A STUDY OF THE GENETIC VARIATIONS IN N-TERMINUS, C-TERMINUS AND PROMOTER OF *RB1* GENE AMONG RETINOBLASTOMA PATIENTS IN MALAYSIA

by

SITI NORULHUDA BT HAJI DRAMAN @ HASHIM

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

May 2012

MUTATIONAL ANALYSIS IN N-AND C- TERMINI OF RB1 GENE AMONG SPORADIC RETINOBLASTOMA PATIENTS IN MALAYSIA

Siti Norulhuda Hashim
MSc Human Genetics

Human Genome Center
School of Medical Sciences, Universiti Sains Malaysia
Health Campus, 16150 Kelantan, Malaysia

Background: Retinoblastoma (RB) is childhood malignant tumors with majority of cases affect children under 5 years old. RB is caused by mutation in *RB1*, a tumor suppressor gene which is located on chromosome 13q14. The aim of this study is to detect mutations and single nucleotide polymorphisms (SNPs) in N- and C-terminus of *RB1* in Malaysian children with RB, as well as its association with laterality and staging of this disease.

Methods: Peripheral blood leukocytes were collected from 66 retinoblastoma patients and 66 healthy volunteers. Tumor tissue specimen also collected from three available patients. The DNA was extracted using commercially available extraction kit. PCR were conducted using self-designed primers. The mutational analysis was performed using DHPLC and direct sequencing method.

Results: Two nonsense mutations and four SNPs were detected in N-terminus, but none in C-terminus. The nonsense mutations; Arg320X and Glu323X were detected in 3 children with RB. A novel SNP, IVS1-3T>G was identified in this study. There was a significant different in allele frequency of IVS4-77G>A between patients and healthy control group (p=0.044). There was no significant association between polymorphisms with laterality and staging of RB.

Discussion and conclusion: Arg320X and Glu323X play as important role in the predisposition to RB. SNP IVS4-77G>A also have a potential for genetic defect RB in our population. However, a larger sample size is needed for more confirmation. The other SNPs have more potential to be used as genetic variant markers for population studies

Assoc Prof Dr Liza Sharmini Ahmad Tajudin: Supervisor
Prof Zilfalil Alwi: Co-supervisor
Assoc Prof Dr Ariffin Nasir: Co-supervisor

ACKNOWLEDGEMENT

First of all, my biggest gratitude to Allah S.W.T., The Most Gracious and Most Merciful, for the strength, patience and healthiness to finish my thesis within the stipulated time. My deep appreciation to Associate Professor Dr. Liza Sharmini Ahmad Tajudin (main-supervisor), Professor Dr. Zilfalil Alwi (co-supervisor) and Associate Professor Dr. Ariffin Nasir (co-supervisor) for giving me guidance, precious time, lessons, comments and supports from the beginning to the final of my study. I would like to acknowledge Universiti Sains Malaysia (USM) for the Research University (RU) grant (1001/PPSP/811080) and USM Fellowship Scheme for the scholarship during my study. I also want to deliver my appreciation to Dr. Jamalia Rahmat (HKL) and Dr. Nurliza Khaliddin (UMMC) who help me recruiting samples from their institution. I would like to acknowledge all the staff at Human Genome Center, Ophthalmology clinics of USM, HKL and UMMC for their helps and cooperation. Special thanks to Kak Cik for all the knowledges I gained from you. I also would like to show my gratitude to my fantastic friends in the lab; Siti, Sathiya, Ina, Kak Mar, Kak Hatin, Kak Mareen, Kak Sha, Kak Arfah, Kak Tasya, Kak Hasnah, Yan Yan, Ju, Syibli, Rani, Aizat, Amin and others, not forget my beloved friends from PPSK; Amalina and Gisel, thank you for all the cheerful times and your help. Last but not least, my gratitude to my family, especially my beloved parent, Hashim bin Ismail and Kolsoom binti Abu Bakar, who were always there for me and giving all their supports during my study. May Allah bless us all. Thank you.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

$(NH_4)_2SO_4$	: Ammonium sulphate
µg	: Microgram
µl	: Microlitre
µM	: Micromolar
2N	: Diploid
4N	: Tetraploid
A_{260}/A_{280}	: Ratio of 260 absorbance over 280 absorbance
aa	: Amino acid
ATF	: Activating Transcription Factor
bp	: Base pair
Buffer AE	: Elution buffer
Buffer AL	: Lysis buffer
Buffer AW1	: Wash buffer 1
Buffer AW2	: Wash buffer 2
Buffer BL	: Lysis buffer
BW buffer	: Wash buffer
C-abl	: Type of tyrosine kinase
CDK	: Cyclin dependent kinase
CDK4	: Cyclin dependent kinase 4
CpG	: C-phosphate-G
CT	: Computerized tomography
dATP	: Deoxyadenosine triphosphate
dbSNP	: Single Nucleotide Polymorphism database
dCTP	: Deoxycytidine triphosphate

ddH_2O	: Double distilled water
ddNTP	: Dideoxy nucleoside triphosphate
DGGE	: Denaturing gradient gel electrophoresis
dGTP	: Deoxyguanosine triphosphate
DHPLC	: Denaturing High Performance Liquid Chromatography
DNA	: Deoxyribonucleic acid
dNTP	: Deoxyribonucleotide triphosphate
dsDNA	: Double-stranded DNA
dTTP	: Deoxythymidine triphosphate
E2F	: E2 promoter binding factor
EDTA	: Ethylenediaminetetraacetic acid
ETOH	: Ethanol
ExoSAP	: Exonuclease I - Shrimp Alkaline Phosphatase
g	: Gram
g	: G-force
G_0	: Gap 0
G_1	: Gap 1
G_2	: Gap 2
H1	: Histone I protein
HaeIII	: Restriction endonuclease from Haemophilus aegypticus
HD	: Heteroduplex
Hg	: Mercury (element)
HKL	: Hospital Kuala Lumpur
HPV	: Human papillomavirus
HUSM	: Hospital Universiti Sains Malaysia

ICRB	: International Classification of Retinoblastoma
IR-RP	: Ion pair chromatography reverse phase
IVS	: Intervening sequence
Kb	: Kilobase pair
kDA	: Kilodalton
LOH	: Loss of heterozygosity
M	: Molarity
M phase	: Mitotic phase
MAF	: Minor allele frequency
MCM7	: Minichromosome maintenance deficient 7 gene
MDM2	: Murine double minute 2 gene
$MgCl_2$	: Magnesium chloride
min	: Minute
ml	: Microlitre
mM	: Micromolar
MRI	: Magnetic resonance imaging
mRNA	: Messenger RNA
MSP	: Methylation Specific PCR
$NaHSO_3$	: Sodium bisulfite
NaOH	: Sodium hydroxide
ng	: Nanogram
p16	: Cyclin-dependent kinase inhibitor 2A
PCR	: Polymerase chain reaction
pH	: Power of hydrogen
pRB	: Retinoblastoma protein

