2019

Luiz Gustavo Dufner de Almeida

Mutational and functional study of Tuberous Sclerosis Complex genes 1 and 2 (*TSC1* and *TSC2*)

Estudo mutacional e funcional dos genes 1 e 2 do Complexo da Esclerose Tuberosa (TSC1 e TSC2)

SÃO PAULO

2019

Luiz Gustavo Dufner de Almeida

Mutational and functional study of Tuberous Sclerosis Complex genes 1 and 2 (*TSC1* and *TSC2*)

Estudo mutacional e funcional dos genes 1 e 2 do Complexo da Esclerose Tuberosa (TSC1 e TSC2)

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo como requisito parcial para obtenção do título de Doutor na área de Biologia/Genética

Orientadora: Profa. Dra. Luciana Amaral Haddad

SÃO PAULO

2019

Ficha catalográfica elaborada pelo Serviço de Biblioteca do Instituto de Biociências da USP,
com os dados fornecidos pelo autor no formulário:
http://www.ib.usp.br/biblioteca/ficha-catalografica/ficha.php

Dufner-Almeida, Luiz Gustavo
Mutational and functional study of Tuberous Sclerosis Complex genes 1 and 2 (TSC1 and TSC2) / Luiz Gustavo Dufner-Almeida; orientadora Luciana Amaral Haddad. -- São Paulo, 2019.

154 f.

Tese (Doutorado) - Instituto de Biociências da Universidade de São Paulo, Departamento de Genética e Biologia Evolutiva.

1. Tuberous Sclerosis Complex. 2. TSC1 gene. 3. TSC2 gene. 4. functional study. 5. transcriptome. I. Haddad, Luciana Amaral , orient. II. Título.

Bibliotecária responsável pela estrutura da catalogação da publicação:
Elisabete da Cruz Neves - CRB - 8/6228

_____ _____

Prof(a) Dr(a) Prof(a) Dr(a)

_____ _____

Prof(a) Dr(a) Prof(a) Dr(a)

_____

Prof(a) Dr(a) Luciana Amaral Haddad

Advisor

Ao meu pai e minha mãe, Luiz Sergio de Almeida e Rosely Christina Dufner de Almeida;

meu irmão, Rafael Augusto Dufner de Almeida;

minha querida esposa, Karina Campos Tisovec Dufner, e

minha filha Sofia Tisovec de Almeida.

"Eu não tenho ídolos. Tenho admiração por trabalho, dedicação e competência"

Ayrton Senna da Silva

This work was carried out with financial support from the Foundation for Fundação de Amparo à Pesquisa do Estado de São Paulo / Centro de Pesquisa, Inovação e Difusão coordinated by Prof. Dra Mayana Zatz (FAPESP 13 / 08028-1), the do Conselho de Aperfeiçoamento de Pessoal de Nível Superior - Programa de Excelência Acadêmica (CAPES-Proex, 001) and Programa de Doutorado-sanduíche no Exterior (CAPES - PDSE, 88881.132401 / 2016-01) to the student.

Acknowledgments

I would like to thank Prof. Dr. Luciana Amaral Haddad, for the encouragement, conviviality and learning during the eight years from the orientation of my master's degree starting in 2012 until the conclusion of my doctorate.

I also would like to thank Dr. Mark Nellist for having me at his laboratory at the Functional Unity on Erasmus Medical Center, Rotterdam, The Netherlands, for enabling the development of the functional assays of the variants found in the TSC patients and the RNAseq analysis of the *TSC1* and *TSC2* knockout cell lines in his laboratory.

To Profa. Dr. Maria Rita Passos Bueno of the Department of Genetics and Evolutionary Biology, Institute of Biosciences-USP for collaboration with the next generation sequencing of *TSC1* and *TSC2* genes and funding through FAPESP/CEPID.

I also thank Dr. Cibele Masotti of the Sírio Libanês, Institute of Education and Research for the analysis of RNAseq data through the RPKM data.

Many thanks to Prof. Dr. Sérgio A. Antoniuk, Dra. Ana Paula A. de Pereira, Dra. Laís F. M. Cardozo, Dra. Mariana R. Schwind and Dra. Danielly S. C. Cândido of the Clinical Hospital Complex of the Federal University of Paraná, for attending and in collecting blood samples from patients with TSC. To Prof. Sérgio Rosemberg, Dra. Juliana Paula and co-workers, thank you for receiving patients with TSC, and the collaboration with our study. From GRAACC-SP, thank Dra. Nasjla Saba and her team for the care of patients with TSC and the collection and sending of blood samples for our research.

Finally, thank the Msc. Thiago Alegria from the Department of Genetics and Evolutionary Biology, Institute of Biosciences-USP and Santoescha Nahoe from the Functional Unit of Erasmus Medical Center, Rotterdam, The Netherlands, for helping me in experimental designs and bench work during the final stages of my doctorate.

Abstract

Tuberous sclerosis complex (TSC) is an autosomal dominant disorder caused by pathogenic variants in either *TSC1* or *TSC2* tumor suppressor genes. It affects more often the brain, skin, kidneys, heart, lungs, and retina. The protein products of both genes, TSC1 (hamartin) and TSC2 (tuberin), interact, assembling a complex that inhibits mTORC1. Cells with bi-allelic inactivation of either *TSC1* or *TSC2* genes present hyperactivation of mTORC1, which phosphorylates downstream targets, up-regulating cell proliferation and growth. Moreover, a functional role as heat-shock protein (HSP) co-chaperone has been assigned to TSC1 protein. The first aim of the thesis was to analyze the nature, distribution and functional effects of *TSC1* and *TSC2* DNA variants from 100 patients with definite clinical diagnosis of TSC. We analyzed leukocyte DNA of 115 TSC patients from three Brazilian tertiary referral hospitals. Pathogenic DNA variants were detected in 99 (86,09%) unrelated individuals; 17 (17,17%) in *TSC1* and 82 (82,82%) in *TSC2*. Clear loss-of-function mutations were detected in 87 patients, of which frameshift (29.29%) and nonsense (29.29%) variants were the most common types. In-frame deletions, missense and putative splicing DNA variants with uncertain clinical significance (VUS) have been functionally assessed. Five variants significantly increased phosphorylation of the reporter residue S6K Thr^{389}. Forty-one novel pathogenic DNA variants and 19 novel single nucleotide variants have been detected. Among the 11 individuals with no mutation identified, seven presented rare putative missense, splicing or in-frame deletion DNA VUS. To understand the regulatory relationship of *TSC1/2* gene expression, we aimed to evaluate *TSC1* and *TSC2* mRNA and protein levels in human cell lines with bi-allelic inactivation of each gene. We employed high throughput transcriptome analysis (RNA-Seq) and Western blotting of HEK293T and other six HEK293T-derived cell lines that had the genomic sequence of the *TSC1* and or *TSC2* genes edited by the CRISPR (clustered regularly interspaced short palindromic repeats)-CAS9 system. In lack of either TSC1 or TSC2 protein, a significant reduction of the respective mRNA was observed, inferring no positive transcriptional feedback. Serum-deprived cell lines without TSC1 decreased *TSC2* mRNA levels. Under these conditions, *TSC1* mRNA levels were not negatively affected by the lack of TSC2. In one cell line with loss of TSC1 (1C2) *TSC1/2* mRNA and TSC2 protein levels were consistently decreased independently on serum. RNA-Seq gene ontology analyses comparing 1C2 to HEK293T reference cell line disclosed down-regulation of translational pathways independently on serum; and up-regulation of protein folding and stability pathways upon serum withdrawal. Our data are consistent with the role of TSC1 as HSP co-chaperone, and suggest that *TSC1* mRNA may be regulated at both transcriptional and decay levels.

Resumo

O complexo de esclerose tuberosa (TSC) é um distúrbio autossômico dominante causado por variantes patogênicas em um de dois genes supressores de tumor *TSC1* ou *TSC2*. Afeta mais frequentemente o cérebro, a pele, os rins, o coração, os pulmões e a retina. Os produtos proteicos de ambos os genes, TSC1 (hamartina) e TSC2 (tuberina), interagem formando um complexo que inibe o mTORC1. As células com inactivação bi-alélica dos genes *TSC1* ou *TSC2* apresentam hiperactivação de mTORC1, que fosforila alvos a jusante, regulando positivamente a proliferação e o crescimento celular. Além disso, o papel funcional da co-chaperona da proteína de choque térmico (HSP) foi atribuído à proteína TSC1. O primeiro objetivo dessa tese foi analisar a natureza, distribuição e os efeitos funcionais das variantes de DNA de *TSC1* e *TSC2* de 100 pacientes com diagnóstico clínico definitivo de TSC. Analisamos o DNA de leucócitos de 115 pacientes com TSC de três hospitais brasileiros de referência. Variantes de DNA patogênico foram detectadas em 99 (86,09%) indivíduos não relacionados; 17 (17,17%) em *TSC1* e 82 (82,82%) em *TSC2*. Mutações de perda de função foram detectadas em 87 pacientes, dos quais as variantes frameshift (29,29%) e nonsense (29,29%) foram os tipos mais comuns. Deleções in-frame, variantes missense e variantes de splicing com significância clínica incerta (VUS) foram avaliadas funcionalmente. Cinco variantes aumentaram significativamente a fosforilação do resíduo repórter S6K Thr^{389}. Quarenta e uma novas variantes de DNA patogênico e 19 novas variantes de nucleotídeo único foram detectadas. Entre os 11 indivíduos sem mutação identificada, sete apresentaram variantes raras missense, splicing ou deleções in-frame do tipo VUS. Para entender a relação regulatória da expressão do gene *TSC1*/2, tivemos com segundo objetivo avaliar os níveis de mRNA e proteína de TSC1 e TSC2 em linhagens de células humanas com inativação bi-alélica de cada gene. Empregamos uma análise de transcriptoma de alto rendimento (RNA-Seq) e Western blotting de HEK293T e outras seis linhagens celulares derivadas de HEK293T que possuíam a sequência genomica dos genes *TSC1* e/ou *TSC2* editados por CRISPR-CAS9. Na falta da proteína TSC1 ou TSC2, foi observada uma redução significativa do respectivo mRNA, inferindo ausência de feedback transcricional positivo. Linhagens celulares privadas de soro $TSC1^{-/-}$ diminuíram os níveis de mRNA de TSC2. Sob estas condições, os níveis de mRNA de TSC1 não foram afetados negativamente pela falta de *TSC2*. Numa linhagem celular com a perda de TSC1 (1C2), os RNAm de TSC1/2 e os níveis de proteína de TSC2 foram consistentemente diminuídos independentemente no soro. Análise de RNA-Seq comparando a linhagens celular de 1C2 com a referência HEK293T revelou uma regulação negativa de vias de tradução independentemente no soro; e supra-regulação das vias de enrolamento e estabilidade das proteínas após a retirada do soro. Nossos dados são consistentes com o papel do TSC1 como co-chaperona de HSP, e sugerem que o mRNA de *TSC1* pode ser regulador nos níveis de transcrição.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

CHAPTER 1

List of Acronyms

ADPKD: Autosomal dominant polycystic kidney disease

AML: Angiomyolipomas

eIF4E: Eukaryotic translation initiation factor 4E

GAP: GTPase-activating protein

HSP: Heat-shock proteins

LAM: Lymphangioleiomyomatosis

MLPA: Multiplex ligation-dependent probe amplification

MMPH: Multifocal, micronodular, type II pneumocyte hyperplasia

mTORC1: Mechanistic target of rapamycin complex 1

PECs: Perivascular epithelioid cells

PKD1: Polycystic Kidney Disease 1 gene

RHEB: RAS-homology enriched in brain

S6K1: p70 40S ribosomal protein S6 kinase B1

SEGA: Subependymal giant cell astrocytoma

SEN: Subependymal nodule

TSC: Tuberous sclerosis complex

TSC1: Tuberous sclerosis complex 1 gene

TSC2: Tuberous sclerosis complex 2 gene

UTR: Untranslated region

List of Figures

List of Tables

Summary

I. Background

Tuberous sclerosis complex (TSC) is a clinical disorder characterized as hereditary, multiple system hamartomatosis. TSC hamartomas can affect any organ, although most commonly the brain, heart, kidneys, lungs, retina and skin. It has an estimated incidence between 1:10,000 and 1:6,000 live born infants, and no ethnic bias (Sampson, Scahill *et al.* 1989, O'Callaghan, Shiell *et al.* 1998, Au, Williams *et al.* 2007).

The term hamartoma, suggested by Albrecht in 1904, defines slow-growing benign tumors of cells that normally occur in the tissue of origin, arranged in a disorganized, though circumscribed, non-encapsulated manner (Cooper 1971). TSC specific organ hamartomas classify in infrequent histological types, and in general have onset in a defined developmental period. Clinical manifestations usually occur due to organ dysfunction as a result of tissue architecture rupture or by compressive or vascular phenomena(Roach and Sparagana 2004, Franz, Bissler *et al.* 2010, Grajkowska, Kotulska *et al.* 2010, Krueger, Northrup *et al.* 2013, Northrup, Krueger *et al.* 2013). Although neoplasms and metastases are not TSC hallmarks, they may uncommonly occur at earlier age than in the general population, in specific organs, notably the kidneys (van Slegtenhorst, de Hoogt *et al.* 1997, Telfeian, Judkins *et al.* 2004).

Two characteristics are frequently shared by TSC hamartomas: the increase in size of the affected cells, which are commonly described as giant cells; and the challenge to classify the cell type by immuno-labeling with specific antibodies (Huttenlocher and Wollmann 1991). Hence, it has been proposed that TSC hamartomas stem from deficiencies in cell size, proliferation and differentiation control.

A. TSC genetics

TSC is an autosomal dominant disorder caused by mutation in either the *TSC1* or *TSC2* tumor suppressor genes (Consortium 1993, van Slegtenhorst, de Hoogt *et al.* 1997). Molecular studies of TSC hamartomas resected from mouse models of the disease or from patient surgery specimens have mostly agreed with Knudson's second-hit hypothesis (Niida, Stemmer-Rachamimov *et al.* 2001), which has been established as a genetic model for tumor suppressor genes (Knudson 1971). It currently predicts that a somatic cell harboring a germline mutation will present tumor clonal properties if the wild-type allele is inactivated by a pathogenic DNA variant or epigenetic alteration. The randomness and uncertainty of the second allele-inactivating hit is believed to contribute to decrease the penetrance of the disease. TSC incomplete penetrance has been reported (Connor, Stephenson *et al.* 1986, Webb and Osborne 1991). However, recent technological developments have estimated the penetrance of this disease to be, though incomplete, significantly higher than previously observed (Hasbani and Crino 2018). As presented below, TSC clinical diagnosis is based on lesion identification and not on symptomatic clinical presentation (Northrup et al., 2013). Therefore, the increase in population coverage of medical imaging has allowed for the definite clinical diagnosis in

asymptomatic patients. Moreover, molecular analysis breakthroughs and the availability of parental DNA testing have contributed to increase the estimates of the disease penetrance rates (Caylor, Grote *et al.* 2018). On the other hand, factors that may negatively influence the penetrance of TSC are its clinical heterogeneity, later age at onset for specific organ hamartomas of the adulthood, and *TSC1* and *TSC2* genotype-phenotype correlations (Northrup, Wheless *et al.* 1993, Camposano, Greenberg *et al.* 2009, von Ranke, Faria *et al.* 2017). Remarkably, nearly 70% of TSC cases have been reported as sporadic, due to a novel mutation that may be in somatic mosaicism. Consequently, inconclusive diagnosis may implicate in difficulty to reconcile clinical data and mutation segregation studies in families. Finally, familial cases tend to have higher penetrance (Kwiatkowski 1994).

B. Historical aspects of TSC diagnosis

TSC was first described in 1862 by Friedrick D. von Recklinghausen after the autopsy of a newborn patient disclosed cardiac tumors and malformations of the brain surface. A similar case was later reported in 1864 by Rudolph Virchow. However, it was only in 1880 that the French physician, Désiré Magloire Bourneville, coined the name of the disease as tuberous sclerosis of the cerebral convolutions upon description of autopsied cerebro-cortical malformations that presented tuber-like morphology and consistency firmer than the adjacent parenchyma. The patient had epilepsy, as well as facial and renal lesions (Bourneville 1880, Gomez, Sampson *et al.* 1999, Jansen, van Nieuwenhuizen *et al.* 2004) (Table 1).

Taken together, these descriptions and four other clinical cases reported until 1890 (Table 1) present clinic-pathological aspects of a disorder that appeared (i) to be systemic, affecting the brain, skin, heart and kidneys; (ii) to be presented as a variable combination of signs, not necessarily observed in all patients; and (iii) to have cortical tubers as the commonest lesion. TSC is also classified as a phacomatosis (from the Greek word *phakos*, birthmark), a clearly congenital condition associated with a pathological birthmark(McCall, Chin *et al.* 2006), or as neurocutaneuos syndrome, an inherited condition with major involvement of both, brain and skin.

Table 1: First clinical cases reported with signs of TSC in the 1860s to 1880s.

Physician	Year	Cardiac tumor	Cortical tubers	Periventricular nodules	Angiofibroma	Fibrotic plaques	Kidney injuries
Friedrick D. von Recklinghausen	**1862**	√	√				
Rudolph Virchow	**1864**	√	√				
Désiré Magloire Bourneville	**1880**		√		√		√
Hartdegen	**1881**		√	√			
Bourneville and Brissaud	**1881**	√	√	√			√
François H. Hallopeau and Émile Leredde	**1885**		√		√		
John J. Pringle	**1890**		√		√	√	

√: Observed phenotype

In 1908, Vogt proposed a clinical triad consisting of sebaceous adenoma (angiofibroma), intellectual disability and epilepsy for the diagnosis of TSC. The development and improvement of medical imaging techniques expanded the knowledge on the disease, leading to the first classification of lesions into major and minor criteria for TSC diagnosis by Gómez in 1979. TSC diagnosis criteria were later reviewed twice by medical experts (Roach, Gomez *et al.* 1998, Northrup, Krueger *et al.* 2013) (Roach et al.,1998; Northrup et al., 2013). Those currently adopted are presented in (Northrup, Krueger *et al.* 2013). The combination of major and minor criteria may yield a definite or possible clinical diagnosis. It is important to notice that each lesion defined as a criterion can occur isolate in a patient, due to two somatic mutation hits. Thus, the adherence to these guidelines is of utmost importance in the clinical setting. Moreover, the latest review board has included the genetic diagnostic criteria, by which the detection of a pathogenic variant in either *TSC1* or *TSC2* genes is sufficient for the definite diagnosis of TSC. Genetic variants of *TSC1* or *TSC2* not confirmed as pathogenic are not a criterion for the diagnosis.

