GENOTYPE-PHENOTYPE STUDY OF TUBEROUS SCLEROSIS COMPLEX IN SELECTED COHORT OF MALAYSIAN PATIENTS WITH *TSC2* MUTATIONS

NUR FARRAH DILA BINTI ISMAIL

UNIVERSITI SAINS MALAYSIA

2015

GENOTYPE-PHENOTYPE STUDY OF TUBEROUS SCLEROSIS COMPLEX IN SELECTED COHORT OF MALAYSIAN PATIENTS WITH *TSC2* MUTATIONS

By

NUR FARRAH DILA BINTI ISMAIL

Thesis submitted in fulfilment of the requirement

for the degree of

Master of Science (Human Genetics)

Universiti Sains Malaysia

OKTOBER 2015

ACKNOWLEDGEMENT

First and foremost, I am grateful to God, the Most Generous and the Most Merciful, upon the completion of my thesis, Alhamdulillah.

I would like to express my sincere gratitude to my wise supervisor, Dr. Teguh Haryo Sasongko for his continuous support, his patience and encouragement, his guidance and immense knowledge throughout my study years. He provided the best supervision for me in the time of my research activities as well as thesis writing. I would also like to express my gratitude for both my co-supervisors, Professor Dr. Narazah Yusoff and Dr. Salmi Abdul Razak for their generous attention and time for my study.

I thank my fellow labmates in Human Genome Center, especially Nik Mohd Ariff Nik Abdul Malik, Jaafar Mohseni, Rani Abdulqawee, Fatemeh Hayati, Roslina Rashid, Marjanu Hikmah Elias, Au Zhian Liang, and Sathiya Maran for their time in helping me in the time of my difficulties. Both of their academic and moral supports in all these years are very much appreciated. Special thanks to my friends, Fathiyyah Iberahim, Nur Haifa Ganti, Nadiah Zakariah, Rozalina Roslan, Nazirah Said and Nur Anisah Nur Ihsan for their endless supports.

I would also like to thank the patients and their family for their willingness in participating in this study, not forgetting the hospitals involved; Hospital Universiti Sains Malaysia, Hospital Raja Perempuan Zainab II, Hospital Kuala Lumpur, Hospital Pulau Pinang, Hospital Sultanah Bahiyah Alor Setar, Hospital Sultanah Aminah Johor Bahru; as well as the clinicians; Dr. Salmi Abdul Razak, Professor Dr. Zabidi Azhar Hussin, Dr. Ngu Lock Hock, Dr, Keng Wee Teik and Dr, Silawati

Abdul Rashid. Others who were directly or indirectly involved in this study, whose names were not mentioned, are very much appreciated.

A special acknowledgement is dedicated to my family, especially my father, Ismail Bin Ahmad and my mother, Rohani Binti Romli. I am very grateful for their faith, patience and undying support throughout my years in Universiti Sains Malaysia.

Last but not least, I would like to express my deepest gratitude for Universiti Sains Malaysia for awarding me USM Fellowship. This research project was also financially supported by USM RU Grant (USM RU Grants No.1001/PPSP/812048) and Center for Neuroscience Services & Research USM (Grants No.304.CNEURO.652205.K134). Without these great financial supports, it would be impossible to conduct and complete this study.

NUR FARRAH DILA ISMAIL, P-UM0070/11(R)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS, ACRONYMS AND SYMBOLS

%	percent
(°C)	degree Celsius
µl	microlitre
µM	microMolar
4E (eIF4E)	eukaryotic translation initiation factor
4EBP-1	factor 4E binding protein-1
AMP	adenosine monophosphate
AMPK	AMP-activated kinase
ATP	adenosine triphosphate
bp	base pairs
CCD1	coiled coil domain 1 (CCD1)
CCD2	coiled coil domain 2 (CCD2)
CDK	cyclin-dependant kinases
DDGE	denaturing gradient gel electrophoresis
ddH2O	double distilled water
DHPLC	Denaturing High-performance Liquid Chromatography
DHPLC	denaturing high-performance liquid chromatography
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphates
dsDNA	double stranded DNA
EDTA	Ethylenediaminetetraacetic acid
ERK/2	extracellular-signal-regulated kinase 2
ERK1	extracellular-signal-regulated kinase 1
ERM	ezrin-radixin-moezin (ERM)
et al.,	et alia/ and others
FISH	fluorescent in situ hybridisation
G	gram
GAP	GTPase-activating protein
GDP	guanosine diphosphate
GTP	Guanosine Triphosphate
kb	kilobase pairs
kDa	kiloDalton

kV	kiloVolt
L	litre
LOVD	Leiden Open Variation Databases
LR-PCR	Long-Range Polymerase Chain Reaction
$MgCl_2$	Magnesium Chloride
ml	millilitre
ml/minute	millilitre per minute
MLPA	multiplex ligation-dependant probe amplification
MLPA	Multiplex ligation-dependent probe amplification
mM	milliMolar
mRNA	messenger Ribonucleic acid
mTOR	mammalian target of rapamycin
mTORC1	mammalian target of rapamycin complex 1
mTORC2	mammalian target of rapamycin complex 2
NCBI	National Center for Biotechnology Information
ng	nanogram
ng/µl	nanogram per microlitre
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
PKB/AKT	protein kinase B
PKD	Polycystic Kidney Disease
REDD	regulation of DNA damage response 1 (REDD1)
RHEB	RAS-homolog Enriched in Brain
RNA	ribonucleic acid
rpm	rotations per minute
RSK1	p90 ribosomal S6 kinase 1
S6KI	phosphorylation of ribosomal S6 kinase (S6K1)
SEGA	sub-ependymal giant cell astrocytomas
SEN	sub-ependymal nodules (SEN)
SSCP	single-strand conformational polymorphism
TSC	Tuberous Sclerosis Complex
U/µl	unit per microlitre
UV	ultraviolet

KAJIAN GENOTIP-FENOTIP BAGI *TUBEROUS SCLEROSIS COMPLEX* DALAM KOHORT TERPILIH PESAKIT-PESAKIT MALAYSIA YANG MEMPUNYAI MUTASI PADA GEN *TSC2*