PSEN-1	: Presenilin 1 gene
pUC18	: Plasmid DNA from *Escherichia coli* RRI
RB	: Retinoblastoma
RB1	: Retinoblastoma susceptibility gene
RBF-1	: Retinoblastoma binding factor 1
RE	: Restriction endonuclease
RNA	: Ribonucleic acid
rpm	: Revolutions per minute
rs	: Reference sequence
RT-PCR	: Reverse transcriptase PCR
S phase	: Synthesis phase
sec	: Second
SMN	: Second malignant neoplasm
SNP	: Single nucleotide polymorphism
Sp1	: Specificity Protein 1
SSCP	: Single-strand confirmation polymorphism
ssDNA	: Single-stranded DNA
STR	: Short tandem repeat
TBE	: Tris-Borate EDTA
TEAA	: Triethylammonium acetate
TNM	: Tumor-node-metastasis
Tris-HCl	: (Hydroxymethyl)aminomethane hydrochloride
UMMC	: Universiti Malaya Medical Center
USA	: United States of America
UV	: Ultra violet

v/v	: Chemical solutions by volume
V	: Voltage
ρM	: Picomolar

LIST OF SYMBOLS

°C	: Degree celcius
>	: More than
<	: Less than
%	: Percentage

KAJIAN VARIASI GENETIK DALAM TERMINUS N, TERMINUS C DAN PROMOTER PADA GEN *RB1* DI KALANGAN PESAKIT RETINOBLASTOMA DI MALAYSIA

ABSTRAK

Retinoblastoma (RB) adalah sejenis kanser mata yang sering terjadi di kalangan kanak-kanak yang berumur lima tahun dan ke bawah dengan kekerapan seorang pesakit di kalangan 15,000 hingga 20,000 kelahiran. Retinoblastoma boleh terjadi secara keturunan atau sporadik. Kanser mata ini berlaku disebabkan oleh mutasi atau perubahan pada gen *RB1*, sejenis gen penindas tumor. Tujuan kajian ini dilakukan adalah untuk mengenalpasti mutasi dalam terminus N, terminus C dan promoter pada gen *RB1* pesakit retinoblastoma serta kaitannya dengan peringkat dan lateraliti penyakit tersebut. Seramai 68 kanak-kanak pesakit RB sporadik daripada Hospital Universiti Sains Malaysia (HUSM), Hospital Kuala Lumpur (HKL) dan Pusat Perubatan Universiti Malaya (PPUM) dan 68 sukarelawan normal dengan padanan etnik dengan kumpulan pesakit telah mengambil bahagian didalam kajian ini. Sebanyak 3cc darah telah diambil dari pesakit dan sukarelawan. Sampel tisu tumor juga telah diambil daripada empat pesakit RB. Sebanyak 200µl DNA telah diekstrak menggunakan kit ekstrak darah komersial. PCR telah dijalankan dengan menggunakan primer yang direka sendiri untuk ekson dalam terminus N dan terminus C serta kawasan intron sekitarnya, dan promoter pada gen *RB1*. Nisbah kanak-kanak lelaki kepada kanak-kanak perempuan adalah 2:1, di mana 79% adalah Melayu, 12% Cina dan 9% India. Leukocoria merupakan tanda klinikal awal pada sebahagian besar pesakit dan kebanyakan adalah melibatkan sebelah mata tanpa

kecenderungan terhadap mata kanan atau kiri. Secara purata, umur pesakit ketika didiagnosis adalah 23.42 (17.64) bulan. Diagnosis lebih awal [17.68 (18.33) bulan] melibatkan pesakit RB yang melibatkan kedua-dua belah mata berbanding dengan pesakit sebelah mata [28.65 (15.46) bulan]. Berdasarkan umur ketika diagnosis terdapat perbezaan statistik yang ketara (p=0.011) di antara dua kumpulan ini. Terdapat juga kaitan yang ketara di antara lateraliti dengan bangsa pesakit RB (p=0.025). Dua mutasi tiada makna dan empat polimorfisma nukleotida tunggal (SNPs) telah dikenalpasti dalam terminus N; satu SNP telah dikenalpasti dalam promoter, tetapi tiada sebarang mutasi mahupun polimorfisma yang dijumpai dalam terminus C. Semua mutasi dan polimorfisma yang dijumpai pada 30 pesakit dan kawalan melibatkan penggantian nukleotida tunggal yang heterozigus. Mutasi tiada makna; Arg320X dan Glu323X telah dikesan pada 3 kanak-kanak RB. Satu nobel SNP terbaru, IVS1-3T>G telah ditemui pada seorang pesakit. IVS4-77G>A yang telah ditemui pada empat pesakit RB tapi tidak dijumpai pada kumpulan kawalan. Polimorfisma ini didapati meningkatkan kecenderungan yang ketara untuk penyakit RB (p=0.044). Walaubagaimanapun, tiada kaitan yang ketara ditemui antara semua variasi genetik yang ditemui dengan peringkat dan lateraliti pesakit. Tiada perbezaan variasi genetik yang dijumpai pada darah dan tisu tumor daripada empat pesakit. Bagaimanapun, lebih banyak sampel tisu tumor diperlukan untuk kepastian lebih lanjut. Berdasarkan penemuan ini SNP IVS4-77G>A mungkin memainkan peranan dalam mempengaruhi risiko seseorang untuk menghidap kanser mata ini. Empat SNP yang lain lebih cenderung sebagai penanda variasi genetik dalam populasi. Walaubagaimanapun, bilangan sampel yang lebih besar diperlukan bagi tujuan pengesahan.

A STUDY OF THE GENETIC VARIATIONS IN N-TERMINUS, C-TERMINUS AND PROMOTER OF *RB1* GENE AMONG RETINOBLASTOMA PATIENTS IN MALAYSIA

ABSTRACT

Retinoblastoma (RB) is the most common primary intraocular tumor affecting mainly children under five years of age, with a prevalence of 1 in 15 000 to 20 000 live births. Retinoblastoma is divided into hereditary and sporadic types. It is caused by mutations or changes in *RB1* gene, a tumor suppressor gene. The aim of this study is to detect mutations in N-terminus, C-terminus and promoter region of *RB1* gene in retinoblastoma patients in Malaysia, as well as its association with staging and laterality of the patients. A total of 68 children with sporadic RB from Hospital Universiti Sains Malaysia (HUSM), Hospital Kuala Lumpur (HKL) and University Malaya Medical Centre (UMMC) and 68 healthy ethnic-matched controls were recruited. Venesection was done and 3cc of blood was taken. Tumor tissue samples were also collected from four patients. A total of 200µl DNA was extracted using commercially available extraction kit. PCR was conducted using self-designed primers for exons in N-terminus and C-terminus with its intronic flanking region, and promoter region of *RB1* gene. Screening for mutations or polymorphisms for N-terminus was conducted using DHPLC analysis and direct sequence was done for C-terminus and promoter region. Male to female patient ratio was 2:1, where 79% were Malays, 12% Chinese and 9% Indian. Majority of our patients presented with leukocoria and unilateral involvement with no predilection to right or left eye. The mean age at diagnosis of patients was 23.42 (17.64) months. There was statistical