Table 2: Revised criteria for the diagnosis of TSC according to the International Conference on Consensus on Tuberous Sclerosis Complex (modified from Northrup et al. (2013)).

A. Genetic diagnostic criteria	
The identification of a pathogenic mutation in the ***TSC1*** or ***TSC2*** gene in normal tissue is enough for the definitive diagnosis of TSC *.	
B. Clinical diagnostic criteria	
Major features	
1.	Hypomelanotic macules (>3, at least 5-mm diameter)
2.	Angiofibromas (≥3) or fibrous cephalic plaque
3.	Ungual fibromas (>2)
4.	Shagreen patch
5.	Multiple retinal hamartomas
6.	Cortical dysplasias**
7.	Subependymal nodules (SEN)
8.	Subependymal giant cell astrocytoma (SEGA)
9.	Cardiac Rabdomioma
10.	Lymphangioleiomyomatosis (LAM)***
11.	Angiomyolipomas (>2)*** (AML)
Minor features	
1.	"Confetti" skin lesions
2.	Dental enamel pits (>3)
3.	Intraoral fibromas (≥2)
4.	Retinal achromic patch
5.	Multiple renal cysts
6.	Nonrenal hamartomas
Definite diagnosis:	Two major features; or one major feature with ≥2 minor features.
Possible diagnosis:	One major feature; or ≥2 minor features.

* A pathogenic mutation is defined as a mutation that inactivates the function of the proteins expressed by the *TSC1* or *TSC2*, prevents protein synthesis, **or is**. a missense mutation whose effect on protein function has been demonstrated by a functional assay (www.lovd.nl/TSC1, www.lovd.nl/TSC2).
** Cortical dysplasia includes cortical tubers and or radial migration lines in cerebral white matter.
*** A combination of the two main criteria - LAM and angiomyolipomas - in the absence of another criterion does not fit the definitive diagnosis of TSC, being considered as a single main criterion, due to the co-occurrence of the two lesions in at least 30% of patients with isolated form of LAM[1].

C. TSC lesion classification according to the cell of origin: clinical manifestations and outcomes

TSC lesions can be classified in different ways; according to organ or system affected, histological diagnosis subtyping, following clinical manifestations, etc. Here we divide TSC into five large groups according to the pathogenesis and cell of origin.

1. CNS lesions: TSC as a neurodevelopmental disorder

Nearly 80% of patients with TSC will present subependymal nodules (SEN, Houser and Gomez (1992), 5-mm nodules in the subependymal layer of the lateral cerebral ventricles protruding into their cavity, with a smooth, firm and hardened surface, due to frequent calcification. They are lined by a relatively intact ependymal epithelium. Giant cells, with large, convoluted nuclei tend to be observed inside the nodule. They are highly vascular lesions, whose vessels are prone to hyalinization (Trombley and Mirra 1981, Scheithauer and Reagan 1999). Internal hemorrhage and necrosis are rare but may be observed in larger nodules. In 5% to

15% of patients with TSC, SENs may progress to subependymal giant cell astrocytoma (SEGA), especially in the first two decades of life, a hypothesis supported by the follow-up imaging tests, histological similarities between these hamartomas and their encephalic location (Roach and Sparagana 2004, Grajkowska, Kotulska *et al.* 2010). The majority of SEGAs are solitary mass, discrete, larger than 5 mm in diameter and located near the foramen of Monro.

SENs and SEGAs are hamartomas formed by dysplastic astrocytes and neuronal cells (Gomez, Sampson *et al.* 1999, Jozwiak, Jozwiak *et al.* 2005). Despite debates on their origins (Davidson, Yoshidome *et al.* 1991, Johnson, Yoshidome *et al.* 1991), and although classified as astrocytomas, it is now recommended that due to the neuroglial composition of SEGAs they should be classified as tumors of mixed origin (Buccoliero, Franchi *et al.* 2009). This composition suggests that SEN and SEGA should originate from neuroglial progenitors of the subventricular zone of the brain (Crino 2013). Although SEGA is a tumor classified as grade I astrocytoma(World Health Organization), it may maintain a slow growth rate, of at most 2 mm per year. This growth and its location may contribute to hydrocephalus and acute intracranial hypertension, due to obstructive phenomena. Vascular calcification and hyalinization contribute to the rare occurrence of intramural hemorrhage. Calcification of the stroma is more common in SEN than in SEGA (Huttenlocher and Wollmann 1991, Shepherd, Scheithauer *et al.* 1991, Stefansson 1991, Scheithauer and Reagan 1999).

Cortical tubers are lesions present at birth in most patients with TSC. However, its observation by magnetic resonance imaging may be impaired in the first two years of life, a period necessary for myelination, which increases the contrast for the identification of these lesions (Houser and Gomez 1992, Shepherd, Houser *et al.* 1995). They consist of circumscribed enlargement of cerebral gyri in the cerebral parenchyma, of either cortical or subcortical location, presenting itself firmer than the adjacent parenchyma; therefore, the term tuberous sclerosis named the disease. Its slight protrusion in the cerebral parenchyma creates tubercle-like appearance (tuber).

Neuronal heterotopias result from dissynchrony in neuronal migration along the radial glial fibers (Roach and Sparagana 2004). They are presented as radial migration lines of the white matter, possibly formed by groups of postmitotic neuron soma that did not complete the radial migration process (Northrup, Krueger *et al.* 2013, van Eeghen, Teran *et al.* 2013). Cortical tubers and neuronal heterotopias are collectively named dysplasias and, since they are composed of non-proliferative cells, they are classified as hamartias, and not hamartomas (Huttenlocher and Wollmann 1991, Scheithauer and Reagan 1999).

Cortical tubers and neuronal heterotopias are associated with epilepsy, often refractory to anti-convulsivant medication. They may be associated with intellectual disability, autism and other types of behavioral disorders. In most patients younger than two years, seizures manifest as infantile spasms, which if not fully controlled, may progress to other types of seizures, and associate with autism and intellectual

disability (Chiron, Dumas *et al.* 1997, Hancock and Osborne 1998, Curatolo, Seri *et al.* 2001, Thiele 2004, Adriaensen, Schaefer-Prokop *et al.* 2009). (Chu-Shore, Major *et al.* 2010). The lack of therapeutic control of epilepsy in the first five years of life (60-80% of patients with TSC - Orlova and Crino (2010)) appears to relate to later severity of intellectual disability (Joinson, O'Callaghan *et al.* 2003, Humphrey, Williams *et al.* 2004). Although a correlation between the number and location of cortical tubers and clinical severity has not been convincingly established, patients with temporal distribution of tubers seem develop more severe epilepsy and behavioral problems (Goodman, Lamm *et al.* 1997, Hosoyaa, Naitoa *et al.* 1999, Holmes, Stafstrom *et al.* 2007). The absence of clinical manifestations due to cortical tubers in the first five years of life is related to a better cognitive and social function in this period and in successive phases of development (Joinson, O'Callaghan *et al.* 2003, Humphrey, Williams *et al.* 2004).

The location of these four brain lesions (SEN, SEGA, neuronal heterotopias and cortical tubers), which are commonly observed in the same patient with TSC, suggests that they have a common origin. It has been demonstrated that SEN and SEGA present cell proliferation and growth control deficiency, and potentially originate after the embryonic corticogenesis period from subventricular zone progenitor cells, in which a second, somatic *TSC1* or *TSC2* mutation is expected. Early in embryonic development (end of the first third of human gestation), post-mitotic neurons generated from neuroglial progenitors start migrating to initiate corticogenesis. Post-mitotic migrating neurons, possibly with cell dysfunction in adhesion and migration processes (Lamb, Roy *et al.* 2000, Goncharova, Goncharov *et al.* 2004) may prematurely abrogate cell migration, and accumulate, forming neuronal heterotopias. Neurons that may reach the appropriate cortical layer could form cortical tubers due to synaptogenesis defects (Johnson, Yoshidome *et al.* 1991, Mizuguchi and Takashima 2001, Nishio, Morioka *et al.* 2001, Crino 2013).

2. PEComas

The World Health Organization histological, including immunohistochemical, definition of perivascular epithelioid cell neoplasms (PECs) is mesenchymal tumors of distinct perivascular epithelioid cells. The family of tumors of the PEComa type is diverse. In patients with TSC, PEComas occur in the form of angiomyolipomas (AML) and pulmonary lymphangioleiomyomatosis (LAM) (Martignoni, Pea *et al.* 2008, Folpe and Kwiatkowski 2010).

PEComas are characterized by their perivascular localization and are often arranged radially around the vascular lumen. Careful analysis of PEComas reveals blood vessels compressed by the tumor wall. They present eosinophilic cytoplasm, clear to transparent, different from smooth muscle cells, which are intensely eosinophilic. Perivascular epithelioid cells can accumulate large amounts of lipids, mimicking adipocytes or lipoblasts, cells normally found distant from the vascular wall (Folpe and Kwiatkowski 2010).

AML is the most common PEComa, with a prevalence of 0.13% in adults, being more frequently diagnosed in women than in men. Renal AMLs are seen in 60% to 80% of adult patients with TSC. Displaying a higher frequency in adulthood, AMLs notably have age-dependent frequency (Rakowski, Winterkorn *et al.* 2006). Despite the high frequency among patients with TSC, about 80% of patients with AML do not have TSC; hence, isolate AML is more common than associated with TSC (Eble 1998).

AMLs are highly vascularized hamartomas of mixed composition, whose cells have characteristics of adipocytes and smooth muscle, also presenting dysmorphic blood vessels. It is now believed that AMLs are derived from perivascular epithelioid cells (Weeks, Malott *et al.* 1991, Ashfaq, Weinberg *et al.* 1993), although the nature of these cells is still a matter of debate, suggestive of pericytes or a cell type originating from the neural crest (Fernandez-Flores 2011).

AMLs are usually expansive, infiltrating masses, which can undergo vascular complications such as microthrombi, embolism and bleeding. These tumors are most commonly seen in the kidneys, and may occur in other organs, such as the liver, adrenal glands, pancreas, and pelvis (Crundwell 2004, Orlova and Crino 2010, Yang, Feng *et al.* 2012, Northrup, Krueger *et al.* 2013). Although a large proportion of patients with AML are asymptomatic (Folpe and Kwiatkowski 2010), there is a risk of local painful bleeding due to rupture of tumor aneurysms or compression and/or invasion of the renal parenchyma. The likelihood of progression to chronic renal failure is high if AML is not treated (Northrup, Krueger *et al.* 2013).

Pulmonary LAM is a progressive and often fatal disease caused by proliferation of perivascular epithelioid cells in both lungs, such as muscle fibers, distributed around lymphatic vessels, bronchi, interlobar septa, and pleura (Smolarek, Wessner *et al.* 1998, Folpe and Kwiatkowski 2010). The diffuse and bilateral distribution of LAM in the lungs; the co-occurrence of AML and isolate LAM; the demonstration of a somatic mutation in the *TSC2* gene in both lesions, AML and LAM, of the same patient; and the identification of circulating cells with the ame aspect of perivascular epithelioid cells found in these hamartomas are all evidences that suggest that LAM results from AML metastasis to the lung (Carsillo, Astrinidis *et al.* 2000).

Non-TSC pulmonary LAM (70%) is more common than LAM in association with TSC (30%) (Smolarek, Wessner *et al.* 1998, Carsillo, Astrinidis *et al.* 2000, Franz, Bissler *et al.* 2010). The first studies indicated that 1% to 5% of patients with TSC had LAM (Castro, Shepherd *et al.* 1995). After the increased use of computed tomography of the thorax, it was revealed that between 25% and 40% of female patients with TSC have LAM (Moss, Avila *et al.* 2001). More recent studies indicate that the frequency of AML in patients with TSC increases with age and can reach 80% of female patients at the end of the fourth decade of life. Patients with AML present progressive dyspnea on exertion and recurrent pneumothorax, typically in the third and fourth decades of life (Northrup, Krueger *et al.* 2013).

3. TSC cutaneous signs

Hypomelanotic macules are frequent signs in patients with TSC, since they occur in 90% of the patients, often at birth, and are thus useful for the clinical diagnosis. At least three hypomelanotic macules are with 5 mm or more in diameter consist a major criterion for the clinical diagnosis (Northrup, Krueger *et al.* 2013).

Facial angiofibromas, characterized by three or more erythematous, painful papules, commonly in the malar processes occur in approximately 75% of patients with TSC. Usually these lesions develop around 2 to 5 years of age, being very evident in the pubertal phase (Northrup, Krueger *et al.* 2013). Over time, these lesions become larger, with no significant increase in the total number, and may extend to the nasolabial folds (Roach and Sparagana 2004).

The ungual fibroids are nodules in the nail bed, usually painful, that can appear upon minor trauma. They have a frequency of 20% in young patients, and higher than 80% in adults with TSC (Roach and Sparagana 2004, Northrup, Krueger *et al.* 2013).

Fibrotic plaques usually occur on the scalp or face in 25% of TSC patients, or in the form of Shagreen plaques in 50% of patients. The latter appear as lumbar plaques with an irregular surface that looks like leather or orange peel, and its clinical presentation is almost always related to TSC (Northrup, Krueger *et al.* 2013).

4. Cardiac Rhabdomyomas

Two-thirds of the fetuses and newborns with TSC have one or more cardiac rhabdomyomas. Although these tumors usually regress in the postnatal period, there may be a second peak of frequency in adolescence. These hamartomas can cause cardiac arrhythmia often in newborns or do not present any clinical sign. Their prenatal or neonatal identification is an early sign of TSC, very useful for the diagnosis as it is highly suggestive of TSC (Roach and Sparagana 2004, Northrup, Krueger *et al.* 2013).

5. Epithelial cells hyperplasia

Hyperplasia is defined as proliferation of cells at rates higher than the average expected for the tissue. In patients with TSC, they occur most commonly in the lungs and kidneys, although they may affect other organs, such as the liver.

Multifocal, micronodular, type II pneumocyte hyperplasia (MMPH) is a pulmonary manifestation in TSC. MMPH is characterized by small nodules (1 mm to 10 mm), with clear margins, on the wall of the alveoli in both lungs, at a frequency between 2% and 3% among TSC patients. MMPH can occur independently of pulmonary LAM (Muir, Leslie *et al.* 1998, Maruyama, Seyama *et al.* 2001, Hayashi, Kumasaka *et al.* 2010).

Although renal epithelial cell hyperplasia as single, unilateral or bilateral cysts may occur in patients with TSC, their isolate presentation is very common in the general population, and increases with aging. On the other hand, 5% of patients with TSC present a genomic deletion of chromosome 16 involving the 3 'portions of the contiguous genes, *TSC2* and *PKD1* (Polycystic Kidney Disease 1 gene). In these patients, with the TSC-ADPKD contiguous deletion syndrome, there is coexistence of TSC with ADPKD, a condition characterized by bilateral development of multiple renal cysts (Bastos and Onuchic 2011).

D. TSC molecular diagnosis and genotype-phenotype correlations

The *TSC1* (NM_000368.4, NG_012386.1) gene maps to 9q34.1, has 23 exons and approximately 55 kb. The coding sequence for the 1,164-amino acid TSC1 protein, also known as hamartin (130 kiloDaltons), lies in exons 3 through 23 (van Slegtenhorst et al., 1997). Exon 23 is the largest (5,407 bp) *TSC1* exon, contributing to the long length of the 3'-untranslated region of its transcript. *TSC1* intron 1 is the longest (9,447 bp) intron, followed by introns 8 (8,905 bp) and 2 (6,080 bp). The average exon size is 374 bp, while the average intron size is 798 bp.

The *TSC2* (NM_000548.3, NG_005895.1) gene maps to 16p13.3, has 42 exons and nearly 46 kb in length. Its coding sequence resides from exon 2 to exon 42, and is responsible for the synthesis of a 1164-amino acid protein (tuberin or TSC2, 180 kiloDaltons). The largest *TSC2* exon is 488-bp long (exon 34), and the largest intron has 4,820 bp (intron 16). The average exon size is 133 bp, while the average intron size is 798 bp (Consortium 1993, van Slegtenhorst, de Hoogt et al. 1997, Jones, Shyamsundar et al. 1999, Dabora, Jozwiak et al. 2001, Sancak, Nellist et al. 2005, Au, Williams et al. 2007).

Since the cloning of the *TSC1* and *TSC2* genes (Consortium 1993, van Slegtenhorst, de Hoogt et al. 1997), TSC mutation detection has been performed by Sanger sequencing of PCR products from different exons of each gene (Wilson et al., 1996; Jones et al., 1997; Au et al., 1998; Dabora et al., 1998; Niida et a., 1999; van Slegtenhorst et al., 1999; Yamashite et al., 2000; Dabora et al., 2001). Techniques such as MLPA (multiplex ligation-dependent probe amplification) and microarray comparative genome hybridization have additionally been used to search for *TSC1* and *TSC2* genomic deletions (Redtorff et al., 2005; Kozlowski et al., 2007; Oyazato et al., 2011; van den Ouwenland et al., 2011), which correspond to approximately 7% of TSC-causing mutations (Kozlowski, Roberts *et al.* 2007). In the last decade, next-generation sequencing (NGS) of leukocyte DNA has been employed to search for *TSC1* and *TSC2* mutations. Most commonly, both gene coding sequences have been captured for NGS library constructions of whole exome or customized multigene panels (Qin, Kozlowski *et al.* 2010; Tyburczy, Wang *et al.* 2013; Tyburczy, 2015a; Tyburczy 2015b; Nellist et al., 2015; Cai et al., 2017; Ismail et al., 2017; Yang et al., 2017; Papadopoulou et al., 2018).

Sanger or next-generation DNA sequence efforts generally detect pathogenic DNA variants in 80% to 95% TSC patients with definite clinical diagnosis. Nearly 70% of all pathogenic variants are *de novo* germline

or post-zygotic alterations. These novel mutations occur approximately four times more often in the *TSC2* than the *TSC1* gene. Among the familial cases, *TSC1* and *TSC2* pathogenic DNA variants have similar frequencies (Wilson, Ramesh *et al.* 1996, Au, Rodriguez *et al.* 1998, Jones, Shyamsundar *et al.* 1999, Niida, Lawrence-Smith *et al.* 1999, van Slegtenhorst, Verhoef *et al.* 1999, Astrinidis, Khare *et al.* 2000, Dabora, Jozwiak *et al.* 2001, Rendtorff, Bjerregaard *et al.* 2005, Sancak, Nellist *et al.* 2005, Hung, Su *et al.* 2006, Au, Williams *et al.* 2007, Kozlowski, Roberts *et al.* 2007, Sasongko, Wataya-Kaneda *et al.* 2008).

