

Research Article

Journal of Clinical & Experimental **Ophthalmology**

Open Access

Screening *RB1* Gene in Algerian Patients and Predicting the Pathogenicity of Variations by *In Silico* Analysis

Amina Mama Boubkeur^{1*}, Lotfi Louhibi¹, Meriem Abdi¹, Fatima Zohra Moghtit¹, Nacera Tabet Aoul², Rym Abderrahmane¹, Khadija Mahmoudi ³, Meriem Aberkane^{1,4} and Nadhira Saidi-Mehtar¹

¹Department of Applied Molecular Genetics, Laboratory of Molecular and Cellular Genetics, Faculty of Natural and Life Sciences, University of Science and Technology of Oran, Algeria

²Department of Biotechnology, University of Oran, Algeria

³Pediatric Hospital of Canastel, Department of Ophtalmology, Oran, Algeria

⁴Department of Pharmacy, Faculty of Medicine, University of Oran 1, BP 1524 El M'naouer 31000 Oran, Algeria

*Corresponding author: Amina Mama Boubkeur, Department of Applied Molecular Genetics, Faculty of Natural and Life Sciences, Laboratory of Molecular and Cellular Genetics, University of Science and Technology of Oran- Mohamed Boudiaf-Ustomb-BP 1505 El, M'naouer 31000, Oran, Algeria, Tel: +213 770 50 24 01; E-mail: amina.boubekeur@hotmail.fr

Received date: March 03, 2017; Accepted date: May 02, 2017; Published date: May 08, 2017

Copyright: ©2017 Boubkeur AM, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Retinoblastoma is a pediatric retinal tumor initiated by biallelic inactivation of the retinoblastoma gene (*RB1*). Most of the alteration being unique and randomly distributed throughout the entire coding region. This study aimed firstly to identify mutations that may affect the RB1 gene in constitutional level witch contribute to the understanding of the molecular pathogenesis of this disease and for early detection of subject at risk and asymptomatic carriers. The detection of variations in 30 patient's DNAs was performed by high performance liquid chromatography (HPLC) followed by sequencing. Secondly, we developed a protocol *in silico* analysis using different software based on analytic methods (Align GVGD, Mutation tasting, SIFT, PolyPhen, I-Mutant and KD4V) and structural (Swiss-Pdb Viewer). This protocol was used to predict the deleterious effects of mutations on protein pRb. The spectrum of mutation included 2 missense mutations, 1 nonsense mutation, 1 deletion, 1 mutation affecting splice site and 2 polymorphisms. Among these mutations, some were identified in germinal level for children with no family history of the disease. Therefore, *in silico* analysis identified only three causal mutations, the first intronic mutation in the splice donor site caused probably a frame shift, giving rise to a truncated protein. The second missense mutation c.1903 G>C was predicted to affect splicing processes. The third mutation c.1961T>A located in exon 20 may disrupt protein function and structure.

In this study, *in silico* analysis has shown the effect of mutations on the structure and function of the protein. It is interesting to compare these results with others functional studies in order to evaluate the functionality of the mutations.

Keywords: Retinoblastoma; *RB1*; *In silico*; HPLC; Sequencing; Mutation; Prediction

Introduction

Retinoblastoma (Rb) is an intraocular malignant tumor which occurs in children during the first months of life, with an incidence of 1 in 15,000 live births [1]. It requires the inactivation of both alleles of the tumor suppressor gene RB1. Roughly, 40% of retinoblastoma cases are heritable, with a germline mutation in one allele and the other at the somatic level, while 60% of Rb cases are non-heritable (sporadic), with both mutated alleles at the cellular level [2]. Generally, the hereditary form manifests bilaterally and is inherited as an autosomal dominant disorder with very high penetrance (90%-100%), characterized by early onset (within the first year of life) and presence of multiple tumors (multifocal). The sporadic form of the disease typically involves only one eye (unilateral) and occurs as a single tumor (unifocal) with late onset (after the first year of life). The functional inactivation of RB1 gene is also associated with different malignancies, such as osteosarcoma, breast cancer and small cell lung carcinoma [3]. The protein encoded by the RB1 gene plays a crucial role in regulating

the cell cycle, promoting G1/S arrest and growth restriction by inhibiting the E2F transcription factors [4].