ABSTRAK

Tuberous Sclerosis Complex (TSC) merupakan penyakit dominan autosomal yang dapat dikenalpasti melalui kehadiran hamartomas pada bahagian organ-organ yang terjejas. Ia merupakan sindrom yang disebabkan oleh mutasi pada salah satu gen, *TSC1* atau *TSC2*. Di sini, analisis mutasi dan hubungkait genotip-fenotip telah dilakukan ke atas 30 orang pesakit yang telah dikenalpasti sebagai pengidap TSC berdasarkan kriteria klinikal 2012 (Northrup et al. 2013). Manifestasi klinikal para pesakit ini pelbagai dan tumor kulit dan otak merupakan manifestasi yang paling kerap dijumpai dalam kalangan pesakit. Epilepsi juga merupakan kelaziman dan ianya lebih banyak berlaku pada pesakit lelaki berbanding perempuan manakala pesakit yang mengalami rencatan akal adalah kurang. TSC tidak mempunyai pemilihan terhadap umur, bangsa mahupun jantina. Ianya dapat diperhatikan bahawa bilangan manifestasi TSC dalam kalangan pesakit familial adalah kurang berbanding pesakit sporadik. Walau bagaimanapun, tiada perbezaan dapat diperhatikan dari segi tahap keterukan penyakit tersebut. Kaedah –kaedah yang digunakan dalam kajian ini adalah kromatografi cecair berprestasi tinggi (DHPLC), penjujukan langsung DNA, *multiplex ligation-dependent probe amplification* (MLPA) dan penjujukan amplikon menggunakan pelantar *Miseq*. Mutasi dapat dikenalpasti dalam 22 daripada 30 orang pesakit TSC. Dua puluh dua orang (73%) pesakit dikenalpasti mempunyai mutasi dalam gen *TSC2* manakala lapan (27%) pesakit tidak mempunyai sebarang mutasi pada gen *TSC2*. Terdapat 20 mutasi patogenik berbeza telah ditemui dan sepuluh merupakan mutasi baru. Tiga puluh peratus dalah mutasi karut, 25% adalah mutasi

missens, 25% adalah penambahan dan penghapusan kecil yang menyebabkan mutasi bingkai-ubah, 15% adalah penghapusan besar gen dan 5% adalah mutasi pada tapak sambat. MLPA disarankan sebagai kaedah pengesanan mutasi barisan pertama diikuti dengan penjujukan amplikon menggunakan pelantar *Miseq* Illumina. Tiada mutasi-mutasi tertentu yang dikenalpasti mempengaruhi tahap keterukan dan/atau bilangan manifestasi klinikal TSC. Walau bagaimanapun, penyakit buah pinggang polisistik (PKD) telah dikenalpasti dalam pesakit TSC yang mempunyai mutasi penghapusan besar gen yang melibatkan gen *TSC2* dan berterusan sehingga ke gen *PKD1*. Manakala lebih ramai pesakit rabdomioma jantung yang dikenalpasti mempunyai mutasi dalam exon 33-41 pada gen *TSC2*. Oleh kerana bilangan kes yang kecil, manifestasi klinikal dalam kumpulan pesakit yang tidak mempunyai sebarang mutasi didapati tidak menunjukkan perbezaan ketara jika dibandingkan dengan pesakit yang dikenalpasti mempunyai mutasi pada gen *TSC2*.

GENOTYPE-PHENOTYPE STUDY OF TUBEROUS SCLEROSIS COMPLEX IN SELECTED COHORT OF MALAYSIAN PATIENTS WITH *TSC2* MUTATIONS

ABSTRACT

TSC (Tuberous Sclerosis Complex) is an autosomal dominant disorder characterized by a widespread hamartomatous lesion in multiple affected organs. It is a syndrome caused by mutations in either of these two genes, *TSC1* and *TSC2*. Here, mutation analysis as well as genotype-phenotype correlation assessment were done in 37 TSC patients. Thirty-seven patients, diagnosed as a case of TSC (either definite or possible) based on the 2012 clinical diagnostic criteria (Northrup et al, 2013) were included in the studies. TSC clinical manifestations among patients were broad and the most common were skin and brain tumours. Epilepsy was also common and was seen more in male compared to female patients while frequency of mental retardation is low. There is no age, ethnicity and gender preference of TSC manifestations. It is noticeable that familial patients showed less number of clinical features compared to sporadic patients although no difference in the severity of the manifestations was observed. The method of choice used were denaturing high-performance liquid chromatography (DHPLC), direct sequencing, multiplex ligation-dependent probe amplification (MLPA) and Amplicon Sequencing using MiSeq Platform. *TSC2* mutations were identified in 22 (73%) of 30 TSC patients while eight (27%) were identified with no mutation. Out of 20 different pathogenic mutations, ten were novel. 30% is nonsense mutations, 25% is missense mutations, 25% is small insertion/deletion causing frameshift mutations, 15% is large deletions and 5% is splice site error mutation. MLPA was suggested as the first line detection method for TSC targeting large duplication and deletion mutations. The second line of mutation detection is Illumina MiSeq Amplicon Sequencing platform for detection of small

mutations. No particular mutations were found to influence severity and/or more number of clinical manifestations. However, polycystic kidney disease was identified in one case with extended deletion from *TSC2* to *PKD1* while cardiac rhabdomyoma are found more in patients with mutations in exon 33-41 of *TSC2* gene. Due to small number of study subjects, the clinical manifestations of the group of patients without identifiable mutation were not much different from the group of patients with identifiable mutations.

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction on Tuberous Sclerosis Complex

Tuberous Sclerosis Complex (TSC) is an autosomal dominant disorder characterized by a widespread hamartomatous lesion in multiple affected organs. Hamartomatous lesions or hamartomas are tumour-like growths and they are also referred as tubers. Commonly affected organs include brain, kidney, heart and skin. The genetic multi-system disorder manifests a broad phenotypic spectrum which includes facial angiofibromas, hypomelanotic macules, renal angiomyolipomas, cardiac rhabdomyomas, cortical tubers, sub-ependymal giant cell astrocytomas (SEGAs) and sub-ependymal nodules (SEN) (Roach *et al.*, 1992).