A genotype-phenotype correlation is observed in patients with a contiguous *TSC2* and *PKD1* gene deletion (Brook-Carter et al., 1994; Longa et al., 1997). The two genes lie at 16p13.3 in a tail-to-tail fashion, displaying their polyadenylation signals 60 bp apart (Figure 1). The development of polycystic kidney disease in TSC patients is due to gross deletions involving both genes; thus, the phenotype has been termed *TSC2*/*PKD1* contigous gene deletion syndrome (Brook-Carter et al., 1994).

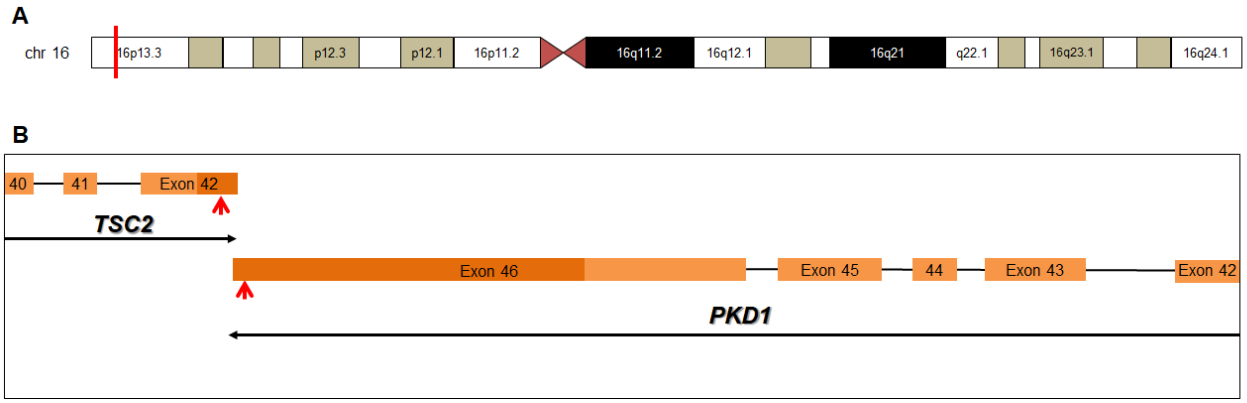


Figure 1: Map of the human chromosome 16p13.3 zoomed into 4,130 bp of *TSC2* and *PKD1* gene 3'-end overlapping region. (**A**) Human chromosome 16. The red bar indicates the location of both genes *TSC2* and *PKD1* on p13.3 band. (**B**) Zoom into 4,130 bp relative to the red bar from (A). The direction of *TSC2* and *PKD1* genes are indicated by thin opposite arrows, as their coding sequences are in different strands of the chromosome 16 DNA. Rectangles represent exons, and lines, introns. Exon numbers from each gene 3'-end are indicated. Dark orange are untranslated region, and light orange rectangles correspond to the coding sequence. Polyadenylation signals (red thick, short arrows) are 60 bp apart. The drawing of the figure was fully based on and in scale to data retrieved from the genome browser at the University of California in Santa Cruz (Santa Cruz, CA, USA; http://genome.ucsc.edu), according to the GRCh38/hg38 human genome version. Accessed in April, 2019.

Although TSC patients without *PKD1* gene alterations may develop renal cysts , the histopathological features appear different in the contiguous gene deletion syndrome (Consugar et al., 2008). Information on kidney imaging data is helpful for genetic service decision on the choice of DNA test, as MLPA analysis has currently a relatively lower cost, and should be sufficient to detect the genetic cause of the contiguous gene deletion syndrome (Consugar et al., 2008).

TSC cases classified as mild phenotypes tend to be diagnosed at an older age, have more frequently normal cognition, less epilepsy episodes, SEN, SEGA, sclerotic bone lesions than patients with *TSC2* pathogenic variants. Additionally, those individuals tend to show more frequent bilateral and larger renal angiomyolipomas than patients with *TSC1* mutations. TSC mild phenotypes are more frequently associated

with patients with no mutation identified (NMI) or with familial cases (Camposano et al., 2009; Boronat et al., 2017; Rosset et al., 2017; Peron et al., 2018).

E. TSC1 and TSC2 proteins

TSC1 (NP_000359.1) protein product of *TSC1*, is a 130-kDa (1164 amino acid) hydrophilic protein with a predicted coiled-coil region of 266 amino acids (residue 730 - 996) close to the C-terminus and an N-terminal domain required for the interaction with TSC2. The TSC1 has the ability to form homodimers, demonstrated by the yeast two-hybrid system. TSC1 homodimers are prone to self-aggregation in detergent resistant complexes in the absence of the TSC2 gene product. In the absence of *TSC2* overexpression, *TSC1* overexpression produces insoluble TSC1, presenting a dotted cytoplasmic distribution at immunofluorescence analysis. If both genes are co-overexpressed, TSC1 becomes soluble and has a more homogeneous cytoplasmic distribution (Nellist, van Slegtenhorst *et al.* 1999).

TSC2 (NP_000539.2) is a 200-kDa (1,807 amino acid) protein with a short 183-residue region close to the C-terminus that is homologous to the GTPase-activating protein (GAP) for the small GTPase RAP1. This C-terminus domain is encoded by exons 34 to 38 (Maheshwar, Cheadle *et al.* 1997) and has been shown to act as a GAP for the GTPase RAS Homolog Enriched in Brain (RHEB), catalyzing the RHEB-dependent hydrolysis of GTP to GDP. The N-terminal domain of TSC2 is required for the interaction with TSC1, enabling the two proteins to form a stable complex (Plank, Yeung *et al.* 1998, van Slegtenhorst, Nellist *et al.* 1998).

Initially, TSC2 was thought to be localized to the Golgi apparatus, while TSC1 was detected in a punctate pattern in the cytoplasm of cultured cells and localized to a membrane fraction (Plank, Yeung *et al.* 1998). However, because TSC1 and TSC2 associate physically *in vivo*, it seemed likely that TSC1 and TSC2 function together in the same complex rather than in separate pathways (van Slegtenhorst, Nellist *et al.* 1998). It had been suggested that TSC2 could act as TSC1 chaperone, preventing its self-aggregation (Nellist, van Slegtenhorst *et al.* 1999) maintaining the TSC1–TSC2 complex in a soluble form. The quaternary structure of the TSC complex is not yet clear, but gel filtration and co-immunoprecipitation experiments suggest that the complex contains multiple subunits of both TSC1 and TSC2 (Hoogeveen-Westerveld, van Unen *et al.* 2012).

1. TSC2 GAP domain and the mechanistic target of rapamycin (mTOR) pathway

TSC1 and TSC2 form a tumor suppressor complex, the TSC complex, which acts as mediator of intracellular and extracellular signals to inhibit the mechanistic target of rapamycin (mTOR, NP_004949.1) complex 1 (mTORC1) kinase (Garami, Zwartkruis *et al.* 2003, Tee, Manning *et al.* 2003). TSC2 acts as GAP toward RHEB (Garami, Zwartkruis *et al.* 2003, Inoki, Li *et al.* 2003). Inactivation of the TSC complex results in increased levels of RHEB-GTP, that is required for mTORC1 to phosphorylate downstream targets, including p70 40S ribosomal protein S6 kinase 1 (S6K1, NP_003152.1) and eukaryotic translation initiation factor 4E

(eIF4E) binding protein 1 (4E-BP1, NP_004086.1). Indeed, co-expression of TSC1 and TSC2 represses S6K1 phosphorylation at Thr^{389}, therefore repressing its activity. mTORC1-dependent phosphorylation of 4E-BP1 causes dissociation of 4E-BP1 from eIF4E, enabling the configuration of the 48S pre-initiation translational complex. Mutants of *TSC2* are defective in repressing phosphorylation of 4E-BP1 (Tee, Fingar *et al.* 2002, Tee, Manning *et al.* 2003). Taken together, the results implicate the TSC complex in the mTORC1-mediated effects on S6, S6K1 and 4E-BP1, leading in the absence of the active TSC complex to cell over-growth and proliferation, frequent aspects of cell dysfunction in TSC patients. Understanding TSC as an mTOR dysfunction has allowed for clinical trials of rapamycin or analogues (rapalogs) to treat SEGA, AML and LAM. Rapamycin or rapalogs have been clinically approved to control tumor size for these indications in Brazil.

The stability of hamartin-tuberin complex is central in the control of mTOR signaling. Understanding the regulatory complex of the TSC1 and TSC2 genes is important and may contribute to the selection of vectors for gene therapies in specific TSC lesions (Prabhakar, Zhang *et al.* 2015). One interesting open question is if the expression regulation of one TSC gene impacts the other TSC gene.

2. Tsc1 as Hsp90 co-chaperone

Misfolded proteins may aggregate due to the exposure of hydrophobic amino acid residues to the cellular environment. Nearly 30% of newly synthesized proteins are targeted for degradation due to improper folding (Schubert, Anton *et al.* 2000). It means that most proteins require assistance to achieve their correct tertiary structure. Among the systems that have evolved to regulate protein turnover, the most well-known are chaperones (Carlisle, Prill *et al.* 2017). To achieve the correct tertiary and quaternary structure, protein folding factors (chaperones) are required for many proteins (Carlisle, Prill *et al.* 2017).

When improper protein folding occurs, the heat shock response is activated, and the synthesis of a class of inducible chaperones known as heat-shock proteins (HSP) is increased (Calderwood and Murshid 2017). The molecular chaperone heat-shock protein HSP 90-alpha (Hsp90, NP_001017963.2) is an essential component of the cellular homeostatic machinery in eukaryotes. *In vitro* and *in vivo,* purified Hsp90 binds to denatured proteins and displays ATP-dependent anti-aggregate properties (Panaretou, Prodromou *et al.* 1998). Most Hsp90 clients are medically relevant, including the nuclear receptors for steroid hormones and several proto-oncogenic kinases and non-kinases.

The ATP-binding pocket in the Hsp90 central domain (residue 272 – 617) is essential for chaperone activity and to bind and hydrolyze ATP, changing the N-terminal “open” state to a “closed” conformation (Prodromou and Pearl 2003). The Aha1 co-chaperone aids in high-energy conformational modulations necessary for Hsp90 ATPase competence (Panaretou, Siligardi *et al.* 2002, Lotz, Lin *et al.* 2003).

Hsp90 (NP_001017963.2) may interact with TSC1 even in the absence of TSC2. The association between TSC2 and TSC1 decreases under thermal shock, which does not affect the association of TSC1-HSP90

(Inoue, Uyama *et al.* 2010). HSP90 is an essential component of the cellular homeostatic machinery in eukaryotes. *In vitro* and *in vivo,* purified Hsp90 binds to denatured proteins and displays ATP-dependent anti-aggregate properties (Panaretou, Prodromou *et al.* 1998).

Tsc1 has a higher binding affinity for Hsp90 than Aha1. The C-terminal domain of Tsc1 (residue 998 - 1,164) forms a homodimer that inhibits ATPase activity and stabilizes the 'closed' conformation of the Hsp90 central domain (residues 284 – 620 and 728 – 732) (Brown, Foreman *et al.* 2015, Woodford, Dunn *et al.* 2016, Woodford, Sager *et al.* 2017). Conversely, phosphorylation of $Aha1^{Y223}$ increases its affinity for Hsp90 and displaces Tsc1. Because Aha1 and Tsc1 compete for the same domain, Tsc1 may prevent the conformational changes in the catalytic loop of Hsp90 that reduces the release of R400 from its retracted, inactivating conformation, consequently inhibiting Hsp90 ATPase activity (Woodford, Sager *et al.* 2017). Thus the Hsp90 system switches between an ATP-bound "closed" conformational state mediated by Tsc1, with low affinity for kinase and non-kinase clients, and an ADP-bound conformational state with high affinity for substrate (Bukau and Horwich 1998, Woodford, Sager *et al.* 2017).

TSC1 p.L117P naturally occurring human variant located within the N-terminal domain of TSC1 was shown to be pathogenic (Nellist, van den Heuvel *et al.* 2009). The L117P substitution destabilizes Tsc1 and therefore prevents Tsc2 binding to Hsp90, leading to ubiquitination and degradation by the proteasome (Woodford, Sager *et al.* 2017), and resulting in increased mTORC1 activity.

II. References

Adriaensen, M. E., C. M. Schaefer-Prokop, T. Stijnen, D. A. Duyndam, B. A. Zonnenberg and M. Prokop (2009). "Prevalence of subependymal giant cell tumors in patients with tuberous sclerosis and a review of the literature." Eur J Neurol **16**(6): 691-696.

Ashfaq, R., A. G. Weinberg and J. Albores-Saavedra (1993). "Renal angiomyolipomas and HMB-45 reactivity." Cancer **71**(10): 3091-3097.

Astrinidis, A., L. Khare, T. Carsillo, T. Smolarek, K. S. Au, H. Northrup and E. P. Henske (2000). "Mutational analysis of the tuberous sclerosis gene TSC2 in patients with pulmonary lymphangioleiomyomatosis." J Med Genet **37**(1): 55-57.

Au, K. S., J. A. Rodriguez, J. L. Finch, K. A. Volcik, E. S. Roach, M. R. Delgado, E. Rodriguez, Jr. and H. Northrup (1998). "Germ-line mutational analysis of the TSC2 gene in 90 tuberous-sclerosis patients." Am J Hum Genet **62**(2): 286-294.

Au, K. S., A. T. Williams, E. S. Roach, L. Batchelor, S. P. Sparagana, M. R. Delgado, J. W. Wheless, J. E. Baumgartner, B. B. Roa, C. M. Wilson, T. K. Smith-Knuppel, M. Y. Cheung, V. H. Whittemore, T. M. King and H. Northrup (2007). "Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States." Genet Med **9**(2): 88-100.

Bastos, A. P. and L. F. Onuchic (2011). "Molecular and cellular pathogenesis of autosomal dominant polycystic kidney disease." Braz J Med Biol Res **44**(7): 606-617.

Bourneville, D. M. (1880). "Sclérose tubéreuse des circonvolutions cérébrales." Arch Neurol **1**: 11.

Brown, M. A., K. Foreman, J. Harriss, C. Das, L. Zhu, M. Edwards, S. Shaaban and H. Tucker (2015). "C-terminal domain of SMYD3 serves as a unique HSP90-regulated motif in oncogenesis." Oncotarget **6**(6): 4005-4019.

Buccoliero, A. M., A. Franchi, F. Castiglione, C. F. Gheri, F. Mussa, F. Giordano, L. Genitori and G. L. Taddei (2009). "Subependymal giant cell astrocytoma (SEGA): Is it an astrocytoma? Morphological, immunohistochemical and ultrastructural study." Neuropathology **29**(1): 25-30.

Bukau, B. and A. L. Horwich (1998). "The Hsp70 and Hsp60 chaperone machines." Cell **92**(3): 351-366.

Calderwood, S. K. and A. Murshid (2017). "Molecular Chaperone Accumulation in Cancer and Decrease in Alzheimer's Disease: The Potential Roles of HSF1." Front Neurosci **11**: 192.

Camposano, S. E., E. Greenberg, D. J. Kwiatkowski and E. A. Thiele (2009). "Distinct clinical characteristics of tuberous sclerosis complex patients with no mutation identified." Ann Hum Genet **73**(2): 141-146.

Carlisle, C., K. Prill and D. Pilgrim (2017). "Chaperones and the Proteasome System: Regulating the Construction and Demolition of Striated Muscle." Int J Mol Sci **19**(1).

Carsillo, T., A. Astrinidis and E. P. Henske (2000). "Mutations in the tuberous sclerosis complex gene TSC2 are a cause of sporadic pulmonary lymphangioleiomyomatosis." Proc Natl Acad Sci U S A **97**(11): 6085-6090.

Castro, M., C. W. Shepherd, M. R. Gomez, J. T. Lie and J. H. Ryu (1995). "Pulmonary tuberous sclerosis." Chest **107**(1): 189-195.

Caylor, R. C., L. Grote, I. Thiffault, E. G. Farrow, L. Willig, S. Soden, S. M. Amudhavalli, A. J. Nopper, K. A. Horii, E. Fleming, J. Jenkins, H. Welsh, M. Ilyas, K. Engleman, A. Abdelmoity and C. J. Saunders (2018). "Incidental diagnosis of tuberous sclerosis complex by exome sequencing in three families with subclinical findings." Neurogenetics **19**(3): 205-213.

Chiron, C., C. Dumas, I. Jambaqué, J. Mumford and O. Dulac (1997). "Randomized trial comparing vigabatrin and hydrocortisone in infantile spasms due to tuberous sclerosis." Epilepsy Research **26**(2): 389-395.

Chu-Shore, C. J., P. Major, S. Camposano, D. Muzykewicz and E. A. Thiele (2010). "The natural history of epilepsy in tuberous sclerosis complex." Epilepsia **51**(7): 1236-1241.

Connor, J. M., J. B. Stephenson and M. D. Hadley (1986). "Non-penetrance in tuberous sclerosis." Lancet **2**(8518): 1275.

Consortium, T. E. C. T. S. (1993). "Identification and characterization of the tuberous sclerosis gene on chromosome 16." Cell **75**(7): 1305-1315.

Cooper, J. R. (1971). "Brain tumors in hereditary multiple system hamartomatosis (tuberous sclerosis)." J Neurosurg **34**(2 Pt 1): 194-202.

Crino, P. B. (2013). "Evolving neurobiology of tuberous sclerosis complex." Acta Neuropathol **125**(3): 317-332.

Crundwell, M. (2004). "Pathology and genetics of tumours of the urinary system and male genital organs." BJU International **94**(4): 675-675.

Curatolo, P., S. Seri, M. Verdecchia and R. Bombardieri (2001). "Infantile spasms in tuberous sclerosis complex." Brain Dev **23**(7): 502-507.

Dabora, S. L., S. Jozwiak, D. N. Franz, P. S. Roberts, A. Nieto, J. Chung, Y. S. Choy, M. P. Reeve, E. Thiele, J. C. Egelhoff, J. Kasprzyk-Obara, D. Domanska-Pakiela and D. J. Kwiatkowski (2001). "Mutational analysis in a cohort of 224 tuberous sclerosis patients indicates increased severity of TSC2, compared with TSC1, disease in multiple organs." Am J Hum Genet **68**(1): 64-80.

Davidson, M., H. Yoshidome, E. Stenroos and W. G. Johnson (1991). "Neuron-like cells in culture of tuberous sclerosis tissue." Ann N Y Acad Sci **615**: 196-210.

Eble, J. N. (1998). "Angiomyolipoma of kidney." Semin Diagn Pathol **15**(1): 21-40.