The *RB1* gene (MIM: 180200),which is located at chromosome 13q14.1, has a high degree of mutational heterogeneity reported in exons, introns and regulatory regions [5] with over 3366 mutations reported until date (http://rb1-lsdb.d-lohmann.de/home.php? select_db=RB1). The deleterious effects of missense mutations and mutations affecting splice sites can be shown to cause the disease. For this reason, the pathogenicity of *RB1* gene variations was predicted by using *in silico* analysis.

In silico analysis is an expression used to mean performed on a computer or via computer simulation. It draws from the vast amounts of biological information available, and applies sophisticated algorithms or simulations to advance scientific understanding. The results of these simulations can be tested experimentally or serve as a guide for future physical experimentation. Many *in silico* bioinformatics tools for the prediction of the effects of mutations on protein function have been developed and became important for their prioritization for experimental characterization.

In this study, we analyzed 30 patient' DNAs by high performance liquid chromatography (HPLC) followed by sequencing. This study is used to identify the different mutations of *RB1* gene and to specify their nature and their potential role in the inactivation this gene. After that we performed a protocol *in silico* analysis using different software based on analytic methods (Align GVGD, Mutation tasting, SIFT, PolyPhen, I-Mutant and KD4V) and structural (Swiss-Pdb Viewer). This study is used to determine the deleterious effects of the mutations found in the different patients with retinoblastoma.

Material and Methods

Study population

In this study, 30 unrelated retinoblastoma patients were recruited from the pediatric hospital of Canastel, Oran, Algeria; and screened for mutations in the *RB1* gene in constitutional level.

Methods

DNA analysis

The genomic DNA from 30 patients was isolated from peripheral blood leucocyte using the salting out precipitation method [6]. Promoter of *RB1* gene, exons and flanking regions were amplified. The detection of deletion or mutation was performed by high performance liquid chromatography (HPLC) [7]. The patients giving particular profiles by HPLC analysis have been systematically sequenced on an ABI 3130 DNA sequencer (Applied Biosystem, New Jersey, USA). In order to determine the nature and the position of mutation, two software were performed in this sequence analysis:

"Seqscanner"(https://products.appliedbiosystems.com/ab/en/US/ adirect/ab?cmd=catNavigate2&catID=600583&tab=Overview) and "Multalin" (http://multalin.toulouse.inra.fr/multalin/) [8].

With all these studies, we performed *in silico* bioinformatics analysis to evaluate the pathogenecity of the detected alterations.

In silico analysis

Different bioinformatics tools exploring impact of mutation on splicing, structure and/or function of the protein were used:

Splicing prediction software

Influence of mutations on splicing efficiency was predicted using two software:

-Human Splicing Finder Version 2.4.1 software "HSF" (http:// www.umd.be/HSF/): evaluates the potential impact of mutations on donor and acceptor splice site and predict if there are any cryptic sites [9].

-Exonic Splicing Enhancer-finder "ESE Finder" (http:// rulai.cshl.edu/cgi- bin/tools/ESE/): facilitates the analysis of regulatory sequences splicing enhancer which represent binding sites of several proteins rich in serine involved in splicing [10].

Structure/function prediction software

I-Mutant 2.0: I-Mutant 2.0 software (http://folding.biofold.org/imutant/i-mutant2.0.html) is a support vector machine based tool for automatic prediction of protein stability changes upon single point mutations. It estimates the free energy difference (Delta DeltaG; DDG) between the wild and mutant protein. A positive value (DDG>0) implies an increase in protein stability, whereas a negative value (DDG<0) suggests a destabilizing mutation [11].

SIFT: Sorting Intolerant from Tolerant software "SIFT" (http:// sift.jcvi.org) is a sequence homology-based tool that sorts intolerant from tolerant amino acid substitutions and predicts whether an amino acid substitution will affect the protein function. To assess their effects, SIFT assumes that important positions in a protein sequence have been conserved throughout evolution and therefore substitutions at these positions may affect protein function [12]. The underlying principle of this program is that SIFT takes a query sequence and uses multiple alignment information to predict tolerated and deleterious substitutions for every position of the query sequence. SIFT scores are classified as intolerant (0.00-0.05), potentially intolerant (0.051-0.10), borderline (0.101-0.20), or tolerant (0.201-1.00) [13].