significant difference between age of diagnosis and laterality of the disease (p=0.011). Bilateral RB was diagnosed earlier [17.68 (18.33) months) than unilateral RB [28.65 (15.46) months). There was also significant association between laterality and racial group of RB patients (p=0.025). Two nonsense mutations and four SNPs were detected in N-terminus; one SNP was detected in the promoter region. However, C-terminus was devoid of any polymorphism or mutation. All mutations and polymorphisms which were found in 30 patients and healthy controls involved a heterozygous type of single nucleotide substitution. The nonsense mutations; Arg320X and Glu323X were detected in 3 children with RB. A novel SNP; IVS1-3T>G found in this study was identified in one patient. IVS4-77G>A identified in four patients was found to significantly increase the susceptibility to develop RB (p=0.044). However, there was no significant association between identified mutations and polymorphisms with laterality and staging of RB. There was no disparity in mutations and polymorphism between peripheral blood leukocytes and available tumor tissue. We postulate that SNP IVS4-77G>A might play a role in the predisposition to retinoblastoma. The other five SNPs are most likely important as genetic variant markers for population studies. However, a larger sample size is needed for further confirmation.

CHAPTER 1

INTRODUCTION

1.1 Retinoblastoma

Retinoblastoma (RB) is the most common primary intraocular malignancy of infancy and childhood (Abramson and Schefler, 2004). It is a relatively rare form of childhood cancer which affects young children under age of 5 years. The tumor develops from the immature retina, which is the part of the eye responsible for detecting light and color **(Figure 1.1)**. Retinoblastoma was a uniformly fatal disease (Albert, 1987). The incidence of RB is estimated to be one in 15,000 and one in 20,000 live births (Lohmann and Gallie, 2004; Shields and Shields, 2004; Rodriguez *et al.*, 2002). Approximately one in three to one in four affected children will have a family history of RB. In United States, approximately 300 children are diagnosed with retinoblastoma each year (Tamboli *et al.*, 1990). In developed countries, over 95% of children survive the malignancy by early detection and immediate treatment of affected eye (Broaddus *et al.*, 2009b). However, the survival rate can be as low as 6.8% in less developed countries mainly caused by late presentation and late treatment as the tumor extended to the orbit, optic nerve and brain (Shields and Shields, 2004). In Malaysia as in most developing countries, parents usually prefer to attempt traditional treatment as they afraid of enucleation (Sinniah *et al.*, 1980), and only seek modern medical treatment at stage when it is incurable (Menon *et al.*, 2009).

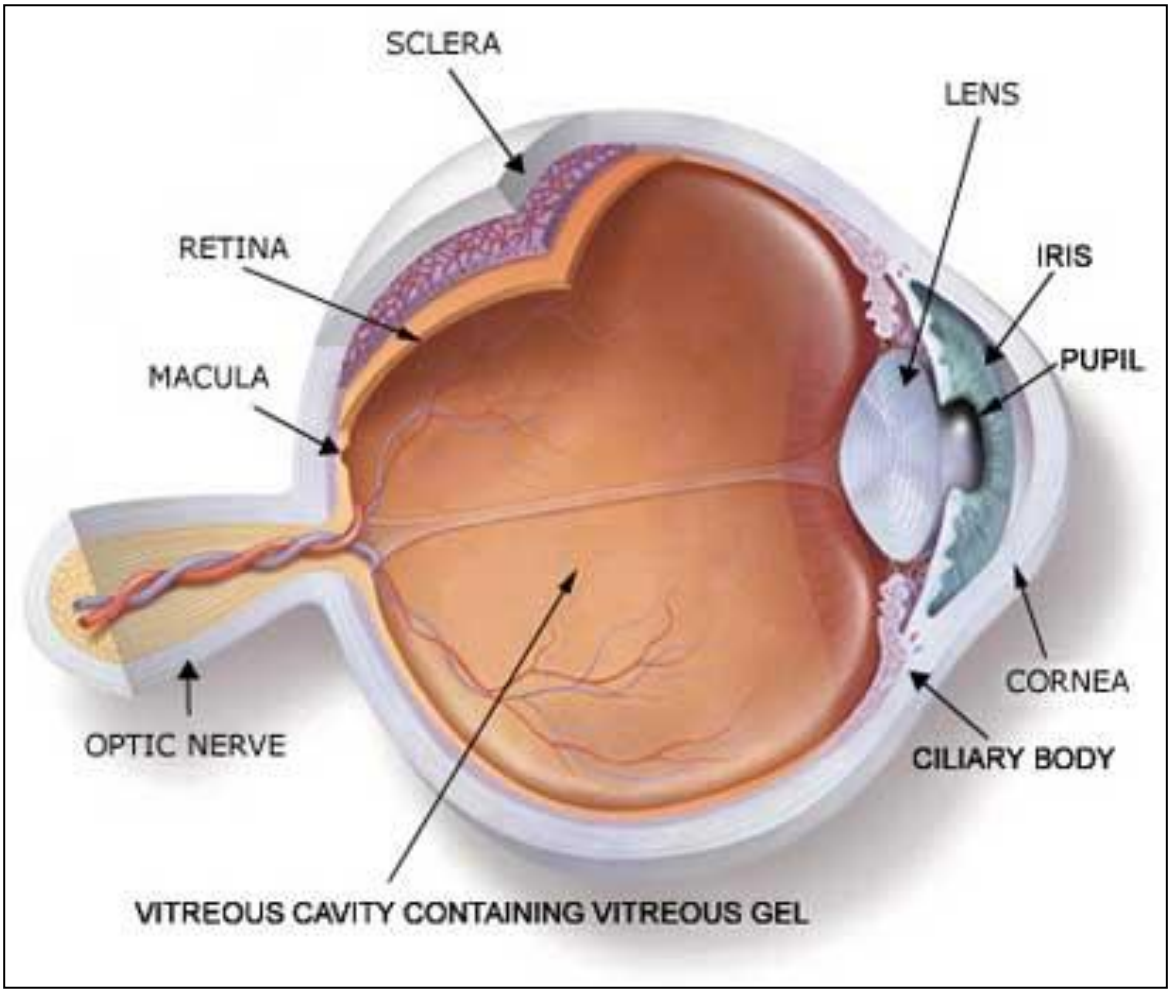

Figure 1.1: Schematic diagram of eye. Retinal cells located at back layer of the eye and connected with the optic nerve to transfer the light rays to electrical impulse and sent to the brain (http://www.clinicaomegalanzarote.com/en/retina.htm)

RB is caused by mutation in *RB1* gene, a tumor suppressor gene. *RB1* gene is a gene which responsible to encode retinoblastoma phosphoprotein (pRB). It is a regulatory protein that functions in cell cycle activity. pRB protein normally suppresses the cell cycle activity when the cell sends the signal to stop the production of new cells. The presence of mutation in *RB1* gene is believed to be the main factor leading to loss of function of pRB protein in cell cycle activity by altering pRB structure, and initiating the tumor development (Knudson, 1971; Lohmann, 1999).

