can occur very early in human embryonic development.

INTRODUCTION

LINE-1 (long interspersed nuclear elements-1) or L1 elements are the most abundant autonomous retrotransposons in the human genome, accounting for ~17% of its mass. Although the vast majority of L1s are functionally inactive, 80 to 100 L1 elements remain capable of retrotransposition in the average human genome (1). Most of these active L1s belong to the transcribed group a (Ta) subset and are polymorphic with respect to their presence or absence and retrotransposition capability (2,3). A functional L1 element measures ~6 kb and consists of a 5'-untranslated region (UTR) with internal promoter activity, two open-reading frames (ORFs) and a 3'-UTR with a polyadenylation signal followed immediately by a poly(A) tail (4). Retrotransposition of L1 elements involves transcription, reverse transcription and integration into a new genomic location. Reverse transcription and integration occur in a coupled reaction termed target-primed reverse transcription (5). L1 retrotransposons have altered the human genome through a variety of mechanisms other than their own mobilization (reviewed in 6). L1 elements can transport downstream flanking DNA to a new genomic location in a process called 3'-transduction. This occurs when the L1 polyadenylation signal is bypassed in favor of a downstream polyadenylation site (7–9). L1s are potential generators of genomic deletions (10,11).

At present, information about either the timing or the cell-type specificity of human L1 retrotransposition events is limited. In vitro studies have demonstrated that a variety of human transformed or immortalized cultured cells can...
accommodate retrotransposition (12–14). In addition, a recent report documented retrotransposition of a human L1 element both in cultured rat neural progenitor cells and in the brain of transgenic mice (15). Also, an L1 insertion disrupting the adenomatous polyposis coli (APC) gene has been identified in a colorectal tumor, but was absent in the surrounding constitutional tissue of the patient (16). It thus seems that L1 retrotransposons can mobilize in some somatic cells. However, L1 elements can only propagate and expand if retrotranspositions occur in primordial germ cells, their derivatives in the germ line or early in embryonic development. A mouse model for human L1 retrotransposition (17) and the investigation of a disease-producing L1 insertion into the cytochrome b(245) beta subunit gene (18) have indicated that L1 retrotransposition can occur during meiosis I in both male and female germ cells. An insertion of an engineered human L1 element at an early stage in development of a transgenic mouse has also been reported (19). It should be noted, however, that transcription of this L1 transgene was under the control of the mouse RNA polymerase II large subunit promoter in addition to its endogenous promoter.

L1 retrotransposons can cause hereditary diseases by integrating into genes (20,21 and references therein). Previously, we identified an L1 insertion in the CHM gene of a Dutch patient diagnosed with choroideremia [CHM (MIM303100)], an X-linked progressive eye disorder. The L1 was integrated into exon 6 in reverse orientation, leading to aberrant splicing of the CHM mRNA (22). Interestingly, this L1 element, designated L1(CHM), appeared to be full-length, which prompted us to characterize it in detail. L1(CHM) contains two 3'-transduced sequences that allowed us to infer its retrotransposition path. We also demonstrate that L1(CHM) is highly active, but more importantly, we present data indicating that the actual L1 insertion into the CHM gene took place during early embryogenesis in the patient’s mother.

RESULTS AND DISCUSSION
Retrotransposition path of L1(CHM)

L1(CHM) was PCR-amplified from genomic DNA of the CHM patient using primers flanking exon 6 of the CHM gene. Subsequent sequence analysis of the PCR product was performed with primers derived from a consensus sequence of active human L1 elements. L1(CHM) has a length of 6688 bp and is surrounded by a perfect 14 bp TSD (5'-AGAAGATCAATTAG-3', designated TSD1 in Fig. 1). The 3'-UTR contains the ACA sequence (positions 5922–5924) diagnostic for the Ta subset (23). At its 3'-end, L1(CHM) contains three poly(A) stretches of 60, 14 and 71 nucleotides, respectively, interrupted by transduced sequences of 119 and 406 bp (Fig. 1 and Supplementary Material, Fig. S1). These 3'-transductions are indicative of multiple rounds of retrotransposition. A transduced sequence acts as a molecular address that permits the identification of the precursor to the transduction event. A BLAST search (24) of the NCBI non-redundant nucleotide sequence database (http://www.ncbi.nih.gov) showed that the 406 bp transduction of L1(CHM) is a precise match to sequences from BX005091.7 (EMBL) and AC004200.1 (GenBank) on chromosome 6p21. BX005091.7 lacks an L1 element and represents the pre-integration site. A 6212 bp L1 is located immediately 5' of the 406 bp sequence in AC004200.1 (Fig. 1). This L1 is identical in sequence to L1(CHM) (1–6211) except for three additional
guanine residues at its 5'-end, a single nucleotide change in ORF2 (T position 28485 versus C in L1CHM position 2439), and a poly(A) stretch of 58 instead of 60 adenosine residues. As in L1CHM, a 119 bp transduced sequence followed by a poly(A) stretch of 14 nucleotides is present at the 3'-end. The L1 element is flanked by a perfect 15 bp TSD (5'-AAAA AACACAGTGAA-3'), designated TSD2 in Fig. 1). Thus, the L1 in AC004200.1 and the precursor to L1CHM are likely to be alleles at a locus on chromosome 6p21, and this locus is polymorphic as to the presence or absence of an L1 element. Indeed, it has previously been established that an L1 element is present at this locus in 0 to 4% of alleles from different human populations (1,2).