PolyPhen-2: Polymorphism Phenotyping v2 "PolyPhen-2" software (http://genetics.bwh.harvard.edu/pph2/) is a prediction tool of possible impact of an amino acid substitution on structure and protein function. This prediction is based on sequence annotations, multiple alignment and structural information from PDB (Protein Data Bank) structure database. Basically, PolyPhen searches for 3D protein structures, multiple alignments of homologous sequences and amino acid contact information in several protein structure databases, calculates position-specific independent counts (PSIC) scores for each of two variants, and then computes the PSIC scores difference of two variants. A PSIC score difference of 1.5 and above is considered to be damaging. PolyPhen-2 classifies substitutions into three levels: probably damaging, possibly damaging or benign.

Align GVGD: Align Grantham Variation Grantham Deviation "Align-GVGD" software (http://agvgd.iarc.fr/agvgd_input.php) combines the biophysical characteristics of amino acids and protein multiple sequence alignments to predict the effect of substitution on protein function [14]. Align-GVGD considers biochemical variation at the relevant position in the multiple sequence alignement (Grantham Variation) and biochemical difference between reference and variant amino acids (Grantham Deviation). Results obtained are determined into seven classes: C0, C15, C25, C35, C45, C55, and C65. The (C45-C65) classes indicate that the variation affect the function while the middle class (C35) and the (C0-C25) classes do not have a functional impact.

Mutation taster: Mutation Taster software (http:// www.mutationtaster.org/) was designed for a rapid assessment of a DNA sequence variants using information from various sources: evolutionary conservation, splice sites changes and loss of protein function. It calculates probabilities for the alteration to be either a *disease mutation* or a harmless *polymorphism* [15]. The probability value is the probability of the prediction, with a value close to 1 indicates a high security of the prediction. The p value used here is **not** the probability of error as used in t-test statistics.

KD4v: Comprehensible system for Knowledge Discovery missense variant software "KD4v" (http://decrypthon.igbmc.fr/kd4v) predicts whether a mutation has a neutral or deleterious effect on a protein [16]. KD4v is based on 3D structure to predict changes in size, charge, polarity, hydrophobicity, accessibility and physico-chemical properties of amino acid resulting from missense mutation.

Swiss PDB viewer software: Missense mutations were performed on the X-ray structure described by PDB, .ID: 4ELJ [17] using Swiss PDB

Page 3 of 9

Viewer Software; Swiss Institute of Bioinformatics, Basel, Switzerland (http://www.expasy.org/spdbv/). The amino acid residues in the native and mutant protein were performed by SWISS PDB viewer to get modeled structures [18].

Results

DNA analysis

Genomic DNA analysis of 30 patients in germline level by HPLC showed different variations of *RB1* gene. Seven mutations were

identified in Table 1. The spectrum of germline alterations given by Seqscanner and Multalin softwares included one nonsense mutations: c.91 G>T (g.2150 G>T) in exon 1; two missense mutations: c.1903 G>C (g.153296G>C) in exon 19 [19] and c.1961T>A (g.156693T>A) in exon 20 [20]; one mutation affecting splice site:c.1215+1 G>A (g. 70330G>A) [21]; a frameshift mutation corresponding to C.630-634 GATC (56875-56879 GATC) in exon 2 [19] and c.1049+58 G>C (g. 64497G>A) in exon 10 [22]. The mutations of exon 1 and 7 were reported only in Algerian population [23] and not described in any databases or in any other population.

Mutations	E/I	Substitution Amino acid	Mutation type	Mutation in Genomic DNA	Mutation in Tumoral DNA
c.1215+1 G>A	Intron 12	1	Donor splice site	-	.+
c.1903 G>C	Exon 19	p.Ala635Pro	Missense	.+	.+
c.1961 T>A	Exon 20	p.Val654Gln	Missense	.+	.+
c.264+75 T>C	Intron 2	1	Polymorphism	.+	.+
c.1049+58 G>C	Intron 10	1	Polymorphism	.+	.+
C.630—634 GATC	Exon 7	p.Asp211-212X	Nonsense	.+	.+
c.91 G>T	Exon 1	p.Glu31X	Nonsense	.+	.+

Table 1: Mutation of RB1 gene identified in the genomic DNA of Algerian patients.

In silico analysis

As known only missense mutation and mutation affecting splice site can be realized by computational analysis, the mutation affecting splice site and the two missense mutations found in our study were analyzed. In addition, one of the two missense mutations was also analyzed for modeled structure of its mutant proteins.