Fernandez-Flores, A. (2011). "Perivascular migration: a clue to the histogenesis of PEComas?" Am J Dermatopathol **33**(5): 528-529.

Folpe, A. L. and D. J. Kwiatkowski (2010). "Perivascular epithelioid cell neoplasms: pathology and pathogenesis." Hum Pathol **41**(1): 1-15.

Franz, D. N., J. J. Bissler and F. X. McCormack (2010). "Tuberous sclerosis complex: neurological, renal and pulmonary manifestations." Neuropediatrics **41**(5): 199-208.

Garami, A., F. J. T. Zwartkruis, T. Nobukuni, M. Joaquin, M. Roccio, H. Stocker, S. C. Kozma, E. Hafen, J. L. Bos and G. Thomas (2003). "Insulin Activation of Rheb, a Mediator of mTOR/S6K/4E-BP Signaling, Is Inhibited by TSC1 and 2." Molecular Cell **11**(6): 1457-1466.

Gomez, M., J. Sampson and V. Holtes-Whittemore (1999). "Tuberous Sclerosis Complex." Oxford University Press **3rd edition**.

Goncharova, E., D. Goncharov, D. Noonan and V. P. Krymskaya (2004). "TSC2 modulates actin cytoskeleton and focal adhesion through TSC1-binding domain and the Rac1 GTPase." J Cell Biol **167**(6): 1171-1182.

Goodman, M., S. H. Lamm, A. Engel, C. W. Shepherd, O. W. Houser and M. R. Gomez (1997). "Cortical tuber count: a biomarker indicating neurologic severity of tuberous sclerosis complex." J Child Neurol **12**(2): 85-90.

Grajkowska, W., K. Kotulska, E. Jurkiewicz and E. Matyja (2010). "Brain lesions in tuberous sclerosis complex. Review." Folia Neuropathol **48**(3): 139-149.

Hancock, E. and J. Osborne (1998). "Treatment of infantile spasms with high-dose oral prednisolone." Dev Med Child Neurol **40**(7): 500.

Hasbani, D. M. and P. B. Crino (2018). "Tuberous sclerosis complex." Handb Clin Neurol **148**: 813-822.

Hayashi, T., T. Kumasaka, K. Mitani, T. Yao, K. Suda and K. Seyama (2010). "Loss of heterozygosity on tuberous sclerosis complex genes in multifocal micronodular pneumocyte hyperplasia." Mod Pathol **23**(9): 1251-1260.

Holmes, G. L., C. E. Stafstrom and G. Tuberous Sclerosis Study (2007). "Tuberous sclerosis complex and epilepsy: recent developments and future challenges." Epilepsia **48**(4): 617-630.

Hoogeveen-Westerveld, M., L. van Unen, A. van den Ouweland, D. Halley, A. Hoogeveen and M. Nellist (2012). "The TSC1-TSC2 complex consists of multiple TSC1 and TSC2 subunits." BMC Biochem **13**: 18.

Hosoyaa, M., H. Naitoa and K. Nihei (1999). "Neurological prognosis correlated with variations over time in the number of subependymal nodules in tuberous sclerosis." Brain & Development **21**(8): 4.

Houser, O. W. and M. R. Gomez (1992). "CT and MR imaging of intracranial tuberous sclerosis." J Dermatol **19**(11): 904-908.

Humphrey, A., J. Williams, E. Pinto and P. F. Bolton (2004). "A prospective longitudinal study of early cognitive development in tuberous sclerosis - a clinic based study." Eur Child Adolesc Psychiatry **13**(3): 159-165.

Hung, C. C., Y. N. Su, S. C. Chien, H. H. Liou, C. C. Chen, P. C. Chen, C. J. Hsieh, C. P. Chen, W. T. Lee, W. L. Lin and C. N. Lee (2006). "Molecular and clinical analyses of 84 patients with tuberous sclerosis complex." BMC Med Genet **7**: 72.

Huttenlocher, P. R. and R. L. Wollmann (1991). "Cellular neuropathology of tuberous sclerosis." Ann N Y Acad Sci **615**: 140-148.

Inoki, K., Y. Li, T. Xu and K. L. Guan (2003). "Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling." Genes Dev **17**(15): 1829-1834.

Inoue, H., T. Uyama, T. Suzuki, M. Kazami, O. Hino, T. Kobayashi, K. Kobayashi, T. Tadokoro and Y. Yamamoto (2010). "Phosphorylated hamartin-Hsp70 complex regulates apoptosis via mitochondrial localization." Biochem Biophys Res Commun **391**(1): 1148-1153.

Jansen, F. E., O. van Nieuwenhuizen and A. C. van Huffelen (2004). "Tuberous sclerosis complex and its founders." J Neurol Neurosurg Psychiatry **75**(5): 770.

Johnson, W. G., H. Yoshidome, E. S. Stenroos and M. M. Davidson (1991). "Origin of the neuron-like cells in tuberous sclerosis tissues." Ann N Y Acad Sci **615**: 211-219.

Joinson, C., F. J. O'Callaghan, J. P. Osborne, C. Martyn, T. Harris and P. F. Bolton (2003). "Learning disability and epilepsy in an epidemiological sample of individuals with tuberous sclerosis complex." Psychol Med **33**(2): 335-344.

Jones, A. C., M. M. Shyamsundar, M. W. Thomas, J. Maynard, S. Idziaszczyk, S. Tomkins, J. R. Sampson and J. P. Cheadle (1999). "Comprehensive mutation analysis of TSC1 and TSC2-and phenotypic correlations in 150 families with tuberous sclerosis." Am J Hum Genet **64**(5): 1305-1315.

Jozwiak, J., S. Jozwiak and P. Skopinski (2005). "Immunohistochemical and microscopic studies on giant cells in tuberous sclerosis." Histol Histopathol **20**(4): 1321-1326.

Knudson, A. G., Jr. (1971). "Mutation and cancer: statistical study of retinoblastoma." Proc Natl Acad Sci U S A **68**(4): 820-823.

Kozlowski, P., P. Roberts, S. Dabora, D. Franz, J. Bissler, H. Northrup, K. S. Au, R. Lazarus, D. Domanska-Pakiela, K. Kotulska, S. Jozwiak and D. J. Kwiatkowski (2007). "Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations." Hum Genet **121**(3-4): 389-400.

Krueger, D. A., H. Northrup and G. International Tuberous Sclerosis Complex Consensus (2013). "Tuberous sclerosis complex surveillance and management: recommendations of the 2012 International Tuberous Sclerosis Complex Consensus Conference." Pediatr Neurol **49**(4): 255-265.

Kwiatkowski, D. J. (1994). "Tuberous Sclerosis." Archives of Dermatology **130**(3).

Lamb, R. F., C. Roy, T. J. Diefenbach, H. V. Vinters, M. W. Johnson, D. G. Jay and A. Hall (2000). "The TSC1 tumour suppressor hamartin regulates cell adhesion through ERM proteins and the GTPase Rho." Nat Cell Biol **2**(5): 281-287.

Lotz, G. P., H. Lin, A. Harst and W. M. Obermann (2003). "Aha1 binds to the middle domain of Hsp90, contributes to client protein activation, and stimulates the ATPase activity of the molecular chaperone." J Biol Chem **278**(19): 17228-17235.

Maheshwar, M. M., J. P. Cheadle, A. C. Jones, J. Myring, A. E. Fryer, P. C. Harris and J. R. Sampson (1997). "The GAP-Related Domain of Tuberin, the Product of the TSC2 Gene, is a Target for Missense Mutations in Tuberous Sclerosis." Human Molecular Genetics **6**(11): 1991-1996.

Martignoni, G., M. Pea, D. Reghellin, G. Zamboni and F. Bonetti (2008). "PEComas: the past, the present and the future." Virchows Arch **452**(2): 119-132.

Maruyama, H., K. Seyama, J. Sobajima, K. Kitamura, T. Sobajima, T. Fukuda, K. Hamada, M. Tsutsumi, O. Hino and Y. Konishi (2001). "Multifocal micronodular pneumocyte hyperplasia and lymphangioleiomyomatosis in tuberous sclerosis with a TSC2 gene." Mod Pathol **14**(6): 609-614.

McCall, T., S. S. Chin, K. L. Salzman and D. W. Fults (2006). "Tuberous sclerosis: a syndrome of incomplete tumor suppression." Neurosurg Focus **20**(1): E3.

Mizuguchi, M. and S. Takashima (2001). "Neuropathology of tuberous sclerosis." Brain Dev **23**(7): 508-515.

Moss, J., N. A. Avila, P. M. Barnes, R. A. Litzenberger, J. Bechtle, P. G. Brooks, C. J. Hedin, S. Hunsberger and A. S. Kristof (2001). "Prevalence and clinical characteristics of lymphangioleiomyomatosis (LAM) in patients with tuberous sclerosis complex." Am J Respir Crit Care Med **164**(4): 669-671.

Muir, T. E., K. O. Leslie, H. Popper, M. Kitaichi, E. Gagné, J. K. Emelin, H. V. Vinters and T. V. Colby (1998). "Micronodular pneumocyte hyperplasia." Am J Surg Pathol **22**(4): 8.

Nellist, M., D. van den Heuvel, D. Schluep, C. Exalto, M. Goedbloed, A. Maat-Kievit, T. van Essen, K. van Spaendonck-Zwarts, F. Jansen, P. Helderman, G. Bartalini, O. Vierimaa, M. Penttinen, J. van den Ende, A. van den Ouweland and D. Halley (2009). "Missense mutations to the TSC1 gene cause tuberous sclerosis complex." Eur J Hum Genet **17**(3): 319-328.

Nellist, M., M. A. van Slegtenhorst, M. Goedbloed, A. M. van den Ouweland, D. J. Halley and P. van der Sluijs (1999). "Characterization of the cytosolic tuberin-hamartin complex. Tuberin is a cytosolic chaperone for hamartin." J Biol Chem **274**(50): 35647-35652.

Niida, Y., N. Lawrence-Smith, A. Banwell, E. Hammer, J. Lewis, R. L. Beauchamp, K. Sims, V. Ramesh and L. Ozelius (1999). "Analysis of bothTSC1 andTSC2 for germline mutations in 126 unrelated patients with tuberous sclerosis." Human Mutation **14**(5): 412-422.

Niida, Y., A. O. Stemmer-Rachamimov, M. Logrip, D. Tapon, R. Perez, D. J. Kwiatkowski, K. Sims, M. MacCollin, D. N. Louis and V. Ramesh (2001). "Survey of somatic mutations in tuberous sclerosis complex (TSC) hamartomas suggests different genetic mechanisms for pathogenesis of TSC lesions." Am J Hum Genet **69**(3): 493-503.

Nishio, S., T. Morioka, S. Suzuki, R. Kira, F. Mihara and M. Fukui (2001). "Subependymal giant cell astrocytoma: clinical and neuroimaging features of four cases." J Clin Neurosci **8**(1): 31-34.

Northrup, H., D. A. Krueger and G. International Tuberous Sclerosis Complex Consensus (2013). "Tuberous sclerosis complex diagnostic criteria update: recommendations of the 2012 Iinternational Tuberous Sclerosis Complex Consensus Conference." Pediatr Neurol **49**(4): 243-254.

Northrup, H., D. A. Krueger and G. International Tuberous Sclerosis Complex Consensus (2013). "Tuberous sclerosis complex diagnostic criteria update: recommendations of the 2012 international tuberous sclerosis complex consensus conference." Pediatr Neurol **49**(4): 243-254.

Northrup, H., J. W. Wheless, T. K. Bertin and R. A. Lewis (1993). "Variability of expression in tuberous sclerosis." J Med Genet **30**(1): 41-43.

O'Callaghan, F. J., A. W. Shiell, J. P. Osborne and C. N. Martyn (1998). "Prevalence of tuberous sclerosis estimated by capture-recapture analysis." Lancet **351**(9114): 1490.

Orlova, K. A. and P. B. Crino (2010). "The tuberous sclerosis complex." Ann N Y Acad Sci **1184**: 87-105.

Panaretou, B., C. Prodromou, S. M. Roe, R. O'Brien, J. E. Ladbury, P. W. Piper and L. H. Pearl (1998). "ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone in vivo." EMBO J **17**(16): 4829-4836.

Panaretou, B., G. Siligardi, P. Meyer, A. Maloney, J. K. Sullivan, S. Singh, S. H. Millson, P. A. Clarke, S. Naaby-Hansen, R. Stein, R. Cramer, M. Mollapour, P. Workman, P. W. Piper, L. H. Pearl and C. Prodromou (2002). "Activation of the ATPase activity of hsp90 by the stress-regulated cochaperone aha1." Mol Cell **10**(6): 1307-1318.

Plank, T. L., R. S. Yeung and E. P. Henske (1998). "Hamartin, the product of the tuberous sclerosis 1 (TSC1) gene, interacts with tuberin and appears to be localized to cytoplasmic vesicles." Cancer Research **58**(21): 4766-4770.

Prabhakar, S., X. Zhang, J. Goto, S. Han, C. Lai, R. Bronson, M. Sena-Esteves, V. Ramesh, A. Stemmer-Rachamimov, D. J. Kwiatkowski and X. O. Breakefield (2015). "Survival benefit and phenotypic improvement by hamartin gene therapy in a tuberous sclerosis mouse brain model." Neurobiol Dis **82**: 22-31.

Prodromou, C. and L. H. Pearl (2003). "Structure and functional relationships of Hsp90." Curr Cancer Drug Targets **3**(5): 301-323.

Rakowski, S. K., E. B. Winterkorn, E. Paul, D. J. Steele, E. F. Halpern and E. A. Thiele (2006). "Renal manifestations of tuberous sclerosis complex: Incidence, prognosis, and predictive factors." Kidney Int **70**(10): 1777-1782.

Rendtorff, N. D., B. Bjerregaard, M. Frodin, S. Kjaergaard, H. Hove, F. Skovby, K. Brondum-Nielsen, M. Schwartz and G. Danish Tuberous Sclerosis (2005). "Analysis of 65 tuberous sclerosis complex (TSC) patients by TSC2 DGGE, TSC1/TSC2 MLPA, and TSC1 long-range PCR sequencing, and report of 28 novel mutations." Hum Mutat **26**(4): 374-383.

Roach, E. S., M. R. Gomez and H. Northrup (1998). "Tuberous sclerosis complex consensus conference: revised clinical diagnostic criteria." J Child Neurol **13**(12): 624-628.

Roach, E. S. and S. P. Sparagana (2004). "Diagnosis of tuberous sclerosis complex." J Child Neurol **19**(9): 643-649.

Sampson, J. R., S. J. Scahill, J. B. Stephenson, L. Mann and J. M. Connor (1989). "Genetic aspects of tuberous sclerosis in the west of Scotland." J Med Genet **26**(1): 28-31.

Sancak, O., M. Nellist, M. Goedbloed, P. Elfferich, C. Wouters, A. Maat-Kievit, B. Zonnenberg, S. Verhoef, D. Halley and A. van den Ouweland (2005). "Mutational analysis of the TSC1 and TSC2 genes in a diagnostic setting: genotype--phenotype correlations and comparison of diagnostic DNA techniques in Tuberous Sclerosis Complex." Eur J Hum Genet **13**(6): 731-741.

Sasongko, T. H., M. Wataya-Kaneda, K. Koterazawa, Gunadi, S. Yusoff, I. S. Harahap, M. J. Lee, M. Matsuo and H. Nishio (2008). "Novel mutations in 21 patients with tuberous sclerosis complex and variation of tandem splice-acceptor sites in TSC1 exon 14." Kobe J Med Sci **54**(1): E73-81.

Scheithauer, B. J. and T. J. Reagan (1999). "In: Gomez, M. R. e Whittemore, V. H. 3a Ed. Tuberous Sclerois Complex." Neurophatology: 44.

Schubert, U., L. C. Anton, J. Gibbs, C. C. Norbury, J. W. Yewdell and J. R. Bennink (2000). "Rapid degradation of a large fraction of newly synthesized proteins by proteasomes." Nature **404**(6779): 770-774.

Shepherd, C. W., O. W. Houser and M. R. Gomez (1995). "MR findings in tuberous sclerosis complex and correlation with seizure development and mental impairment." AJNR Am J Neuroradiol **16**(1): 149-155.

Shepherd, C. W., B. W. Scheithauer, M. R. Gomez, H. J. Altermatt and J. A. Katzmann (1991). "Subependymal giant cell astrocytoma: a clinical, pathological, and flow cytometric study." Neurosurgery **28**(6): 864-868.

Smolarek, T. A., L. L. Wessner, F. X. McCormack, J. C. Mylet, A. G. Menon and E. P. Henske (1998). "Evidence that lymphangiomyomatosis is caused by TSC2 mutations: chromosome 16p13 loss of heterozygosity in angiomyolipomas and lymph nodes from women with lymphangiomyomatosis." Am J Hum Genet **62**(4): 810-815.

Stefansson, K. (1991). "Tuberous sclerosis." Mayo Clin Proc **66**(8): 868-872.

Tee, A. R., D. C. Fingar, B. D. Manning, D. J. Kwiatkowski, L. C. Cantley and J. Blenis (2002). "Tuberous sclerosis complex-1 and -2 gene products function together to inhibit mammalian target of rapamycin (mTOR)-mediated downstream signaling." Proc Natl Acad Sci U S A **99**(21): 13571-13576.

Tee, A. R., B. D. Manning, P. P. Roux, L. C. Cantley and J. Blenis (2003). "Tuberous Sclerosis Complex Gene Products, Tuberin and Hamartin, Control mTOR Signaling by Acting as a GTPase-Activating Protein Complex toward Rheb." Current Biology **13**(15): 1259-1268.

Telfeian, A. E., A. Judkins, D. Younkin, A. N. Pollock and P. Crino (2004). "Subependymal giant cell astrocytoma with cranial and spinal metastases in a patient with tuberous sclerosis. Case report." J Neurosurg **100**(5 Suppl Pediatrics): 498-500.

Thiele, E. A. (2004). "Managing epilepsy in tuberous sclerosis complex." J Child Neurol **19**(9): 680-686.

