Recurrent deletion of a region containing exon 24 of the RB1 gene caused by non-homologous recombination between a LINE-1HS and MER21B element

P Albrecht, J Bode, K Buiting, A K Prashanth, D R Lohmann

Key points

- Retinoblastoma, a malignant tumour of the eye, is caused by mutations in the RB1 gene. About 15–20% of oncogenic mutations in this gene are gross deletions.
- We have identified a recurrent deletion of a region containing exon 24 of the RB1 gene in peripheral blood of three patients with retinoblastoma. The 5′ and 3′ deletion breakpoints of these deletions are clustered in an L1HS and MER21B element, respectively.
- The region of the deletion breakpoints does not show DNA sequences with direct or inverted similarity. However, the breakpoints are localised at the borders of strong scaffold/matrix attachment elements that mark the position of recombinogenic DNA structures.
- Our findings emphasise the role of scaffold/matrix attachment regions for deletion formation in humans.

METHODS

Samples
DNA from peripheral blood and from fresh frozen retinoblastoma samples was extracted as described previously. We investigated constitutional DNA from 60 patients with bilateral or familial retinoblastoma and tumour DNA from 85 patients with isolated unilateral retinoblastoma. Most of these samples had previously been screened for RB1 gene point mutations but with negative results. Using quantitative multiplex polymerase chain reaction (PCR) we identified a spectrum of gross deletions that was heterogeneous with respect to extent and location. In three patients, we found deletions in a region that contains exon 24 of the RB1 gene. Further analysis showed that the 5′ and 3′ deletion breakpoints of these three mutations are located close to each other in an L1HS and MER21B element, respectively. These two DNA elements belong to different classes of interspersed repetitive DNA. The regions surrounding the 5′ and 3′ breakpoints do not show any sequence similarity. However, they are localised at the borders of strong scaffold/matrix attachment elements that mark the position of recombinogenic DNA structures.

Quantitative multiplex PCR
Quantitative multiplex PCR was performed using the multiplex PCR assay described by Richter et al with only a few minor adaptations of primer sequences and controls. Products were analysed on an ABI 3100 Genetic Analyzer (Applera). Peak integrals were determined using GeneScan and Genotyper software (Applera).

Long range PCR and sequencing
Long range PCR of a fragment spanning intron 23 to intron 26 of the RB1 gene was performed as described by Brenmer et al. To obtain smaller fragments the deleted region, two long range PCRs were performed with primers GCT GGT TCAATA CAA ATG (RBg166248se), TCA TCT GAT CCT TCA AAG CGT AGT (RBg173837as), and AAA AAT ATA AGC TCT GAG AAT GGA GAT (RBg171908as) (fig 1). To detect both normal and mutant alleles in heterozygous samples, an additional long range PCR was performed using three primers in one reaction: RBg166248se, RBg171908as, and AGT TGA ATT TAT GCT CAT TCT TGC (RBg170286se) (fig 2). RBg170286se binds to a sequence that is deleted in all three mutant alleles investigated here (fig 3A). PCRs with these three primers on samples heterozygous for one of the exon 24 deletions identified result in a product specific for the deleted allele and in a 1.6 kb product from the normal allele. For long range PCRs we used the Roche Expand Long Template kit (Roche, Mannheim, Germany).

For sequencing, the PCR products obtained with primers RBg166248se and RBg171908as were separated on a 1% agarose gel, cut out, and eluted. As sequencing primers we used markers that mark the position of recombinogenic DNA structures.

Abbreviations: LINE, long interspersed nuclear elements; LOH, loss of constitutional heterozygosity; PCR, polymerase chain reaction; SIDD, stress induced duplex destabilisation; S/MARs, scaffold/matrix attachment regions
RESULTS AND DISCUSSION

Gross deletions were identified in 41 patients with retinoblastoma. In line with the results reported by Richter et al. and Houdayer et al., the spectrum of gross deletions was heterogeneous with respect to extent and localisation. In both our data set and the data sets published by Richter et al. and Houdayer et al., recurrent deletions that affect only a single exon are infrequent. Richter et al. identified two mutations of this kind: a deletion of exon 17 in three patients and of exon 13 in two patients; Houdayer et al. identified deletions of exon 2 and 8 in one patient each and a deletion of exon 3 in two patients. The breakpoint sequences of these deletions were not reported in the study of either group. In our series, three samples (peripheral blood DNA from M4716, tumour DNA from M14400 and M3625) also showed smaller junction products (labelled with asterisks) from RB1 alleles with deletion mutations.

We used RBg166248se, RBg171908as, AGG ACA CAA ACA AAT GGA AG (RBg166914se), and AAT TTA GGA TGG AAG CTG GT (RBg171317as). The products of sequencing reactions were analysed on an ABI 3100 Genetic Analyzer. The results of sequence analysis were compared using Lasergene sequence analysis software (DNASTAR).

