A recurrent translocation is mediated by homologous recombination between HERV-H elements

Karen E. Hermetz, Emory University
Urvashi Surti, University of Pittsburgh School of Medicine
Jannine D. Cody, University of Texas Health Science Center at San Antonio
Katie Rudd, Emory University

Journal Title: Molecular Cytogenetics
Volume: Volume 5, Number 6
Publisher: BioMed Central | 2012-01-19, Pages 1-8
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1186/1755-8166-5-6
Permanent URL: http://pid.emory.edu/ark:/25593/dwqrd

Final published version: http://www.molecularcytogenetics.org/content/5/1/6

Copyright information:
© 2012 Hermetz et al; licensee BioMed Central Ltd.
This is an Open Access work distributed under the terms of the Creative Commons Attribution 2.0 Generic License (http://creativecommons.org/licenses/by/2.0/).

Accessed September 26, 2019 1:28 PM EDT
A recurrent translocation is mediated by homologous recombination between HERV-H elements

Karen E Hermetz1, Urvashi Surti2, Jannine D Cody3 and M Katharine Rudd1*

Abstract

Background: Chromosome rearrangements are caused by many mutational mechanisms; of these, recurrent rearrangements can be particularly informative for teasing apart DNA sequence-specific factors. Some recurrent translocations are mediated by homologous recombination between large blocks of segmental duplications on different chromosomes. Here we describe a recurrent unbalanced translocation caused by recombination between shorter homologous regions on chromosomes 4 and 18 in two unrelated children with intellectual disability.

Results: Array CGH resolved the breakpoints of the 6.97-Megabase (Mb) loss of 18q and the 7.30-Mb gain of 4q. Sequencing across the translocation breakpoints revealed that both translocations occurred between 92%-identical human endogenous retrovirus (HERV) elements in the same orientation on chromosomes 4 and 18. In addition, we find sequence variation in the chromosome 4 HERV that makes one allele more like the chromosome 18 HERV.

Conclusions: Homologous recombination between HERVs on the same chromosome is known to cause chromosome deletions, but this is the first report of interchromosomal HERV-HERV recombination leading to a translocation. It is possible that normal sequence variation in substrates of non-allelic homologous recombination (NAHR) affects the alignment of recombining segments and influences the propensity to chromosome rearrangement.

Keywords: HERV-H, HERV, NAHR, translocation, t(4;18), recurrent translocation, 4q35.1, 18q22.3, 18q
breakpoints lack significant sequence homology that would suggest NAHR [18-20].

NAHR between homologous interspersed repeats like LINE, Alu, and human endogenous retrovirus (HERV) elements can also lead to genomic rearrangements [21,22]. Alu-Alu recombination events have been described at multiple loci, giving rise to pathogenic deletions, duplications, and translocations [6,23-26]. Homologous recombination between HERV15 elements underlies the recurrent Y chromosome microdeletion that removes the azoospermia factor a (AZFa) region and causes male infertility [27-29]. Recently, a HERV-H-mediated deletion of chromosome 8q13.3 has been described in a child with heterozygous loss of the EYA1 gene and branchio-oto-renal syndrome [30]. Nevertheless, homologous recombination between HERV elements on different chromosomes has not been previously described as a mechanism for recurrent translocations.

HERVs make up ~3% of the human genome [31,32]; however, most copies have mutations and/or deletions that disrupt one or more of the ORFs, rendering the retrovirus inactive [33-35]. Though most HERVs are inactive as retrotransposons, they may spread through the genome via ectopic recombination processes. Phylogenetic studies of HERV-K elements have revealed signatures of intraelement gene conversion and recombination [36,37]. Thus, the density of HERVs and the sequence homogenization between copies make them ideal substrates for NAHR in the human genome.

Here we describe a recurrent translocation mediated by NAHR between HERVs on chromosomes 4q and 18q. Sequencing the breakpoint junctions in two unrelated individuals with similar translocations [38,39] revealed breakpoints within a few hundred basepairs (bp) of each other. Both sets of translocation breakpoints are located in HERV-H elements, and the orientation and sequence homology between recombining segments on 4q and 18q are consistent with a NAHR rearrangement mechanism.