Trombley, I. K. and S. S. Mirra (1981). "Ultrastructure of tuberous sclerosis: cortical tuber and subependymal tumor." Ann Neurol **9**(2): 174-181.

van Eeghen, A. M., L. O. Teran, J. Johnson, M. B. Pulsifer, E. A. Thiele and P. Caruso (2013). "The neuroanatomical phenotype of tuberous sclerosis complex: focus on radial migration lines." Neuroradiology **55**(8): 1007-1014.

van Slegtenhorst, M., R. de Hoogt, C. Hermans, M. Nellist, B. Janssen, S. Verhoef, D. Lindhout, A. van den Ouweland, D. Halley, J. Young, M. Burley, S. Jeremiah, K. Woodward, J. Nahmias, M. Fox, R. Ekong, J. Osborne, J. Wolfe, S. Povey, R. G. Snell, J. P. Cheadle, A. C. Jones, M. Tachataki, D. Ravine, J. R. Sampson, M. P. Reeve, P. Richardson, F. Wilmer, C. Munro, T. L. Hawkins, T. Sepp, J. B. Ali, S. Ward, A. J. Green, J. R. Yates, J. Kwiatkowska, E. P. Henske, M. P. Short, J. H. Haines, S. Jozwiak and D. J. Kwiatkowski (1997). "Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34." Science **277**(5327): 805-808.

van Slegtenhorst, M., M. Nellist, B. Nagelkerken, J. Cheadle, R. Snell, A. van den Ouweland, A. Reuser, J. Sampson, D. Halley and P. van der Sluijs (1998). "Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products." Human molecular genetics **7**(6): 1053-1057.

van Slegtenhorst, M. A., A. Verhoef, A. Tempelaars, L. Bakker, Q. Wang, M. Wessels, R. Bakker, M. Nellist, D. Lindhout, D. Halley and A. van den Ouweland (1999). "Mutational spectrum of the TSC1 gene in a cohort of 225 tuberous sclerosis complex patients: no evidence for genotype-phenotype correlation." J Med Genet **36**(4): 5.

von Ranke, F. M., I. M. Faria, G. Zanetti, B. Hochhegger, A. S. Souza, Jr. and E. Marchiori (2017). "Imaging of tuberous sclerosis complex: a pictorial review." Radiol Bras **50**(1): 48-54.

Webb, D. W. and J. P. Osborne (1991). "Non-penetrance in tuberous sclerosis." J Med Genet **28**(6): 417-419.

Weeks, D. A., R. L. Malott, M. Arnesen, C. Zuppan, D. Aitken and G. Mierau (1991). "Hepatic angiomyolipoma with striated granules and positivity with melanoma--specific antibody (HMB-45): a report of two cases." Ultrastruct Pathol **15**(4-5): 563-571.

Wilson, P. J., V. Ramesh, A. Kristiansen, C. Bove, S. Jozwiak, D. J. Kwiatkowski, M. P. Short and J. L. Haines (1996). "Novel Mutations Detected in the TSC2 Hene From Both Sporadic and Familial TSC Patients." Human Molecular Genetics **5**(2): 249-256.

Woodford, M. R., D. M. Dunn, A. R. Blanden, D. Capriotti, D. Loiselle, C. Prodromou, B. Panaretou, P. F. Hughes, A. Smith, W. Ackerman, T. A. Haystead, S. N. Loh, D. Bourboulia, L. S. Schmidt, W. Marston Linehan, G. Bratslavsky and M. Mollapour (2016). "The FNIP co-chaperones decelerate the Hsp90 chaperone cycle and enhance drug binding." Nat Commun **7**: 12037.

Woodford, M. R., R. A. Sager, E. Marris, D. M. Dunn, A. R. Blanden, R. L. Murphy, N. Rensing, O. Shapiro, B. Panaretou, C. Prodromou, S. N. Loh, D. H. Gutmann, D. Bourboulia, G. Bratslavsky, M. Wong and M. Mollapour (2017). "Tumor suppressor Tsc1 is a new Hsp90 co-chaperone that facilitates folding of kinase and non-kinase clients." EMBO J **36**(24): 3650-3665.

Yang, L., X. L. Feng, S. Shen, L. Shan, H. F. Zhang, X. Y. Liu and N. Lv (2012). "Clinicopathological analysis of 156 patients with angiomyolipoma originating from different organs." Oncol Lett **3**(3): 586-590.

CHAPTER 2

List of Acronyms

ACMG: American College of Medical Genetics and Genomics

DMSO: Dimethyl sulfoxide

GADPH: Glyceraldehyde-3-phosphate dehydrogenase

GFI1B: Growth Factor-Independent 1B

GRAACC: Grupo de Apoio ao Adolescente e à Criança com Câncer

HC-UFPR: Hospital de Clínicas of the Universidade Federal do Paraná

LOVD: Leiden Open Variation Database

mTORC1: Mechanistic target of rapamycin kinase complex 1

MLPA: Multiplex ligation-dependent probe amplification

NGS: Next generation sequencing

panel-NGS: Next Generation Sequencing on a customized gene panel chip

NMI: No mutation identified

ORF: Open reading frame

PKD1: Polycystic Kidney Disease 1 gene

PKDTS: Polycystic Kidney Disease, Infantile Severe, with tuberous sclerosis

PCR: Polymerase chain reaction

qPCR: Quantitative PCR

RHEB: Ras homologue enriched in brain

S6K1: S6 kinase 1

HSCM-SP: Santa Casa de Misericórdia

SFs: Secondary findings discoveries

SINE: Short interspersed nuclear element

SNV: Single nucleotide variants

TSC: Tuberous sclerosis complex

VUS: Variant of unclear significance

WT: Wild-type

List of Tables

List of Figures

Summary

I. Introduction

Tuberous sclerosis complex (TSC) is a genetic disease affecting 1 in every 6,000 to 10,000 live-born infants, showing no ethnic bias. It is characterized by hamartoma growth more often in the brain, skin, kidneys, heart, lungs, and retina (Sampson, Scahill *et al.* 1989, O'Callaghan, Shiell *et al.* 1998, Au, Williams *et al.* 2007, Northrup, Krueger *et al.* 2013). Renal manifestations and brain lesions such as SEGA, are an important source of morbidity and mortality, after severe intellectual disability leading to epilepsy (Rakowski, Winterkorn *et al.* 2006, Moavero, Pinci *et al.* 2011, Northrup, Krueger *et al.* 2013, Kwiatkowski, Palmer *et al.* 2015). Northrup et al (2013) presented an up-to-date classification of TSC lesions considered for the clinical diagnosis of the disease.

TSC is an autosomal dominant disorder due to pathogenic variants in either *TSC1* or *TSC2* tumour suppressor genes (Consortium 1993, van Slegtenhorst, de Hoogt *et al.* 1997). The *TSC1* gene (NG_012386.1, Ch:9q34.1, MIM#605284), with approximately 55 kb comprising 23 exons, codes for the TSC1 protein (hamartin, NP_000359.1) with 130 kDa (van Slegtenhorst, de Hoogt *et al.* 1997). The *TSC2* (NG_005895.1, Ch:16p13.3, MIM#191092) gene has 42 exons and nearly 40 kb, and encodes the TSC2 protein (tuberin, NP_000539.2) with 180 kDa (Consortium 1993, van Slegtenhorst, de Hoogt *et al.* 1997, Jones, Shyamsundar *et al.* 1999, Dabora, Jozwiak *et al.* 2001, Sancak, Nellist *et al.* 2005, Au, Williams *et al.* 2007). TSC1 and TSC2 physically interact as a cytosolic heterodimer (TSC1/2 complex), which acts as mediator of intracellular signals to inhibit the mechanistic target of rapamycin (mTOR) kinase complex 1 (mTORC1) through the small GTPase RHEB (Ras homologue enriched in brain; (Garami, Zwartkruis *et al.* 2003, Tee, Manning *et al.* 2003). Inactivation of the TSC1/2 complex results in increased levels of RHEB-GTP, required for activation of mTORC1, which phosphorylates downstream targets. Inactivation of the TSC1/2 complex increases mTOR phosphorylation of, for instance, the S6 kinase 1 (S6K1) at residue Thr^{389} (Goncharova, Goncharov *et al.* 2002, Tee, Fingar *et al.* 2002) causing excessive cell growth and augmented cell proliferation.

Classified as tumor suppressor genes, biallelic inactivation of the *TSC1* or *TSC2* genes is expected for the development of hamartomas and hamartias leading to manifestation of TSC. In this way, a second mutational event must occur, causing biallelic inactivation of one or the other gene. Thus, in addition to the germ mutation, which can be detected in leukocyte DNA from patients, the second mutational event must be somatic and necessary for the clonal development of the tumors, according to the hypothesis developed by Knudson for tumor suppressor genes Knudson (1971). On the other hand, the most frequent lesion and early diagnosis for patients with TSC, i.e. cortical tuberosities, are considered hamartias and have the somatic mutation generally unidentified, being justified by the haploinsufficiency of TSC1/2 (Niida, Stemmer-Rachamimov *et al.* 2001, Qin, Chan *et al.* 2010).Gómez (1979) first established a broad set of major and minor clinical criteria for the definite, probable or possible diagnosis of TSC. Since then TSC diagnosis criteria have

been reviewed and updated (Roach, Gomez *et al.* 1998, Northrup, Krueger *et al.* 2013). Besides updating the TSC clinical criteria, the 2012 working group implemented the genetic diagnostic criterion, by which the detection of a pathogenic variant in ether *TSC1* or *TSC2* genes is sufficient for the definite molecular diagnosis of TSC (Northrup, Krueger *et al.* 2013), following the American College of Medical Genetics and Genomics (ACMG) guidelines for pathogenicity report of DNA variants (Richards, Aziz *et al.* 2015). Nearly 90% of patients with definite clinical diagnosis of TSC have a *TSC1* or *TSC2* phatogenic DNA variant identified by genetic testing (Peron, Au *et al.* 2018). The ratio between *TSC1* and *TSC2* mutation frequencies appears close to 1:1 among TSC familial cases, and may vary from 1:3 to 1:6 among sporadic cases (Jones, Shyamsundar *et al.* 1999, Dabora, Jozwiak *et al.* 2001, Rendtorff, Bjerregaard *et al.* 2005, Au, Williams *et al.* 2007). Sporadic TSC cases with more severe angiomyolipomas and or intellectual disability tend to have germline mutations more frequently on the *TSC2* gene than the *TSC1* gene (Dabora, Jozwiak *et al.* 2001, Sancak, Nellist *et al.* 2005, Au, Williams *et al.* 2007). In approximately 10% of patients with definite clinical diagnosis of TSC, no pathogenic DNA variant is identified in *TSC1* or *TSC2* by current molecular techniques (Kozlowski, Roberts *et al.* 2007, Camposano, Greenberg *et al.* 2009, Qin, Kozlowski *et al.* 2010, Tyburczy, Wang *et al.* 2013, Nellist, Brouwer *et al.* 2015). *TSC1* or *TSC2* somatic mosaicism and deep intronic DNA variants have recently been reported in those patients with no mutation identified (NMI) in the coding region of either gene (Nellist, Brouwer *et al.* 2015, Tyburczy, Dies *et al.* 2015).

In a previous work, we have evaluated by Sanger sequencing the *TSC1* coding sequence in a cohort of 28 TSC patients (Dufner-Almeida 2014). In the present study, we extended the cohort to a total of 115 patients from the states of São Paulo and Paraná, Brazil, and employed their leukocyte DNA to search for *TSC1* and *TSC2* germline pathogenic DNA variants and segmental deletions/duplications. We disclosed a mutation detection rate of 86,09% (99/115), with 82 (82.82%) pathogenic variants in the *TSC2* and 17 (17.17%) in the *TSC1*.

II. Aim and Objectives

A. Aim

To analyze the nature, distribution and functional effects of *TSC1* and *TSC2* DNA variants from 115 patients with clinical definite diagnosis of TSC.

B. Objectives

1) To sequence the DNA coding sequence of *TSC1* and *TSC2* by the Sanger method or next generation sequencing (NGS);
2) To perform multiplex ligation-dependent probe amplification (MLPA) of TSC NMI patient DNA;
3) To confirm every MLPA alteration on *TSC1* or *TSC2* by quantitative PCR (qPCR);
4) To functionally evaluate the effect of identified missense *TSC1* or *TSC2* DNA variants of uncertain significance on *in vitro* hyperactivation of mTOR having S6K1 Thr^{389} phosphorylation as reporter;
5) To functionally assess putative splicing DNA variants detected outside canonical splice sites, by overexpressing a minigene harboring the DNA variant and reporting its splicing output.

III. Subjects and Methods

A. Patients

This study was approved by the Institutional Ethics Review Board from the three participating institutions (Protocol Number CAAE: 125729112.3.000.5464): *Santa Casa de Misericórdia* (HSCM-SP), São Paulo, Brazil; *Hospital de Clínicas* of the *Universidade Federal do Paraná* (HC-UFPR), Curitiba, Brazil, and *Grupo de Apoio ao Adolescente e à Criança com Câncer* (GRAACC), São Paulo, Brazil. All 115 patients had informed consent signed by a parent or tutor. A definite clinical diagnosis of TSC (Northrup, Krueger *et al.* 2013) was the only inclusion criterion. Twenty-three patients were referred from the HSCM-SP, 75 from the HC-UFPR and 19 from GRAACC. Peripheral blood samples (4 mL) were harvested by venipuncture and sent to the Functional Genetics laboratory at the University of São Paulo (Instituto de Biociências, IBUSP, São Paulo, Brazil) for *TSC1* and *TSC2* testing. Genetic counselling has been offered to all patients after the genetic analysis had been concluded.

B. Sanger Sequencing and Next Generation Sequencing

Genomic DNA (DNA) was extracted from peripheral blood leukocytes using *Puregene chemistry* (QIAGEN, Hilden, Germany) and *Autopure LS* devices (Thermo Fisher Scientific, Waltham, Massachusetts, USA), at *Centro de Estudos do Genoma Humano e Células Tronco* (CEGH-CEL-IBUSP) or with *QIAamp DNA Blood Midi* and *Maxi Kits* (QIAGEN, Hilden, Germany) at the Functional Genetics laboratory, according to manufacturer recommendation. Quantification and quality control were measured on a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) for Sanger sequencing procedures, as well as agarose gel and Qubit 2.0 (Invitrogen, Carlsbad, California, USA) for Next Generation Sequencing on a customized gene panel chip (panel-NGS).

Oligonucleotides for polymerase chain reaction (PCR) and Sanger sequencing were designed with *Primer3* (v. 0.4.0 (Rozen and Skaletsky 2000) or Primer-BLAST (NCBI), with 20-22 nucleotides (optimum at 21 nt), melting temperatures of 59-61°C (optimum at 60°C, and 1°C of maximal difference between pair oligonucleotides) and guanine and cytosine percentage of 40-60% (optimum at 50%). The ensemble of PCR amplicons covered all exons, intron boundaries and the core promoter regions of both *TSC1* (NG_012386.1) and *TSC2* (NG_005895.1) genes (Supplementary Table 1 and Supplementary Table 2). For *TSC1,* the sequenced region consisted of ~12.5 kb of the total 64 kb genomic locus (19.5%), including 517 bp of the core promoter and upstream sequence (Gerstein, Bruce *et al.* 2007), 3.5 kb (100%) of exonic sequences, and 5.4 kb (12.3%) of intronic sequences. For *TSC2*, the sequenced region consisted of ~21 kb of the total 46 kb genomic locus (52%), 485 bp of the core promoter and upstream sequence, 5.6 kb (100%) of exonic sequences and 11.7 kb (33.4%) of intronic sequences. Coding sequence annotation was according to transcript reference sequences NM_000368.4 (*TSC1*) and NM_000548.3 (*TSC2*). Primer specificity was tested against the human genome build

GRCh37/hg19 with the BLAT program (UCSC Genome Browser, Santa Cruz CA, USA) and nucleotide BLAST (NCBI, Bethesda MD, USA). Conditions for 25 µL of PCR reaction were standardized with DNA samples from three unrelated non-TSC individuals (non-TSC): 1X PCR buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl; Invitrogen, Carlsbad, California, USA), 2 mM $MgCl_2$ (Invitrogen, Carlsbad, California, USA), 5% dimethyl sulfoxide (DMSO), each deoxyribonucleotide at 0.2 mM, primer oligonucleotide at 0.2 µM, 80 ng of DNA and 0.2 U of *Taq* polymerase (Invitrogen, Carlsbad, California, USA). For *TSC1* exon 1 plus the promoter segment, a region highly enriched in cytosines and guanines (66%), we adopted the slowdown PCR protocol (Supplementary Table 3, Frey, Bachmann *et al.* (2008)); a standard PCR protocol was employed for the remaining reactions (Supplementary Table 4). PCR products were electrophoresed on 1.5% agarose gels with 5 µL of SYBRsafe (Thermo Fisher Scientific, Waltham, Massachusetts, USA) in 50 mL of 90 mM Tris and boric acid, 2 mM EDTA (TBE), and band images were captured on a Gel Doc™ EZ System using Image Lab™ software (Bio-Rad, Hercules, California, USA). Eight microliters of PCR products were treated with 1.5 µL of exonuclease I : shrimp alkaline phosphatase (5U:1U; Affymetrix, Santa Clara, California, USA) for 30-60 minutes at 37°C, followed by enzymatic inactivation for 15 minutes at 60°C.

Sanger sequencing was performed on purified PCR products following the ABI BigDye terminator protocol (Applied Biosystems; Foster City, California, USA) having 4 µL of enzyme-treated PCR product, the sequencing oligonucleotide primer at 0.3 µM (Supplementary Table 1 and Supplementary Table 2), 2 µL of BigDye terminator solution (v3.1), 1.5 µL 5X BigDye buffer (BigDyeno kit® - Applied Biosystems, Carlsbad, CA) and 6.5 µL of ultrapure water to a final reaction volume of 15 µL. Products were purified in a *Sephadex* column (Amersham Biosciences; Little Chalfont, United Kingdom), and submitted to capillary electrophoresis on an ABI 3730xl DNA Analyzer (Amersham Biosciences; Little Chalfont, United Kingdom). The results were analyzed using Sequencher V.4.10.1 or early version (Gene Codes Corporation; Ann Arbor, Michigan, USA).

