An overview of *RB1* transcript alterations detected during retinoblastoma genetic screening.

Elizabeth A. Price[1], Mandeep S. Sagoo[2,3], M. Ashwin Reddy[2,4], Zerrin Onadim[1,4].

Author affiliations:

[1]Retinoblastoma Genetic Screening Unit, Barts Health NHS Trust, London E1 2ES, U.K.

[2]Retinoblastoma Service, Royal London Hospital, Barts Health NHS Trust, London E1 1FR, U.K.

[3]NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital, Institute of Ophthalmology, University College London, London EC1V 2PD, UK

[4]Faculty of Medicine, Queen Mary University of London, London E1 4NS, UK

Corresponding Author:

Zerrin Onadim

Retinoblastoma Genetic Screening Unit

Barts Health NHS Trust

NHS ESEL

3rd Floor, Pathology & Pharmacy Building

80 Newark Street

London E1 2ES

Tel: 020 3246 0265

E-mail: z.onadim@nhs.net

An overview of *RB1* transcript alterations detected during retinoblastoma genetic screening.

Abstract

Identification of pathogenic *RB1* variants aids in the clinical management of families with retinoblastoma. We routinely screen DNA for *RB1* variants, but transcript analysis can also be used for variant screening, and to help decide variant pathogenicity. DNA was screened by conformation analysis followed by Sanger sequencing. Large deletion/insertions were detected by polymorphism analysis, MLPA and quantitative-PCR. Methylation specific PCR was used to detect hypermethylation. RNA screening was performed when a DNA pathogenic variant was missing, or to determine effects on splicing.

Two hundred and thirteen small coding variants were predicted to affect splicing in 207 patients. Splice donor (sd) variants were nearly twice as frequent as splice acceptor (sa) with the most affected positions being sd+1 and sa-1. Some missense and nonsense codons altered splicing, while some splice consensus variants did not. Large deletion/insertions can disrupt splicing, but RNA analysis showed that some of these are more complex than indicated by DNA testing. RNA screening found pathogenic variants in 53.8% of samples where DNA analysis did not. *RB1* splicing is altered by changes at consensus splice sites, some missense and nonsense codons, deep intronic changes and large deletion/insertions. Common alternatively spliced transcripts may complicate analysis. An effective molecular screening strategy would include RNA analysis to help determine pathogenicity.

Key Words Retinoblastoma, *RB1*, splicing, pathogenic, genetic, transcript

Introduction

Retinoblastoma (Rb, MIM:#180200) is a childhood cancer (European incidence ~1 in 14,000 live births)(1) that develops from retinal cells at an early age (birth to around five). Predisposition is primarily caused by pathogenic variants in the *RB1* tumor suppressor gene (MIM:614041), with inherited variants leading to high risks of eye tumors and later second primaries. It is autosomal, dominant with variable penetrance depending on the type of variant. The penetrance of splicing alterations depends upon the exon involved, whether the products remain in-frame, and the variants' parental origin (2, 3, 4). Expression can be affected by parental origin due to imprinting in intron 2 (5, 6). Around half of cases are due to heritable germline variants, and around 10% of sporadic cases involve mosaic *RB1* variants in the fetus or a parent. If Rb is detected promptly there is a high cure rate, so molecular genetic screening is performed as part of family management. Identification of pathogenic variants can inform treatment, clinical screening of potential carriers and family planning. Effective molecular screening should detect over 95% of variants.

High penetrance variants generally cause bilateral, multifocal tumors due to loss of functional protein. Low penetrance variants may produce lower protein levels, or protein that retains some function. This can lead to unaffected carriers or less severe disease (unilateral, unifocal Rb)(2). Around 60% of variants are substitutions giving rise to nonsense or missense codons, or splice site changes, and around 25% are small insertions/deletions (7). Around 70% of Rb tumors display loss of heterozygosity (LOH) and about 12% display promoter hypermethylation which is also seen in a few blood samples with chromosomal rearrangements (8). Up to10% of bloods have large alterations of chromosome 13 that may be detected by cytogenetic or array analysis

(9). Chromosomal rearrangements, chromothripsis, LINE insertions, or deep intronic variants occur in a small fraction of cases (10, 11, 12, 13). Another route to Rb is high level, focal amplification of the *MYCN* proto-oncogene in the absence of *RB1* variants as seen in less than 2% of sporadic, unilateral cases (14).

RB1 is a relatively large gene (180kb) with 26 introns where the average human transcript contains around 11. In an analysis of the Human Gene Mutation Database (HGMD) around 9.5% of variants in hereditary disease alleles were classified as single base pair substitutions in splicing relevant regions (intronic and exonic) (15). *RB1* was reported to have a large excess of splicing variants with 46% of hereditary disease alleles listed in HGMD (point variants) mapping to canonical splice sites compared to ~13% on average (16). We previously found that out of 428 pathogenic *RB1* variants identified in Rb patient blood and tumor samples (not including LOH), 19.2% mapped to splice sites, while ~29% of 'small' changes affected splice sites (7). This analysis did not consider changes at sites other than the consensus splice acceptor (sa) and donor (sd) motifs, such as nonsense and missense codons which could also affect splicing (17). Cygan *et al* (16) reported that 27% of *RB1* coding variants tested in their combined *in vitro*/*in silico* assay could affect splicing. Identifying DNA changes may not give the full implications of possible transcript alterations. *In silico* analysis and RNA studies help to give a better picture of possible consequences, especially for Variants of Uncertain Significance (VUS) which often complicate patient counselling and clinical management (18, 19, 20, 21). We therefore looked at a set of *RB1* variants to assess how they could affect splicing. For some cases which failed to yield pathogenic variants after DNA screening, RNA screening was performed to detect deep

intronic variants, which can cause the inclusion/skipping of exons by using alternative splice sites, and to detect some large rearrangements.

Materials and Methods

This audit was approved by the Barts Health Clinical Effectiveness Unit (audit no. 12614). It included blood and Rb tumor samples collected from 1993 – March 2021. Patients were referred to the Retinoblastoma Genetic Screening Unit (RGSU) for *RB1* molecular analysis by clinical geneticists, genetic counsellors, or ophthalmologists. Consent for screening was obtained from parents/guardians. Chart 1 in supplementary material shows the RGSU testing flow.

DNA from peripheral blood, flash frozen Rb tumor, or fixed tumorwas initiall screened. Genomic DNA was extracted by a variety of methods including the phenol/chloroform method and manual kits (Quick-DNA Plus Kits from Zymo) (7). DNA was screened by conformation analysis covering the *RB1* promoter, and each exon with associated splice sites (50 bp upstream and 30 bp downstream). This was followed by Sanger sequencing of candidate exons. Dosage analysis was performed using Multiplex Ligation-dependent Probe Amplification (SALSA MLPA *RB1* probe mix P047, MRC-Holland) and Quantitative Fluorescent PCR (in-house QF-PCR) with LOH testing by polymorphism analysis. Methylation-specific PCR was used to detect promoter hypermethylation (7, 8). Ten tumor samples where routine screening failed to identify two pathogenic variants were also analyzed by whole genome sequencing (WGS) on a research basis (10). DNA variants were identified by comparison to *RB1* reference sequences (Genbank L11910.1, NCBI RefSeq NG_009009.1, LRG_517) and normal control samples. HGVS nomenclature is used except for large rearrangements and complex changes where description is

acceptable. Nomenclature was checked using an online tool (VariantValidator.org). Variants were classified according to their putative effects as described by Hülsenbeck *et al* (22).

RNA screening was performed by RT-PCR (Reverse Transcriptase Polymerase Chain Reaction) and Sanger sequencing of the transcript (NCBI RefSeq NM_000321.3, c.-67 to c.2842). Total RNA was extracted using the QIAamp RNA Blood Mini Kit (Qiagen) for blood samples. Tri-Reagent (Trizol, Invitrogen) was used for fresh tumors and PAXgene (Qiagen) stabilized blood samples. *RB1* cDNA (complementary DNA) was generated and amplified using the One-Step RT-PCR Kit (Qiagen) in four overlapping fragments covering c.-67–727, c.524-1400, c.1114-2149 and c.1798-2842. These were cycle sequenced using the BigDye v.1.1 kit (Applied Biosystems) and run on an ABI3730 with a 50cm capillary carrying POP7. Variants were identified by comparison to the *RB1* reference sequence (NM_000321.3; LRG_517t1) and normal control samples. Testing VUS involved sequencing restricted regions of interest (three or more exons). Putative VUS consequences were analyzed using Alamut Visual Software (SOPHiA Genetics).

Some bloods were pretreated to inhibit Nonsense Mediated Decay (NMD) of transcripts carrying premature termination codons. 300ul blood (collected into LiHep) was added to 5ml RPMI1640 media (Life Tech) containing 20% Fetal Bovine Serum (FBS; Sigma Aldrich), 0.5mg/ml Penicillin Streptomycin Solution (Thermo Fisher) and 1% v/v Phytohemagglutinin M (Thermo Fisher). It was incubated at 37^{o}C for 2 nights prior to the addition of Puromycin (Sigma Aldrich) to a final concentration of 200ug/ml. Control tubes received no Puromycin. Samples were then incubated at 37^{o}C for 5 hours before RNA extraction.