1.2 Clinical presentation

The clinical presentation of retinoblastoma is dependent on tumor growth pattern, duration of disease, degree of tumor vascularity, the presence of calcifications, vitreous seeding, and retinal detachment (Kiss *et al.*, 2008). The most common presentation of retinoblastoma is leukocoria, an abnormal white discoloration in one or both pupils. In United States, almost 60% of all cases of retinoblastoma are diagnosed after observation of leukocoria in the affected eye, which initially noticed by a parent or relative (Shields *et al.*, 1991). Leukocoria is a white reflection in the pupil and only visible at certain angles and under certain light conditions especially on flash photography. The survival rate in retinoblastoma patients with leukocoria as initial presentation is high (88% at 5 years) but poor prediction for preserving the globe (Balmer *et al.*, 2006; Shields and Shields, 2004). The second most common sign of retinoblastoma is strabismus, which detected approximately 1 in 5 retinoblastoma patients in the United States (Abramson *et al.*, 1998). Strabismus, or squint, is the impairment of the vision resulting from a loss of central vision in one or both eyes. Tumors presenting with strabismus as the initial sign are associated with a

higher survival rate (similar to leukocoria), but with a higher chance of preservation of the globe (Balmer *et al.*, 2006; Shields and Shields, 2004). The remaining 20% of the cases of retinoblastoma present with atypical signs and symptoms, including painful red eye, glaucoma, cloudy cornea, vitreous hemorrhage, or signs of orbital inflammation mimicking orbital cellulitis (Abramson *et al.*, 1998). These uncommon presentations are usually late signs associated with more advanced disease and poor salvation of the globe (Balmer *et al.*, 2006; Butros *et al.*, 2002). Some children with retinoblastoma may be asymptomatic (Aerts *et al.*, 2006). The presenting signs and symptoms in advanced stage usually correlate with the degree of extraocular invasion, which can result in orbital swelling and proptosis (Chintagumpala *et al.*, 2007).

Retinoblastoma growth patterns are subdivided into endophytic and exophytic. Endophytic retinoblastoma accounting for nearly 60% of cases, exhibit growth toward the vitreous cavity as a result of cell division within the inner retinal layers (Palazzi *et al.*, 1990). It can be unifocal or multiple foci, isolated or fused, variable in size, round or oval-shaped, and appears as calcified creamy-white mass or vascularized pink retinal mass. Endophytic lesion has higher tendency for vitreous seeding (Balmer *et al.*, 2006). The location of endophytic retinoblastoma is tumor age-dependent. The earliest tumors usually develope closer to the posterior pole and later progresses anteriorly (Abramson and Gombos, 1996). Approximately 39% of retinoblastoma presented as exophytic growth pattern with tumor development in the outer retinal layers and expansion beneath the retina into the subretinal space (Palazzi *et al.*, 1990). Retinal detachment is the common presentation in exophytic retinoblastoma (Palazzi *et al.*, 1990). Choroidal invasion also occur significantly

more often in patients with exophytic retinoblastoma than in those with endophytic retinoblastoma (Palazzi *et al.*, 1990).

Retinoblastoma can involve unilateral or bilateral. Unilateral involvement is seen in 60% of retinoblastoma cases, the disease is unilateral and the median age at diagnosis is two years. 15% of unilateral cases are hereditary (Aerts *et al.*, 2006). Patients with bilateral tumors have one or more tumors in both eyes. It is presumed that these patients have the hereditary form of the disease even in the absence of a positive family history. The severity of tumors may be variable in the two eyes, and this becomes important in weighing potential treatment options. Retinoblastoma is bilateral in about 40% of cases with a median age at diagnosis of one year (Lohmann and Gallie, 2004). Trilateral retinoblastoma refers to bilateral retinoblastoma associated with an intracranial neuroblastic tumor of the pineal gland called pinealoblastoma (Paulino, 1999). Trilateral retinoblastoma was recognized to be the most frequent cause of death among children with hereditary RB in the first decade of life, accounting for 50% of all deaths (Blach *et al.*, 1994), and occurs in approximately 8% of patients with hereditary retinoblastoma (De Potter *et al.*, 1994). The median survival time after diagnosis of trilateral retinoblastoma was 9 months (Kivela, 1999).

1.2.1 Staging

Retinoblastoma can be classified according to three classifications; Tumor-node-metastasis (TNM) classification, Reese-Ellsworth (RE) classification, and the new

International Classification of Retinoblastoma (ICRB). The classification or staging helps to prognosticate the survival, treatment options as well as visual potential.

As retinoblastoma has spread beyond the eye, the TNM classification used to stage disease. T describes the size of the tumor and whether it has invaded nearby tissue; N describes regional lymph nodes that are involved; M describes distant metastasis. Groups T1 and T2 staged the tumor contained in the eye, however, the International Intraocular Retinoblastoma Classification (ICRB) and Reese-Ellsworth (RE) Classification are globally preferred for staging of retinoblastoma, as they are better predictors of treatment outcome.

The Reese-Ellsworth classification was introduced in 1963 following the presentation by Reese and Ellsworth during 67th annual meeting of the American Academy of Ophthalmology. The classification was based on intraocular tumor staging and prediction on globe salvation following external beam radiation (Reese and Ellsworth, 1963). External beam radiation was the main treatment during this era. It has gained wide popularity for the last four decades. However, the Reese-Ellsworth classification failed to address the problem of subretinal seeding and did not differentiate between focal and diffuse vitreous seeding. For these reasons, the Reese–Ellsworth classification was found to be a poor predictor of chemoreduction success (Shields and Shields, 2004). Therefore, a new classification for retinoblastoma was designed (Linn Murphree, 2005).

The new classification, the International Classification of Retinoblastoma (ICRB) **(Table 1.1)**, was developed to predict the outcome from combination chemotherapy

and focal therapy. It based mainly on the extent of tumor seeding in the vitreous and subretinal space, with consideration for tumor size and location (Linn Murphree, 2005). ICRB was later accepted as standard staging for intraocular tumor globally in the early 2000s. The chances of maintaining useful vision decreased exponentially from groups A to E. Each eye is staged independently.