Twenty-three DNA samples submission to Panel-NGS was justified by lack of identification of pathogenic DNA variants by Sanger sequencing (N=18), missesnse variant with unclear significance (N=4) or due to low concentration of the sample (N=1). Gene probes had been designed to specifically capture the coding sequence of 95 selected genes that compose a customized gene panel. This strategy comprehends capture of coding and intronic boundary sequences of both *TSC1* and *TSC2* genes with Nextera rapid capture (Illumina; San Diego, California, USA) The library was quantified using Bioanalyzer 2100 (Agilent Technologies; Santa Clara, California, USA), and sequencing was performed on the MiSeq platform (Illumina; San Diego, California, USA). All reads were aligned to the human genome (build GRCh37/hg19) using the Burrows-Wheeler Aligner (BWA) algorithm to efficiently align short sequencing reads against the human genome (Li and Durbin 2009), followed by Genome Analysis Toolkit (GATK - McKenna, Hanna *et al.* (2010)) and ANNOVAR variant calling and annotation (Open Bioinformatics - Wang, Li *et al.* (2010)). For *TSC1* and *TSC2* genes, the coverage for 10x and 20x was 100% with

C. Multiplex ligation-dependent probe amplification (MLPA)

The SALSA MLPA kits P124-C1 *TSC1* and P046-C1 *TSC2* (MRC-Holland; Amsterdam, The Netherlands) were used for detection of duplications and deletions affecting *TSC1* and *TSC2.* (Rendtorff, Bjerregaard *et al.* 2005, Kozlowski, Roberts *et al.* 2007, Padma Priya and Dalal 2012). The three non-TSC control samples used for Sanger sequencing conditions standardization were employed as reference for all procedures, following the manufacturer protocol. Ligated products were submitted to capillary electrophoresis on an ABI 3730xl DNA Analyzer (Amersham Biosciences; Little Chalfont, United Kingdom), and analyzed using GeneMarker (SoftGenetics; State College, Pennsylvania, USA) and Coffalyser.NET (MRC-Holland; Amsterdam, The Netherlands). Signal intensities between deleted or duplicated and normal genomic segments from the same subject were compared using Student's *t* test and considered significant if p-value < 0.05.

D. Quantitative polymerase chain reaction

To validate the copy number variations detected by MLPA, quantitative PCR (Q-PCR) was performed. Primers were designed using Primer-BLAST (NCBI) and tested *in silico* with Beacon Designer Free Edition (Premier Biosoft; Pallo Alto, California, USA), BLAT program (UCSC Genome Browser, Santa Cruz CA, USA) and nucleotide BLAST (NCBI, Bethesda, MD, USA) (Supplementary Table 5). Prior to Q-PCR, all primers were tested by standard PCR and submitted to agarose gel electrophoresis, according to section III.B. Q-PCR was carried out with the SYBR Green system (Applied Biosystems; Foster City, California, USA) on a 7500 Fast Real-Time PCR System apparatus (Applied Biosystems; Foster City, California, USA), as described previously (Ambar and Chiavegatto 2009). Optimal DNA and primer concentrations were determined by titration: for a 24-µL reaction, 25 ng DNA and 12.5 pmol of each primer were used with 12 µL PCR SYBR Green master mix. A non-TSC control DNA sample was used as a reference. Amplification and melting step (dissociation curve) were performed following standard parameters from the manufacturer program (Supplementary Table 6). The melting step followed each run to confirm product specificity and the absence of primer dimers (Supplementary Figure 1).

For copy number calibration, the reaction was normalized to the amplification of an amplicom for the *GADPH* gene (Glyceraldehyde-3-phosphate dehydrogenase, NM_002046.1, human chr:12). All samples were run in triplicates at least three times, and the data analyzed using the comparative $\Delta\Delta C_t$ cycle threshold method (Applied Biosystems; Foster City, California, EUA). Student's *t* test was applied to compare test and reference samples.

E. *In silico* analysis of DNA variants

DNA variant nomenclature was according to Human Genome Variation Society (HGVS, Melbourne, Australia) and were verified using both Variant Validator (University of Leicester, Freeman, Hart *et al.* (2018)) and Mutalyzer (LUMC, Wildeman, van Ophuizen *et al.* (2008)) web tools. Pathogenic DNA variants were defined according to ACMG standards and guidelines. *In silico* tools that predict damaging effects of missense

change over protein structures and those that predict effect on splicing output were considered as a single piece of evidence in sequence interpretation (Richards, Aziz *et al.* 2015).

To determine the pathogenicity of the DNA variants, the following data banks, web tools and software were consulted: Leiden Open Variation Database (Version 2; Leiden, The Netherlands), Ensembl (release 93 - July 2018; Wellcome Trust Sanger Institute; European Bioinformatics Institute, Hinxton, Cambridgeshire, United Kingdom) (Zerbino, Achuthan *et al.* 2018), PolyPhen-2 (Adzhubei, Schmidt *et al.* 2010), Mutation Taster (Schwarz, Rodelsperger *et al.* 2010), SIFT and PROVEAN (Ng and Henikoff 2006, Kumar, Henikoff *et al.* 2009, Choi and Chan 2015), PhosphositePlus (Hornbeck, Zhang *et al.* 2015) and Alamut 2,7-1 (Sophia Genetics, Saint Sulpice, Switzerland).

In silico analysis of DNA variants of uncertain clinical significance (VUS) potentially regulating pre-mRNA splicing was performed using the bioinformatics web tools Acescan2, SpliceAid2 (Piva, Giulietti *et al.* 2012), Human Splicing Finder 3.1 (HSF -Desmet, Hamroun *et al.* (2009)) and Alamut 2,7-1 (Sophia Genetics, Saint Sulpice, Switzerland). To identify repetitive elements in the *TSC2* genomic sequence the RepeatMasker (Version open-3.0, Institute for Systems Biology, Seattle WA) web tool was used. Information on all identified DNA variants has been deposited at the *TSC1* and *TSC2* Leiden Open Variation Database (LOVD).

Global alignment of *Homo sapiens* (NP_000539.2), *Rattus norvegicus* (P49816), and *Mus musculus* (Q61037) TSC2 C-terminus, and Rap1GAP catalytic domain (Homo sapiens, pdb|1SRQ|A) (Daumke *et al.*, 2004) were performed using UniProtKB BLAST (EMBL-EBI,). The alignments among TSC2 paralogues from the three organisms and Rap1GAP domain were set as reference to identify the position of *TSC2* variants. Structural predictions and spatial position of *TSC2* variant were generated with PyMOL molecular Graphic System Version 2.1.1 software (Schrödinger, Inc., Cambridge, Massachusetts, USA). Phosphorylation sites of the TSC2 GAP domain was accessed using PhosphoSite Plus (NIH,) web tool. Supplementary Table 7contains a list of softwares and websites consulted.

F. Functional assessment of *TSC2* VUS missense and in-frame deletion

TSC2 missense and in-frame deletion functional assessment was performed according to Hoogeveen-Westerveld, Wentink *et al.* (2011) and Hoogeveen-Westerveld, Ekong *et al.* (2012), except for the usage in the present work of the cell line HEK293T that had been genetically modified to have both alleles from *TSC1* and *TSC2* genes inactivated by CRISPR/Cas9 genome editing, generating the 3H9-1B1 cell line (Dufner-Almeida et al., submitted). Briefly, 3H9-1B1 cell line overexpressing full-length *TSC2* harboring the assayed DNA variant was tested against the control according to its ability to inhibit mTOR and thus activate S6K phosphorylation.

Full-length *TSC2* expression construct encoding the variant of interest was derived by site-directed mutagenesis of the DNA of the human wild-type (WT) *TSC2* (NM_000548.3) expression construct previously cloned in the pcDNA3 vector (Hoogeveen-Westerveld, Ekong *et al.* 2013). Procedures followed the

manufacturer protocol (QuikChange II Site-Directed Mutagenesis Kit, Agilent Technologies; Santa Clara, California, USA). Plasmidial DNA that had been submitted to site-directed mutagenesis was isolated and digested with *Eco*RI, in order to check the presence of right plasmodia DNA. Clone sequence identity was verified by Sanger sequencing, and clones with successful mutagenesis were subcloned and had the entire open reading frame (ORF) sequenced. Other constructs employed in the functional assay have been described before (Nellist, Sancak *et al.* 2005), including wild-type (WT) *TSC2, TSC2-GFP, TSC1-Myc* and WT *RPS6KB1* (*S6K*). Because of the low levels of endogenous expression of *S6K*, every transfection included the *S6K* clone DNA. Transfections were performed with 0.5 x 10^6 to 1.0 x 10^6 cells per well of 24-well dishes, and contained a combination of 0.2 µg WT-*TSC2,* 0.2 µg *TSC2-GFP,* 0.2 µg mutant *TSC2*, 0.4 µg *TSC1-Myc* and 0.2 µg *S6K* plasmidial DNA, as well as 4.0 µg polyethyleneimine (PEI; Sigma-Aldrich, St. Louis, Missouri, USA). Eighteen hours after transfection, the transfection efficiency was verified according to the fluorescence emitted by *TSC2-GFP* cells on confocal microscopy. Only assays yielding transfection rates higher than 10% were considered for Western blotting.

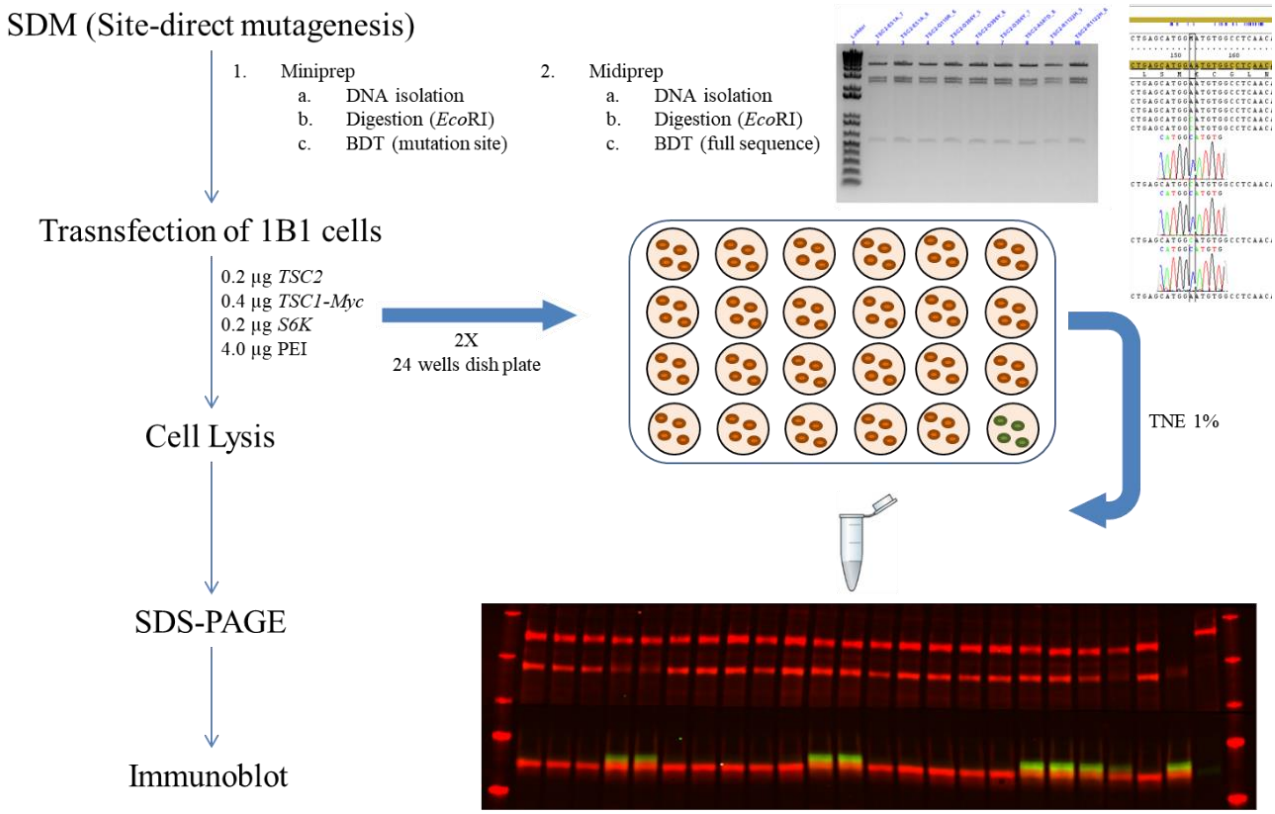


Figure 1: TSC2 missense and in-frame deletion functional assessment workflow.

For immunoblotting, cells were lysed in 50 mM Tris-HCl (pH 7.6), 100 mM NaCl, 50 mM NaF, 1% Triton-X-100, 1 mM EDTA and complete protease inhibitor cocktail (Roche, Basileia, Switzerland) on ice. After 10000 g centrifugation for 10 minutes at 4°C, proteins were submitted to SDS-PAGE in a Criterion™ TGX™ Precast Gel

(4 -15%) (Bio-Rad, Hercules, California, USA) and transferred to nitrocellulose filters (Bio-Rad, Hercules, California, USA). The membrane was blocked with 5% powder milk in PBS (Dulbecco's Phosphate Buffer Saline (0.0095M) – BioWhittaker - Radnor, Pennsylvania, USA) for 30 minutes. Antibodies were from Cell Signaling Technology (1A5 anti-Thr^{389} phospho- p70 S6 kinase (S6K) mouse monoclonal; 9B11 anti-myc tag mouse monoclonal) or have been described previously (anti-TSC1 and -TSC2; van Slegtenhorst, Verhoef *et al.* (1999)). Secondary antibodies were from Li-Cor Biosciences (Lincoln, Nebraska, USA). Antibody dilutions were in PBS containing 0.1% Tween 20 (PBST) (Sigma-Aldrich Fine Chemicals, Poole, United Kingdom) according to Table 1. Primary or secondary antibody incubations were for 60 minutes. All washes were in PBST with tween. Blots were scanned using either the Odyssey scanner (Li-Cor Biosciences, Lincoln, Nebraska, USA) or the Gel Doc™ EZ System and Image Lab™ software (Bio-Rad, Hercules, California, USA). Signal intensities were measured and normalized to the signals corresponding to wild-type *TSC2*. The experiment was repeated at least three times and Student's *t* test was used for pairwise comparison of each variant to the WT-*TSC2* and the *TSC2* p.R611Q variant, a positive control acknowledged to disrupt the TSC1-TSC2 dimer (Nellist, Sancak *et al.* 2005). Tested DNA variant effect was considered positive if the ratio of S6K-Thr^{389}/total S6K is at least a four-fold compared to WT-*TSC2*.

Table 1: Monoclonal and polyclonal antibodies used to functional assessment of TSC2 DNA variants.

Color on Blot	Detected Protein	Mass (kDa)	Primary antibody/species in which it was generated	Dilution fold
Red	TSC2-GFP	225	Anti-TSC2 (1895*)/rabbit	1:10,000
Red	TSC2	200	Anti-TSC2 (1895*)/rabbit	1:10,000
Red	TSC1	130	Anti TSC1 (2197**)/rabbit	1:5,000
Green	Myc-Tag	130	9B11 anti-myc tag/mouse	1:5,000
Green	S6K-Thr389	70	1A5 anti-Thr389 phospho- S6K/mouse	1:2,000
Red	Total S6K	70	Anti-myc tag/rabbit	1:2,000

* Primary antibodies 1/15,000 dilution of 1895 (rabbit polyclonal against TSC2) from (van Slegtenhorst, Nellist *et al.* 1998)
** Primary antibodies 1/5,000 dilution of 2197 (rabbit polyclonal against TSC1) from (van Slegtenhorst, Nellist *et al.* 1998)

G. Splicing effect of *TSC2* missense and deep intronic variants

Partial genomic segments containing a DNA variant potentially affecting *TSC2* pre-mRNA splicing were cloned in the pCDNA3.1 vector, determining a mini-gene for overexpression in HEK293 cells that had not been genetically modified. The two *TSC2* variants tested, c.1361+54_1361+57del in intron 13 and c.4493G>T (p.S1498I) in exon 34, defined PCR amplicons containing genomic segments from exon 12 through exon 14, and from exon 33 through exon 35, respectively. Both mutated and wild-type *TSC2* alleles were separately cloned in the pcDNA3.1 vector following PCR amplification of DNA from the heterozygous individual for each DNA variant. Forward and reverse primers contained *BamH*I and *Xho*I or *NotI* restriction sites (Supplementary Table 8). Sanger sequencing of recombinant plasmidial DNA confirmed cloning of both wild-type alleles, the 4-bp deletion and c.4493G>T variant. HEK293 cells were transfected using lipofectamine 2000 (Thermo Fisher

Scientific, Waltham MA, USA). Eighteen hours after transfection, total RNA was isolated with TRIzol, treated with AmpGrade DNase I (1U/µL; Invitrogen, Carlsbad, CA, USA), and reversely transcribed (SuperScript III; Invitrogen, Carlsbad CA, USA). PCR amplification of the cDNA from the mini-gene transcript, with primers specific for the pcDNA3.1 vector (Supplementary Table 8), yielded products with expected sizes of 210 bp (non-recombinant clone); 596 bp (upon exon 12-14 mini-gene transcript constitutive splicing) or 458 bp (if internal exon is skipped due to the 4-bp deletion); and 896 bp (upon exon 33-35 mini-gene transcript constitutive splicing) or 408 bp (if internal exon is skipped due to the c.4493T mutant allele). Additionally, a nested PCR approach was adopted with primers annealing to *TSC2* exons 12 and 14, and 33 and 35 (Supplementary Table 8). Moreover, primers annealing to exons 8 and 10 were used to amplify *TSC2* endogenous transcript (Supplementary Table 8) as an internal control (275 bp). RT-PCR products were verified after electrophoresis in 1.5% agarose gel, and the images captured by Gel Doc™ EZ System and Image Lab™ software (BioRad, Hercules, California, EUA). All bands were isolated from agarose gel and submitted to Sanger sequencing in order to identify the sequence alteration caused by any splicing defect.

IV. Results

A. Pathogenic DNA variants

Our cohort of 115 patients with definite clinical diagnosis of TSC, consisted of 63 male and 54 female individuals. Seventy-five patients were from HC-UFPR, 23 patients from HSCM-SP, and 19 patients from GRAACC-SP. Age at the time of enrolment varied between five months and 19 years (mean: 10.4 years, median: 11 years).

Among the 115 patients, 28 had DNA previously analyzed by Sanger sequencing of the *TSC1* gene, having pathogenic alterations detected in seven of them (Dufner-Almeida, 2014). In the present work, the patient cohort and experimental approach have been significantly extended. Overall pathogenic DNA variants were identified in 99 patients (86.09%), 17 affecting *TSC1* (17.17%), and 82 on *TSC2* (82.82%), and a *TSC1*:*TSC2* mutation ratio of 1:4.8 (Figure 2A). The number of TSC pathogenic DNA variants is presented on Table 2 according to the hospital of patient origin.

Table 2: *TSC1* and *TSC2* pathogenic DNA variants according to hospital of patient origin.