In 1835, TSC was first found in a male patient who had multiple dots of numerous small erythematous papules on his face, depicting the facial angiofibromas (Rayer, 1835), one of well-known TSC clinical manifestations revised nowadays. Later on, cardiac myomas and cerebral sclerosis were found in a newborn baby who died a few minutes after birth (von Recklinghausen, 1862). Only in 1880, TSC was given its name by Bourneville who had provided a detailed description of the cerebral pathology of TSC which he had seen in a patient with seizures, hemiplegia, mental disorders and renal cysts (Bourneville, 1880). Later on in 1913, the hereditary nature of TSC was reported (Berg, 1913). Since then, numerous studies have been carried out and have contributed to more detailed description and characterization of the disease.

1.2 Diagnostic Criteria

TSC has no pathognomonic feature which has made it hard to be diagnosed. The clinical manifestations and findings of TSC are also highly variable. Some patients may suffer from severe symptoms while other patients may have milder manifestations or can even appear asymptomatic even between the closely related family members. Even so, it is crucial to accurately diagnose the disease for the sake of implementation of appropriate medical attention and treatment as well as genetic counselling for patients.

Historically, the first attempt on guideline for diagnostic criteria of TSC was made by Campbell in 1906 and Vogt in 1908. They proposed diagnostic triad of epilepsy, mental retardation and adenoma sebaceum. However, it was too basic that they would miss out half of the real number of people with TSC. Since then, a few more revisions were made including revision in 1979 by Gomez and another attempt by Roach in 1992. Only in 1998, Roach managed to put up systematic and comprehensive diagnostic criteria which consisted of a set of major and minor features. These features can be clinically diagnosed based on combination of clinical, radiological and histopathological findings. It was used since then for diagnosis of TSC patients until more recent revision was done in 2012 which has made clinical diagnostic criteria of TSC to become even more stringent than before (Northrup *et al.*, 2013).

There are slight differences between the revision made in 1998 and 2012. In 2012, they added other diagnostic criteria by means of genetic testing which was not included in previous TSC diagnosis in 1998. It is the most significant change

recommended to the diagnosis criteria. On the genetic testing basis, definite TSC diagnosis can be made by the identification of pathogenic mutation either in *TSC1* or *TSC2* gene in DNA from normal tissue. A pathogenic mutation is defined as a mutation that causes the inactivation of the *TSC1* or *TSC2* protein and hence contributes to the disease development. *TSC1* or *TSC2* non-pathogenic and undecided genetic variations are not included as definite TSC diagnosis (Northrup *et al.*, 2013).

There are also a few differences in the major and minor features. Most of the major features were refined and put into more detail. The revision in 2012 was made focusing on the sensitivity and specificity of TSC diagnosis based on clinical presentations of a patient. Some features were rarely identified in TSC patients and lack of specificity for TSC. Such features like bone cysts and hamartomatous rectal polyps were removed from the diagnosis in the 2012 revision. Cerebral white matter radial migration line was included into the major features under "cortical dysplasia" with cortical tubers because both are commonly associated with intractable seizures and mental disability in TSC. The comparisons of the two revisions are shown in Table 1.1.

Table 1.1: Comparison of revision in 1998 by Roach and revision in 2012 by Northrup.

	Roach 1998	**Northrup 2013**
Diagnosis		
Definite	Presence of either two major features or one major feature with two or more minor features	Presence of two major features or one major feature with two or more minor features Addition: identification of pathogenic mutation either in *TSC1* or *TSC2* gene in DNA from normal tissue
Probable	Presence of one major feature and one minor feature	Removed
Possible	Presence of either one major feature or two minor features or more minor features	Presence of either one major feature or two or more minor features
Clinical manifestations		
Major features	Hypomelanotic macules (three or more) Facial angiofibromas Forehead plaque Non-traumatic ungula or periungual fibroma Shagreen patch Multiple retinal nodular hamartomas Cortical tuber SEN SEGA Cardiac rhabdomyoma, single or multiple Lymphangiomyomatosis Renal angiomyolipoma	Hypomelanotic macules (3 or more, at least 5 mm in diameter) Angiofibromas (3 or more) or fibrous cephalic plaque Ungual fibromas (2 or more) Shagreen patch Multiple retinal hamartomas Cortical dysplasias (includes tubers and cerebral white radial migration lines) SEN SEGA Cardiac rhabdomyoma Lymphangioleiomyomatosis (LAM)(without other features does not meet criteria for definite diagnosis) Angiomyolipomas (2 or more) (without other features does not meet criteria for definite diagnosis)
Minor features	Multiple randomly distributed pits in dental enamel Hamartomatous rectal polyps Bone cysts Cerebral white matter radial "migration tracts" Gingival fibromas Non-renal hamartomas Retinal acromic patch Confetti-like skin lesion Multiple renal cysts	Confetti-like skin lesions Dental enamel pits (3 or more) Intraoral fibromas (2 or more) Retinal achromic patch Multiple renal cysts Non-renal hamartomas

1.3 Molecular pathogenesis of Tuberous Sclerosis Complex

1.3.1 Knudson's two-hit theory

The pathogenesis of TSC tumours can be explained by Knudson's two-hit hypothesis. The hypothesis mentions that two mutations or 'hits' is necessary for the tumour development. The hypothesis predicted that the chance is greater for a carrier of the germ line mutation to get the second hit as compared to non-carrier to get the same two hits in the same location (Knudson, 1971). Pathogenic germline mutation on either one of the TSC tumour suppressor genes may activate the tumour progression caused by the second random somatic mutation (eg: loss of heterozigosity). Based on Knudson's two-hit hypothesis, the second somatic mutation abrogates TSC tumour suppressor genes function completely by accelerating the effect of the pathogenic germline mutation in the first place.

1.3.2 Hamartin and tuberin functions

The function of hamartin and tuberin alone is still a big mystery. In earlier studies, hamartin and tuberin have been shown to be expressed together in most human cell types and tissues, including in brain, liver, cardiac muscle, kidney, gut, prostate and testes even though they have not been identified to function specifically (Johnson *et al.*, 2001). Most studies have figured out that both hamartin and tuberin play role in cell proliferation, cell growth, cell adhesion, cell migration as well as protein transportation in the cell.