Results and discussion
In an earlier study, we identified two individuals (18q-82C and 18q-146C) with unbalanced translocations between the ends of the long arms of chromosomes 4 and 18. Both carry a derivative chromosome missing the end of chromosome 18q, with an additional copy of the end of 4q, as shown previously by array comparative genomic hybridization (CGH) [38,39]. Although two children are not enough for us to evaluate the phenotype associated with this unbalanced translocation comprehensively, some shared clinical features are worth noting. Patients 18q-82C and 18q-146C both exhibited developmental delays, auditory canal atresia, midface hypoplasia, microcephaly, and a broad nasal bridge. Parental studies revealed that 18q-82C carries a de novo translocation, derived from paternal chromosomes 4 and 18 [39]. Parents of 18q-146C were not available for study.

Using high-resolution array CGH, we resolved the 4q and 18q breakpoints in both subjects. We designed a custom oligonucleotide array targeting the 5 Mb spanning the 4q and 18q breakpoints with a mean probe spacing of one oligonucleotide per 100 bp. Array CGH revealed the same breakpoints in both individuals: a 6.97-Mb loss of 18q and a 7.30-Mb gain of 4q (Figure 1).

Based on the array CGH data, we designed PCR primers to amplify across the breakpoint junctions of the derivative chromosomes 18 from patients 18q-82C and 18q-146C. We cloned and sequenced the breakpoint junctions to generate complete sequence across the two independent junction fragments (GenBank sequences 18q82C_junction and 18q146C_junction). In both translocations, the junction between chromosomes 18 and 4 lies in a HERV-H element (Figure 2). As represented in the reference genome (Build 36.1, hg18), the HERV-H elements on chromosomes 4q and 18q are 4.6 kilobases (kb) and 5.7 kb, respectively, and are ~92% identical overall.

We aligned our junction sequences from patients 18q-82C and 18q-146C to sequence from the 4q and 18q HERV-H elements in the human genome assembly (Build 36.1, hg18). Recombination for both translocations occurred in an ~three-kb region that is 91% identical between the 4q and 18q HERV-Hs (Figure 2). Though the HERV-Hs are highly identical, chromosome-specific SNPs distinguish the 18q and 4q sides of the translocation junctions and allow us to further resolve the sites of recombination. The sites of recombination mediating the translocations in patients 18q-82C and 18q-146C are ~150 bp apart (Figure 3).

It is possible that some copies of the 4q and 18q HERV-Hs share more homology than represented in the human genome assembly, which may affect the propensity of the two chromosomes to recombine. There are reports of sequence variation in HERVs at several loci, which may occur via gene conversion or transposition processes [29,36,37]. To capture the actual HERV-H sequences that recombined to form the translocations in patients 18q-82C and 18q-146C, it was necessary to sequence the parental 4q and 18q alleles. Microsatellite analysis of 18q-82C and his parents revealed that the de novo translocation was paternal in origin [39]. We sequenced both alleles of the HERV-H elements on 4q and 18q in 18q-82C’s father (18q-82P). The two 18q HERV-H alleles in 18q-82P were 99.5% identical to the 18q HERV-H in the
reference genome, with no significant differences between the two alleles. However, 18q-82P has two different 4q HERV-H alleles, described here as 4qA (GenBank sequence 82Pchr4HERVHA) and 4qB (GenBank sequence 82Pchr4HERVHB).

The HERV-H elements on 4qA and 4qB are 99.0% identical, but 4qB has a 122-bp duplication not present in 4qA (Figure 2). The duplication is made up of a 61-bp tandem repeat present in two copies and four copies on the 4qA and 4qB alleles, respectively. The human

---

**Figure 1** High-resolution array CGH reveals the breakpoints of the unbalanced translocation from patient 18q-146C. (a) The recurrent translocation occurs between 4q35.1 and 18q22.3, as indicated by the boxed chromosome bands. (b) The 6.97-Mb loss of 18q and (c) the 7.30-Mb gain of the end of 4q are shown. Averaged log₂ ratios of probe signal intensities are shown (black dots). Dashed lines indicate log₂ ratios of -1, 0, and +1. Similar array CGH results were obtained from patient 18q-82C's translocation.
reference genome (Build 36.1, hg18) corresponds to the 4qB allele, including the 122-bp duplication. The 122-bp duplication is not present in the 18q HERV-H or in the HERV-H consensus sequence [40], consistent with the duplication being a new event that arose on the 4qB allele.