Results

DNA variants

After initial DNA screening there was *RB1* variant data available for 462 bloods and 396 tumors (336 fresh; 60 Formalin Fixed Paraffin Embedded FFPE). One blood had 2 pathogenic variants, while 3 tumors had 3 pathogenic changes. Seven tumors with *MYCN* amplification but no *RB1* pathogenic variants (1.77% $MYCN^{A}$ $RB^{+/+}$) were excluded from further analysis. We removed other tumor specific variants (LOH and missing second pathogenic variants) for a better comparison of blood and tumor variations. Table 1 lists the subsets of *RB1* variants found.

In bloods, 27% (99/366) of pathogenic small variants (transition, transversion, small del/ins) in coding regions and associated splice sites were expected to affect splicing, compared to 18.5% (70/378) in tumors which have more alternative alterations such as promoter hypermethylation and complex rearrangements. All large del/ins (whole exons or more), complex rearrangements and pathogenic deep intronic changes are expected to alter splicing.

207 patients carried 213 small DNA coding variants (126 blood, 87 tumors) which could potentially alter splicing (including pathogenic changes and VUS). These included variants which were not situated at splice consensus motifs. They occurred at 118 different *RB1* positions, with 28 variants occurring multiple times. Splice donor (sd) site variants were nearly twice as common as splice acceptor (sa) - Fig.1 shows their positions around splicing motifs. The most frequently affected consensus positions were the intronic bases, sd+1 and sa-1. Sd+1 comprised 53.5% (68/127) of donor consensus variants (between sd-3 to sd+8). Sa-1 position comprised

48.9% (22/45) of acceptor consensus variants (between sa-8 to sa+3). The consensus sequence ranges were taken as those given by Wai *et al* (21).

Out of the 213 small, coding alterations the most frequently affected exons were 12 (16%), exons 15 and 16 which have a short dividing intron and are often both affected by the same change (10.8%), exon 24 (8%), and exon 6 (6.6%). The most frequent pathogenic change was c.1215+1G>A at the exon 12 sd site (also had G>C and G>T changes). Changes here comprised 13.6% (29/213) of small variants and were nearly three times more common than the next variant (exon 6 sd c.607+1 at 5.2%).

RT-PCR of small variants

RT-PCR analysis was performed for 53 different small variants (Tables 2 and 3: 36 bloods and 20 tumors). These included 39 substitutions (4 nonsense; 12 missense); 8 deletions; one insertion; one duplication; 2 del/ins; 2 complex del+sub. RNA analysis was performed for variants at consensus splice sites, missense changes in patients with no other pathogenic change (additional missense variants that were not expected to be pathogenic are listed in supplementary material Table 1), deep intronic changes, and VUS to confirm predicted splicing outcomes. Analysis was performed for fresh frozen tumors (FT) and blood samples where RNA could be extracted from residual stored sample, or bloods where PAXgene RNA and/or LiHep samples could be obtained upon request.

27/53 variants were potential sd changes with 18 involving the core consensus motif (sd-3 to sd+8). Of the 9 outside the core motif, 4 were deep intronic (≥sd+18) and altered splicing. All 3

missense changes ≤ 4bp from the intron altered splicing. No nonsense codons altered a sd site. There were 25 potential sa changes (Tables 2 and 3) with 10 in the core consensus motif (sa-8 to sd+3). Of the 15 outside the core motif, 7 were deep intronic (≥sa-10) and 5 altered splicing. 2/3 missense codons altered sa splicing (sa+68, sa+91). Of 4 nonsense changes, only that within a core consensus (exon 7 sa+3) was shown to alter splicing. All core sd changes led to skipping of the adjacent, upstream exon. Changes at sa sites were more variable. They caused skipping of the adjacent 3' exon, the use of cryptic sites within exons leading to the loss of coding sequence, or the creation of cryptic splice sites with the inclusion of intronic sequences.

We also performed transcript analysis for 37 small variants/polymorphisms that were expected to be non-pathogenic after literature searches and family studies (28 intronic; 9 exonic). These confirmed predictions, showing no detectable changes in transcripts (supplementary material Table 1). It should be noted that if a variant causes low levels of altered splicing, or is subject to NMD, then Sanger sequencing of cDNA products may fail to detect altered transcripts (limit of detection is ~15% heterozygous variant depending on sequence context).

Large insertion/deletions

Of 81 bloods (Table 1) with large del/ins, most had whole gene deletions (46.9%), or deletions from outside *RB1* to/from common breakpoint regions within introns 2 or 17 (21%). The remainder lost smaller regions (of one or more whole exons). Three (3.7%) had duplications/insertions (exon 3-23dup, 8-23dup, and a 180bp insertion into exon 8 that showed as a deletion in dosage analysis). Of 77 tumors, 28.6% had whole *RB1* deletions, while 11.7% had deletions of both copies (*RB1* doubly deleted). 26% had large deletions from outside of *RB1*

involving the common breakpoints of intron 2 or intron 17. The remainder carried deletions from outside *RB1* to other introns (9.1%) or smaller regions of loss (15.6%). 9.1% tumors contained insertions/duplications (two 7-17dup, one 12-17dup, single exon gains of 12, 13 or 24, and one ~3kb insertion in intron 6).

Table 4 shows samples where RT-PCR was performed across 26 large deletions/amplifications. Not all changes detected by dosage analysis were confirmed at the transcript level. For instance, a deletion of exon 8 proved to be due to a 180bp insertion into that exon causing primer/probe dropout. Some changes were only detected by in-house QF-PCR, and not MLPA, due to the positions of kit probes. One case of possible chromothripsis showed alternating regions of heterozygous and LOH results across chromosome 13, with no other changes seen in DNA or transcript analyses. A second chromothripsis candidate had a deletion covering exons 25-26, alongside regions of LOH/no LOH. 4/6 amplifications proved to be tandem duplications. The others were exon 7-17 amplifications, most likely due to complex rearrangements in tumor samples as previously reported (10). These exon 7-17 amplifications, in two separate patients, had different transcripts and were apparently caused by different rearrangements.

RNA screening for pathogenic variants

We sequenced the full *RB1* cDNA for 13 samples (6 FT; 7 blood) where pathogenic variants were missing after complete DNA testing. In 7 (53.8%; 2 blood and 5 FT) we found pathogenic variants at the transcript level (Table 5). Most changes (71.4%) were deep intronic and/or complex rearrangements. All expected changes in these 7 samples were identified after RNA screening.

Alternative RB1 transcripts

We often observe alternative splicing of exons 2, 8, 9 and 21 which can complicate interpretation of variants. We analyzed 49 samples (29 EDTA, 14 PAX stabilized blood, and 6 FT), with no previously identified variants in those exons, by RT-PCR and sizing on 1.75% agarose to determine frequency of alternative transcripts. Exon 2 is the region most often skipped at moderate (~25%) to high levels (up to 50%), especially in EDTA bloods. Overall, nearly 98% of samples showed this alternative splicing at some level. It is seen at lower levels in fresh tumor or PAX stabilized blood samples, which also displayed the lowest levels of exon 21 skipping (65% of all samples). 98% of samples showed a low level of exon 8 skipping, and all types except PAX showed exon 9 skipping (50% of samples). See Tables 2, 3 and 4 in supplementary material.

Discussion

RNA analysis can increase variant detection rates, resolve VUS, and help to explain variant penetrance. For 13 cases with no pathogenic variants after DNA screening, *RB1* cDNA sequencing was performed. In 7 of these cases we identified extra variants missed by DNA testing with 5 being deep intronic and/or complex rearrangements (Table 5). Two exonic changes had not altered previous HRM melt traces, probably due to their position in the screened products (near ends) or low-level presence. Transcript changes can be identified in blood samples so that obtaining RNA from bilateral, familial, or early onset sporadic cases is worthwhile; one low level mosaic deletion of exons 25-26 was detected in blood cDNA but not DNA (Table 5). Transcript analysis was used to investigate splicing alterations caused by some

VUS, missense and nonsense codons. Eight of ten non-core intronic VUS were found to alter splicing and could be re-classified as pathogenic/likely pathogenic, while only three of nine deep exonic missense VUS codons did so. One nonsense codon at a core site led to exon skipping which may reduce penetrance. We assessed 37 variants (28 intronic; 9 exonic) that were likely benign after literature and family studies, supporting predictions by finding no transcript changes (Table S1).