Table 1.1: The International Classification of Retinoblastoma

Group	Quick reference	Specific features
A	Small tumor	RB <3 mm*
B	Larger tumor	RB >3 mm* in size or
	Macula	Macular RB location (<3 mm to foveola)
	Juxtapapillary	Juxtapapillary RB location (<1.5 mm to disc)
	Subretinal fluid	Additional subretinal fluid <3 mm from margin
C	Focal seeds	RB with: Subretinal seeds <3 mm from RB Vitreous seeds <3 mm from RB Both subretinal and vitreous seeds <3 mm from RB
D	Diffuse seeds	RB with: Subretinal fluid >3 mm from RB Subretinal seeds > mm from RB Vitreous seeds >3 mm from RB Both subretinal and vitreous seeds >3 mm from RB
E	Extensive RB	Extensive RB occupying >50% globe or Neovascular glaucoma Opaque media from hemorrhage in anterior chamber, vitreous, or subretinal space Invasion of postlaminar optic nerve, choroid (>2 mm), sclera, orbit, anterior chamber

RB – retinoblastoma
*refers to 3 mm in basal dimension or thickness

1.2.2 Second malignant neoplasm (SMN)

Hereditary retinoblastoma is associated with higher risk of developing second malignancies throughout life (Margo *et al.,* 1998). More children with RB die from the second malignant neoplasms than from the retinoblastoma itself (Abramson, 1999). The cumulative incidence of second malignancy in hereditary retinoblastoma after 50 years of diagnosis is 51%, 10 times higher compared to 5% in nonhereditary cases (Abramson and Frank, 1998; Wong *et al.*, 1997). The mean response time between retinoblastoma and second malignancy is approximately 13 years (Wong *et al.*, 1997; Margo *et al.,* 1983; Draper *et al.,* 1986). The most common SMNs are osteosarcoma (37%), followed by soft-tissue sarcoma (7%) and melanoma (7%) (Moll *et al.,* 1997a). Other reported malignancies including neuroblastoma, rhabdomyosarcoma, glioma, leukemia, squamous cell carcinoma and cutaneous melanoma (Wong *et al.,* 1997; Margo *et al.,* 1983; Draper *et al.,* 1986). External-beam radiotherapy plays a major role in pathogenesis of second malignant neoplasms, and shows a radiation dose-response relationship for all sarcomas (Wong *et al.,* 1997). With longer follow-up of diagnosis of retinoblastoma, the incidence of SMN in cases that were treated with external-beam radiotherapy is more than doubled when compared with those not receiving radiotherapy (58% and 27% respectively) (Abramson and Frank, 1998).

1.3 Knudson two-hit hypothesis

The genetics of retinoblastoma initiated the paradigm of understanding a genetically inherited cancer and provides the basis for the two-hit hypothesis of carcinogenesis

(Friedman *et al.*, 2000). Deletion, mutation, inactivation or total loss of both copies of the *RB1* gene at the 13q14 locus need to occur in order for retinoblastoma to develop (Shields and Shields, 2004). Knudson (1971) proposed the "two hit" hypothesis to explain the events that are necessary for both hereditary and non-hereditary retinoblastoma. He used the knowledge about the time of clinical presentation of retinoblastoma and the number of cell divisions in human retina. His theory was based on a comparative analysis of unilateral and bilateral retinoblastoma cases. Hereditary retinoblastoma occurs when the first mutational event or "hit" is at germinal level of retinal cell which is transmitted from the parent, followed by the second "hit" which occurs in somatic cell during development. This possibly explained the hereditary retinoblastoma; all cells in the body are predisposed to possible tumor development inherited from germline mutation ("first hit"). This explains the high incidence of second nonocular tumors, such as osteosarcoma. By contrast, sporadic or non-hereditary retinoblastoma, the "two hits" occur during development of the retina and both "hits" are somatic mutations. The rest of the body theoretically carries no higher risk to develop other tumors because these patients presumably have normal chromosomal structure elsewhere in the body (Shields and Shields, 2004).

1.4 *RB1*

A tumor suppressor gene is a gene that encodes a protein to inhibit cell division. Tumor suppressor gene and their gene products must be inactive or absent for the cell to divide. Permanent inactivation through mutation causes the disruption of division and induces uncontrolled proliferation of tumor cell.

RB1 was the first tumor suppressor gene to be identified which is responsible for retinoblastoma (Friend *et al.*, 1986). *RB1* is located on the long arm of chromosome 13 in region 13q14 **(Figure 1.2)**. The genomic locus of *RB1* consists of 27 exons distributed over 200 kilobases (Wiggs *et al.*, 1988) **(Figure 1.3)**. The sizes of exons range from 31 to 1889 base pairs, while the 26 introns span 80 to 71712 base pairs (Hong *et al.*, 1989). Promoter region of *RB1* is identified at 186 to 206 bp upstream from initiation codon. It contains binding domain for various transcription factor protein such as RBF-1, Sp1, ATF and E2F (Ohtani-Fujita *et al.*, 1993). The *RB1* gene transcribed into a 4.7 kb messenger RNA (mRNA), which encodes a 110 kDa nuclear phosphoprotein of 928 amino acids. It is known as pRB (Lee *et al.*, 1987). pRB acts as a cell-cycle regulated growth suppressor. The absence of pRB has been confirmed in retinoblastoma cell lines (Goodrich and Lee, 1993).

It is postulated that the development of cancer occurs when there is mutation of both alleles of this tumor suppressor gene. Apart from retinoblastoma, mutations involving *RB1* gene also predisposes to a variety of other tumors including osteosarcoma (Dryja *et al.*, 1986; Toguchida *et al.*, 1988), breast cancer (Eyfjord and Thorlacius, 1992; Ceccarelli *et al.*, 1998), acute leukemia (Sauerbrey *et al.*, 1998; Kornblau *et al.*, 1998; Ahuja *et al.*, 1991), bladder cancer (Grossman *et al.*, 1998; Niehans *et al.*, 1999), lung cancer (Otterson *et al.*, 1994; Harbour *et al.*, 1988) and other cancers in later life (Sellers and Kaelin, 1997).

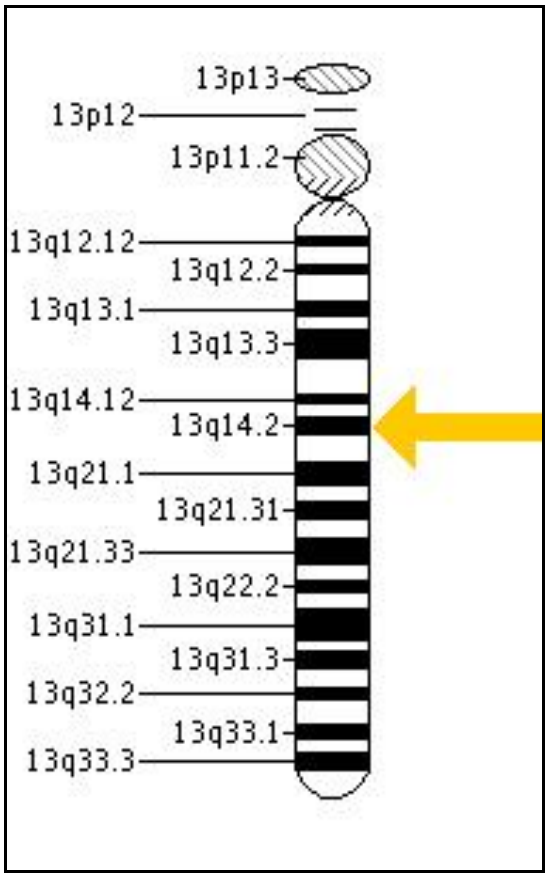

Figure 1.2: Location of *RB1* gene at q-arm (long arm) of chromosome 13, at sub-band 14.2 (http://ghr.nlm.nih.gov/dynamicImages/chromomap/RB1.jpeg).

