A bout 10% of the mutations that are known to cause human monogenic disease are deletions of more than 20 bp (gross deletions). To date, 2127 disease causing gross deletions have been listed in the Human Gene Mutation Database (www.hgmd.org). For only few of these mutations, breakpoints have been determined at the sequence level because this information is often of little relevance in clinical diagnostics. Those gross deletions that have been analysed in greater detail most often show DNA sequences with direct or inverted similarity at 5' and 3' breakpoints. The presence of repeated sequences has suggested models that help to explain the formation of these mutations.2

Mutations in the RB1 gene can cause retinoblastoma, a childhood tumour of the eye. We have searched for gross deletions in samples of constitutional and tumour DNA from patients with this tumour. 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.3 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.

METHODS

Samples

DNA from peripheral blood and from fresh frozen retinoblastoma samples was extracted as described previously.4 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 searched for point mutations by Single-strand conformation polymorphism, heteroduplex analysis, or denaturing high performance liquid chromatography.5–7 DNA from tumour samples had also been analysed for hypermethylation of the CpG island associated with the promoter of the RB1 gene and for loss of constitutional heterozygosity (LOH) at STR loci RBi2 and RB1.20.

Quantitative multiplex PCR

Quantitative multiplex PCR was performed using the multiplex PCR assay described by Richter et al8 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). Peak integrals were determined using GeneScan and Genotyper software (Applera).

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.

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 Bremner et al.9 To obtain smaller fragments spanning the deleted region, two long range PCRs were performed with primers GCT GGT TCA ATA TAC CA AAT CAA TA ATG (RBg166248se), TCA TCT GAT CCT TCA ATA AAG CGT AGT (RBg173837as), and AAA AAT ATA AGC TCT CAG GGG AAT GGA GAT (RBg171908as) (fig 2). 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 TCA GAA TGT ATT TAT GCT CAT CTC TGC (RBg170286se) (fig 3A). 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

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
used RBg166248se, RBg171908as, AGG ACA CAA ACA AAT GGA AG (RBg166945se), 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, Madison WI, USA).

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 in three patients 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 patient M4716, tumour DNA from patient M14400, and tumour DNA from patient M3625) showed a deletion of exon 24 in quantitative multiplex PCR. All three samples had been screened for point mutations. The tumour samples from patient M14400 and from patient M3625, which showed no LOH at intragenic STR loci, were found to be heterozygous for a g.162237C→T (R787X) and a g.76999A→G (IVS15-2A→G) transition, respectively. To confirm and further characterise the gross deletions identified in the three samples, we performed long range PCR. Agarose gel electrophoresis of products spanning intron 23 to intron 26 and intron 23 to intron 24 showed smaller than expected DNA bands in all three samples (fig 1). Because of preferential amplification of the smaller products from mutant alleles, there was poor amplification of the products specific for the normal allele. We designed a three-primer PCR to obtain a more balanced product ratio in a single PCR assay. The result of the three-primer PCR indicated that, in constitutional DNA, all three patients are heterozygous for the mutant allele (fig 2). In patient M3625 this finding was unexpected because he had isolated unilateral retinoblastoma and less than 15% of these patients are carriers of RB1 gene mutations. All three deletions occurred de novo because analysis of peripheral blood DNA from parents showed only normal RB1 alleles. On the level of the transcript the most probable consequence of all three gross deletions is loss of the 31 bp that correspond to exon 24, which results in a shift of the reading frame with premature termination eight triplets downstream. Splice donor site mutations in intron 24 that result in skipping of exon 24 lead to identical mutant transcripts. Mutations that involve this site have been reported in patients with bilateral and familial retinoblastoma. Therefore, the gross deletions involving exon 24 identified here most likely cause predisposition to retinoblastoma in these patients. We further analysed the products of long range PCR by direct sequencing. Sequencing results were compared to the published sequence of the RB1 gene (GenBank accession number L11910) and to the results of sequence analysis of a sample without deletion. Patient M14400 showed a deletion of 3605 bp (fig 3). The 5’ breakpoint of the mutation is located in intron 23 at nucleotide 167157 (RBg167157) in a region of 6017 bp with high sequence similarity to an L1HS, 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 (del167988^170975) 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 only to one sequence (in the opposite direction) that is located in intron 17 of the RB1 gene (RBg79218 to RBg79123) (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 a single exon of the RB1 gene is rare. Intriguingly, there is a clustering of the locations of 3’ and 5’ breakpoint deletions of the exon 24 deletions identified here. The 5’ breakpoints have a maximum distance of 830 bp and are located in an L1HS LINE. This L1HS has 97% similarity to the full length retrotransposition competent L1 (L1.3, GenBank accession no. L19092) but has acquired...
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.
While these facts have been elaborated some time ago, more thorough recent analyses have shown that sites of genomic fragility mainly arise at the edges of a S/MAR, that 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 example, the “MAR-finder” (now known as MARwiz, 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 (0.1–0.3). 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 σ = −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. 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, GBF, 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).
REFERENCES