Hospital	**HC-UFPR**	**HSCM-SP**	**GRAACC-SP**	**Total**
Number of patients enrolled	75	23	17	115
Patients with pathogenic DNA variants detected (percentage of total)	65 (86.67%)	22 (95.65%)	12 (63.16%)	99 (86,08%)
Patients with *TSC1* pathogenic DNA variants (percentage of total identified)	6 (9.37%)	9 (40.91%)	2 (16.67%)	17 (17.17%)
Patients with *TSC2* pathogenic DNA variants (percentage of total identified)	59 (90.76%)	13 (59.09%)	10 (83.33%)	82 (82.82%)

Frameshift (29 patients) and nonsense (29 patients) were the most frequent pathogenic variant categories. Complex mutation and large segmental gene duplication were observed once each in our cohort (Figure 2B).

Distinct *TSC1* mutations comprehended eight frameshift variants, five nonsense and one large deletion (Figure 2C and D, Table 3 and 4). *TSC1* exon 15 exhibited four nonsense and three frameshift variants; thus 41.18% of this gene pathogenic DNA variants (7/17). Two variants, Glu478Lysfs*53 and Arg500*, were identified in two unrelated patients (Table 3).

Distinct pathogenic *TSC2* DNA variants included 19 nonsense, 20 frameshift, 16 splicing and ten missense DNA variants, six large deletions, four in-frame deletions, one large duplication and one complex mutation. To our knowledge, 52 variants have not been previously reported. Although four of these have been deposited at LOVD *TSC2* database by others, no further description is available (Table 4). We observed a cluster of pathogenic variants in the GAP domain (GAP-Rheb residues 1562 to 1764) of TSC2, where 20 out of 82

pathogenic variants (24.69%) were identified on this segment, as previously studies (Jones, Shyamsundar *et al.* 1999, Niida, Lawrence-Smith *et al.* 1999)(Figure 4).

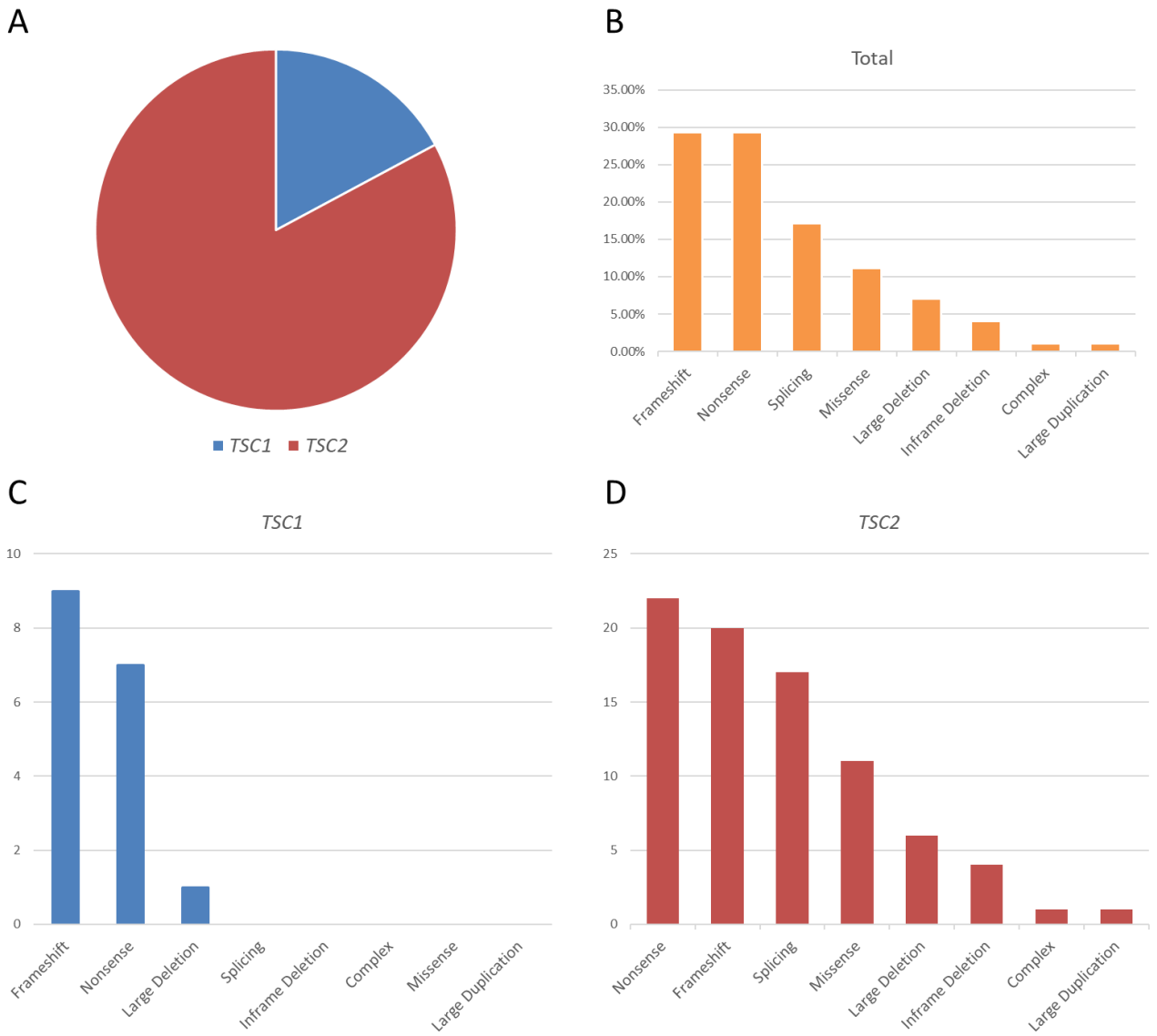


Figure 2: Overview of the distribution of *TSC1* and *TSC2* pathogenic DNA variants according to frequency and mutation type.

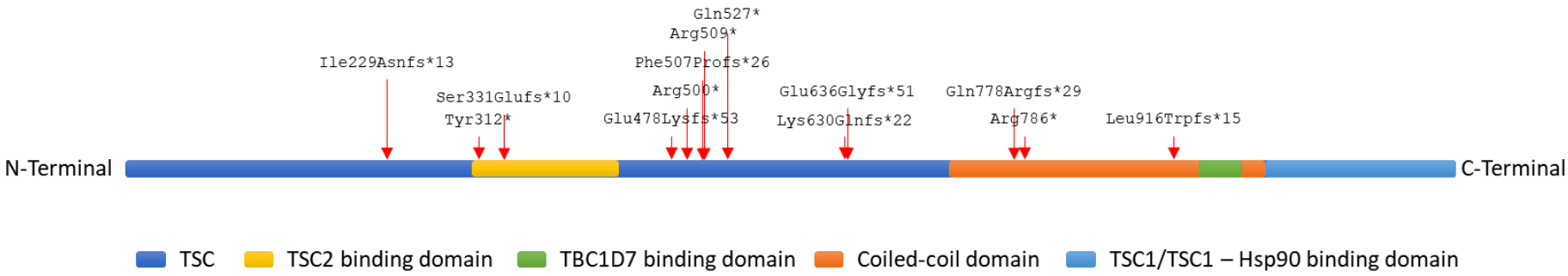


Figure 3: Schematic representation of TSC1 protein segments and distribution of all *TSC1* pathogenic variants according to the amino acid number.

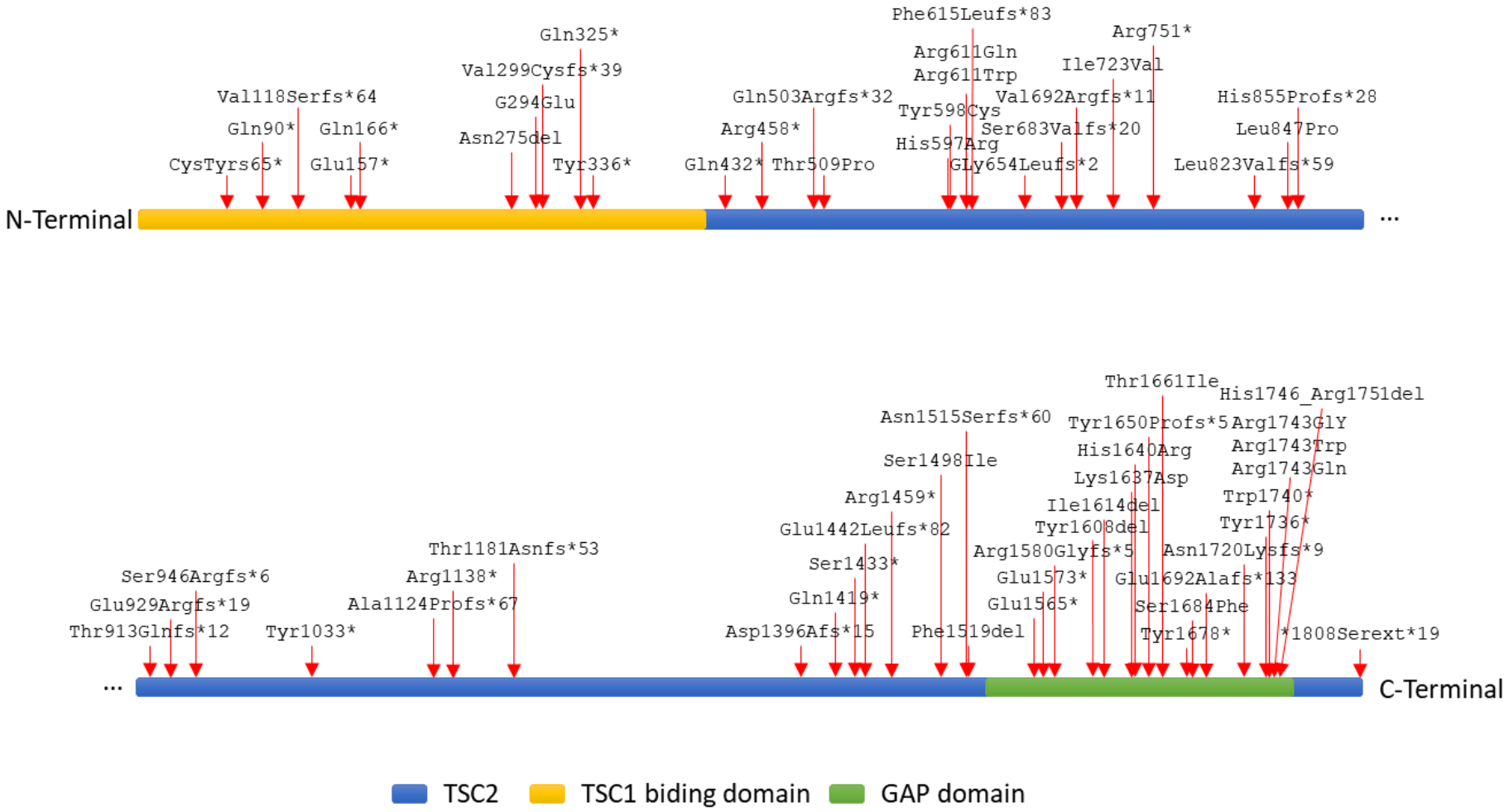


Figure 4: Schematic representation of TSC2 protein segments and distribution of all *TSC2* pathogenic variants according to the amino acid number. Splicing variants are not represented.

Table 3: TSC1 pathogenic DNA variants in a cohort of 115 TSC patients.

DNA variant type	DNA Variant	Location	Patient #	References [f]
Frameshift	c.683dupG (p.(Ile229Asnfs*13)) [a]	Exon 08	36	-
Frameshift	c.989dupT (p.(Ser331Glufs*10)) [a, g]	Exon 10	19	-
Frameshift	c.1431_1434del (p.(Glu478Lysfs*53)) [c, d, g]	Exon 14	13, 78	Au, 2007
Frameshift	c.1517_1518dupCC (p.(Phe507Profs*26)) [a, g]	Exon 15	17	-
Frameshift	c.1888_1891del (p.(Lys630Glnfs*22)) [c, d]	Exon 15	111	Rosset, 2017; Tyburczy, 2014; Tsai, 2011; Au, 2007; Hung, 2006
Frameshift	c.1907_1908del (p.(Glu636Glyfs*51))	Exon 15	122	Iyer, 2013; Muzykewicz, 2009; Dabora, 1998; Kwiatkowska, 1998; vanSlegtenhorst, 1997
Frameshift	c.2332del (p.(Gln778Argfs*29)) [a, c, d]	Exon 18	79	-
Frameshift	c.2746delC (p.(Leu916Trpfs*15)) [a, g]	Exon 21	10	-
Nonsense	c.936C>G (p.(Tyr312*)) [a, c, d]	Exon 10	93	-
Nonsense	c.1525C>T (p.(Arg509*)) [g]	Exon 15	5	Kwiatkowska, 1998 and 2015; Suspitsin, 2018; Ismail, 2017; Yang, 2016; Lee, 2014
Nonsense	c.1498C>T (p.(Arg500*)) [g]	Exon 15	8, 35	Tan, 2017; Tyburczy, 2015; Jansen, 2008; Au, 2007; Sancak, 2005
Nonsense	c.1579C>T (p.(Gln527*)) [c, d]	Exon 15	106	Voss, 2014; Iyer, 2013
Nonsense	c.2356C>T (p.(Arg786*)) [g]	Exon 18	16	Suspitsin, 2018; Lee, 2014; Au, 2007; Hung, 2006; Sancak, 2005
Nonsense	c.2626G>T (p.(E876*))	Exon 21	82	-
Large Deletion	c.(737+1_738-1)_(*1+_?)del [a, b, d]	Exon 9-23	29	-

[a] First description of this mutation.
[b] Next-generation sequencing and MLPA performed for both genes, besides Sanger sequencing.
[c] MLPA performed for both genes, besides Sanger sequencing
[d] Deletion confirmed by qPCR.
[e] On LOVD database without further information of submission.
[f] Five most recent references.
[g] Previously described by (Dufner-Almeida 2014).

Table 4: TSC2 pathogenic DNA variants in a cohort of 115 TSC patients.

	DNA variant type	DNA Variant	Location	Patient #	References [h]
1.	Nonsense	c.195T>A (p.(Cys65*)) [a]	Exon 03	4	-
2.	Nonsense	c.268C>T (p.(Gln90*))	Exon 04	105, 120	Kwiatkowski, 2015; Jansen, 2008; Au, 2007; Yuan, 2007; Hung, 2006
3.	Nonsense	c.469G>T(p.(Asp157*))		42	
4.	Nonsense	c.496C>T (p.(Gln166*)) [c]	Exon 06	94, 99	Niida, 1999
5.	Nonsense	c.973C>T (p.(Gln325*))	Exon 10	52	Jansen, 2008
6.	Nonsense	c.1008T>G (p.(Tyr336*))	Exon 11	115	Rosset, 2017
7.	Nonsense	c.1294C>T (p.(Gln432*)) [a]	Exon 13	28	-
8.	Nonsense	c.1372C>T (p.(Arg458*))	Exon 14	108	Cai, 2017; Kwiatkowski, 2015; Tyburczy, 2015; Tsai, 2011; Au, 2007
9.	Nonsense	c.2251C>T (p.(Arg751*))	Exon 21	1	Cai, 2017; Yang, 2016; Kwiatkowski, 2015; Yang, 2014; Kacerovska, 2012
10.	Nonsense	c.3099C>G (p.(Tyr1033*)) [a]	Exon 27	9	-
11.	Nonsense	c.3412C>T (p.(Arg1138*))	Exon 30	92	Cai, 2017; Kwiatkowski, 2015; Tyburczy, 2014; Hung, 2006; Rendtorff, 2005
12.	Nonsense	c.4298C>A (p.(Ser1433*)) [a]	Exon 34	26	-
13.	Nonsense	c.4375C>T (p.(Arg1459*)) [a]	Exon 34	33	-
14.	Nonsense	c.4255C>T (p.(Gln1419*))	Exon 34	67	Cai, 2017; Au, 2007; Roberts, 2004; Langkau, 2002; Jones, 1999
15.	Nonsense	c.4693G>T (p.(Glu1565*)) [a]	Exon 37	32	-
16.	Nonsense	c.4717G>T (p.(Glu1573*)) [a]	Exon 37	101	-
17.	Nonsense	c.5034C>A (p.(Tyr1678*)) [a]	Exon 39	18	-
18.	Nonsense	c.5208C>G (p.(Tyr1736*)) [a]	Exon 41	48	-
19.	Nonsense	c.5220G>A (p.(Trp1740*))	Exon 41	90	Kwiatkowski, 2015; Roberts, 2002
20.	Frameshift	c.352del (p.(Val118Serfs*64)) [a, c]	Exon 05	83	-
21.	Frameshift	c.894dup (p.(Val299Cysfs*39))	Exon 10	123	Giannikou, 2016
22.	Frameshift	c.1507del (p.(Gln503Argfs*32)) [a, c]	Exon 15	91	-
23.	Frameshift	c.1842del (p.(Phe615Leufs*83)) [a]	Exon 18	22	-
24.	Frameshift	c.2073dup (p.(Val692Argfs*11)) [a]	Exon 19	41	-
25.	Frameshift	c.2046dup (p.(Ser683Valfs*20)) [a]	Exon 19	71	-
26.	Frameshift	c.1959_1960del (p.(Gly654Leufs*2))	Exon 19	103	Li, 2011

27.	Frameshift	c.2467_2476delinsGTGGATGA (p.(Leu823Valfs*59)) [a, b]	Exon 21	20	-
28.	Frameshift	c.2563dup (p.(His855Profs*28)) [a]	Exon 23	73	-
29.	Frameshift	c.2737_2739delinsC (p.(Thr913Glnfs*12)) [a]	Exon 24	3	-
30.	Frameshift	c.2784del (p.(Glu929Argfs*19)) [a]	Exon 25	39	-
31.	Frameshift	c.3370del (p.(Ala1124Profs*67)) [a]	Exon 29	110	-
32.	Frameshift	c.3541dup (p.(Thr1181Asnfs*53)) [a]	Exon 30	86	-
33.	Frameshift	c.4324_4327delinsCTTCT (p.(Glu1442Leufs*82)) [a]	Exon 34	57	-
34.	Frameshift	c.4187del (p.(Asp1396Alafs*15)) [a, c]	Exon 34	95	-
35.	Frameshift	c.4544_4547del (p.(Asn1515Serfs*60))	Exon 35	69	Kwiatkowski, 2015; Li, 2011; Au, 2007; Hung, 2006; Ali, 2005
36.	Frameshift	c.4738del (p.(Arg1580Glyfs*5)) [a]	Exon 37	112	-
37.	Frameshift	c.4947_4948insCCATTGT (p.(Tyr1650Profs*5)) [a]	Exon 38	31	-
38.	Frameshift	c.5159dup (p.(Asn1720Lysfs*9)) [a]	Exon 39	11	-
39.	Frameshift	c.5075_5078del (p.(Glu1692Alafs*133))	Exon 40	49	Roberts, 2004
40.	Missense	c.1790A>G (p.(His597Arg))	Exon 17	68	Wang, 2016; Hardy, 2012; vanEeghen, 2012; Dunlop, 2011;
41.	Missense	c.1793A>G (p.(Tyr598Cys))	Exon 17	109	Hoogeveen-Westerveld, 2013
42.	Missense	c.1832G>A (p.(Arg611Gln)) [c]	Exon 17	38	Suspitsin, 2018; Cai, 2017; Ishikawa, 2017; Kwiatkowski, 2015; Liu, 2015
43.	Missense	c.1831C>T (p.(Arg611Trp))	Exon 17	58	Kwiatkowski, 2015; Tyburczy, 2015; You, 2013; vanEeghen, 2012; Hoogeveen-Westerveld, 2011
44.	Missense	c.4909_4911delinsGAC (p.(Lys1637Asp)) [a]	Exon 38	85	-
45.	Missense	c.4919A>G (p.(His1640Arg)) [a]	Exon 38	72	-
46.	Missense	c.5024C>T (p.P1675L)	Exon 39	64, 65	van Eeghen, 2012; Hoogeveen-Westerveld, 2013; Sancak, 2005
47.	Missense	c.5227C>G (p.(Arg1743Gly)) [a, c]	Exon 41	43	Zhang, 2013; Kwiatkowski,2015
48.	Missense	c.5227C>T (p.(Arg1743Trp))	Exon 41	77	Papadopoulou, 2018; Suspitsin, 2018; Cai, 2017; Yu, 2017; Bai, 2016
49.	Missense	c.5228G>A (p.(Arg1743Gln))	Exon 41	124	Papadopoulou, 2018; Ismail, 2017; Bai, 2016;