1.4.1 pRB

The protein product of *RB1* is a nuclear phosphoprotein (pRB) with DNA binding properties (Lee *et al.*, 1987). It acts as a cell cycle regulator, which blocks the transition of normal cells from G_0/G_1 into S phase of the cell cycle (Ludlow *et al.*, 1990). pRB has the ability to bind and inhibit the transcription factor; E2F, that activate the genes involved in cell division. E2Fs bind to pRB in the pocket domain and the carboxy-terminal region of the protein. Thus, mutations at either of these locations could reduce the binding potential to E2F and allows for uncontrolled entry into the S-phase and contribute to rapid cell division (Harbour and Dean, 2000).

pRB consist of four domains; N terminus, pocket A and B domain and C terminus **(Figure 1.3 (c)).** Combination of small pocket A and B and part of C terminus make a complex which is known as large pocket domain (Harbour, 1998). Pocket domain A and B of pRB protein acting as the binding site for other proteins in transcription process (Chow and Dean, 1996; Chellappan *et al.*, 1991). Frequent mutations have been identified within this region that is associated with non-functioning of pRB (Braggio *et al.*, 2004; Lohmann, 1999). N and C termini were less studied, hence their function is not clearly known. Some studies found interaction of pRB binding protein through this residue (Inoue *et al.*, 1995). A study by Xu *et al.*, (1994) suggested that N-terminus of pRB protein may play a role in regulation of cell growth or differentiation. The C-terminus is part of large pocket domain, which contains some important domains for protein-protein interaction (Shew *et al.*, 1990). For example c-Abl tyrosin kinase which might have the potential to regulate the transcriptional activity in S-phase of cell cyle (Welch and Wang, 1993).

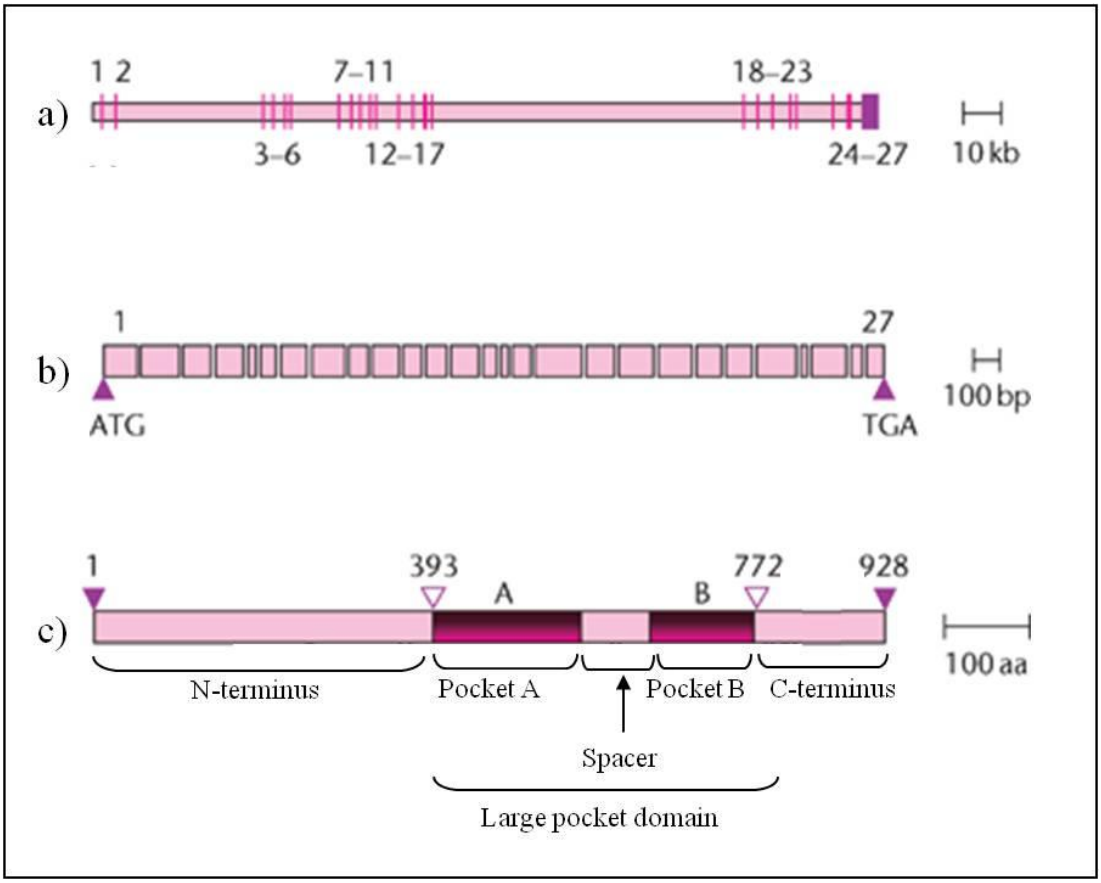

Figure 1.3: Organisation of the *RB1* gene and protein, pRB. (a) Genomic organisation of the *RB1* gene. (b) Organisation of the 27 exons containing the open reading frame. (c) Structure of the pRb protein. kb-kilobase pair; bp-base pair; aa-amino acid. (Modified from http://www.els.net/WileyCDA/ElsArticle/refId-a0006053.html)

1.4.1 (a) N-terminus

The amino-terminal (N-terminus) region of *RB1* comprised of exon 1 to exon 11. Although the N-terminus is not structurally well defined, yet it is still known to be an important essential part of the protein. Deletion mutants of N-terminus have been found to cause retinoblastoma in human (DiCiommo *et al.*, 2000). There are 6 consensus CDK phosphorylation sites in N-terminus, which may play a role in regulating pRB in the cell cycle (Knudsen and Wang, 1997). In addition, this region contains binding sites for MCM7 (a replication licensing factor), a kinase that phosphorylates both pRB and histone H1 in the G_2/M phase of the cell cycle and other proteins (Sterner *et al.*, 1998). Moreover, mice that have been genetically engineered to lack of pRB die before embryonic day 16 with severe developmental and apoptotic defects (Lee *et al.*, 1992). However, expression of pRB lacking the N-terminus region delayed this embryonic mortality, but did not prevent it (Riley *et al.*, 1997). Similarly, this N-terminus deficient mutant delayed but did not prevent the pituitary tumors seen in mice heterozygous for *RB1* deletion (Riley *et al.*, 1997). These results suggest that the N-terminus region is important but not completely essential for pRB function and tumor suppression.