Analysis of 462 bloods and 396 tumors showed 213 small variants at 118 different *RB1* positions. Sd variants were nearly twice as common as sa. Most frequently affected were exons 12, exons 15 /16, exon 24, and exon 6 where sd+1 changes can vary in penetrance according to parental origin (3). This agrees with Aggarwala *et al* (23) where an enrichment of exon 6 and 12 sd mutations was demonstrated compared to the predictions of a modeling algorithm. They also observed less *RB1* intronic changes outside of essential splice sites than predicted. Cygan *et al* (16) reported that in a dual *in vitro/in-vivo* assay for splicing disruptions, ~27% exonic *RB1* substitutions affected splicing, which was the highest fraction from diseases studied in HGMD. Also, they found an excess (46%) of hereditary pathogenic point variants mapped to *RB1* canonical splice sites. We did not find levels quite that high in DNA screening; in blood samples (hereditary changes) 31.8% of transitions/transversions were predicted to affect splicing (Table 1). Overall, 27% (99/366) of all pathogenic small blood variants (transitions/transversions, small del/ins) in coding regions and associated splice sites were predicted to affect splicing. This is around double the average of ~13% of all hereditary disease alleles (single nucleotide variants) classified as splicing variants by Cygan *et al* (looked at all such variants in 2314 intron-containing genes in HGMD) (16). In this study 27.8% missense codons altered splicing (5/18;

Tables 2, 3 and supplementary Table 1). Three were in consensus sites (sd-2 or sd-4) so were expected to be pathogenic.

The splicing analysis tool previously reported to be most accurate for sensitivity and specificity was SpliceAI, but such software is not reliable (21, 24). For instance, a late onset (42 months) unilateral patient had a germline exon 23 missense c.2393G>A (sa+68). SpliceAI predicted a non-significant sa change. However, RNA analysis showed three abnormal transcripts, mainly involving exon 23 skipping. Only one nonsense codon altered splicing. A homozygous exon 7 variant c.610G>T (sa+3) was seen in tumor DNA and RNA. This produced two transcripts with exon 7 skipping. One transcript produced a frameshift to stop, while another very faint product had just exon 7 skipping which remained in-frame. Such alternative splicing could reduce the penetrance of nonsense codons. SpliceAI correctly predicted that this variant would lead to sa site loss. However, in exon 25 a nonsense c.2536C>T (sa+16) did not display a transcript alteration although this had been suggested by SpliceAI. Another nonsense in exon 14 (c.1333C>T sa+1) did not alter splicing despite being in the core motif (SpliceAI predicted a non-significant effect).

Deletions of whole exons produce abnormal transcripts due to the loss of large protein sections. Those over 1Mb can be low penetrance if the breakpoints are outside *RB1* and include the *MED4* gene (~ 0.3 Mb 5'of *RB1*) (9, 22, 25, 26). 86.8% of bloods (germline cases) with large deletions encompassed *MED4*. Of those, 57.5% were bilateral showing that these changes can still be high penetrance. Hülsenbeck *et al* reported that 70.4% of *MED4* deleted patients were bilateral (22). Single exon deletions seen in DNA can be false positives due to probe/primer dropout in testing. One exon 8 deletion was detected in DNA by dosage analysis and exon skipping was confirmed

in the transcript. However, sequencing DNA showed an insertion into exon 8 of ~180bp from intron 23 alongside part of a LINE element. It was previously shown that exon 24 deletions can be caused by the insertion of a LINE element with a 5' breakpoint in intron 23 (12). One tumor was normal with DNA analysis, but RNA screening identified a 291 bp insertion between exons 6 and 7. Subsequent long-range PCR and sequencing of DNA showed an ~3 kb insert derived from chromosome 8. In the matched blood there was an additional, 38bp shorter version of the insertion, with both versions predicted to lead to downstream stop codons. As dosage analysis does not efficiently detect mosaic deletions, cDNA sequencing can help to confirm loss of exons (unless deletions originate outside of *RB1*).

DNA level amplification of exons 7-17 in two tumors gave different transcripts. One tumor with LOH (copy neutral) had a possible duplication of exons 7-17 with exons 9, 10, 11 deleted in the second copy. The other tumor had two different transcripts (exons 7-17 del and 13 -17 del) which may be due to a complex re-arrangement. The latter also had an exon 14 nonsense variant (r.1333c>t; sa+1) in the second allele causing faint, in-frame exon 14 skipping. This may have been an illegitimate transcript as it was also seen in some normal controls. However, it is also possible that a putative correction pathway, Nonsense Associated Altered Splicing (NAS) upregulated the alternatively spliced transcript, while NMD acted to downregulate it (27). If this variant caused no significant splicing effect, despite involving first base of exon 14, it confirms that VUS at canonical splice sites cannot always be assumed to alter splicing (20).

Complex genomic rearrangements cause discrepant results where DNA deletion/insertions cannot be verified by RNA analysis. WGS can identify complex rearrangements, and

chromosome shattering events, but does not always determine whether both tumor *RB1* alleles are affected. In one case QF-PCR of tumor DNA showed heterozygous deletions of both exon 1 and exon 4, whereas MLPA showed an exon 4 deletion. No other variants were identified in DNA or RNA. WGS eventually found 4 large rearrangements transecting *RB1,* but it remains unclear whether both alleles are involved (10). In 8/10 tumors analyzed by WGS due to missing pathogenic variants, there were complex rearrangements, or a translocation, that would not be detectable in our DNA/RNA screening. However, for 6 of these it could not be determined whether both alleles were involved. Single strand, long range sequencing may resolve such rearrangements.

Alternatively spliced *RB1* products (skipping of exon 2 and/or 8) in normal human breast and prostate tissue were reported (28), and exon 8 skipping in lymphocytes with normal *RB1* was seen after inhibition of NMD (29). We often see these variations with exon 2 skipping being present in most samples analyzed, sometimes reaching almost heterozygous levels. Skipping exon 2 or 8 could be interpreted as pathogenic as this would produce downstream terminations. We have also seen low level skipping of exons 9 and 21 which would produce in-frame, truncated proteins. The presence of alternative transcripts complicates analysis, and results have to be interpreted with caution as it is unknown why this occurs in normal samples. WGS studies could possibly identify modifying factors such as intronic variants that might be involved. Fresh frozen tumors displayed less alternative splicing than bloods so tumor samples might be better for RNA screening if available. PAX stabilized and EDTA blood samples displayed similar levels of alternative splicing although the numbers of samples tested were low. Analysis may be improved by using controls extracted from the same tissue type and collection medium. We also

saw low level skipping of exon 14 in some normal controls. This complicated analysis of a nonsense codon (c.1363C>T; sa+31) in two tumors which had low level exon 14 skipping, possibly due to illegitimate transcription.

Splice site variants can cause skipping of extra adjacent exons (27), especially around exon 15 where exons 14, 15 and 16 are separated by short introns. We saw skipping of exon 15, exon 16, and 15 plus 16 in some samples. Another short intron is 22 and this contained a variant (c.2326-8T>A; sa-8) that led to multiple transcripts in a tumor with LOH. It created a splice acceptor site with three transcripts involving exon 23 skipping, plus inclusions from intron 22. Where Zhang *et al* (29) reported skipping of both exons 10 and 11 in a bilateral patient's blood with c.1049+3A>G (sd+3), we saw only exon 10 skipping in a homozygous tumor, showing how results can vary with sample and patient. Results may also vary with patient age as we have observed that some mosaic variants in blood DNA appeared to be lower in patients resampled after several years.

NMD inhibition is not always required to detect transcript changes. NMD downregulates most mRNAs by 3-10 fold so that inhibition may not cause a significant increase in altered transcripts (30). We detect variants in RNA extracted from fresh EDTA samples without NMD inhibition, especially if they are under a day old and have not been frozen. NMD depends upon the location of the terminating codon and the tissue studied, and can also target normal, or alternatively spliced, transcripts (31). This could explain the variable levels of alternative splicing seen in normal bloods. Screening tumor RNA is difficult if the sample is of poor quality due to treatment prior to enucleation, has heterogeneity, or contains normal cells. Tumors cannot be treated to

inhibit NMD but have high rates of *RB1* transcription from mutant alleles (32), and RNA can be extracted from flash frozen tumor that has been stored for several years. Currently, tumor DNA/RNA is tested if the eye is removed during treatment and is not too heavily treated. However, tumor cell free DNA (cfDNA) in aqueous humor is a promising alternative (33, 34). Future cell-free RNA studies may also be feasible (35).

RNA studies can be complicated by a differentially methylated CpG island in intron 2 (5, 36) which leads to maternally derived variants being expressed more strongly than paternal. An exon 6 splice donor change (c.607+1G>T) which causes exon skipping shows higher penetrance when paternally inherited (3). It therefore helps to obtain parental samples to determine inheritance patterns and predict skewing of expression with altered penetrance. Low level mosaicism in bloods (or Rb allele imbalance in multifocal or heterogeneous tumors) can make altered splicing difficult to detect, although high depth next generation sequencing may help. However, it is not clear that expression levels of different transcripts correlate to protein levels and activity.

VUS can remain unresolved after RNA testing. This also applies to predicting a variants' penetrance as in instances where a missense codon causes exon skipping (increasing penetrance), or a transcript is still obtained from an allele with a large deletion. In one tumor with a deletion from *RB1* exon 26 to outside the gene there was still a transcript from the deleted allele as shown by the low-level presence of the second tumor variant (exon 23 stop). When a 3' primer was placed within the deletion, the level of the exon 23 stop codon increased, showing that there was a transcript from the deleted area. A cryptic 3' poly A/UTR may be used to produce an exon 26-27del transcript, with a corresponding truncated protein that avoids NMD. Two families had a 2

bp deletion in exon 11 (c.1064_1065del, sa+15) that should produce a terminating frameshift with high penetrance. However, one family displayed a low penetrance pattern. The other family had a unilateral mother (variant on paternal allele), who had a bilateral child: paternal/maternal inheritance does not explain the altered penetrance. RNA analysis did not detect any transcript alterations that could explain the pattern seen. In such cases WGS might be able to identify possible modifying factors.