A BLAT search (25) of the UCSC genome database [http://genome.ucsc.edu, Human May 2004 (hg17) assembly] using the 119 bp transduction of L1CHM and its precursor produced exact matches near the telomeres of the short arms of chromosome 10 (10p15.3 nucleotides g.60462–g.60344) and chromosome 18 (18p11.32 nucleotides g.14856–g.14738). No full-length L1 element is present. Likewise, a BLAST search (24) of the NCBI non-redundant nucleotide sequence database (http://www.ncbi.nih.gov) revealed no sequences in which the 119 bp segment was preceded by an L1. Thus, the precursor to the L1CHM precursor may reside on either chromosome 10pter or 18pter. PCR analysis of these regions in 200 Dutch individuals failed to detect an L1 element. This indicates that the presumed L1 element at 10p15 or 18p11 is very rare or no longer present in this population, although a technical problem cannot be excluded. Taken together, these findings indicate a scenario in which an L1 on chromosome 6p21 retrotransposed into the CHM gene on chromosome Xq21. The L1 element on chromosome 6p21 in turn is derived from either chromosome 10p15 or chromosome 18p11.

**L1CHM displays high retrotransposition activity**

L1CHM was tested in an established cultured cell assay that uses an enhanced green fluorescent protein (EGFP) marker to determine retrotransposition activity (1,26). L1RP, an L1 element that was discovered as a *de novo* insertion into the retinitis pigmentosa 2 gene (27,28), served as a standard for comparison in this assay. The retrotransposition activity of L1CHM was about one-third that of L1RP, classifying L1CHM as a highly active or ‘hot’ L1 (1). Hot L1s comprise only a
small proportion of active L1s but are responsible for the majority of retrotransposition in the human population.

Somatic and germ-line mosaicism for L1CHM

In order to determine when the L1 insertion into the CHM gene took place, the relatives of the CHM patient were studied. CHM exhibits an X-linked mode of inheritance. Males affected with CHM experience night blindness, followed by progressive visual field loss leading to tunnel vision and often blindness. These symptoms are paralleled by degeneration of the choroid and retina, beginning in the midperiphery and gradually progressing toward the macula (29). Figure 2 shows a fundus photograph of the patient (II.1) demonstrating the extensive chorioretinal hypopigmentation with preservation of pigment in the central macula that is typical for CHM. Female CHM carriers generally have no serious visual impairment. They can, however, be diagnosed by fundoscopic examination as they manifest patchy areas of chorioretinal degeneration resulting from X chromosome inactivation. Ophthalmological examination of the patient’s mother (I.2) revealed three hyperpigmented radial lines in the temporal inferior part of the right fundus but was otherwise unremarkable (Fig. 2). One of the patient’s sisters (II.2) showed normal optic discs, marked peripapillary atrophy in zones alpha and beta, a hypopigmented fundus with some white flecks in the periphery, slight pepper and salt pigment alterations in the peripheral retina and hyperpigmented equatorial rings and streaks (Fig. 2). She was diagnosed as a carrier. Apart from changes consistent with her myopia, the patient’s other sister (II.3) had a normal fundus appearance (not shown).

Initial genotyping of the family for two polymorphic markers within the CHM gene, i.e. a single-nucleotide polymorphism (c.351G/A) in exon 5 (30) and a microsatellite marker in intron 9 (31), showed that the patient and his two sisters share the disease haplotype (Fig. 3A). For female II.3, this was inconsistent with her clinical examination. Therefore, the presence of the L1CHM insertion in the family members was investigated using a PCR assay that amplified the mutant CHM allele containing the L1CHM insertion and the normal CHM allele. Both alleles were found in lymphocytederived DNA from the mother but the mutant PCR product had a significantly reduced level compared with the normal PCR product (Fig. 3B). These results indicate that the mother is a somatic mosaic for the L1 insertion. PCR analysis further showed that female II.2 is heterozygous for the L1 insertion, whereas female II.3 has only the normal CHM allele (Fig. 3B). The finding that female II.3 inherited the at-risk haplotype but not the L1CHM insertion proves that the mother is a germ-line mosaic. Thus, these data demonstrate that the L1 retrotransposition took place in the patient’s mother. A precursor to L1CHM is actually present on one of her 6p21 alleles (Supplementary Material, Fig. S2). Since the patient’s mother shows both somatic and germ-line mosaicism for the L1 insertion into the CHM gene, the L1 retrotransposition event must have occurred during early embryogenesis, before the germ-line segregated from the somatic lineages.