Hamartin, a protein product of *TSC1* gene, does not have similarity with any known vertebrate protein but hamartin does share significant homology to a *Schizosaccharomyces pombe* putative protein (Sampson, 2003). It is widely

expressed and has been identified to have four functioning domains which are a putative transmembrane domain (at amino acids 127-144), coiled-coil domain (CCD) (at amino acids 719-998), amino acid residues as rho-activating domain (amino acids 145-510) and amino acid residues as ezrin-radixin-moezin (ERM) family of actin binding proteins domain (amino acids 881-1084). The coiled-coil domain is necessary for its interaction with tuberin.

Tuberin, the product of *TSC2* gene on the other hand consists of seven functioning domains which are leucine zipper domain (amino acids 81-98), coiled-coil domain 1 (CCD1) (amino acids 346-371), coiled-coil domain 2 (CCD2) (amino acids 1008-1021), transcription activation domain 1 (amino acids 1163-1259), GTPase-activating protein domain (GAP) (1517-1674), transcription activation domain 2 (1690-1744) and a calmodulin-binding domain (amino acids 1740-1755) (Povey *et al.,* 1994; Krymskaya, 2013; Napolioni and Curatolo, 2008).

The binding domains of these two proteins remained unclear until 2003 (Rosner *et al.,* 2003). Individually, hamartin inhibits interaction of tuberin with *HERC1* ubiquitin ligase and stabilizes the protein. The presence of mutation in *TSC2* permits tuberin interaction with *HERC1* ubiquitin ligase even with the presence of hamartin, preventing their interaction to form tumour suppressor complex (Chong-Kopera *et al.,* 2006). Hamartin has also been demonstrated to stabilize and improve the expression of tuberin (Benvenuto *et al.,* 2000).

In *Drosophila*, presence of both hamartin and tuberin had been shown to inhibit cell proliferation and eventually cell growth. Abnormal expression of these two proteins

on the other hand, caused cell growth and size increment. Both regulate cell cycle and hamartin has been found to be highly expressed in the G_0 phase of cell cycle. Too high hamartin level in the cell will reduce cell proliferation. This inhibition process can be carried out by hamartin alone without the presence of tuberin (Miloloza *et al.,* 2000). It has also been demonstrated that loss in either hamartin or tuberin shortens the G1 phase of cell cycle, eventually causing the cell cycle to progress into the S phase, mitosis and finally cell proliferation (Potter *et al.,* 2001; Gao and Pan, 2001 and Tapon *et al.,* 2001). Cell cycle progression is initiated by cyclin-dependant kinases (CDKs) via protein phosphorylation and the activity of CDKs are controlled by CDK inhibitors. The presence of hamartin and tuberin stabilized the level of these inhibitors, indirectly causing inhibition of cell cycle progression (Miloloza *et al.,* 2002).

Hamartin was found to organise cytoskeleton by its interaction with radixin and meiosin. These proteins of ERM families link the actin cytoskeleton to the plasma membrane and take part in forming adhesion contacts, lamellipodia and microvilli. Hamartin also showed interaction with intermediate neurofilament-L. In a cultured cell, cells adhesion was no longer seen when hamartin was inhibited. When hamartin was overexpressed, focal adhesions was seen as the result of Rho G-protein activation (Lamb *et al.,* 2000). These findings suggest that it has multiple functions in adhesion and extension of neurons and ganglia process (Haddad *et al.,* 2002).

There is a recent finding on the function of hamartin by Yasui and colleagues in 2007. They identified interaction of NADE, cell death executor that is associated with p75NTR (low-affinity neutrophin receptor p75), with hamartin coiled-coil domain in

a yeast two-hybrid system. Interaction of the two prevents NADE from proteosomal degradation and indirectly induces neuron cells apoptosis. Reduction of NADE was seen due to low expression of hamartin, leading to disregulation of neuronal cell apoptosis. This finding is likely to be the cause of brain pathology in TSC (Yasui *et al.*, 2007).

As for tuberin, it may function in few different signalling pathways due to its interesting feature that contains multiple phosphorylation sites of Serine, Threonine and Tyrosine. A specific association of a protein, 14-3-3, with phosphorylated tuberin has been ruled out. Due to 14-3-3 ability to bind to phosphorylated protein, it has been known to be involved in various biological events. Tuberin contains several putative binding sites for 14-3-3 protein, AKT-phosphorylated dependent sites on Serine residues at position 939, 981 and 1341, and a direct binding site on Serine residues at 1210. Interestingly, 14-3-3 protein only binds to phosphorylated tuberin and not hamartin. Phosphorylated tuberin bind to 14-3-3 protein thus indirectly regulates cell growth by AKT phosphorylation inhibition. They found that overexpression of 14-3-3 protein increased phosporylation of S6K1 and 4E-BP1, resulting in unusual cell growth (Liu *et al.*, 2002).

Wienecke and Xiao have found that tuberin has specific GAP activity towards Rap1 and Rab5 (Wienecke *et al.*, 1995 and Xiao *et al.*, 2001). Rap1 and Rab5 are a small GTPase cytosolic protein. Rap1 promotes cell division while Rab5 involves in early endocytic pathway. GAP activity of tuberin towards Rap1 may inhibit the GTPase protein thus inactivate cell division. The interaction of tuberin and Rab5 is consistent

with the GAP domain found in tuberin. In cell lacking tuberin, the rate of fluid-phase endocytosis was increased and the process was reversed by re-expression of tuberin.

1.3.3 Hamartin-tuberin tumour suppressor complex

Hamartin's transmembrane bound domain and two coiled-coil domains are also necessary for its interaction with tuberin (van Slegtenhorst *et al.,* 1997 and van Selgtenhorst *et al.,* 1998). There is also a report suggesting phosphorylation side outside the putative interaction domains are also crucial to possibly initiate the interaction of hamartin and tuberin (Aicher *et al.,* 2001).

The actual binding site of hamartin and tuberin was first described in 2003 by Rosner and colleagues. The finding illustrated that amino acids 302-430 of hamartin and amino acids 1-418 of tuberin are the interacting domains between these two gene products (Rosner *et al.,* 2003). Presence of physiological stimuli such as insulin triggers the phosphorylation of tuberin at its multiple sites involving Serine and Tyrosine residues. This is also said to be one of the key mechanisms for its interaction with hamartin thus regulating the formation of tuberin-hamartin protein complex. (Johnson *et al.,* 2001 and Mizuguchi *et al.,* 1997).