Where a tumor is positive for *MYCN* amplification, it should still be fully screened for *RB1* variants (DNA and RNA) as a case with $MYCN^{A}$ and no *RB1* pathogenic variant would be considered non-heritable (37). A variety of VUS (synonymous, non-coding and 3'UTR) in multiple genes, may be involved in processes that contribute to Rb progression (38). Our testing does not assess changes in *RB1* RNA abundance, processing, or stability so that there could be missed contributions to pathogenicity. This also applies to alterations in Rb protein structure, phosphorylation levels and function (39). Despite difficulties in obtaining samples and interpreting results, RNA analysis remains valuable for detecting pathogenic variants missed by DNA screening, refining VUS classifications, and investigating penetrance. We now routinely extract RNA from the blood samples of high risk early onset, bilateral and familial cases where transcript analysis could aid variant detection or classification. RNA can be obtained from fresh frozen tumors for analysis, but matched blood samples should be obtained to determine the germline status of any variants (Chart 1 in supplementary material).

Acknowledgements

The authors gratefully acknowledge the contributions of the late Mr. John Hungerford (Consultant Ophthalmologist) and Dr. Judith Kingston (Consultant Pediatric Oncologist) for the referral, provision of clinical information and material for analysis. We also thank the Barts Health Rb clinical nurse specialists and pathologists for their help in collecting samples, and RGSU genetic technologists for sample processing and testing.

Author contributors

EAP and ZO performed *RB1* variant analysis, interpreted the data and drafted the article. MSS and MAR referred cases, provided clinical information and material for analysis, and revised/approved the article. All authors critically read, revised, and approved the final document.

Disclosure statement

The authors report no conflict of interest. The authors alone are responsible for the content and writing of this article.

Funding

The authors reported there is no funding associated with the work featured in this article.

Ethics approval

Barts Health Clinical Effectiveness Unit (audit no. 12614).

ORCID

Elizabeth Price. http://orcid.org/0000-0002-2544-202X

Mandeep S. Sagoo. http://orcid.org/0000-0003-1530-3824

Zerrin Onadim. http://orcid.org/0000-0002-8594-9586

References

1 Stacey AW, Bowman R, Foster A, Kivelä TT, Munier FL, Cassoux N, Fabian ID; Global Retinoblastoma Study Group. Incidence of Retinoblastoma Has Increased: Results from 40 European Countries. Ophthalmology. 2021;128(9):1369-1371. DOI: 10.1016/j.ophtha.2021.01.024.

2 Onadim Z, Hogg A, Baird PN, Cowell JK. Oncogenic point mutations in exon 20 of the *RB1* gene in families showing incomplete penetrance and mild expression of the retinoblastoma phenotype. Proc Natl Acad Sci U S A. 1992; 89:6177-6181.

3 Klutz M, Brockmann D, Lohmann DR. A parent-of-origin effect in two families with retinoblastoma is associated with a distinct splice mutation in the *RB1* gene. Am J Hum Genet. 2002;71(1):174-179. DOI: 10.1086/341284.

4 Sánchez-Sánchez F, Ramírez-Castillejo C, Weekes DB, Beneyto M, Prieto F, Nájera C, Mittnacht S. Attenuation of disease phenotype through alternative translation initiation in low-penetrance retinoblastoma. Hum Mutat. 2007; 28:159-67.

5 Kanber D, Berulava T, Ammerpohl O, Mitter D, Richter J, Siebert R, Horsthemke B, Lohmann D, Buiting K. The human retinoblastoma gene is imprinted. PLoS Genet 2009; 5:e1000790. DOI: 10.1371/journal.pgen.1000790.

6 Eloy P, Dehainault C, Sefta M, Aerts I, Doz F, Cassoux N, Lumbroso le Rouic L, Stoppa-Lyonnet D, Radvanyi F, Millot GA, et al. A Parent-of-Origin Effect Impacts the Phenotype in Low Penetrance Retinoblastoma Families Segregating the c.1981C>T/p.Arg661Trp Mutation of *RB1*. PLOS Genet. 2016;12(2): e1005888. DOI:10.1371/journal.pgen.1005888.

7 Price EA, Price K, Kolkiewicz K, Hack S, Reddy MA, Hungerford JL, Kingston JE, Onadim Z. Spectrum of *RB1* mutations identified in 403 retinoblastoma patients. J Med Genet. 2014;51:208-214.

8 Price EA, Kolkiewicz K, Patel R, Hashim S, Karaa E, Scheimberg I, Sagoo MS, Reddy MA, Onadim Z. Detection and reporting of *RB1* promoter hypermethylation in diagnostic screening. Ophthalmic Genet. 2018;39(4):526-531.

9 Mitter D, Ullmann R, Muradyan A, Klein-Hitpass L, Kanber D, Ounap K, Kaulisch M, Lohmann D. Genotype-phenotype correlations in patients with retinoblastoma and interstitial 13q deletions Eur J Hum Genet. 2011;19(9):947-958. DOI: 10.1038/ejhg.2011.58.

10 Davies HR, Broad KD, Onadim Z, Price EA, Zou X, Sheriff I, Karaa E, Scheimberg I, Reddy MA, Sagoo MS, et al. Whole-genome sequencing of retinoblastoma reveals the diversity of rearrangements disrupting *RB1* and uncovers a treatment-related mutational signature. Cancers. 2021;13(4):754-773. doi: 10.3390/cancers13040754.

11 McEvoy J, Nagahawatte P, Finkelstein D, Richards-Yutz J, Valentine M, Ma J, Mullighan C, Song G, Chen X, Wilson M, Brennan R, et al. RB1 gene inactivation by chromothripsis in human retinoblastoma. Oncotarget. 2014;5:438-50. DOI:10.18632/oncotarget.1686.

12 Albrecht P, Bode J, K Buiting, Prashanth AK, Lohmann DR. Recurrent deletion of a region containing exon 24 of the RB1 gene caused by non-homologous recombination between a LINE-1HS and MER21B element. J Med Genet. 2004;41(12):e122. DOI: 10.1136/jmg.2004.021923.

13 Dehainault C, Michaux D, Pagès-Berhouet S, Caux-Moncoutier V, Doz F, Desjardins L, Couturier J, Parent P, Stoppa-Lyonnet D, Gauthier-Villars M, Houdayer C. A deep intronic mutation in the RB1 gene leads to intronic sequence exonisation. Eur J Hum Genet. 2007;15:473-7. DOI: 10.1038/sj.ejhg.5201787.

14 Rushlow DE, Mol BM, Kennett JY, Yee S, Pajovic S, Thériault BL, Prigoda-Lee NL, Spencer C, Dimaras H, Corson TW, et al. Characterisation of retinoblastomas without *RB1* mutations: genomic, gene expression and clinical studies. Lancet Oncol. 2013;14:327-334.

15 Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. Hum Genet. 2014;133(1):1-9. DOI: 10.1007/s00439-013-1358-4.

16 Cygan KJ, Soemedi R, Rhine CL, Profeta A, Murphy EL, Murray MF, Fairbrother WG. Defective splicing of the *RB1* transcript is the dominant cause of retinoblastomas. Hum Genet. 2017;136(9):1303-1312.

17 Chabot B, Shkreta L. Defective control of pre-messenger RNA splicing in human disease J Cell Biol. 2016;212(1):13-27. DOI: 10.1083/jcb.201510032.

18 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-424. DOI: 10.1038/gim.2015.30.

19 Ellard S, Baple EL, Callaway A, Berry I, Forrester N, Turnbull C, Owens M, Eccles DM, Stephen Abbs S, Scott R, et al. ACGS Best practice guidelines for variant classification in rare disease 2020. The Association for Clinical Genomic Science (ACGS); 2020 April 04 [accessed 2023 Jan 18]. https://www.acgs.uk.com/media/11631/uk-practice-guidelines-for-variant-classification-v4-01-2020.pdf

20 Garrett A, Durkie M, Callaway A, Burghel GJ, Robinson R, Drummond J, Torr B, Cubuk C, Berry IR, Wallace AJ et al. Combining evidence for and against pathogenicity for variants in cancer susceptibility genes: CanVIG-UK consensus recommendations. J Med Genet 2021; 58297-304 DOI:10.1136/jmedgenet-2020-107248

21 Wai HA, Lord J, Lyon M, Gunning A, Kelly H, Cibin P, Seaby EG, Spiers-Fitzgerald K, Lye J, Ellard S, et al. Splicing and disease working group.

22 Hülsenbeck I, Frank M, Biewald E, Kanber D, Lohmann DR, Ketteler P. Introduction of a Variant Classification System for Analysis of Genotype-Phenotype Relationships in Heritable Retinoblastoma. Cancers. 2021;13(7):1605-1626. DOI: 10.3390/cancers13071605.