1.4.1 (b) C-terminus

The carboxy-terminal (C-terminus) region of *RB1* is composed of exon 22 to exon 27. This region has non-specific DNA-binding activity (Wang *et al.*, 1990). However, C-terminus domain was found to play crucial role for growth suppression of retinoblastoma (Qian *et al.*, 1992). In addition, this region was also found to be

important for function and regulation of *RB1* gene (Durfee *et al.*, 1994). The cyclin binding motif of C-terminus that is important for phosphorylation of RB protein (Adams *et al.*, 1999). The C- and N-termini are appear to bind with the pocket domain of RB. This domain is important for phosphorylation of the *RB1* protein, which is a control point for the activation or deactivation of the protein. The C-terminal region also binds the oncoproteins c-Abl and MDM2 (Whitaker *et al.*, 1998). Deletion of exons 24 and 25 within the C-terminus of *RB1* gene was associated with low-penetrance retinoblastoma (Bremner *et al.*, 1997).

1.4.1 (c) Promoter of *RB1*

Inactivation of the *RB1* gene is mostly caused by mutations affecting the coding region and promoter region (Shimizu *et al.*, 1994). Bookstein *et al*., (1990) reported an association between prostate cancer and 103-bp deletion of the promoter region that abrogated RB promoter activity. In addition to genetic changes such as sequence variations or deletions, epigenetic changes such as CpG hypermethylation can inactivate a promoter. Ohtani-Fujita *et al.* (1993) reported 9 unilateral, sporadic retinoblastoma cases associated with hypermethylation in the 5' region of the *RB1*. CpG methylation in the RB promoter inhibits binding of the retinoblastoma binding factor 1 (RBF-1) and the activating transcription factor (ATF)-like factors, that cause reduction in RB promoter activity (Ohtani-Fujita *et al.*, 1993).

1.4.2 pRB and cell cycle

In mammals, the normal process of cell division consists of DNA replication (S phase) and mitosis (M phase), where a cell divides into two daughter cells. After mitosis, the diploid cell (2N) enters G_1 stage, proceeded to S phase in which the DNA replication occurs. During this time, the diploid cell becomes tetraploid (4N). There is a gap between S and M phases known as G_2. The cells in G_1 stage will enter a quiescent state called G_0 if there is absence of appropriate growth factor, without entering the rest of cell-cycle division.

pRB is a tumor suppressor protein that controls the G_1/S cell-cycle checkpoint. pRB protein is found in the nuclei of all cell types and at all stages of the cell cycle. Several studies found that phosphorylation of pRB occur at G_0 phase of the cell cycle (Chen *et al.*, 1989; DeCaprio *et al.*, 1992; Ludlow *et al.*, 1990). Nonphosphorylated form of pRB has the binding ability to transcription factor E2F and forms pRB/E2F complex. This complex blocks the transition of normal cells from G_1 into S stage in the cycle (Ludlow *et al.*, 1990). Growth factors stimulate the cells to enter G_1 and approaches S phase. Throughout the G_1 phase, pRB is phosphorylated by the CDK4/cyclin D1 complex. Phosphorylated pRB releases its bound regulatory protein. When E2F and other regulators are released by pRB, they can induce the transition of cells from G_1 to S phase. After cells passed by S, G_2 and M phases, pRB return back to its nonphosphorylated state and binds to regulatory protein such as E2F, and keeps them apart, until they required for the next cell cycle. The presence of pRB protein prevents the transition into S phase in normal quiescent cells. In

many cancer cells including retinoblastoma, both copies of *RB1* can be defective, inactive, or absent that cause unregulated cell cycle (Klug, 2010).

pRB can bind with several viral oncoproteins such as adenovirus E1A protein, polyoma large T antigens, and human papillomavirus (HPV) E7. However, these viruses may use a similar mechanism in transformation, and this may cause inactivation of retinoblastoma. The inactivation of pRB may be due to the binding with these viral oncoproteins. The binding may disrupts the pRB/E2F complexes and leads to the activation of E2F-responsive genes that cause progression of the cell-cycle from G1 to S phase (Sellers and Kaelin, 1997).

1.5 Spectrum of mutations in *RB1*

Mutation is defined as changes in the sequence of genomic DNA. It may comprise a single base-pair substitution, deletion or insertion of one or more base pairs to major alteration in the structure of a chromosome. Mutations may occur within regions of a gene that code for protein or within noncoding regions of a gene, such as introns and regulatory sequences. Mutations may or may not bring changes in phenotype. The level of a mutation can change the characteristics of an organism depends on where the mutation occurs, and the degree to which mutation alters the function of the gene product.

The severity of mutations depends on the site of mutation; in somatic cells or germ cells. Germ cell mutations are heritable and become the basis for the transmission of genetic diversity and evolution, as well as genetic diseases. Somatic cell mutations

are usually non-transmittable to the next generation, but may lead to altered cellular functions of tumors (Klug *et al*., 2010).

Geneticists usually classify the gene mutations in terms of nucleotide changes. A change of one base pair to another base pair in DNA molecule is known as point mutation, or base substitution. A change of one nucleotide of a triplet within a protein-coding portion of a gene may result in a creation of a new triplet which codes for different amino acid in protein product; known as a missense mutation. However, if the triplet amino acid changed into a stop codon, the translation of the protein will terminate, and this is known as nonsense mutation. If the point mutation alters a codon but does not result in a change in the amino acid at that position in the protein (due to degeneracy of genetic code), it can be considered as a silent mutation.

Another type of change is the insertion or deletion of one or more nucleotides at any point of the gene. The loss or addition of a single nucleotide causes the triplet codon to be changed. These are called frameshift mutations because the frame of triplet reading during translation has altered. A frameshift mutation will occur when any number of bases are added or deleted, except the triplet, which would develop the initial frame for reading. It is possible that one of the many altered triplets will be UAA, UAG or UGA, the translation termination codons. When one of these triplets is encountered during translation, the synthesis of polypeptide will be terminated at that point (Klug *et al.,* 2010).

Inactivation of *RB1* is associated with retinoblastoma. Loss of function of *RB1* is also found in other tumor types such as lung carcinoma, non-small cell lung cancer and

esophageal cancer, based on loss of heterozygosity (LOH) and absence of *RB1* expression in these tumors (Xing *et al.*, 1999; Tamura *et al.*, 1997). Low expression of *RB1* was also found with loss of function mutations or epigenetic phenomenon, where hypermethylations occur within 5' promoter region (Lohmann, 1999).

RB1 mutations are found scattered along the whole sequence of genomic DNA and accumulated at discrete region with high recurrence mutations. Majority of the mutations affect coding sequence with higher number at globular domain A and B that interrupt binding function with *RB1*-associated protein (Lohmann, 1999; Valverde *et al.*, 2005).