Analysis of mutation affecting splice site

The mutation affecting splice site c.1215+1 G>A was checked for splicing effect using two *in silico* splice prediction programs: Human

Splicing Finder Version 2.4.1 (HSF) and Exonic Splicing Enhancer-finder (ESE Finder).

Analysis by human splicing finder version 2.4.1 (HSF)

The analysis of mutation c.1215 + 1 G>A by comparison between the normal and mutated sequence of RB1 gene by HSF software predicted the abolition of splicing donor site (Table 2). HSF also predicted the absence of cryptic site activation in this region (Table 2).

Sequence Position	cDNA Position	Split site type	Motif	New splite site	Wild Type	Mutant	If cryptic site use exon length variation	Variation (%)
117	C.1213	Donor	AACgtaac	AACataagc	85.86	59.02	0	WT site broken -31.26

Table 2: Predictive table obtained by HSF software.

Analysis by enhancer-finder (ESE Finder)

The analysis of the mutation c.1215 + 1G>A by the ESE Finder software is shown in Figure 1. The comparison between normal and mutated sequence of the RB1 gene by ESE Finder predicted abolition of fixing site of the two SR proteins: SRp40 and SRp55.

Analysis of missense mutations

Two missense mutations were identified in two patients. The missense mutation c.1903 G>C, (p.Ala635Pro) was located in the

spacer region, previously assumed to be non-essential to pRb protein function [24]. The second missense mutation c.1961T>A, (p.Val654Gln) is located in the B box of the pRb pocket. The missense variants were analyzed-using 6 online in silico protein analysis softwares. The results were shown in Table 3.

Page 4 of 9



Figure 1: Distribution of proteins SR in reference sequence of RB1 gene and mutated sequence (c.1215 + 1G>A).

Analysis by I-mutant

The results obtained by the software I Mutant 2.0 is in favor of a destabilizing effect of the two mutations in pRb protein with a negative DDG as described in Table 3.

Variant	I-Mutant 2.0 DDG	SIFT Score	PolyPhe n-2 PSICD	Align GVGD Class	Mutation Taster Probability	KD4V
C.1903 G>C	1.91	Tolerate d 0.3	Benign 0.001	Class C0	Polymorphis me 0.64	Neutral
C.1961 T>A	3.41	Intolerat ed 0.01	Probably Damagi ng 1	Class C35	Disease Causing 1	Deleterio us

 Table 3: Effect of the pathogenicity of RB1 gene missense mutation by in silico analysis.

Analysis by SIFT software

The protein sequences of the 2 missense mutations were submitted independently to SIFT program to check its tolerance index. The mutation c.1903 G>C (p.Ala635Pro) was found tolerant with the tolerant index score of>0,05 while the mutation c.1961T>A (p.Val654Gln) was found deleterious with the tolerant score of <0,05. The results are shown in Table 3. Also according to the SIFT results, the both mutations, showed no change in physicochemical property of amino acid, all them are non-polar.

Analysis by Polyphen software

The two mutations were submitted as input to the PolyPhen software the mutation c.1903 G>C (p.Ala635Pro) was considered to be begnin (Table 3) and the second missense mutation c.1961T>A (p.Val654Gln) were considered to be probably damaging (Table 3).

Analysis by Align GVGD software

The result of the analysis by the Align GVGD software predicted that mutation c.1903 G>C (p.Ala635Pro) corresponds to the class C0 and the mutation c.1961T>A (p.Val654Gln) correspond to the class C35 (Table 3). The class C0 can less likely to interfere with the function of protein while the class C35 can interfere with the function of pRb. Analysis of the two mutations by Align GVGD software has calculated GV values and GD for each mutation. The variation values of Grantham (GV) is 215,31 for mutation c.1903 G>C (p.Ala635Pro) which means that the substitution site is located in not conserved region between the species analyzed. For the mutation c.1961T>A (p.Val654Gln), the GV value is 30,92. This means that the mutation is situated in the conserved region. This software calculates also the standard of Grantham (GD) between wild-type and mutated amino acid. GD value for mutation c.1903 G>C is 0 which means that the physicochemical properties between alanine and proline are conserved. While the predicted value for GD mutation c.1961T>A is 121.34, the replacement of valine by a glutamine is predicted as change.