23 Aggarwala v, Ganguly A, Voight BF. *De novo* mutational profile in *RB1* clarified using a mutation rate modeling algorithm. BMC Genomics 2017; 18:155 doi::10.1186/s12864-017-3522-z

24 Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, et al. Predicting Splicing from Primary Sequence with Deep Learning. Cell. 2019;176(3):535-548.e24. DOI: 10.1016/j.cell.2018.12.015.

25 Albrecht P, Ansperger-Rescher B, Schüler A, Zeschnigk M, Gallie B, Lohmann DR. Spectrum of gross deletions and insertions in the *RB1* gene in patients with retinoblastoma and association with phenotypic expression. Hum Mutat. 2005;26(5):437-445. DOI: 10.1002/humu.20234.

26 Dehainault C, Garancher A, Castéra L, Cassoux N, Aerts I, Doz F, Desjardins L, Lumbroso L, RoMontes de Oca R, Almouzni G, et al. The survival gene MED4 explains low penetrance retinoblastoma in patients with large RB1 deletion. Hum Mol Genet. 2014;23(19):5243-5250. DOI: 10.1093/hmg/ddu245.

27 Wang J, Chang Y-F, Hamilton JI, Wilkinson MF. Nonsense-associated altered splicing: a frame-dependent response distinct from nonsense-mediated decay. Mol Cell. 2002;10:951-957

28 Latil A, Pesche S, Valeri A, Cussenot O, Fournier G, Lidereau R. Common Δ2 and Δ8 *RB1* variants in human breast and prostate tissues. Int J Cancer. 1999;80(4):632-633.

29 Zhang K, Nowak I, Rushlow D, Gallie BL, Lohmann DR. Patterns of missplicing caused by RB1 gene mutations in patients with retinoblastoma and association with phenotypic expression. Hum Mut. 2008;29: 475-484.

30 Kurosaki T, Li W, Hoque M, Popp MW, Ermolenko DN, Tian B, Maquat LE. A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev. 2014 Sep 1;28(17):1900-16. doi: 10.1101/gad.245506.114. PMID: 25184677; PMCID: PMC4197951.

31 Tan K, Stupack DG, Wilkinson MF. Nonsense-mediated RNA decay: an emerging modulator of malignancy. Nat Rev Cancer. 2022;22(8):437-451. DOI: 10.1038/s41568-022-00481-2.

32 Dunn JM, Phillips RA, Zhu X, Becker A, Gallie BL. Mutations in the RB1 gene and their effects on transcription. Mol Cell Biol. 1989; 9(11): 4596–4604. DOI: 10.1128/mcb.9.11.4596.

33 Berry JL, Xu L, Polski A, Jubran R, Kuhn P, Kim JW, Hicks J. Aqueous Humor is Superior to Blood as a Liquid Biopsy for Retinoblastoma. Ophthalmology. 2020;127(4):552-554. DOI: 10.1016/j.ophtha.2019.10.026.

34 Gerrish A, Jenkinson H, Cole T. The Impact of Cell-Free DNA Analysis on the Management of Retinoblastoma. Cancers;13(7):1570-1585. DOI: 10.3390/cancers13071570.

35 Hassan S, Shehzad A, Khan SA, Miran W, Khan S, Lee Y-S. Diagnostic and Therapeutic Potential of Circulating-Free DNA and Cell-Free RNA in Cancer Management. Biomedicines. 2022;10(8):2047-2068. DOI: 10.3390/biomedicines10082047.

36 Kanber D, Buiting K., Roos C, Gromoll J, Kaya S, Horsthemke B, Lohmann D. The Origin of the *RB1* imprint. PLoS ONE 2013;8:e81502. DOI: 10.1371/journal.pone.0081502.

37 Price EA, Patel R, Scheimberg I, Kotiloglu Karaa E, Sagoo MS, Reddy MA, Onadim Z. MYCN amplification levels in primary retinoblastoma tumors analyzed by Multiple Ligation-dependent Probe Amplification. Ophthalmic Genet. 2021;42:604-611. DOI: 10.1080/13816810.2021.1923038.

38 Stachelek K, Harutyunyan N, Lee S, Beck A, Kim J, Xu L, Berry JL, Nagiel A, Reynolds CP, Murphree AL, Lee TC, Aparicio JG, Cobrinik D. Non-synonymous, synonymous, and non-coding nucleotide variants contribute to recurrently altered biological processes during retinoblastoma progression. Genes Chromosomes Cancer. 2022 Dec 22. doi: 10.1002/gcc.23120. Online ahead of print.

39 Ewens KG, Bhatti TR, Moran KA, Richards-Yutz J, Shields CL, Eagle RC, Ganguly A. Phosphorylation of pRb: mechanism for RB pathway inactivation in MYCN-amplified retinoblastoma. Cancer Med. 2017;6(3):619-630.

Table 1. *RB1* pathogenic variants in 462 blood and 389 tumor samples from 851 unrelated Rb patients. *Splice* – missense/nonsense or small del/ins variants that obviously affect splice consensus sites. Splice Consensus – transitions and transversions only. Complex - chromothripsis, translocations. Epigenetic – hypermethylation of *RB1* promoter. One blood sample had 2 variants. Tumor variants – excluded LOH, and $MYCN^{A}$ $RB^{+/+}$ cases.

	Blood n=462 1 Pathogenic variant		**Tumor n=389 2 Pathogenic variants**	
Variant Type	**No. Found**	**%**	**No. Found**	**%**
Promoter	6	1.30%	3	0.57%
Small del/ins	91	19.65%	74	14.09%
Small del/ins; *Splice*	14	3.02%	9	1.71%
Nonsense	152	32.83%	215	41.00%
Nonsense; *Splice*	3	0.65%	0	0.00%
Missense	24	5.18%	19	3.61%
Missense; *Splice*	8	1.73%	3	0.57%
Splice Consensus	74	15.98%	58	11.04%
Large del/ins	81	17.06%	77	14.66%
Deep Intronic	8	1.73%	4	0.76%
Complex	0	0.00%	15	2.85%
Epigenetic	2 (X:13 translocations)	0.43%	48	9.14%
Total	**463**	**100%**	**525**	**100%**

Table 2: RT-PCR testing that gave positive results for 38 different small variants (single bp or small del/ins) in 42 patients (27 blood and 15 tumor samples). Variants are listed according to start position of change. Core splice regions are in bold. Pathogenicity classification is according to ACMG/ACGS guidelines (18-20). Classification for genotype-phenotype analysis is according to Hulsenbeck *et al* 2021 (22). Variants are defined according to NM_000321.3; genomic ref LRG_517; NG_009009.1.

Position	DNA (c.)	RNA Change	Putative Consequence	Variant Class
SA-19	c.1961-19_1964del abolishes sa intron19/exon 20	r.1961_2106del exon 20 skip downstream	p.(V654Dfs*18)	REC-I
SA-15	c.2212-15_2221-14delins11 abolishes sa intron 21	r.2212_2325del Exon 22 skip downstream	p.(T738_R775del	REC-III
SA-13	c.1333-23_1333-13del Intron 13	r.1333_1389del In-frame del of 19 amino acids from downstream exon 14.	p.(R445_S463del)	REC-III
SA-12	c.608-12T>G creates cryptic sa intron 6. Seen twice.	r.607_608ins608-11_608-1 Insert 11bp from intron 6	p.(G203Vfs*15)	REC-I
SA-10	c.540-10T>A creates cryptic sa intron 5	r.539_540ins540-8_540-1 Insert 8bp from intron 5	p.(I181Lfs*8)	REC-I (leaky)?
SA-8 core intron	**c.2326-8T>A 3 variant transcripts. Abolishes usual sa and creates cryptic sa. Intron 22**	**r.2326_2489del Exon 23 skip downstream r.2326_2489delins2325+1_2325+11Exon 23 skip with 11bp insert from intron 22 r.2325_2326ins2326-6_2326-1 Insert 6bp from intron 22 at start of exon 23**	**p.(P776Nfs*7) p.(P776_R830delinsVGQK) p.(R775_P776insL*)**	**REC-III/-I (leaky)**
SA-3 core intron	**c.1961-3C>G creates cryptic sa intron 19**	**r.1960_1961ins1961-2_1961-1 Insert from intron 19**	**p.(V654Efs*5)**	**REC-I**
SA-2 core intron	**c.265-2A>G abolishes sa intron 2 c.1696-2A>G abolishes sa intron 17**	**r.265_380del Exon 3 skip downstream r.1698_1733del Use of cryptic sa in downstream exon 18 with in-frame deletion**	**p.(G89Cfs*3) p.(D566_K577del)**	**REC-I REC-III**
SA-1 core intron	**c. [1050-3del; 1050-1G>A] abolishes sa intron10 c.1696-1G>A abolishes sa intron 17 c.2490-1G>A abolishes sa intron 23**	**r.1050_1062del First 13 bases of exon 11 deleted as internal AG used as sa r.1696del First base of exon 17 (G) used as sa r.2490_2520del Exon 24 skip downstream.**	**p.(S350Rfs*13) p.(D566Ifs*45) p.(I831Lfs*8)**	**All REC-I**
SA+3 core exon	**c.610G>T nonsense abolishes sa exon 7. Seen twice.**	**r.609_718del Multiple products involving exon 7 skip**	**p.(E204Nfs*7) / p.(G203_K240del) / p.(E204*)**	**REC-I**