1.5.1 Chromosomal aberration

Based on karyotyping, previous studies showed that less than 10% of patients with RB had constitutional chromosome 13q abnormalities, and majority with deletions (Bunin *et al.*, 1989a). Cytogenetic or chromosome aberration is disruption of chromosomal content in a cell producing either abnormal number or structure of chromosomes. Other karyotpye abnormalities including mosaic and nonmosaic deletions, *de novo* translocation, deletion and insertion involving 13q14 in peripheral blood lymphocytes of 7.5% bilateral RB and 4.9% among unilateral RB (Bunin *et al.*, 1989b). High frequency of chromosomal aberration was found in Indian RB patients, where abnormal karyotypes were seen within all tumor samples. Abnormalities were seen in all chromosomes except sex chromosome with half (50%) of abnormalities seen in chromosome 13. Most frequent changes are structural aberration such as 13q14 deletions (Harini *et al.*, 2001), translocations and

aneusomies. The study had suggested that locus of 13q14 as fragile site that might play role in retinoblastoma carcinogenesis (Amare Kadam *et al.*, 2004).

1.5.2 Small-length mutation

A review by Lohmann (1999) found that most of the mutations in *RB1* gene were dominated by small length mutation such as single base substitution, small insertion and deletion (Lohmann, 1999). The most common single base substitution is nonsense mutations, with small frequency of missense mutations (Lohmann, 1999; Brichard *et al.*, 2006; Abouzeid *et al.*, 2007). A C to T transition involving CGA-arginine codon is the most common nonsense mutations (Lohmann *et al.*, 1996). It is generally assumed that hypermutability at CGA codon are dependent on methylation status at CpG island and deamination of 5-methyl cytosine to thymidine (Mancini *et al.*, 1997). In addition to nonsense mutations, frameshift and splicing mutations are also found to scatter along the genomic sequence of *RB1* (Valverde *et al.*, 2005). Frameshift and splicing mutation exert deleterious effect due to alteration of final product of pRB protein. (Tsai *et al.*, 2004) demonstrated that *RB1* frameshift mutation demonstrated causes earlier stop codon within exon 19 and producing truncated pRB protein. *RB1* frameshift distributed unequally within exonic and intronic region of *RB1* (Valverde *et al.*, 2005; Braggio *et al.*, 2004). In contrast, splicing mutation commonly occur in adjacent intronic within intron-exon junction or splicing site of *RB1* (Valverde *et al.*, 2005). There is uneven distribution of small mutation at splice site of open reading frame of *RB1* in both acceptor and donor sites. Most frequent splice site mutation affected intron-exon 19 and 24 (Lohmann, 1999). Researchers found that splice site mutation leads to exon skipping and producing

defective protein with lack of amino acid sequence (Nichols *et al.*, 2005; Valverde *et al.*, 2005).

1.5.3 Hypermethylation of CpG Island

CpG island is a short stretch of DNA with higher frequency of CG sequence compared to other regions. CpG islands are usually located around the promoters of housekeeping genes or other genes frequently expressed in a cell. CG is normally not methylated at these locations. By contrast, the CG sequences in inactive genes are usually methylated to suppress their expression. Methylation of cytosine (C) nucleotide converts the cytosine to 5-methylcytosine. Therefore, the transcription machinery fails to recognize 5-methylcytosine, halt the transcription and turn off the gene.

Hypermethylation at CpG island of 5'-promoter region was found to decrease the expression of *RB1* (Kishi *et al.*, 2005). Hypermethylation of *RB1* in tumor tissues is believed to cause allelic inactivation of the gene (Sakai *et al.*, 1991; Ohtani-Fujita *et al.*, 1993; Ohtani-Fujita *et al.*, 1997). Specific hypermethylation at promoter region of *RB1* reduced expression of the gene. Hypermethylation also was found at 5'-end of the gene without any chromosomal aberrations. This provides evidence to support the hypothesis that CpG methylation of the human tumor-suppressor gene causes the inactivation of the gene and lead to tumor formation (Ohtani-Fujita *et al.*, 1997). In addition, hypermethylation at CpG island have been suggested to interrupt the binding site of transcription factor (ATF)-like factor and retinoblastoma binding

factor 1 (RBF-1) that are important for activities of promoter (Ohtani-Fujita *et al.*, 1993).

1.5.4 Single nucleotide polymorphism

Single nucleotide polymorphism (SNP) is genomic DNA variation that occurs when single nucleotides A, T, C or G differs within population of organism. SNPs are the most common type of genetic variation among people. A variation can be considered as SNP if it presents in at least 1% within population in a time (Carlson, 2008). SNPs normally occur throughout individual's DNA. SNPs can act as biological markers that can help scientists to locate genes that are associated with disease. SNPs can be nucleotide substitution, insertion or deletion. It may present in coding sequence, noncoding sequence, and promoter region. SNPs within coding sequence that change amino acid may cause deleterious effect (Lohmann, 1999). The importance of SNPs are spread in a broad range of cases since it can show up to 80% differences between two individual (Carlson, 2008). Certain SNPs may help to predict individual's response to certain drugs, susceptibility to environmental factors such as toxins, and risk of developing particular diseases (Collins *et al.*, 1998). SNPs can also be used to track the inheritance of disease genes within families (Carlson, 2008). Although 99.9% of genetic profile in all human are similar but variation in 0.1% of the human genome have high impact to each individual in term of response to different diseases, environment or drug treatment (Orphanides and Kimber, 2003). Almost 90% of variation in human genetics is SNP (Collins *et al.*, 1998). Therefore, SNPs are valuable for medical diagnostic development especially for pharmacogenetics study. Common SNPs in *RB1* such as IVS3+45C>T (rs520342), IVS4+22G>T (rs198617)

and IVS4-77G>A (rs198616) were found to significantly associate with various diseases such as ovarian cancer, breast cancer, bladder cancer and retinoblastoma (Song *et al.*, 2006; Lesueur *et al.*, 2006; Majewski *et al.*, 2008; Kim *et al.*, 2006; Lohmann, 1999). Most of reported SNPs in *RB1* were found in intronic region (Leone *et al.*, 2003; Lohmann, 1999; Schubert and Hansen, 1996).

1.6 Mutation detection technique

Various techniques for DNA mutational analysis are available (Cotton, 1997). The example of mutation detection technique is denaturing gradient gel electrophoresis (DGGE). This method was developed based on the principles that different sequences of double-stranded DNA (dsDNA) melt under different condition. However, this method was formatted for manual use only, technically challenging, limited by the sensitivity of detection as well as time consuming (Xiao and Oefner, 2001; Vidal-Puig and Moller, 1994; Couch and Weber, 1996).

Another method is single strand conformational polymorphism (SSCP). This method used the principle of difference conformation of secondary structure from single stranded (ss) DNA under certain condition. The secondary structure of single strand DNA is corresponding to the base composition in the DNA sequence. Any nucleotide changes, even one nucleotide may caused the differences of the structure, thus interrupt the electrophoretic mobility under non-denaturing condition. Under optimal condition, with sequence region <200bp, SSCP can detect up to 80% to 90% of all mutations in gene (Sheffield *et al.*, 1993). However, SSCP method cannot be successfully applied on long stretch of DNA.