To determine the frequency of the 4qA and 4qB alleles in the human population, we designed a genotyping assay that distinguishes the two alleles. We performed a nested PCR that specifically amplifies a six-kb region, including the entire 4q HERV-H in the first PCR, followed by a second PCR that amplifies the region around the 122-bp duplication. This results in a 326-bp band for the 4qA allele and a 446-bp band for the 4qB allele. We genotyped DNA from 99 Caucasians obtained from the Coriell Cell Repository (Human Variation Panel HD100CAU), representing 198 4q alleles, of which 37 (18.7%) were 4qA and 161 (81.3%) were 4qB (Additional file 1). We also genotyped 62 samples from 10 populations included in the Human Genome Diversity Panel (HGDP). Allele frequencies were not significantly different between geographic populations, and as a group, there were 26 (21.0%) 4qA alleles and 98 (79.0%) 4qB alleles (Additional file 1). Thus, the 4qA

![Figure 2 HERV-H elements from chromosomes 4q and 18q, as compared to the intact Herv-H consensus sequence.](image)

![Figure 3 Sequence alignment of the translocation junction sequences from patients 18q-82C and 18q-146C.](image)
allele carried by 18q-82P is the minor allele, and the
4qB allele in the reference genome is the major allele.

Conclusions
Recurrent copy number variations (e.g., deletions, duplica-
tions, and translocations) provide mechanistic insight
into the etiology of chromosome rearrangements. We
sequenced the breakpoint junctions of two independent
translocations with nearly identical breakpoints on chro-
mosomes 4q and 18q. Analysis of the recombining seg-
ments revealed that the breakpoints lie in HERV-H
elements that are 92% identical. NAHR between other
HERVs has been found to underlie interstitial deletions
of the Y chromosome and chromosome 8; in both cases,
the recombining HERVs were ~94% identical [27-30].

NAHR between larger segmental duplications causes
the most common microdeletion and microduplication
syndromes. In these cases, recombining segmental dupli-
cations range from 10 kb to hundreds of kilobases in
size and are typically 95% identical or greater [10,41-43].
NAHR between smaller substrates with greater sequence
divergence, such as the HERV-Hs described here, is
probably less frequent than NAHR between large seg-
mental duplications. However, sequence variation in
recombining segments could also impact rearrangement
frequency. Polymorphism in the HERV15 elements that
mediate the recurrent AZFa microdeletion leads to bet-
ter sequence alignment between the recombining
regions on the Y chromosome and is predicted to pre-
dispose to microdeletion [29]. It is possible that
sequence variation in the 4q HERV-H also affects its
propensity to recombine with the 18q HERV-H. Our
study of two recurrent translocations is not comprehen-
sive enough to draw conclusions about HERV-H
sequence variation and recombination frequency. How-
ever, the lack of the 122-bp duplication in the 4qA
HERV-H makes it more similar to the 18q HERV-H,
which may be important for recombination between the
two chromosomes. Furthermore, the fact that 18q-82P
carries a 4qA allele is consistent with recombination
between the 4qA HERV-H and the 18q HERV-H to give
rise to patient 18q-82C's translocation. Translocations in
patients 18q-82C and 18q-146C both occurred in the
same region of the HERV-H that is polymorphic
between the 4q alleles (Figure 2), and the 122-bp duplica-
tion alters the sequence alignment between this part of
chromosomes 4q and 18q.