SA+68	c.2393G>A missense p.(R798Q) exon 23	r.2393G>A and multiple products, primarily exon skips. r.2326_2489delins11 Exon 23 skip plus 11bp inclusion from intron 22 r.2326_2489del Exon 23 skip. SpliceAI and other in silico analysis predicted sa loss and gain of cryptic site.	p.(R798Q); p.(P776_R830delinsVGQK); p.(P776Nfs*7)	RECIII / I (leaky)
SA+91	c.1589A>C missense p.(K530T) exon 17	r.1589a>c and r.1499_1591del Deletion of 93bp from exon 17	p.(R500_V531 delinsM)	REC-III
SD-6	c.1492_1493ins66 exon 16	r.1422_1498del Predicted p.(Y498Lfs*13) but exon 16 skip seen	p.(S474Rfs*8)	REC-I
SD-4	c.1811A>G missense p.(D604G) creates sd site in exon 18	r.1811_1814del Deletion of last 4bp of exon 18.	p.(D604Gfs*6)	REC-I_L3
SD-1 core exon	**c.1421G>A missense p.(S474N) Exon 15 (looked at twice)** **c.2663G>A missense p.(S888N) Exon 25**	**r.1390_1421del Skip of upstream exon 15 and r.1421g>a** **r.2521_2663del Skip of upstream exon 25**	**p.(E464Qfs*5) / p.(S474N)** **p.(T841*)**	**REC-I / III** **REC-IV**
SD+1 core intron	**c.380+1G>T abolishes sd intron 3** **c.607+1G>T abolishes sd intron 6. (looked at twice). Seen 7 times.** **c.1215+1G>A abolishes sd intron 12. (looked at three times). Seen 27 times.** **c.1814+1G>C seen twice.** **c.1814+1G>A abolishes sd intron 18** **c.1960+1_1960+6delins22 abolishes sd intron 19** **c.2489+1G>C intron 23**	**r.265_380del Skip of upstream exon 3.** **r.540_607del Skip of upstream exon 6** **r.1128_1215del Skip of upstream exon 12** **r.1696_1814del Skip of upstream exon 18** **r.1815_1960del Skip of upstream exon 19** **r.2326_2489del Skip of upstream exon 23**	**p.(G89Cfs*3)** **p.(I181Gfs*8)** **p.(V378Afs*3)** **p.(D566Vfs*47)** **p.(M605Ifs*14)** **p.(P776Nfs*7)**	**REC-I** **REC-I** **REC-I_L2** **REC-I** **REC-I_L3** **REC-I**
SD+2 core intron	**c.861+2T>G abolishes sd intron 8** **c.1049+2T>C abolishes sd intron 10** **c.1421+2T>C abolishes sd intron 15**	**r.719_861del Skip of upstream exon 8** **r.940_1049del Skip of upstream exon 10** **r.1390_1421del Skip of upstream exon 15**	**p.(K240Sfs*22)** **p.(V314Ffs*2)** **p.(E464Qfs*5)**	**All REC-I**
SD+3 core intron	**c.1049+3A>G abolishes sd intron 10** **c.1215+3A>T abolishes sd intron 12. Seen twice.** **c.1421+3A>T abolishes sd intron 15**	**r.940_1049del Skip upstream exon 10** **r.1128_1215del Skip upstream exon 12** **r.1390_1421del Skip upstream exon 15** **r.2490_2520del Skip upstream exon 24**	**p.(V314Ffs*2)** **p.(V378Afs*3)** **p.(E464Qfs*5)** **p.(I831Lfs*8)**	**REC-I** **REC-I_L2** **REC-I** **REC-I**

	c.2520+3_2520+6del abolishes sd intron 24. Seen 6 times.			
SD+4 core intron	**c.2520+4A>C abolishes sd intron 24**	**r.2490_2520del Skip upstream exon 24**	**p.(I831Lfs*8)**	**REC-I**
SD+18	c.1421+18_1421+32del c.1421+18_1421+33del intron 15	r. [1390_1421del, 1422_1498del] for both. Exon15 skip (major) plus exon 16 skip (minor). *Note: intron 15 only 80 bp long.*	p.(E464Qfs*5) p.(S474Rfs*8)	REC-I REC-I
SD+40; +43	c.[1421+40A>T; 1421+43_1421+48del] same sample. Abolishes sd intron 15	r.1390_1421del Skip upstream exon 15	p.(E464Qfs*5)	REC-I (leaky)
SD+63	c.1215+63T>G cryptic sd intron 12	r.1215_1216ins62 Insert from intron 12 between exons 12 and 13	p.(N406Vfs*5)	REC-I (leaky)

Table 3: RT-PCR testing that gave negative results for 15 small *RB1* variants (single base pair or small del/ins) in 14 different people (9 blood and 5 tumor samples). Variants are listed according to start position of change. Core splice regions are in bold. Variants are defined according to NM_000321.3; genomic ref LRG_517; NG_009009.1.

Position	**DNA (c.)**	**RNA Change**	**Putative Consequence**	**Variant Class**
Initiation codon core exon	**c.3G>C exon 1 initiation start site/missense**	**r.3g>c no splicing effect. May initiate translation at Met113.**	**p. (M1I)**	**REC-III**
SA-49	c.608-52_608-49del intron 6	None.	-	VUS in *cis* with a missense VUS c.1543C>T
SA-15	c.540-15C>G intron 5	None.	-	VUS
SA+1 core exon	**c.1333C>T nonsense at sa exon 14**	**r.1333c>t no splicing effect despite change in first base of exon 14**	**p.(R445*)**	**REC-I**
SA+3 core exon	**c.1963dup exon 20**	**r.1963dup no splicing effect**	**p.(Y655Lfs*13)**	**REC-I**
SA+15	c.1064_1065del exon 11	r.1064_1065del no splicing effect. An exon skip would have explained low penetrance	p.(R355Nfs*6)	REC-I_L2 (leaky)?
SA+16	c.2536C>T nonsense exon 25	r.2536c>t no splicing effect.	p.(Q846*)	REC-IV
SA+31	c.1363C>T nonsense exon 14	r.[1363c>t, 1333_1389del] faint exon 14 skip observed but could be illegitimate transcript as also seen in normal control	p. (R455*)	REC-I

SA+33	c.1454C>T missense exon16	r.1454c>t no splicing effect.	p.(S485F)	VUS
SA+45	c.1543C>T missense exon 17	r.1543c>t no splicing effect	p.(P515S)	VUS in *cis* with c.608-52_608-49del
SA+47	c.1861C>A missense exon 19	r.1861c>a no splicing effect	p.(R621S)	VUS
SD-68	c.1892A>G missense exon 19	r.1892a>g no splicing effect	p.(Q631R)	VUS
SD-27	c.474G>C missense exon 4	r.474g>c no splicing effect	p.(L158F)	VUS
SD-23	c.2084T>G novel exon 20 missense in tumor, not seen in blood	r.2084t>g no splicing effect.	p.(M695R)	VUS rs727504122
SD-2 core exon	**c.1959del from a poly-A tract. Exon 19**	**r.1959del no splicing effect**	**p.(V654Cfs*4)**	**REC-I**

Table 4. RT-PCR analysis of 26 large deletions/amplifications (29 samples from 28 patients: 20 tumor and 9 blood). Variants are defined according to NM_000321.3; genomic ref LRG_517; NG_009009.1. UTR – Untranslated Region.

DNA Result	**RNA (r.)**	**Putative Consequence**	**Variant Class**	**Comments**
Del exon 2	r.138_264del	p.(L47Efs*22)	REC-I	FFPE fixed tumor.
Del exon 2	tumor RNA poor amp	p.(L47Efs*22)	REC-I	Tumor double del/LOH. Somatic rearrangement resulting in del of exon 2 + LOH found by WGS.
Del exon 3	r.265_380del	p.(G89Cfs*3)	REC-I	1 blood and 1 tumor.
Del exon 3-19	normal	p.(G89Cfs*4)	REC-I	Tumor with LOH. MLPA showed del as hemizygous exons 3-19. QF-PCR showed whole *RB1* del with exons 3-19 as a double del.
Del exon 4	normal	Not determined/ no functional protein?	?	Complex: WGS showed 4 large rearrangements transecting *RB1* in tumor.
Del exon 7 ***plus*** Del intron 17 to beyond 3' UTR	normal	No functional protein	?	Could be part of a large rearrangement in tumor.
Del exon 8	r.719_861del	p.(K240Sfs*22)	REC-I	~180bp insert (from intron 23 and LINE element) into exon 8 causes exon skip. Sporadic bilateral blood.
Del exon 9-11	r.862_1127del	p.(V288Dfs*18)	REC-I	Seen twice - both sporadic bilateral bloods. Analysis of one sample.
Del exon 12	r.1128_1215del	p.(V378Afs*3)	REC-I_L2	Sporadic bilateral blood.
Del exon 13	r.1216_1332del	p.(N406_Q444del)	REC-III	Bilateral blood: mosaic mother affected.