The first discovery on direct interaction of hamartin and tuberin with one another was demonstrated in 1998 by van Slegtenhorst and colleagues. In the study, they used three independent methods that showed the same finding of hamartin and tuberin relationship in mammalian cells. The interaction was mediated in between the two coiled-coil domains of these two proteins. They found that these two proteins

co-localized more generally in the cytoplasm when co-transfected in mammalian cells (van Slegtenhorst *et al.*, 1997).

The findings later on leads to abundant evidences that both genes act as tumour suppressor genes and recent investigation of somatic mutation in a variety of TSC hamartomas had strengthened the evidence (Cheadle *et al.*, 2000). Hamartin and tuberin binds directly with one another, forming a cytoplasmic protein complex that has a stable interaction with stoichiometry of 1:1. This tumour suppressor heterodimer has an inhibitory effect on cell growth. The need of both proteins to be working together in regulating cell growth has also been demonstrated in *Drosophila* (Potter *et al.*, 2001). The figurative structure of both hamartin and tuberin are shown in Figure 1.1.

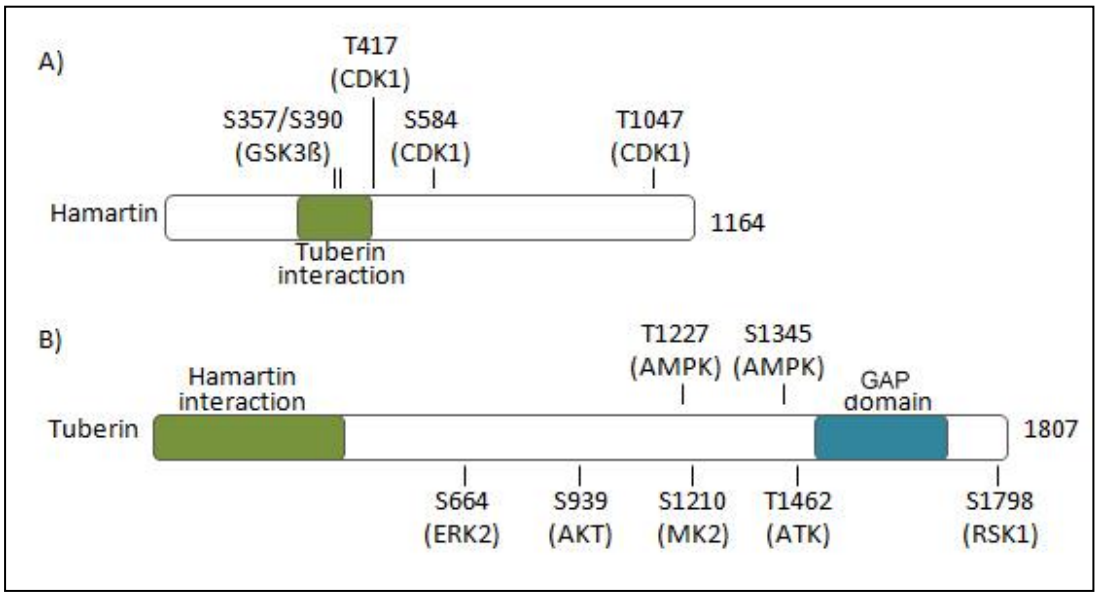

Figure 1.1: Illustrative figure on structure of (A) hamartin and (B) tuberin. Shown together are inhibitory phosphorylation site on both protein as well as its respective kinase. Site where interaction of both proteins takes place is also shown. GAP domain is the domain for RHEB which interacts with mTORC1 pathway (Adapted from Astrinidis and Henske *et al.*, 2005, Crino *et al.*, 2006).

The most exciting and promising discovery is the function of the heterodimer complex formed from the interaction of two proteins as the major regulator for the mammalian target of rapamycin (mTOR). mTOR is a major signalling pathway that monitor the cellular nutrients and energy level of cells. It regulates various cellular processes, including metabolism, growth and proliferation of cells. There are at least two different multi-proteins complexes which are mTORC1 and mTORC2. TSC complex involves mainly in mTORC1 pathways.

The mechanism by which hamartin-tuberin complex regulates mTORC1 pathway lies in the discovery of RHEB (a small G protein of the Ras family, enriched in brain) as physiological target for tuberin GTPase-activating protein activity, *in vitro* and *in vivo* (Zhang *et al.,* 2003). It has been proven that RHEB is important in progression of cell cycle as well as cell growth in *Drosophila* (Plank *et al.,* 1998). RHEB stimulates. TSC complex serves as GTPase-activating protein (GAP) for RHEB and activation of RHEB into its GTP-bound state triggers the interaction with mTORC1 thus stimulating its activity in cells (Long *et al.,* 2005 and Sancak *et al.,* 2007).

The Ras-like GTPase is activated when bound to GTP. In the presence of hamartin-tuberin intracellular complex, GTP on RHEB is hydrolised to GDP, inactivating the RHEB thus downregulates the mTOR and the downstream pathways. Therefore, malfunction of TSC complex due to inactivating mutations will contribute to TSC development.

There are four key signals for mTORC1 which are growth factors, amino acids, oxygen and energy statuses. However, stimulation of mTORC1 by amino acids is known to be independent of hamartin-tuberin complex. Even in cells lacking hamartin-tuberin complex, stimulation of mTORC1 by amino acids remained possible (Nobukuni *et al.,* 2005). General overview of mTOR pathways is shown in Figure 1.2.

In the presence of growth factor, canonical insulin and Ras signaling pathways are activated, increasing tuberin phosphorylation via three factors which are protein kinase B (PKB, also known as AKT), extracellular-signal-regulated kinase 1/2 (ERK1/2) and p90 ribosomal S6 kinase 1 (RSK1) (Inoki *et al.*, 2002, Potter *et al.,* 2002, Ma *et al.,* 2005, Roux et a., 2004). The phosphorylation on tuberin will inactivate hamartin-tuberin complex thus leading to activation of mTORC1.