Analysis by mutation taster software

The two missense variation was submitted independently to mutation taster software to check its potential disease; the results are shown in Table 3. The mutation c.1903 G>C (p.Ala635Pro) is considered as a polymorphism while the mutation c.1961T>A (p.Val654Gln) can be the cause of the disease.

Analysis by KD4v software

The analysis by KD4v software predicts that the two missense mutations were deleterious (Table 3). The mutation c.1903 G>C is responsible for the increase in size and a decrease in the hydrophobicity of amino acid mutant. The second mutation (c. 1961T>A) is responsible for the increase in size, charge, polarity and a decrease in the hydrophobicity of amino acid mutant.

Modeling and analysis of modeling structure

The main criteria in determining the result of a mutation is its location in the 3D structure. The available structure for RB1 gene has a PDB id 4ELJ. The mutation c.1903 G>C (p.Ala635Pro) is located in the spacer. This region connects the A box and the B box of the pRb pocket. This spacer has not been Crystallographed because it assumed to be non-essential to pRb protein function. For this reason, we could not realize the modeled structure of the pRb containing this mutation by SWISS PDB. The mutational position and amino acid variant associated with mutation c.1961T>A (p.Val654Gln) was mapped in the B box of 4ELJ native structure. The mutation is performed by SWISS PDB viewer to get modeled structures. The valine 654 in the native protein interacts with the leucine 657 (Figure 2). In silico structural analysis of the pRb revealed the Gln residue to form a hydrogen-bond network with four backbone carbonyls. Hydrogen bonds to backbone carbonyls of residues Phe650, Tyr655, Leu657 and Ala658 (Figure 3). The Val654Gln predicted to disrupt this intramolecular hydrogen bond network with a possible structural and functional consequence on the pRb protein.





We analyzed the secondary structure of each amino acid residue of the native and mutant protein structures. We found that the exposure of these two amino acids were different (Figure 4). This may be due to their structural differences that make each of them a well specific position.



Figure 3: Structure of the pRb protein. The model was made from the structure PDB 4ELJ. The Valine 654 that is mutated to Glutamine is located in the B box of the pRb pocket, and forms a hydrogen bond network with four backbone carbonyls. Hydrogen bonds to backbone carbonyls of residues Phe650, Tyr655, Leu657 and Ala658 are shown by green dotted line.

Comparison of physicochemical properties between normal and mutated amino acid

A comparison of the physicochemical properties between normal and mutated residues was performed. The variation c.1903 G>C is responsible for the substitution of alanine into a proline at position 635; while the physicochemical properties of each residue are reported in Table 4. The point mutation c.1961T>A is responsible for the substitution of valine into a glutamine at position 635; while the physicochemical properties of this two residues are reported in Table 5.



Figure 4: The comparison between the exposition of amino acid Alanine in the native protein **(A)** and amino acid Glutamine in the mutate protein **(B)**.

Page 6 of 9

Residue	Туре	Taille	Accessibility	Hydrophobicité	Charge
Alanine	Acide	89Da	Accessible	Hydrophobic	Neutre
Proline	Neutre	115Da	Accessible	Hydrophobic	Neutre

Table 4: The physicochemical properties of alanine and proline.



Figure 5: Close-up of the mutations (Ala635Pro and Val654Gln) (seen from a slightly different angle left and right). The protein is gray, the side chains of both the wild type and the mutant residue are presented in red and green respectively.

Residue	Туре	Taille	Accessibility	Hydrophobicity	Charge
Valine	Neutral	89Da	Buried	Very Hydrophobic	Positive

Page 7 of 9

Glutamine	Neutral	115Da	Buried	Neutral	Neutral

Table 5: the physicochemical properties of valine and glutamine.

Regarding the mutation c.1903 G>C (p.Ala635Pro), the wild type and mutant amino acids differ in size. The proline is bigger than alanine. The wild-type residue is not conserved at this position. This sometimes suggests that this variant is not damaging for the protein's structure and function. The mutated residue is located on the surface of a spacer domain with unknown function.

For the other mutation c.1961T>A (p.Val654Gln), the residue glutamine is bigger than the valine. The point mutation will cause loss of hydrophobic interactions in the core of the protein because the wild-type residue is more hydrophobic than the mutant residue.