Del exon 13	r.1216_1332del	p.(N406_Q444del)	REC-III	At DNA level c.1220_1221ins ~700bp (5 bp after sa) causes exon 13 skipping. Tumor with LOH.
Del exon 14	r.1333_1389del	p.(R445_S463del)	REC-III	Tumor with whole *RB1* del.
Del exon 20	r.1961_2106del	p.(V654Dfs*18)	REC-I	Exon 20 skip due to 160-180bp del which starts inside the exon. Sporadic bilateral blood.
Del exons 21-23	r.2107_2489del	p.(I703Nfs*7)	REC-I	Tumor with LOH
Del exons 23-24	normal	Not determined/ no functional protein?	?	WGS: 2 large rearrangements transecting *RB1* in tumor.
Del exon 24	r.2490_2520del	p.(I831Lfs*8)	REC-I	2 bloods (1 had matched tumor): sporadic unilateral and sporadic bilateral.
Del exons 24-26	r.2490_2713del	p.(R830Sfs*14)	REC-I	Tumor with LOH.
Del exons 25-26 (mainly)	r.2521_2713del	p.(T841Lfs*12)	REC-IV?	Tumor chromothripsis? Complex rearrangement with multiple *RB1* gains and losses.
Del exons 26-beyond 3' UTR	had a transcript from deleted area	?	REC-IV	3' del includes exons 26 and 27. NMD avoided so truncated protein produced? Tumor also has exon 23 nonsense variant.
Amp ex 3-23	r.265_2489dup	p.(I831Efs*22)	REC-I	Tandem dup. Sporadic unilateral retinoma blood.
Amp exons 7-17	insertion of multiple exons between exons 17 and 18	p.(D566Gfs*11)	REC-I	Ins of exons 7,8,12,13,14,15,16,17. Tumor with LOH (or dup 7-17 with exons 9, 10, 11 deleted in the second copy)
Amp exons 7-17	r. [608_1695del, 1216_1695del] r.[1363c>t, 1333_1389del]	p.[(E204Ffs*5), (N406_S565del)]	REC-I: REC-III	Three transcripts: exons 7-17del; 13-17 del; exon 14 del. May be part of a complex re-arrangement. Tumor also has exon 14 c>t nonsense variant which was observed at the RNA level with exon 14 skip.
Amp exon 12	r.1128_1215dup	p.(N406Dfs*18)	REC-I	Tandem dup of exon 12 in transcript. Tumor also has 13 bp deletion in exon 21.
Amp exons 12-17	RNA poor	p.(S567Cfs*17)	REC-I	Tandem dup of exons 12-17 confirmed by WGS. Tumor also has exon 8 nonsense variant.
Amp exon 24	r.2490_2520dup	p.(T841Nfs*7)	REC-I	Tandem dup of exon 24. Tumor with LOH.
Alternating loh/no loh	Numerous non-specific transcripts	?	?	Tumor chromothripsis? Neither DNA nor RNA screen identified another variant.

Table 5. Variants that were not detected in DNA screening but found at the transcript level (5 tumor and 2 blood). Variants are defined according to NM_000321.3; genomic ref LRG_517; NG_009009.1.

RNA (r.)	**DNA (c.)**	**Putative Consequence**	**Variant Class**	**Comments**
r.1184_1215del32 ins128	c.1184_1215+755delinsTG Covers part of exon and intron.	p.(Q395_N405delins43)	REC-III	In DNA this is a 787bp del and ins of 2bp. Sporadic bilateral blood (PAX stabilized sample for RNA screen).
r.939_940ins42 creates cryptic sa site	c.939+541A>G cryptic sa intron 9. Deep intronic.	p.(V314Mfs*5)	REC-I	Transition in intron 9 leads to an exon 9A (42bp long) in tumor. Present in matched blood (bilateral case).
r.939_940ins52 creates cryptic sa site	c.939+531T>G cryptic sa intron 9. Deep intronic	p.(V314Ffs*12)	REC-I	Transition in intron 9 leads to an exon 9A (52bp long) in bilateral blood (PAX stabilized, also NMD inhibited EDTA for RNA screen).
r.795del exon 8	c.795del Exonic	p.(K265Nfs*3)	REC-I	Homozygous 1 bp somatic del in tumor. HRM missed this.
r.1390_1421del Skip of upstream exon 15	c.1421+2T>A abolishes sd SD+2 core intron	p.(E464Qfs*5)	REC-I	Somatic variant at low level in tumor which had imbalance. HRM missed this.
r.607_608ins291	c.56853-3001_56853-3000ins~3kb Deep intronic rearrangement	p.(G203Dfs*14)	REC-I	~3kb deep intronic insert into intron 6 (sa-3000) from chromosome 8 in tumor. Blood RNA had a 38bp shorter insert (mosaic) than tumor.
r.2521_2713del	Complex. Various *RB1* gains and losses with low level ex 25 and 26.	p.(T841Lfs*12)	REC-IV?	Deletion of ex 25 and 26 as part of complex rearrangement in tumor with LOH. Low level mosaic presence in blood detectable in cDNA not DNA.

Supplementary Table 1. 37 *RB1* variants of uncertain significance (VUS), or likely benign, which showed no change at the RNA level. Variants are defined according to NM_000321.3; genomic ref LRG_517; NG_009009.1.

***RB1* (c.)**	**Region Altered**	**Reference dbSNP cluster ID / Comments**
c.45_53del	Exon 1 p.(A16_A18del)	rs572454921 reported to be non-pathogenic
c.68C>T	Exon 1 (sa+68) p.(P23L) missense	ClinVar 1418753 tumor had two other pathogenic changes - not present in patient's blood
c.137+94C>T	Intron 1 (sd+94)	rs191994935 Another patient tumor lost the 'T' allele and had a homozygous pathogenic variant
c.137+110C>T	Intron 1 (sd+110)	Present in unaffected parent and grandparent - patient mosaic for a pathogenic variant
c.138-20T>C	Intron 1 (sa-20)	rs376775524 also present in unaffected parent - patient mosaic for a pathogenic variant
c.352A>T	Exon 3 (sd-2) p.(T118S) missense	rs1429553692 found in unaffected parent and sibling – not present in affected parent and proband who have pathogenic variants
c.380+12T>C	Intron 3 (sd+12)	rs3092881 patient also heterozygous for a pathogenic change. Tumor of another patient had homozygous pathogenic change in addition to this variant
c.539+43A>G	Intron 5 (sd+43)	rs376032950 unilateral patient
c.540-28T>G	Intron 5 (sa-28)	rs368358330 present in unaffected parent
c.571C>T	Exon 6 (sa+32) p.(L91=)	rs538578527 on the same allele as a pathogenic variant. Present in unaffected parent.
c.607+55G>C	Intron 6 (sd+55)	Present in unaffected parent
c.607+55G>T	Intron 6 (sd+55)	rs553934083 present in unaffected parent
c.862-15C>A	Intron 8 (sa-15)	rs115108608 present alongside a likely pathogenic variant. Present in an unaffected parent
c.990T>G	Exon 10 (sa+51) p.(D330E) missense	Present in an unaffected parent. Not present in proband
c.1127+74C>T	Intron 11 (sa+74)	rs3092888 present with two pathogenic variants in tumor, benign variant
c.1128-72G>T	Intron 11 (sa-72)	rs185587 benign variant
c.1156A>G	Exon 12 (sa+29) p.(M386V) missense	rs564780653 present in an unaffected parent.
c.1215+941A>G	Intron 12 (sa+941)	rs399413 benign variant
c.1216-29A>G	Intron 12 (sa-29)	rs3092886 present with two pathogenic variants in tumor, benign variant
c.1313G>A	Exon 13(sd-20) p.(C438Y) missense	ClinVar 935171 patient's tumor had two other pathogenic variants not detected in blood
c.1389+40G>A	Intron 14 (sd+40)	rs187166242 present in unaffected parent - patient has another germline pathogenic variant
c.1389+104C>T	Intron 14 (sd+104)	rs191504668 patient's rb tumor has two pathogenic variants
c.1390-11A>G	Intron 14 (sa-11)	rs200658795 present in unaffected parent. Present in a family and segregating with a pathogenic variant.
c.1421+9T>C	Intron 15 (sd+9)	rs183417081 present in an unaffected parent
c.1421+27T>A	Intron 15 (sd+27)	rs527912551 lost from a rb tumor. Present in unaffected parents and sibling in other cases

c.1498+73G>A	Intron 16 (sd+73)	rs112189207 benign variant
c.1499-57del	Intron 16 (sa-57)	rs11351399 benign variant
c.1695+16357G>A	Intron 17 (sa+16357)	rs9568036 benign variant
c.1753C>G	Exon 18 (sa+58) p.(H585D) missense	ClinVar 577923 a missense inherited from unaffected parent. In *trans* with a pathogenic variant.
c.1815-113A>G	Intron 18 (sa-113)	rs3092897 in *cis* with a pathogenic variant in tumor
c.1815-104A>G	Intron 18 (sa-104)	rs3092898 benign variant
c.1961-12T>C	Intron 19 (sa-12)	rs201697122 in *trans* with a pathogenic familial variant - present in unaffected parent
c.2212-16T>C	Intron 21 (sa-16)	Present in unaffected parent - patient has another pathogenic variant
c.2212-15dup	Intron 21 (sa-15)	rs201258424 c.2212-16T>A present in unaffected parent
c.2455C>T	Exon 23 (sd-35) p.(L819=)	rs375751988 present in unaffected parent
c.2490-1471C>T	Intron 23 (sa-1471)	rs561092656 unilateral patient
c.*703G>T	3' Untranslated Region	rs139023385 present with two pathogenic variants in tumor - present in unaffected parent and sibling

Supplementary Table 2. Levels of *RB1* alternative transcripts displaying exon 2 skipping in samples with no variant detected in that region.