In cells with low energy status, AMP-activated kinase (AMPK), a key sensor of intracellular energy status, is activated and tuberin is phosphorylated, increasing its GAP activity towards RHEB thus leading to inactivation of mTORC1 (Inoki *et al.,* 2003).

In cells that lack oxygen, AMPK is activated by reduction in the level of ATP. AMPK activation leads to hamartin-tuberin complex stimulation thus inhibiting mTORC1 (Arsham *et al.,* 2003, Liu *et al.,* 2006). In another mechanism, hamartin-tuberin complex is activated through transcriptional regulation of DNA damage response 1 (REDD1) which releases tuberin from its association with 14-3-3 protein

thus inhibits mTORC1 (Brugarolas *et al.,* 2004, Reiling and Hafen, 2004 and DeYoung *et al.,* 2008).

In protein synthesis, the stimulation triggers the phosphorylation of ribosomal S6 kinase (S6K1) that activates ribosomal subunit protein S6 which leads to the recruitment of ribosome. mTORC1 phosphorylates factor 4E binding protein-1 (4EBP-1) causing it to permit the activity of eukaryotic translation initiation factor 4E (eIF4E). These activations will increase protein synthesis and eventually cell growth (Jozwiak et al, 2005; Fingar *et al.,* 2002 and Gingras *et al.,* 1998).

However, constant phosphorylation of S6K1 and 4EBP-1 of mTOR will cause unregulated cell growth and possible tumour formation or progression. This is when hamartin-tuberin protein complex roles into picture. In studies which investigate the function of hamartin-tuberin complex, findings demonstrated increased level of phosphorylated S6K1 and 4EBP-1 in cells that carry abnormal gene of either *TSC1* or *TSC2* (Kwiatkowski *et al.,* 2003 and Tee *et al.,* 2002).

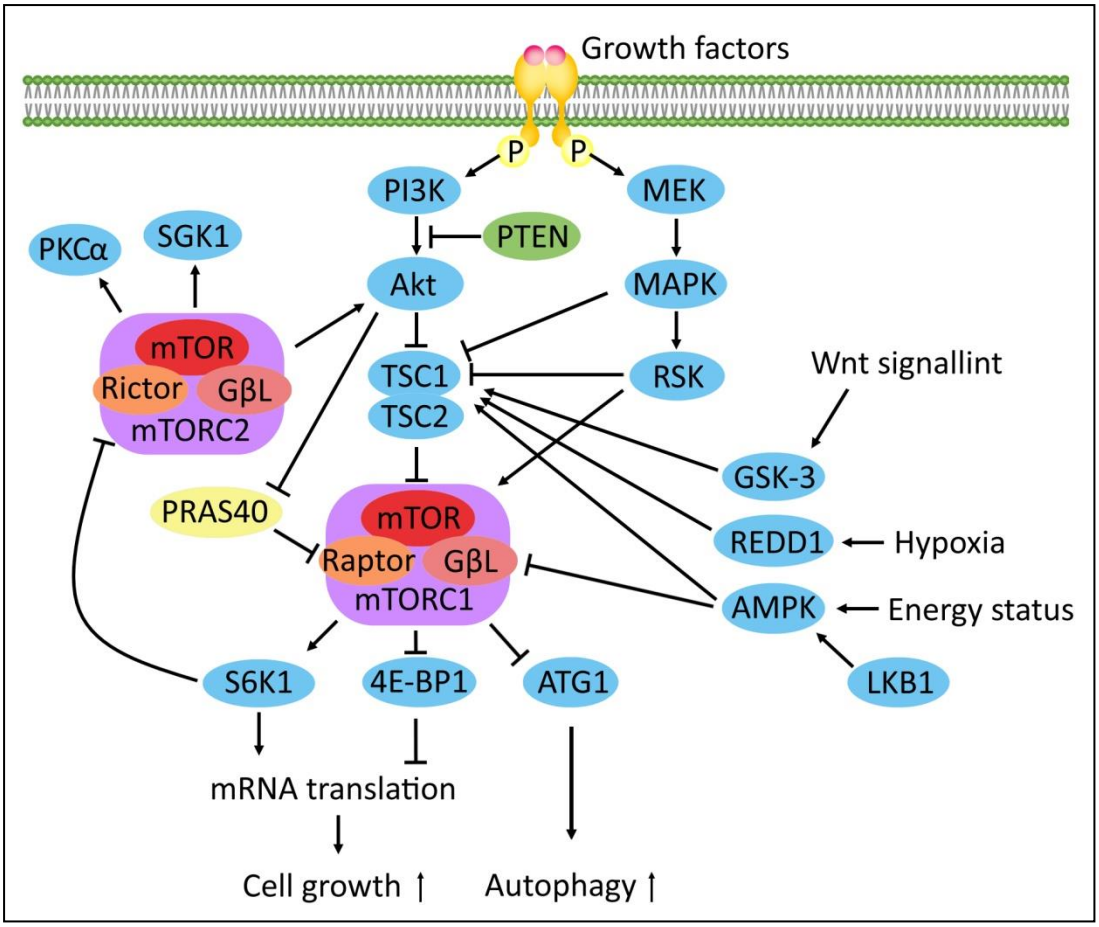

Figure 1.2: General overview of mTOR pathways. There are two distinct complexes which are mTORC1 and mTORC2. *TSC1-TSC2* complex involves mainly in mTORC1 pathway (Adapted from Stern, 2010).

Loss of *TSC1* or *TSC2* had been demonstrated in cell lines and human or mouse tumours to indirectly increase cells vulnerability towards apoptosis. The loss caused stress to endoplasmic reticulum and activated the unfolded protein response which then mediated the negative feedback of the insulin signalling and apoptosis through the mTOR pathway (Ozcan *et al.*, 2008).

1.4 Genetic basis of Tuberous Sclerosis Complex

1.4.1 *TSC1* and *TSC2* genes

75-85% of the TSC cases are caused by abnormalities in either *TSC1* gene or *TSC2* gene which has been identified to be disease causing and were linked to two different loci. These two genes have been identified by positional cloning and have been studied by multigenerational linkage analysis (Curatolo *et al.*, 2003; Gomez *et al.*, 1999; Kandt *et al.*, 1992 and Fryer *et al.*, 1987).