Figure 1. Long range PCR with primers RBg166248se and RBg171908as (for primer localisation also see fig 3A). PCR from DNA from a normal control shows the expected 5.7 kb product only. All three patients (blood DNA from M4716, and tumour DNA from M14400 and M3625) also show smaller junction products (labelled with asterisks) from RB1 alleles with deletion mutations.

The 5' breakpoint is located in intron 23 at nucleotide 167157 (RBg167157) in a region of 6017 bp with high sequence similarity to an LIHS, a member of the family of long interspersed nuclear elements (LINE). The 3' breakpoint of the mutation is located in intron 24 at RBg170762 within a 764 bp region with similarity to the DNA transposon MER21B. Interestingly, the 5' and 3' breakpoints of the deletions identified in the other two patients are located in the same interspersed repetitive DNA elements. In patient M4716, the 5' and 3' breakpoints are located after nucleotide position RBg167672 and RBg171215, respectively. In addition to loss of bases RBg167673 to 171214, the mutant allele also shows an insertion of 19 bp. A BLASTn sequence comparison of this inserted sequence to the human genome showed no significant similarities to any sequence mapped on chromosome 13. The deletion identified in the sample of patient M3625 is more complex. The mutant allele may be described as a combination of a 2988 bp deletion (del167988^170979) and an insertion of 3 bp (ins167988insACT). However, this interpretation cannot account for four mismatches located 95 bp (RBg167893), 68 bp (RBg167920), 51 bp (RBg16737), and 1 bp (RBg167988) upstream of the most distal possible 5' breakpoint location. A BLASTn similarity search against the human genome showed that the 95 bp sequence ranging from the first mismatch (RBg167893) to the 5' breakpoint (RBg167988) is identical to only one sequence (in the opposite direction) that is located in intron 17 of the RB1 gene (RBg79218 to RBg 79123) (fig 3). This region in intron 17 is situated in a 1.3 kb long L1PA5 element (RBg78540 to RBg79873), a truncated member of the LINE family. Recurrent loss of only one single exon of the RB1 gene is rare. Intriguingly, there is a clustering of the locations of 3' and 5' deletion breakpoints of the exon 24 deletions identified here.

The 5' breakpoints have a maximum distance of 830 bp and are located in an LIHS LINE. This LIHS has 97% similarity to the full length retrotransposition competent L1 (L1.3, GenBank accession no. L19092) but has acquired several
Figure 3  (A) Genomic organisation of a part of the RB1 gene and location of repetitive elements L1PA5, L1HS, and MER21B. The locations of exon 17 and exons 21 to 25 are indicated by vertical grey bars. Sequence similarity between L1HS and L1PA5 is indicated by light and dark blue colour. The 95 bp region identical to the sequence at the 5' breakpoint of patient M3625 is indicated by a vertical red bar in intron 17. (B) Nucleotide sequence alignment of the deletion mutants (A^D) and the normal allele. Normal sequences from the 5' breakpoints (A^B) within the L1HS are displayed in blue, normal sequences from the 3' breakpoints (C^D) within the MER21B in green, and the sequence from intron 17 in red.
sequence changes that truncate the two open reading frames that are part of a functional L1. ORF1 and ORF2 of an L1 element encode an RNA binding protein and a protein with endonuclease and reverse transcriptase activities, respectively. Both proteins are necessary for the retrotransposition activity of these autonomous transposable elements. As sequence changes have destroyed the coding potential of the L1HS in intron 23 of the RB1 gene, this element is most likely not an active LINE. Therefore, a deletional mechanism involving active retrotransposition of this transpon is unlikely. The 3' breakpoints of the three deletions have a maximum distance of only 235 bp and are located in a MER21B DNA transposon fossil. L1HS and MER21B belong to different classes of interspersed repetitive DNA and show no significant sequence similarity. A search for direct and reverted repeats as well as symmetric elements (using DNASTAR software) identified no significant sequence similarity between 5' and 3' breakpoint regions. Therefore, the most common pattern for the occurrence of deletions including the formation of deletion facilitating secondary structures due to sequence similarities between proximal and distal breakpoint areas, as described by Chuzhanova et al., is not a likely explanation for the deletions observed in this work. In the mammalian nucleus the genome is subdivided into about 60 000 chromatin loops by periodic attachments to a nuclear substructure called the nuclear matrix or nuclear scaffold. These attachment points are provided by S/MARs (scaffold/matrix attachment regions), that is, sequences of a minimum length of 300 bp without an apparent sequence consensus. Although prototype elements are AT rich, a distinct distribution of AT rich tracts rather than overall base composition determines their activity. This pattern mediates a propensity for local strand unpairing which is, at least in part, utilised in the living cell. Owing to this base-unpairing nature is, at transition structures between B-type and ss-DNA which are frequently characterised by increased accessibility for DNase. The matrix binding sequence itself is largely protected from recombination events, probably because they are associated with proteins of ss-DNA recognition potential. Our present study suggests that also in case of the RB1 gene S/MARs flanking regions are the preferred sites triggering deletion events.