Exon 2 RT-PCR exons 1-5	**No Skip**	**Low Level**	**Moderate (~25% skip)**	**High Level (~50% skip)**
Fresh Rb Tumor (n=6)	0	6 (100%)	0	0
PAX Blood (n=14)	0	9 (64.3%)	5 (35.7%)	0
Fresh EDTA (n=23)	1 (4.3%)	5 (21.7%)	7 (30.4%)	10 (43.5%)
Frozen EDTA (n=6)	0	2 (33.3%)	0	4 (66.7%)
Total (n=49)	1 (2%)	22 (44.9%)	12 (24.5%)	14 (28.6%)

Supplementary Table 3. Levels of *RB1* alternative transcripts displaying exon 8 skipping in samples with no variant detected in that region.

Exon 8 RT-PCR exons 7-10	**No Skip**	**Low Level**	**Moderate (~25% skip)**	**Exon 9 skip also seen in product**
Fresh Rb Tumor (n=6)	0	6 (100%)	0	3 faint (50%)
PAX Blood (n=14)	0	14 (100%)	0	0
Fresh EDTA (n=23)	0	23 (100%)	0	14 faint; 6 obvious (87%)
Frozen EDTA (n=6)	1 (16.7%)	5 (83.3%)	0	5 faint (83%)
Total (n=49)	1 (2%)	48 (98%)	0	28 (50%)

Supplementary Table 4. Levels of *RB1* alternative transcripts displaying exon 21 skipping in samples with no variant detected in that region.

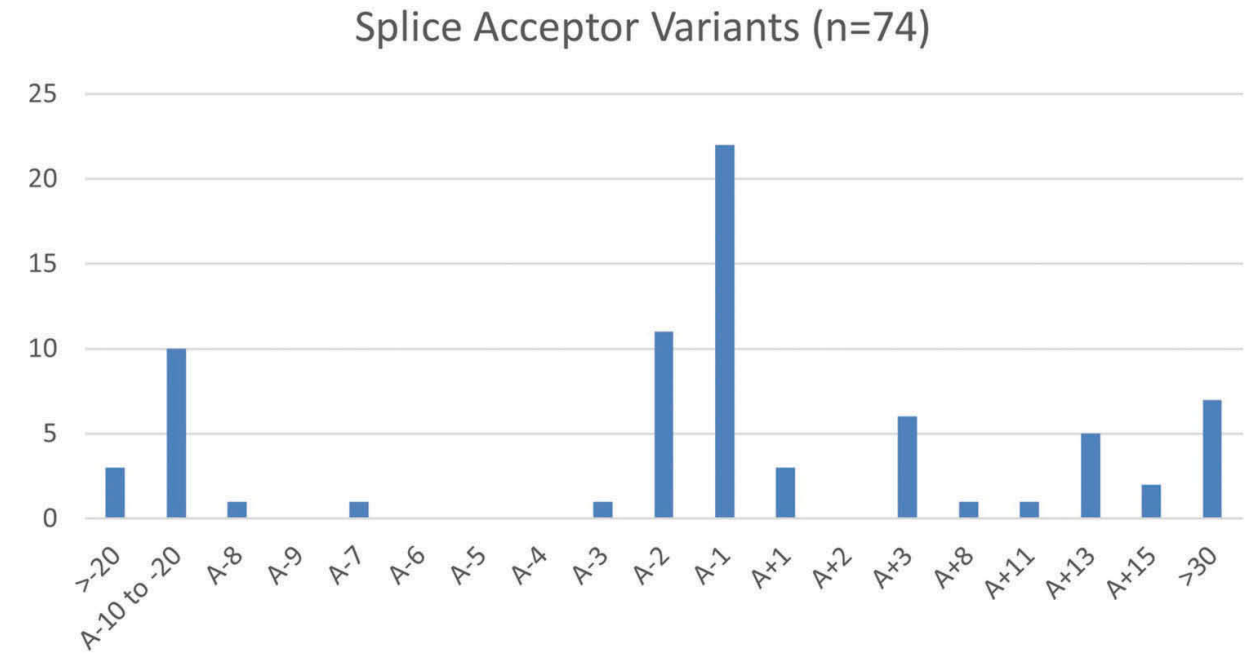
Exon 21 RT-PCR exons 20-23	No Skip	Low Level	Moderate (~25% skip)
Fresh Rb Tumor (n=6)	6 (100%)	0	0
PAX Blood (n=14)	10 (71.4%)	4 (28.6%)	0
Fresh EDTA (n=23)	0	17 (74%)	6 (26%)
Frozen EDTA (n=6)	1 (16.7%)	5 (83.3%)	0
Total (n=49)	17 (34.7%)	26 (53%)	6 (12.2%)

Supplementary Chart Caption (Supplementary Screening Flow Chart 1 submitted as JPG)

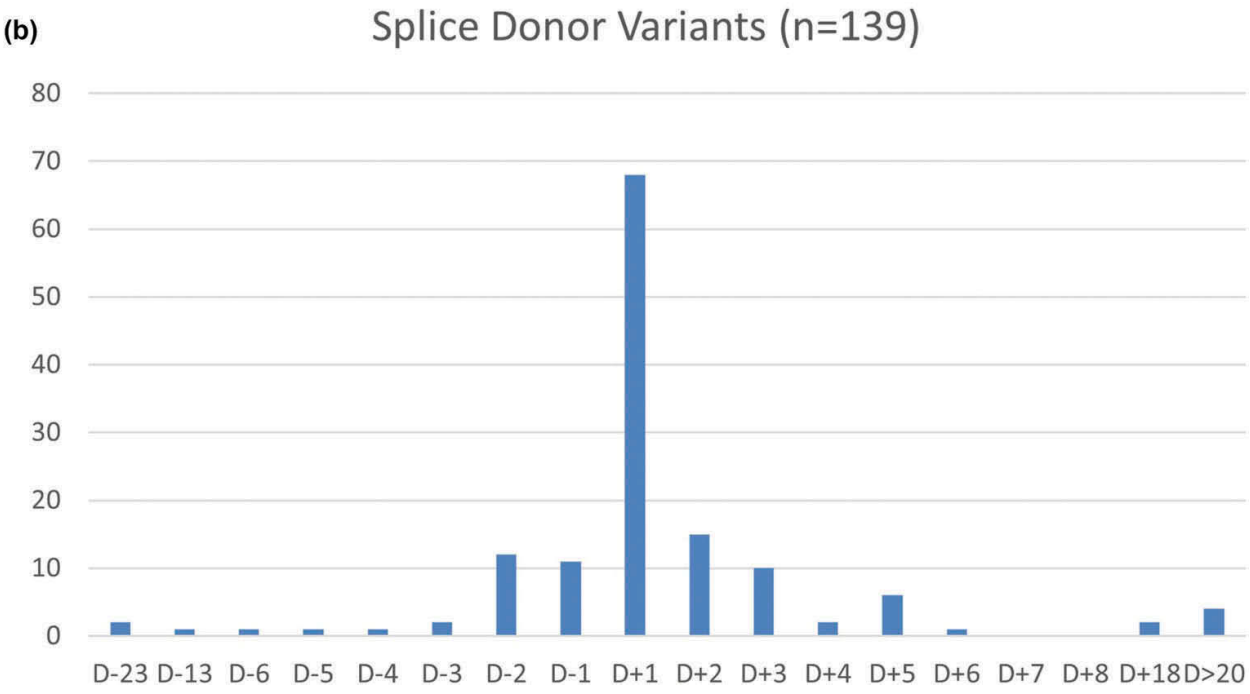
Supplementary Chart 1. Flow chart for RGSU retinoblastoma molecular screening.

Figure 1. The spread of 213 *RB1* small coding splice variants (transitions/transversions and small del/ins). Numbers describe positions in the consensus sequences. Figure 1A – altered acceptor site positions. Figure 1B – altered donor site positions.

(a)


Splice Acceptor Variants (n=74)
25
20
15
10
5
0
>-20
A-10 to -20
A-8
A-9
A-7
A-6
A-5
A-4
A-3
A-2
A-1
A+1
A+2
A+3
A+8
A+11
A+13
A+15
>30

(b)


Splice Donor Variants (n=139)
80
70
60
50
40
30
20
10
0
D-23
D-13
D-6
D-5
D-4
D-3
D-2
D-1
D+1
D+2
D+3
D+4
D+5
D+6
D+7
D+8
D+18
D>20