The other gene, *TSC2*, is located on chromosome 16p13.3 and consist of 41 exons which encodes a 5.5 kb mRNA. It encompasses 40 kb of genomic DNA. The gene encodes a different protein named tuberin (1198 kDa, 1807 amino acids) (Eur Chr 16 TS Cons, 1993). Exons 25, 26 and 31 have been found to be the alternate splice site for tuberin isoform (Cheadle *et al.*, 2000).

1.4.2 Mutations Spectrum

There are no mutation hot spots for both genes. The mutations happened at random location. Up to 90% of them are small mutations involving one to several nucleotides while the rest 10% are gross changes in the genes. Mutations that have been reported to be found in TSC are various, including deletion, insertion, frame-shift, missense

and splice-site. According to the Human Gene Mutation Database (HGMD), up until now there are more than 700 unique mutations have been reported in *TSC2* gene (Stenson *et al.*, 2014).

The most common type of mutation is point mutation which contributes nearly 90% all mutations. Point mutation is defined as alteration in only single nucleotide that occurs either by deletion, substitution or insertion. Based on the functional characterization, point mutations are classified into nonsense mutation, missense mutation and silent mutation (https://www.genome.gov/glossary).

Nonsense mutation introduces stop codon prematurely into the DNA sequence hence produces a truncated protein of either hamartin or tuberin depending upon the location. Missense mutation codes for different amino acid which gives rise to a slightly different protein with different properties of the hamartin or tuberin. Silent mutation codes for the same amino acid. Therefore, the protein remains unaffected and can retain its function normally. Insertion or deletion of nucleotides which is not in triplets will give more severe effect to the normal protein translation as the result of disturbance in the open reading frame of the gene. This type of mutation is specifically called frameshift mutation since it shifts the reading frame of protein translation hence producing a completely different type of protein from the original one. Insertion or deletion of nucleotide which is divisible by three is called inframe deletion or insertion. A codon may be missing or added but it does not cause shift in the triplet reading thus the protein remains the same with only one amino acid lost or gain (https://www.genome.gov/glossary).

Gross changes of the gene can either be deletion or duplication of nucleotides in larger scale which some time involves one or more exons. There are also cases of inversion mutation which involved approximately 600 kb of nucleotides of *TSC2* exon which disrupted the gene (Sampson *et al.,* 1997). Exonic deletion is more common than exonic duplication. Due to deletion of some of the exons, the hamartin or tuberin protein will be shortened and truncated, losing the capability to carry out their function properly. These types of mutations occur more frequently in *TSC2* compared to *TSC1* (Nellist *et al.,* 2005 and Longa *et al.,* 2001).

A study reported that mutations occurred more in the region of exon 16 of *TSC2* (Au *et al.,* 2007). While in large studies, exon 16 of *TSC2* gene is one of the famous spot along exon 32, exon 39 and exon 40, which has recorded more number of mutations compared to other exons (Sancak *et al.,* 2005 and Dabora *et al.,* 2001).

As mentioned, there is no specific mutation hot-spot for the genes. However, certain type of mutations does occur more frequently in some domain of the gene. According to large studies, GAP domain on *TSC2* gene is the favoured spot for missense type mutations (Maheswar *et al.,* 1997 and Sancak *et al.,* 2005). Notably, in-frame deletion is the most common type of mutation found in this region. However another report showed that only small percentage of the missense mutations found were located at the *TSC2* GAP-domain (Au *et al.,* 2007). These conflicting findings show that TSC mutations are highly variable.

Different studies reported different findings on the most common mutation occurrence. Study by Dabora and colleagues (Dabora *et al.,* 2001) have reported the

most common mutation is c.5238_5255del which occurred on *TSC2* gene exon 40. The mutation has also been reported to be the most common in later study by Au and colleagues (Au *et al.,* 2007) and a number of studies have also found the same mutation in their samples (Wang *et al.,* 2013; Yao *et al.,* 2008; Chong *et al.,* 2006; Rendtorff *et al.,* 2005 and Beauchamp *et al.,* 1998). These six amino acids (18 nucleotides) in-frame deletion has been identified to significantly repress ligand-induced steroid/nuclear receptor-mediated transcription activity that results in uncontrolled cell growth and cell proliferation which gives rise to tissue malformations such as facial angiofibromas (Noonan *et al.,* 2002).

Similar 18-nucleotides deletions have also been found in two different studies that occurred in the same exon 40 and were located closed to one another. One happened at nucleotide 5256 which was reported to be the most common mutation in the study by Jones and colleagues while the other one happened at nucleotide 5227 (Choi *et al.,* 2006; Martin *et al.,* 2003 and Jones *et al.,* 1999). The possible cause of this type of deletion is slipped mispairing during replication (Cooper and Krawczak, 1993). Deletions which were located in the rabaptin binding site of exon 40 may disrupt the endocytic pathway thus contributing to TSC (Xiao *et al.,* 1997). These data has shown that this type of 18 nucleotides in-frame deletion only attacked and is specific to exon 40 or *TSC2* GAP-domain in general.

Another single most common site for mutation reported in the same study by Au and colleagues as well as another large study by Jones and colleagues (1999), is on codon p.R611. There are two mutations that occurred adjacently on this site which are p.R611W (c.1831C>T) and p.R611Q (c.1832G>A). These mutations were likely

caused by spontaneous deamination of methylated cytosines that resulted in substitution at CpG dinucleotides in the gene sequence (Jones *et al.,* 1999). According to Nellist and colleagues, both substitution mutations result in major conformational changes on tuberin thus interrupting its interaction with hamartin (Nellist *et al.,* 2005). The same finding was previously reported in a mutational study by Sancak and colleagues in 2005. It seems that the higher the number of samples used, the higher the probability to discover mutations distribution pattern as well as the most common site of mutation occurrence.