Comparison of the size and orientation of side chain of the normal and mutated amino acid

The comparison between the alanine's side chain with that of proline clearly shows the size difference between the two chains. Indeed, the

alanine's side chain is shorter than proline's side (Figure 5). In the other hand, the superposition of the two 3D structures of normal pRb protein and that containing the substitution (Val654Gln) clearly demonstrated that the side chain of valine is shorter than that of glutamine (Figure 5). This difference could prevent or create new molecular interactions within the structure of the pRb protein.

The result obtained by the six online *in silico* analysis (I-Mutant 2.0, SIFT, Polyphen2, Align GVGD, Mutation taster and KD4V) predict that the variant c.1903 G>C (p.Ala635Pro) is considered as a polymorphism. For this reason, the missense variant was checked for splicing effect using the software Enhancer-finder (ESE Finder). The program predicted the abolition of the protein binding site SC35 and the apparition of the fixation sit of the protein SF2/ASF (Figure 6). The real deleterious effect of this mutation appears to be disrupting the process of splicing.



Figure 6: Position of proteins SR in reference sequence of RB1 gene and mutated sequence "c.1903 G>C(p.Ala635Pro)".

Discussion

In general, the variations of bases in the genes, most often are responsible for functional disorders in the proteins leading to the appearance of several diseases. These variations may occur at a coding or non-coding regions, also can therefore affect the splicing process and/or structure/function of the protein.

In this study, the mutations of RB1 gene at exon 1 and 7 give a premature arrest of the translation resulting in a functionally altered short protein. This is due to the absence of the important domains

downstream which are required for the activity of the pRb protein. These mutations were identified only in Algerian population; these differences might be related to the different ethnic composition of the Algerian population.

Other changes found in intron 2 and 10 correspond to polymorphisms already described. The distribution of these variations along the RB1 gene shows that there are no preferential regions of mutations. Through this study, we identified 5 cases out of 30 studied children with retinoblastoma, with no family history of the disease. These mutations can be transmitted from a parent with asymptomatic

carrier or due to neo-mutation during gametogenesis. An investigation of parents still required to confirm or refute a genetic predisposition.

In silico programs and 3D molecular modeling was performed to determine the deleterious effects of 2 missense mutations and mutation affecting splice site at the protein level.

The impact of the mutation c.1215+1G> C on splicing was predicted using two *in silico* prediction softwares HSF and ESE Finder.

As known, splicing is the complex process with which is produced mature mRNA from the pre-mRNA. This requires donor and acceptor sites splice [25]. The Splicing is catalysed by the "spliceosome", which consists of many proteins [26]. Among these include the splicing factor binding to family sequences cis regulatory: the amino acid-rich proteins serine and arginine, named SR proteins (rich serine) which consist of 50 to 300 amino acid residues. There are 9 types of SR protein: SF2/ASF, SC35, SRp20, SRp40 and SRp55, SRp75, 9G8, SRp30c and p54 [27]. This protein binds to the sequences activating (SA) which is located in exons and which are called "exonic splicing enhancers". These proteins promote the recognition of splice sites and facilitate recruitment of U1 and U2 snRNP (small nuclear ribonucleoproteins) allowing through other protein elimination of the intron and exon ligation [28].

The analysis of the mutation c.1215+1G>C by HSF predicted the abolition of the splice donor site. ESE finder also predicts that this mutation is responsible for the abolition of binding sites of the two SR proteins: SRp40 and SF2/ASF.

The aberrant splicing associated with mutation c.1215+1G>A is due to the abolition of the splice donor site and the abolition of binding sites of two SR proteins: SRp40, SRp55. The consequence of this mutation is skipping exon 12 by use the splice donor site of exon 11 and the splice acceptor site of exon 13. The mutation is located in A box of the pRb pocket which contains distinct sequences required for the binding of mediator of pRb (E2F) induced cell cycle arrest and binds directly to histone deactylase 1 (HDAC1), wich is required for several chromatin remodeling factors that participate in transcriptional repression of E2F target genes [29]. Skipping exon 12 causes probably a frame shift, giving rise to a truncated protein with a loss of a large encoding fragment of pRb protein and will generate a non-functional protein. This result will not inhibit E2F factor and will cause increased proliferation of retinal cells thus leading to the appearance of retinoblastoma. Although mRNA analysis for this mutation has not been carried out, this approach would be required, in the future, to elucidate the exact defect of this mutation in splicing. More functional and structural studies are still required to confirm the exact role of this mutation on splicing.