It is important to point out that we only recognized
the signatures of HERV-HERV NAHR at the transloca-
tion breakpoints of patients 18q-82C and 18q-146C
after sequencing the breakpoint junctions. Inferring
chromosomal rearrangement mechanisms from lower-
resolution approaches (e.g., array CGH only) is likely to
underestimate the frequency of NAHR-mediated events
between shorter homologous segments and only detect
NAHR between large blocks of segmental duplication.
Furthermore, NAHR between homologous interspersed
repeats, such as Alus, LINEs, and HERVs, would be
overlooked by strategies focusing on recombination
between segmental duplications [15]. We have pre-
viously detected a translocation that is the product of
interchromosomal LINE-LINE recombination between
L1PA2s on chromosomes 6 and 16 that are 96% identi-
cal over ~six kb by sequencing across the translocation
breakpoint junction [6]. In addition, sequencing of nor-
mal copy number variation breakpoints has uncovered
signatures of intrachromosomal NAHR [44,45]. Future
sequencing-based studies of other chromosome rearran-
gements will likely capture more NAHR events between
shorter homologous segments, which would give us a
better understanding of the requirements for interchro-
mosomal and intrachromosomal NAHR in the human
genome.

Methods
Participants
We obtained informed consent from individuals with
chromosome 18 abnormalities and their families. The
human subjects protocol was approved by the Institu-
tional Review Board of the University of Texas Health
Science Center at San Antonio.

Array CGH
Using a 244K platform from Agilent Technologies
(Santa Clara, CA), we designed a custom two-plex array
covering the five-Mb regions spanning the previously
described breakpoints on chromosomes 4q and 18q
[39], with a mean probe spacing of one oligonucleotide
per 100 basepairs (bp). Oligonucleotides were designed
using Agilent’s eArray program (https://earray.chem.agi-
lent.com/earray/). To minimize non-unique oligonucleo-
tides that would not be informative in array CGH, we
performed a high definition (HD) probe search to prefer
existing “catalog probes” and used the most stringent
“similarity score filter” designed to select probes that
hybridize to only one genomic location. The unique
identifier (AMADID) for the array design is 021748; this
design is available upon request.

Lymphoblastoid cell lines derived from 18q-82C, 18q-
82P, and 18q-146C were established previously [39]. We
extracted genomic DNA from cell lines using the Gentra
Puregene DNA Extraction Kit (Qiagen, Valencia, CA).
Subject DNA was co-hybridized with reference DNA
from the GM15510 cell line (Coriell Cell Repositories,
Camden, NJ). Arrays were scanned using a GenePix
4000B scanner (Molecular Devices, Sunnyvale, CA), and
signal intensities were evaluated using Feature Extrac-
tion Version 9.5.1.1 software (Agilent Technologies,
Santa Clara, CA). We used DNA Analytics Version 4.0 software (Agilent Technologies, Santa Clara, CA) to analyze the array data and call breakpoints.

**Breakpoint PCR**
Starting with breakpoints identified by array CGH, we designed PCR primers to amplify across the translocation breakpoint junctions in 18q-82C and 18q-146C. The chromosome 18-specific primer is 5’-TCAACTGTAGAAGGAGCTAGCTCCCCCTA-3’, and the chromosome 4-specific primer is 5’-GGTCAATGATCCGGGGTTCTGGATG-3’. We performed PCR using TaKaRa Ex Taq polymerase (Clontech Laboratories, Inc., Madison, WI) with 1× PCR buffer, 0.2 mM dNTP, 8 pmol of each primer, and 50-100 ng of DNA template. PCR conditions were as follows. 94°C for 1 min; 10 cycles at 94°C for 30 s, 66°C for 1 min, decreasing 0.5°C per cycle, and 72°C for 7 min; 20 cycles at 94°C for 30 s, 61°C for 1 min, and 72°C for 7 min; and a final extension at 72°C for 10 min. Bands were visualized via gel electrophoresis on a 1% agarose gel.