During recent years, several software tools have been developed for the prediction of S/MARs. An early example, the “MAR-finder” (now known as MARviz, http://www.futuresoft.org/MarFinder/) is based on the statistical occurrence of S/MAR motifs typical for replication origins, TG rich motifs, curved/kinked DNA, topoisomerase II consensuses, and AT rich sequences. The “stress induced duplex destabilisation” (SIDD) approach identifies regions of DNA unwinding associated with nuclear matrix binding using a statistical mechanical procedure. There are many cases where the prediction of these two approaches is in close agreement. Figure 4 shows a superimposition of both analysis schemes. Here the MAR-finder positions the two breakpoint cluster region (BCR) genes in locations of minimum MAR potential. In the case of the SIDD profile, the centre (approximate map positions 1100–3500) is marked by a regular succession of destabilised sites which are analysed at a standard superhelical density of \( \sigma = -0.055 \). This architecture is typical of strong S/MAR elements as it allows the mass binding mode for prototypical matrix proteins such as the lamins and scaffold attachment factor A (see Bode et al. 1)). On both sides of this base-unpairing region there are segments of DNA that are stabilised at about 8–10 kcal/mol and, based on the combined evidence of both analysis schemes, do not have a S/MAR character. It is clearly seen that these S/MAR associated regions are the primary locations of the deletion endpoints thus confirming the concept that emerged from previous studies.

In conclusion, we have identified two deletion breakpoint clusters in the RB1 gene that are involved in recurrent gross deletions in patients with retinoblastoma. The putative mechanism underlying these mutations highlights the importance of S/MARs for deletion formation in humans.
ELECTRONIC-DATABASE INFORMATION

The URL of the Human Gene Mutation Database is www.hgmd.org and that of the MAR-finder is http://www.futuresoft.org/MarFinder/.

AUTHORS’ AFFILIATIONS

P Albrecht, K Buiting, D R Lohmann, Institut für Humangenetik, Universitätsklinikum Essen, Essen, Germany
J Bode, GFB, German Research Center for Biotechnology/Epigenetic Regulation, D-38124 Braunschweig, Germany
A K Prashanth, UC Davis Genome Center Davis, CA 95616-8816, USA

This study was supported by the Deutsche Forschungsgemeinschaft (Klinische Forschergruppe Ophthalmologische Onkologie und Genetik, DFG KFO 109/1-1, TP II-1).

Conflict of interest: none declared.

Ethics approval: Diagnostic mutation analysis was performed conforming to national guidelines. Scientific analysis of the data was approved by the ethics committee of the Medizinische Fakultät der Universität Essen (file # 01-100-1700).

Correspondence to: Dietmar Lohmann, Institut für Humangenetik, Universitätsklinikum Essen, Hufelandstrasse 55, IG1, 45122 Essen, Germany, dr.lohmann@uni-essen.de

Revised version received 30 June 2004
Accepted for publication 2 July 2004

REFERENCES


Dear Author,

During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof. Many thanks for your assistance.

<table>
<thead>
<tr>
<th>Query Reference</th>
<th>Query</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Here and elsewhere where there was no noun after ‘RB1’ (in italics), the word ‘gene’ has been inserted – is this correct?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>‘LINE-1HS’ only occurs once (in the title) – is this correct?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>The sentence beginning ‘Most of...’ has been changed – please check</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Please cite location of manufacturer – is Applera correct?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Please cite location for Applera if not cited above</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Please cite location of ‘DNASTAR’</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>The word ‘transition’ has been inserted – is this OK?</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>‘Gene Bank’ here and below has been changed to ‘GenBank’ – is this OK?</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>‘BLASTN’ has been changed to ‘BLASTn’ to agree with the capitalisation of the word later in the paragraph – is this correct?</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>The word ‘of’ has been inserted after ‘upstream’ – is this OK?</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>The sentence beginning ‘Normal sequences’ has been changed slightly – please check</td>
<td></td>
</tr>
</tbody>
</table>

The Authors Queries

Journal: Journal of Medical Genetics
Paper: mg21923
Title: Recurrent deletion of a region containing exon 24 of the RB1 gene caused by non-homologous recombination between a LINE-1HS and MER21B element

www.jmedgenet.com
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>The sentence beginning ‘Here the MAR-finder’ has been changed – please check</td>
</tr>
<tr>
<td>13</td>
<td>Should ‘G(x)’ in panel B be explained?</td>
</tr>
<tr>
<td>14</td>
<td>Please check that author affiliations are correct</td>
</tr>
<tr>
<td>15</td>
<td>All author names have been inserted for those references with et al - please check</td>
</tr>
</tbody>
</table>