Several *in silico* prediction softwares have evaluated the pathogenicity of the two missense mutations (c.1903 G>C; c. 1961T>A). The deleterious effects of missense mutations on stability were evaluated by I-mutant 2.0. While three different predicted the effects of this mutation on the protein function softwares (SIFT, Polyphen, Align GVGD and mutation taster).

For the two missense RB1 gene mutations identified, the c.1903 G>C was considered to be polymorphism and the c.1961T>A to be mutation causing the disease (probably pathogenic). The analysis of effect of the mutation c.1903 G>C, (p.Ala635Pro) using in silico methods predicted this mutation as destabilizing to the structure of the protein pRb by I-Mutant software.

Furthermore, this mutation was predicted as tolerated, benign and classify as a polymorphism using SIFT, Polyphen2 and Mutation taster softwares. Analysis with Align GVGD shows that this mutation was considered as "less likely interfere with function". However, the effect of substitution of amino acid alanine by proline was clearly shows the increase in size and a decrease in the hydrophobicity of mutated amino acid. This difference can have an impact on the structure and function of pRb. The missense mutation c.1903 G>C, (p.Ala635Pro) was considered like a polymorphism without any impact on the protein function by SIFT, Polyphen, Align GVGD and mutation taster softwares but in the mutation databases of RB1 gene (http://rb1-lsdb.dlohmann.de/home.php?select_db=RB1) this variation was considered to be causal mutation for the retinoblastoma. This variation is located in the spacer region, previously assumed to be non-essential to peptide-binding activity of the pRb pocket by many authors [27]. Recent study employing Clustal X using default parameters [30], a multiple sequence alignment of the pRb spacer region including sequences from eight different species (cow, mouse, chicken, newt..) showed that the spacer region is fairly well conserved. This finding indicates that the spacer region is important. In addition to this study demonstrated that the mutation Leu607Ile and Arg621Cys located to the spacer region was considered the mutant protein less stable compared to pRb wild type and expressed a reduced apoptotic function. These findings, indicates a functional role for this region of pRb pocket. The c.1903 G>C, (p.Ala635Pro) mutation also Predicted by in silico splice prediction program: Enhancer-finder (ESE) affect splicing is the consequence of the skipping the exon 19 by using the splice donor site of exon 18 and the splice acceptor site of exon 20. The deletion of exon 19 can disrupt the three-dimensional conformation of the protein resulting in a perturbation of function.

The missense mutation c.1961T>A, (p.Val654Gln) was also scored for predicted pathogenicity using seven online in silico protein analysis programs. This mutation was predicted to be destabilizing the structure of the pRb protein by I-Mutant software. In addition, it was categorized as not tolerated, "probably damaging", and causing the disease by SIFT, Polyphen2 and Mutation taster softwares. The analysis by Align GVGD showed that the p.Val654Gln was considered as "can interfere with the function. However, the KD4V software predicted that this substitution might significantly affect pRb function (change in charge and hydrophobicity).

The major significance of our results is to consider *in silico* analysis a suitable protocol for predict deleterious effect of mutations before wet laboratory experimentation and also showing optimal path for further clinical and experimental studies to characterize pRb mutant protein.

Conclusion

The detection of neo-mutations is great interest in the case of genetic counseling for members of a family where there is only one sporadic case of disease. It makes it possible to resolve definitively the branch of the family not concerned by the disease. In retinoblastoma as in most cancers, the prognosis depends on the precocity of the diagnosis. This early diagnosis, possibly prenatal, of the mutation will allow the rapid management of these children by continuous monitoring and will contribute to the prevention of the disease.

The *in silico* software use a variety of approaches to achieve a prediction: sequence and evolutionary conservation-based methods, protein sequence, structure-based methods and machine learning

methods. Our study demonstrated the efficacy of the use *in silico* analysis, combined of several predictive tools to evaluate the deleterious effects of missense mutations and understand the impact of these mutations on protein level. It is interesting to compare the *in silico* study results with functional studies to evaluate the functionality of mutations because in retinoblastoma as in most cancers, understanding the molecular mechanisms associated with different dysfunctions of the pRb protein is essential for the diagnosis, prognosis, and development of new therapeutic strategies.

Conflict of Interest

The authors have declared that no conflict of interests exists.