Based on some of these frequently occurring mutation, distinct effects of single amino acid changes to tuberin on the function of the hamartin-tuberin has been studied. It includes the effect of tuberin amino acid changes on the tuberin-hamartin complex, effect on tuberin phosphorylation, effect on tuberin-dependent inhibition of S6K and S6 phosphorylation and effect of tuberin truncation on RHEB GTPase activity *in vitro*. The single amino acid changes studied include p.R367Q, p.N525S, p.K599M, p.A607T, p.609insS, p.R611Q, p.R611W, p.A614D, p.F615S, p.C696Y, p.V769E, p.L826M, p.R905Q, p.P1202H and p.G1556S. Only three of the amino acid changes, p.R367Q, p.A607T and p.L826M did not give any effect on tuberin-hamartin formation, inhibition S6K or S6 phosporylation or the stimulation of RHEB GTPase activity while the rest of the changes appeared pathogenic. p.R611Q, p.R611W, p.A614D, p.C696Y and p.V769E have been identified to inactivate tuberin completely in all of the essays conducted. These substitutions are said to cause major conformational changes to the protein (Nellist *et al.,* 2005).

Nellist and colleagues studied two more changes which are p.609insS and p.F615S and found out that these changes destroyed the interaction of hamartin and tuberin, prevented tuberin phosphorylation by PKB and prevented S6K phosphorylation inhibiton. However, the tuberin inactivation by these changes was incomplete and they still have some RHEB GAP activity along partial inhibition of S6 phosphorylation. Changes denoted as p.R905Q, p.P1202H, p.G1556S, p.N525S and p.K599M which were also studied did not show any pathogenic effects (Nellist *et al.,* 2005).

Some studies have also reported a small portion of somatic mosaicism in TSC cases. Somatic mosaicism happens when some of the cells in the body carry the TSC mutation while some other cells carry the normal TSC gene. A large study by Kozlowski and colleagues in 2006 involved 261 patients has reported eight cases of somatic mosaicism (Kozlowski *et al.,* 2006). In another large study involving 224 patients, two patients have been suspected to have been the case of somatic mosaicism (Dabora *et al.,* 2001).

Another common phenomenon in TSC cases is loss of heterozygosity. Every somatic cell has two normal copies of TSC genes. The genes are said to be in heterozygous state when one copy of the allele has already been affected (either due to inherited or *de novo* mutation), leaving only one functional copy of the gene. In this state, the remaining normal copy of the gene is capable of compensating the loss of the mutant copy of the gene. Loss of heterozygosity comes into picture when the only normal copy of the gene is also affected, losing both normal copies, thus resulting in the

inactivation of the gene. Without the presence of the normal gene, TSC disease will develop.

A number of studies have shown that frequency of loss of heterozygosity is high in TSC tubers, especially in renal angiomyolipomas than neurological tumours (Niida *et al.,* 2001; Au *et al.,* 1999 and Henske *et al.,* 1997). One possible reason is that inactivation of both copies of the gene may not be necessary for the pathogenesis of some TSC tumours (Tucker and Friedman, 2002). Loss of heterozygosity is also evidence that *TSC1* and *TSC2* work together as a tumour suppressor gene where loss of heterozygosity can be on either one gene for the disease to be developed.

Despite the numerous mutations found, all studies have reported that there are a small portion of the clinically diagnosed TSC patients that showed negative result for the mutational analysis. These patients with no identifiable mutation generally developed milder symptom compared to patients that have been detected to have mutation either in *TSC1* or *TSC2* gene (Dabora et al., 2001, Sancak et al., 2005 and Au et al., 2008).

1.5 Epidemiology of Tuberous Sclerosis Complex

The prevalence of TSC was reported to be one case in approximately 6000 births (Osborne *et al.,* 1991). Worldwide, TSC affects about one to two million people while in United States, it affects as many as 25 000 to 40 000 people. TSC has no ethnic, gender and race preference (National Institute of Neurological Disorders and Stroke, 2014). Up to two-thirds of the cases are sporadic while the rest of the cases are familial. Being autosomal dominant in nature, only one affected parent with

mutant gene is needed to pass down TSC to an offspring with 50% chance of developing TSC. However, the offspring may not present the same manifestations as the parent even with the same type of mutation.

Many cases remain undiagnosed in previous years. However, more cases even with milder manifestations are now being diagnosed attributed to significant progress in TSC researches as well as development of TSC diagnosis technologies. The prevalence was expected to go higher in near future.

In many developing countries like in Malaysia, data on the prevalence of TSC is absent. However, there has been no indication that TSC occur more frequently in certain geographic locations over the other, nor within certain ethnicities over the other. It is therefore reasonable to suggest that TSC may occur in the same prevalence rate as that reported above.

1.5.1 Tuberous Sclerosis Complex and ethnicity

Mutational studies have been carried out in European, American and Asian counterparts (Taiwanese, Chinese Han, Japanese and Korean). All the studies have reported almost similar mutations distribution among the populations involved in the studies. Some population such as Korean represented lower mutations distribution compared to Western countries and Japan but no correlation between ethnicity and higher frequency of TSC mutations was observed (Choi *et al.*, 2006).

1.5.2 Tuberous Sclerosis Complex and gender

The chance for male TSC patient to have mental retardation is higher compared to female patient. The male patients are likely to have renal cysts, retinal and skin lesions more. The observable different manifestations seen in male and female patients are best explained by the different hormones produced by male and female body. Sex hormones have been illustrated to influence the progression of TSC manifestations in human as well as in animal models (Sancak *et al.,* 2005).

1.5.3 Familial versus sporadic cases

One of the aspects to be looked at is the comparison between familial cases and sporadic cases. Clinical manifestations in familial cases were reported to be lower and milder than in the sporadic cases. Skin manifestations, mental retardation and renal cysts are significantly more common in sporadic cases (Au *et al.,* 2007; Choi *et al.,* 2006 and Jones *et al.,* 1999). Seizures and onset age of seizures, SEGAs, and number of cortical tubers were also low in familial cases (Choi *et al.,* 2006). A large study by Sancak and colleagues has also reported the same findings and in addition, they found a significantly higher occurrence of SEN and retinal phakomas. They also found out that hypomelanotic macules was significantly higher in familial cases (Sancak *et al.,* 2005).

In familial cases, the affected parent (either mother or father) of the affected patient are highly likely to have the same mutation in the gene though the clinical manifestations might be different. There was a case with the same mutation found in monozygotic twins but only one of them was diagnosed to have shagreen patch and cardiac rhabdomyoma (Martin *et al.,* 2003). Another study by Sasongko and