**Nested PCR**
We designed a nested PCR to specifically amplify the HERV-Hs on 4qA and 4qB. The first PCR amplifies a 6-kb region larger than the 4q HERV-Hs using the following primers: 5’-GATCATTTGTCAAATGAAATCTCA-CAAGAGGC-3’ and 5’-GGTCAATGATCCGGA GGTTCTGGATG-3’. The PCR reagents were the same as described above, except for the addition of betaine (0.7 M final concentration in 50 μl PCR). Conditions for the first PCR were: 94°C for 1 min; 10 cycles at 94°C for 30 s, 65°C for 1 min, decreasing 0.5°C per cycle, and 72°C for 7.5 min; 30 cycles at 94°C for 30 s, 60°C for 1 min, and 72°C for 7.5 min; and a final extension at 72°C for 10 min.

We diluted amplicons from the first PCR 1:1000 to use as template in the second PCR. The second PCR amplifies a 326-bp or a 446-bp product from the 4qA and 4qB alleles, respectively, using the following primers: 5’-CACCTGTTTGTGGTCATCAC-3’ and 5’-ACTTTCCCCTCTCCAGAAA-3’. Conditions for the second PCR were: 94°C for 1 min; 35 cycles at 94°C for 20 s, 55°C for 10 s, and 72°C for 10 s; and a final extension at 72°C for 1 min. Bands were visualized via gel electrophoresis on a 1% agarose gel.

**Sequence analysis**
We purified PCR products from agarose gels using the QIAquick gel extraction kit (Qiagen, Valencia, CA), and cloned them into a TOPO-TA vector following the manufacturer’s protocol (Invitrogen, Carlsbad, CA). We transformed the ligated construct into SURE 2 Supercompetent Cells (Agilent Technologies, Cedar Creek, TX) following the manufacturer’s protocol. We propagated plasmids in recombination-deficient SURE 2 *Escherichia coli* to prevent rearrangement of the cloned insert.

We purified plasmid DNA (Qiagen Miniprep kit, Valencia, CA) and submitted plasmids for Sanger sequencing (Beckman Coulter Genomics, Danvers, MA). DNA sequences were analyzed by comparing reads to the human genome reference assembly (NCBI36/hg18) using the BLAT tool [46] on the UCSC Genome Browser (http://genome.ucsc.edu/). Junction sequences from 18q-82C (18q82C_junction) and 18q-146C (18q146C_junction) have been deposited in GenBank. The breakpoint junctions on chromosome 4 correspond to positions chr4:183974152-183974212 (18q-146C) and chr4:183974285-183974380 (18q-82C) and the breakpoint junctions on chromosome 18 correspond to positions chr18:69144350-69144410 (18q-146C) and chr18:69144483-691444578 (18q-82C) of the NCBI36/hg18 build of the human genome.

We also cloned and sequenced the products of the first round of nested PCR from 82P to characterize the entire 6-kb HERV-H regions from the 4qA and 4qB alleles. These sequences have been deposited in GenBank as 82Pchr4HERVHA and 82Pchr4HERVHB. We aligned the intact HERV-H sequences from 4qA and 4qB to the HERV-H consensus as represented in RepeatMasker [40] and analyzed the 61-bp tandem repeat within the HERV-H using Tandem Repeats Finder (TRF), http://tandem.bu.edu/trf/trf.html[47].

**Additional material**

**Additional file 1: Gentoype results of the 4qA and 4qB alleles in the population**

**Acknowledgements**
The authors wish to thank families and members of the Chromosome 18 Registry for their participation in this study. This study was supported by a grant from the National Institute of Mental Health (1R01MH092902 to MKR). We thank Madhuri Hegde for comments on the manuscript and Cheryl Strauss for editorial assistance.

**Author details**
1Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA. 2Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. 3Department of Pediatrics, Chromosome 18 Registry and Research Society, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA.

**Authors’ contributions**
KEH performed the array CGH, breakpoint sequencing, and genotyping experiments. KEH and MKR performed the sequence alignments and analyzed the data. MKR drafted the manuscript. US and JDC participated in the design of the study and the manuscript. All authors have read and approved the manuscript.

**Competing interests**
The authors declare that they have no competing interests.


References


doi:10.1186/1755-8166-5-6
Cite this article as: Hermetz et al. A recurrent translocation is mediated by homologous recombination between HERV-H elements. Molecular Cytogenetics 2012 5:6.