References

- 1. Aerts I, Lumbroso-Le Rouic L, Gauthier-Villars M, Brisse H, Doz F, et al. (2006) Retinoblastoma. Orphanet J Rare Dis 1: 31.
- Bunin GR, Meadows AT, Emanuel BS, Buckley JD, Woods WG, et al. (1989) Pre- and postconception factors associated with sporadic heritable and nonheritable retinoblastoma. Cancer Res 49: 5730-5.
- Fletcher O, Easton D, Anderson K, Gilham C, Jay M, et al. (2004) Lifetime risks of common cancers among retinoblastoma survivors. J Natl Cancer Inst 96: 357-363.
- 4. Weinberg RA (1995) The retinoblastoma protein and cell cycle control. Cell 81: 323-330.
- Valverde JR, Alonso J, Palacios I, Pestaña A (2005) RB1 gene mutation up-date, a meta-analysis based on 932 reported mutations available in a searchable database. BMC Genet 6: 53.
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16: 1215.
- Houdayer C, Gauthier-Villars M, Laugé A, Pagès-Berhouet S, Dehainault C, et al. (2004) Comprehensive screening for constitutional RB1 mutations by DHPLC and QMPSF. Hum Mutat 23: 193-202.
- 8. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16: 10881-10890.
- 9. Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, et al. (2009) Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37: 1.
- Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR (2003) ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 31: 3568-3571.
- 11. Capriotti E, Fariselli P, Casadio R (2005) I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 33: W306-310.
- 12. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31: 3812-3814.
- Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding nonsynonymous variants on protein function using the SIFT algorithm. Nat Protoc 4: 1073-1081.

- 14. Tavtigian SV, Byrnes GB, Goldgar DE, Thomas A (2008) Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. Hum Mutat 29: 1342-54.
- 15. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods.
- Luu TD, Rusu A, Walter V, Linard B, Poidevin L, et al. (2012) KD4v: Comprehensible Knowledge Discovery System for Missense Variant. Nucleic Acids Res 40: 27.
- 17. Burke JR, Hura GL, Rubin SM (2012) Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control. Genes Dev 26: 1156-66.
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18: 2714-23.
- Blanquet V, Turleau C, Gross-Morand MS, Sénamaud-Beaufort C, Doz F, et al. (1995) Spectrum of germline mutations in the RB1 gene: a study of 232 patients with hereditary and non hereditary retinoblastoma. Hum Mol Genet 4: 383-8.
- Corden J, Wasylyk B, Buchwalder A, Sassone-Corsi P, Kedinger C, et al. (1980) Promoter sequences of eukaryotic protein-coding genes. Science 209: 1406-14.
- 21. Lohmann DR, Gerick M, Brandt B, Oelschläger U, Lorenz B, et al. (1997) Constitutional RB1 gene mutations in patients with isolated unilateral retinoblastoma. Am J Hum Genet 61: 282-294.
- 22. Blanquet V, Turleau C, Gross MS, Goossens M, Besmond C (1993) Identification of germline mutations in the RB1 gene by denaturant gradient gel electrophoresis and polymerase chain reaction direct sequencing. Hum Mol Genet 2: 975-9.
- Boubekeur A, Louhibi L, Mahmoudi K, Boudjema A, Mehtar N (2012) [Molecular study of retinoblastoma in the Algerian population. Screening of Rb gene in constitutional and tumoral level]. Bull Cancer 99: 127-35.
- Lee JO, Russo AA, Pavletich NP (1998) Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7. Nature 391: 859-865.
- 25. Houdayer C, Stoppa-Lyonnet D (2005) [Transcriptional abnormalities and genetic testing]. Med Sci (Paris) 21: 170-174.
- Corcos L, Solier S (2005) [Alternative mRNA splicing, pathology and molecular therapeutics]. Med Sci (Paris) 21: 253-260.
- 27. Liu H-X, Zhang M, Krainer AR (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev 12: 1998-2012.
- Nilsen TW (2003) The spliceosome: the most complex macromolecular machine in the cell? Bioessays 25: 1147-1149.
- Sachdeva UM, O'Brien JM (2012) Understanding pRb: toward the necessary development of targeted treatments for retinoblastoma. J Clin Invest 122: 425-434.
- Berge EO, Knappskog S, Geisler S, Staalesen V, Pacal M, et al. (2010) Identification and characterization of retinoblastoma gene mutations disturbing apoptosis in human breast cancers. Molecular Cancer 9: 173-173.

Page 9 of 9