In an accompanying paper, Garcia-Perez et al. (32) report that human embryonic stem cells express endogenous L1 elements and can accommodate L1 retrotransposition in vitro. In reinforcement of these data, our detailed analysis of a disease-causing L1 insertion provides in vivo evidence that L1 retrotransposition can occur very early in human embryonic development.

MATERIALS AND METHODS

Molecular characterization of L1CHM

L1CHM was amplified from genomic DNA of the CHM patient with the Expand Long Template PCR system (Roche) using primers flanking exon 6 (5'-CGGAGGACTGGAATTTACGCTTTATGA-3' and 5'-GAATTTCAGCTAATCAATTCTGAGCCTG-3') as described (22). The PCR product was purified with Qiaquick columns (Qiagen). Sequence analysis was performed using primers derived from a consensus sequence of active human L1 elements (adapted from 28). Primer sequences are available upon request. Sequencing reactions were carried out by the use of BigDye Terminator chemistry according to the guidelines provided by the manufacturer (PE Applied Biosystems) and analyzed on an ABI 3700 capillary DNA sequencer (PE Applied Biosystems).
Molecular analysis of possible L1<sub>CHM</sub> precursor loci

To investigate the presence of a precursor L1 element at 10p15/18p1, a primer derived from the 3'-UTR of L1<sub>CHM</sub> (5'-CCTAATGCTAGTACACA-3') and a primer specific for both 10p15 and 18p1 (5'-AAATGTCTAGAGATGG-3' and 5'-TGGGATGAGCATGAATAAC-3' (272 bp product)).

Retrotransposition assay

Primer 5'-ATTTTCGGCAGGGTTAGATAAGATCATATT AACCTCTCCGGTTC-3' containing a NotI site and primer 5'-ATACATATGTACACGGTACGACCACGCA-3' containing a BstZ1 site were used for PCR amplification of L1<sub>CHM</sub> from the patient's genomic DNA by Expand Long Template PCR (Roche) following the protocol provided by the manufacturer. The PCR product was digested with NotI and BstZ1 and swapped into plL1RP-EGFP, a previously published construct containing L1RP tagged with an EGFP reporter. The PCR reaction mixture contained 1 ng of genomic DNA and 1 unit of Taq polymerase buffer, 1.5 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 400 nM of primers, 250 ng of genomic DNA and 1 unit of Taq polymerase (Invitrogen). After 2 min at 94°C, 35 cycles were carried out consisting of 20 s at 94°C, 20 s at 56°C and 1 min at 72°C, followed by 2 min at 72°C. PCR products were analyzed on a 1.5% agarose gel stained with ethidium bromide.

Clinical and molecular analysis of the CHM family

The patient and his relatives were examined by fundoscopy and functional ophthalmological methods. Genomic DNA was isolated from blood lymphocytes by a salt extraction procedure (33). Genotyping for a single-nucleotide polymorphism containing a Z17I site were used for PCR amplification of the wild-type allele and sequencing of the variant allele. For the PCR amplification of mutant alleles flanking exon 6 (5'-TGCTTTATGA-3'), the following primers were used: Primer 5'-ATACATATGTACACGGTACGACCACGCA-3' and 5'-TGCTCTCTTCAAAGCTGTCAGACAG-3'. The PCR reaction mixture contained 1 ng of genomic DNA and 1 unit of Taq polymerase buffer, 1.5 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 400 nM of primers, and 0.25 U of Taq polymerase. The PCR products were purified with Qiaquick columns (Qiagen) and sequenced as described above. Analysis of a microsatellite marker in intron 9 was performed using a 6-FAM-labeled forward primer (5'-TCAGTACAG TGCCCTTGTAGTGGA-3') and an unlabeled reverse primer (5'-GGTAGATATTTAATACCAAGGAG-3') as described previously (31). To investigate the presence of the L1<sub>CHM</sub> insertion, a PCR assay amplifying both the mutant allele containing L1<sub>CHM</sub> and the normal CHM allele was designed. For the PCR amplification of mutant CHM alleles, primers derived from intron 6 (5'-CAGCTGGGTGT TTATATTATATT-3') and intron 6 (5'-CAGAAGATGGGT AAAATTAGT-3') were used in a 1.5% agarose gel stained with ethidium bromide.

ACKNOWLEDGEMENTS

We thank Dorien van de Pol, Krista Voesenek and Ellen Blokland for excellent technical assistance. We are also grateful to the patient and his family for participating in this study. Our work was supported by the kind donations of several Dutch organizations, i.e. Algemene Nederlandse Vereniging ter Verkoening van Blindheid, Gelderse Blindenvereniging, Rotterdams Vereniging Blindenbelangen, Stichting Blindenhulp, Stichting OOG and Stichting voor Ooglijders. This work was also supported by the European Union Research Training Network Grant RETNET MRTN-CT-2003-504003 (K.N., A.M.).

Note added in proof: Real-time Q-PCR analysis of genomic DNA from the mosaic mother I.2 and the choroideremia patient I.1 indicated that I.2 carries the L1<sub>CHM</sub> insertion in 3% of her lymphocytes (Supplementary Material, Fig. S3).

REFERENCES