						Kwiatkowski, 2015; Tyburczy, 2015
50.	In-frame deletion	c.824_826del (p.(Asn275del))		Exon 09	88	Hoogeveen-Westerveld, 2011; Coevoets, 2009; Sancak, 2005
51.	In-frame deletion	c.4842_4844del (p.(Ile1614del)) [a]		Exon 37	37	Kwiatkowski, 2015; van Eeghen, 2012, Sancak, 2005, Dabora, 2001
52.	In-frame deletion	c.4823_4825del (p.(Tyr1608del)) [a]		Exon 37	40	Kwiatkowski, 2015; Hayashi, 2012; van Eeghen; 2012 Au, 1998 and 2007
53.	In-frame deletion	c.5238_5255del (p.(His1746_Arg1751del))		Exon 41	81	Hoogeveen-Westerveld, 2011; Qin, 2011; Tsai, 2011; Au, 2007; Choi, 2006
54.	Large Deletion	c.(774+1_775-1)_(848+1_849-1)del [a, d]		Exon 9	97	-
55.	Large Deletion	c.(975+628)_(1716+41)del [a, b, d]		Exon 11-16	12	-
56.	Large Deletion	c.(1599+1_1600-1)_(2545+1_2546-1)del [d]		Exon 16 - 22	59	Longa, 2001
57.	Large Deletion	c.(1716+1_1717-1)_(2545+1_2546-1)del [a, d]		Exon 17 - 22	84	-
58.	Large Deletion	c.(5068+1_5069-1)_(*102_?)del [d, g]		Exon 39 - *PKD1*	51	-
59.	Large Deletion	c.(?_-30)_(*102_?)del [a, d]		*TSC2*	61	-
60.	Large Duplication	c.(2355+1_2356-1)_(*102_?)dup [d]		Exon 22 - 42	89	-
61.	Complex	c.5423G>C (p.(*1808Serext*19)) [c]		Exon 42	76	-
	Splicing variants	**DNA Variant**	**Predicted effects: [r.spl?] [e] or [r.(spl?)] [f]**	**Location [g]**	**Patient #**	**References**
62.	Splicing	c.337-1G>A [b]	[r.spl?]: exon 5 skipping, p.(Gly113Leufs*20)	Intron 04	2	Niida, 2013
63.	Splicing	c.481+2T>C [a]	[r.spl?]: exon 5 skipping, p.(Gly113Leufs*20)	Intron 05	27	-
64.	Splicing	c.482-1G>A + c.482-3C>T [a, i]	[r.spl?]: exon 6 skipping, p.(Gly113Leufs*20)	Intron 05	114	-
65.	Splicing	c.600-2A>G	[r.spl?]: exon 7 skipping, p.(Met201Glyfs*18)	Intron 06	74	Suspitsin, 2018; Yu, 2017; Rendtorff, 2005; Sancak, 2005; Langkau, 2002
66.	Splicing	c.775-2A>G [a]	[r.spl?]: exon 9 skipping, p.()	Intron 08	75	-
67.	Splicing	c.848+4_848+9del [a]	[r.(spl?)]: exon 9 skipping, p.(Leu259Serfs*52)	Intron 09	6	-
68.	Splicing	c.1717-1G>A [a]	[r.spl?]: exon 17 skipping, p.()	Intron 16	102	-
69.	Splicing	c.1947-2_ 1947-1delinsCC [a]	[r.spl?]: exon 19 skipping, p.(Glu650Glyfs*3)	Intron 18	45	-
70.	Splicing	c.2355+2_2355+5del	[r.spl?]: exon 21skipping, p.()	Intron 21	98	Rosset, 2017; Kwiatkowski, 2015; LeCaignec, 2009; Sancak, 2005; Jones, 1999
71.	Splicing	c.2545+1G>A [g]	[r.spl?]: exon 22 skipping, p.()	Intron 22	62	-

72.	Splicing	c.2639+1G>A [g]	[r.spl?]: exon 23 skipping, p.()	Intron 23	121	-
73.	Splicing	c.2838-122G>A (p.(S946Rfs*6)) [c]	[r.(2837_2838ins2838-120_2838-1)], p.(Ser946Argfs*6)	Intron 25	34, 47	Nellist,2015
74.	Splicing	c.3132-3T>G (p.(1044del50) [a, c]	[r.(spl?)]: in-frame exon 28 skipping, p.(1044del50)	Intron 27	7	-
75.	Splicing	c.4493+1G>C [a]	[r.spl?]: exon 34 skipping, p.()	Intron 34	70	-
76.	Splicing	c.5160+1G>C [a]	[r.spl?]: exon 40 skipping, p.(Asp1690Aspfs*6)	Intron 40	23	-
77.	Splicing	c.5160+1G>A	[r.spl?]: exon 40 skipping, p.(Asp1690Aspfs*6)	Intron 40	24	Sancak, 2005; Dabora, 2001
			Variants of unclear significance (VUS)			
78.	Missense	c.1525A>C (p.(Thr509Pro)) [g, j]		Exon 15	118	-
79.	Missense	c.1663G>C (p.(Ala555Pro)) [j]		Exon 16	50	
80.	Missense	c.2167A>G (p.(Ile723Val)) [a, j]		Exon 20	53	-
81.	Missense	c.2540T>C (p.(Leu847Pro)) [j]		Exon 22	46	Suspitsin, 2018; Hoogeveen-Westerveld, 2012; Li, 2011
82.	In-frame deletion	c.4527_4529del (p.(Phe1510del)) [j]		Exon 35	96	
83.	Splicing	c.4493G>T [j]		Exon34	30	-

[a] First description of this mutation.
[b] Next-generation sequencing and MLPA performed for both genes, besides Sanger sequencing.
[c] MLPA performed for both genes, besides Sanger sequencing.
[d] Deletion or duplication confirmed by qPCR.
[e] [r.spl?]: RNA was not analysed but the change is expected to affect splicing.
[f] [r.(spl?)]: RNA was not analysed but the change might affect splicing.
[g] On LOVD database without further information of submission.
[h] Five most recent references.
[I] Found in *trans*
[j] Under functional analysis

B. *TSC2* missense and in-frame deletion variants

Initially, 18 *TSC2* putative missense variants and five in-frame deletions had general population frequency lower than 0,1, and were predicted to be disease-causing mutations by Mutation Taster, PolyPhen, PROVEAN, SIFT and Humas Splicing Finder bioinformatics web tools (Table 5). Among the missense variants, eight have been previously reported by other groups: His597Arg, Tyr598Cys, Arg611Trp, Arg611Gln, Leu847Pro, Arg1743Gly, Arg1743Trp and Arg1743Gln (Table 4). Arg611Trp, Arg611Gln, Arg1743Trp and Arg1743Gln have been reported in family cases (Wilson, Ramesh *et al.* 1996, Beauchamp, Banwell *et al.* 1998, Niida, Lawrence-Smith *et al.* 1999, Nellist, Sancak *et al.* 2005, Au, Williams *et al.* 2007, Hoogeveen-Westerveld, Wentink *et al.* 2011). Pathogenicity had been confirmed by functional assays for six cases: His597Arg, Arg611Trp, Arg611Gln, Arg1743Trp and Arg1743Gln (Nellist, Sancak *et al.* 2005, Hoogeveen-Westerveld, Wentink *et al.* 2011); and Lys1637Asp (M. Nellist, Erasmus Medical Centre, The Netherlands, personal communication). Moreover, TSC2 Tyr598Cys has been reported as likely pathogenic because its functional assay had yielded a T389/S6K ratio of intermediate value. It was significantly higher than that of wild-type TSC2, but significantly lower than that of pathogenic TSC2 Arg611Gln (Hoogeveen-Westerveld, Ekong *et al.* 2013). DNA variants Leu847Pro and Arg1743Gly had been reported as pathogenic without further information (Li, Zhou *et al.* 2011, Hoogeveen-Westerveld, Ekong *et al.* 2012, Kwiatkowski, Palmer *et al.* 2015). Finally, variants Thr1661Ile and Ser1684Phe had yielded non-pathogenic evidence in the mTOR-based functional assay (M. Nellist, Erasmus Medical Centre, The Netherlands, personal communication). In summary, nine out of 18 missense DNA variants identified in our cohort with significantly low or null frequency in population databases had been functionally assessed, yielding six pathogenic and two non-pathogenic results (Figure 5).

Table 5: 18 putative missense variants and five in-frame deletions on TSC2 analysis by Mutation Taster, PolyPhen, PROVEAN, and SIFT and Human Splicing Finder bioinformatics web tools

Variante	Pacientes	Mutation Taster	PolyPhen	PROVEAN	Sift	Human Splicing Finder	
						Interpretation	Prediction algorithm
c.52C>G (p.L18V)	30	**Disease causing**, protein features (might be) affected (prob: 0.9914)	**Benign** (pph2_prob = 0.089; pph2_FPR = 0.149; pph2_TPR = 0.932)	**Neutral** (score: -1.08)	**Damaging** (score: 0.044) Median sequence conservation: 3.20	**New ESS Site.** Creation of an exonic ESS site. **Potential alteration of splicing.** **ESE Site Broken**. Alteration of an exonic ESE site. **Potential alteration of splicing.**	IIEs from Zhang et al., Sironi et al., Motif 2, ESE-Finder - SC35, ESR Sequences from Goren et al.
c.499T>C (p.W167R)	33	**Disease causing**, protein features (might be) affected, splice site changes (prob: 0.9999)	**Probably damaging** (pph2_prob = 0.998; pph2_FPR = 0.0112; pph2_TPR = 0.273)	**Deleterious** (score: -11.61)	**Damaging** (score: 0.002) Median sequence conservation: 2.93	**ESE Site Broken.** Alteration of an exonic ESE site. **Potential alteration of splicing.**	ESE-Finder - SF2/ASF(Ig), EIEs from Zhang et al.
c.1525A>C (p.T509P)	118	**Disease causing**, protein features (might be) affected, splice site changes (prob: 0.9999)	**Probably damaging** (pph2_prob = 0.998; pph2_FPR = 0.0112; pph2_TPR = 0.273)	**Deleterious** (score: -3.04)	**Damaging** (score: 0.014) Median sequence conservation: 2.93	**ESE Site Broken.** Alteration of an exonic ESE site. Potential alteration of splicing.	HSF Matrices - 9G8, ESR Sequences from Goren et al., ESE-Finder - SF2/ASF
c.1663G>C (p.A555P)	50	**Disease causing**, protein features (might be) affected, splice site changes (prob: 0.9995)	**Probably damaging** (pph2_prob = 1; pph2_FPR = 0.00026; pph2_TPR = 0.00018)	**Neutral** (score: -1.76)	**Tolerated** (score: 0.056) Median sequence conservation: 3.45	**No significant splicing motif alteration detected.** This mutation has probably no impact on splicing.	
c.1790A>G (p.H597R)	68, 80	**Disease causing**, known disease mutation at this position (HGMD CM114651), protein features (might be) affected, splice site changes (prob: 0.9999)	**Probably damaging** (pph2_prob = 0.99; pph2_FPR = 0.0338; pph2_TPR = 0.719)	**Deleterious** (score: -6.77)	**Damaging** (score: 0.003) Median sequence conservation: 2.93	**ESE Site Broken.** Alteration of an exonic ESE site. Potential alteration of splicing.	ESE-Finder - SRp40, HSF Matrices - 9G8, RESCUE ESE Hexamers
c.1793A>G (p.Y598C)	109	**Disease causing**, (prob: 0.9999)	**Probably damaging** (pph2_prob = 1; pph2_FPR = 0.00026; pph2_TPR = 0.00018)	**Deleterious** (score: -7.93)	**Damaging** (score: 0.003) Median sequence conservation: 2.93	**ESE Site Broken.** Alteration of an exonic ESE site. Potential alteration of splicing.	HSF Matrices - 9G8, EIEs from Zhang et al., RESCUE ESE Hexamers, ESE-Finder - SRp40, ESR Sequences from Goren et al.
c.1831C>T (p.R611W)	58	**Disease causing**, known disease mutation at this position (HGMD CM961387) (prob: 0.9999)	**Probably damaging** (pph2_prob = 1; pph2_FPR = 0.00026; pph2_TPR = 0.00018)	**Deleterious** (score: -7.26)	**Damaging** (score: 0.003) Median sequence conservation: 2.93	**No significant splicing motif alteration detected.** This mutation has probably no impact on splicing.	

c.1832G>A (p.R611Q)	38	**Disease causing**, known disease mutation at this position (HGMD CM981945), known disease mutation: rs28934872 (pathogenic), protein features (might be) affected, splice site changes (prob: 0.9999)	**Probably damaging** (pph2_prob = 0.999; pph2_FPR = 0.00574; pph2_TPR = 0.136)	**Deleterious** (score: -3.52)	**Damaging** (score: 0.003) Median sequence conservation: 2.93	**New ESS Site.** Creation of an exonic ESS site. **Potential alteration of splicing.** **ESE Site Broken.** Alteration of an exonic ESE site. **Potential alteration of splicing.**	Sironi et al. - Motif 3, ESR Sequences from Goren et al., ESE-Finder - SF2/ASF, ESE-Finder - SF2/ASF(Ig), PESE Octamers from Zhang & Chasin, ESE-Finder - SRp55, ESR Sequences from Goren et al.
c.2167A>G (p.I723V)	53	**Disease causing**, protein features (might be) affected, splice site changes (prob: 0.9642)	**Benign** (pph2_prob = 0.07; pph2_FPR = 0.157; pph2_TPR = 0.935)	**Neutral** (score: -0.38)	**Tolerated** (score: 0.379) Median sequence conservation: 2.93	**No significant splicing motif alteration detected.** This mutation has probably no impact on splicing.	
c.2540T>C (p.L847P)	46	**Disease causing**, known disease mutation at this position (HGMD CM091031), protein features (might be) affected, splice site changes (prob: 0.9999)	**Probably damaging** (pph2_prob = 0.999; pph2_FPR = 0.00574; pph2_TPR = 0.136)	**Deleterious** (score: -6.22)	**Damaging** (score: 0.001) Median sequence conservation: 2.93	**ESE Site Broken.** Alteration of an exonic ESE site. **Potential alteration of splicing.**	ESR Sequences from Goren et al., EIEs from Zhang et al.
c.4493G>T (p.S1498I)	30	**Disease causing**, amino acid sequence changed, known disease mutation at this position (HGMD CM991216), protein features (might be) affected, splice site changes (prob:0.9999)	**Probably damaging** (pph2_prob = 0.995; pph2_FPR = 0.0277; pph2_TPR = 0.681)	**Deleterious** (score: -5.07)	**Damaging** (score: 0.001) Median sequence conservation: 2.92	**Broken WT Donor Site.** Alteration of the WT donor site, **most probably affecting splicing.** **ESE Site Broken.** Alteration of an exonic ESE site. **Potential alteration of splicing.**	HSF Matrices, MaxEnt, ESE-Finder - SC35, ESE-Finder - SF2/ASF(Ig), ESE-Finder - SRp40, EIEs from Zhang et al.
c.4909_4911delinsGAC (p.K1637D)	85	**Disease causing**, known disease mutation at this position (HGMD CD067216), protein features (might be) affected (prob:0.9999)	**Probably damaging** (pph2_prob = 0.999; pph2_FPR = 0.00574; pph2_TPR = 0.136)	**Deleterious** (score: -5.15)	**Damaging** (score: 0.001) Median sequence conservation: 2.92		
c.4919A>G (p.H1640R)	72	**Disease causing**, protein features (might be) affected, splice site changes (prob:0.9999)	**Probably damaging** (pph2_prob = 0.995; pph2_FPR = 0.0277; pph2_TPR = 0.681)	**Deleterious** (score: -6.87)	**Damaging** (score: 0.001) Median sequence conservation: 2.92	**ESE Site Broken.** Alteration of an exonic ESE site. **Potential alteration of splicing.**	ESR Sequences from Goren et al., ESE-Finder - SF2/ASF
c.4982C>T (p.T1661I)	113	**Disease causing**, protein features (might be) affected (prob:0.9999)	**Probably damaging** (pph2_prob = 1; pph2_FPR = 0.00026; pph2_TPR = 0.00018)	**Deleterious** (score: -5.09)	**Damaging** (score: 0.001) Median sequence conservation: 2.92	**ESE Site Broken.** Alteration of an exonic ESE site. **Potential alteration of splicing.**	ESE-Finder - SF2/ASF, EIEs from Zhang et al.

c.5051C>T (p.S1684F)	116	**Disease causing**, protein features (might be) affected (prob:0.9973)	**Possibly** (pph2_prob = 0.808; pph2_FPR = 0.0715; pph2_TPR = 0.841)	**Neutral** (score: -1.08)	**Damaging** (score: 0.002) Median sequence conservation: 2.92	**New ESS Site.** Creation of an exonic ESS site. **Potential alteration of splicing.** **ESE Site Broken.** Alteration of an exonic ESE site. **Potential alteration of splicing.**	Fas-ESS hexamers, IIEs from Zhang et al., ESR Sequences from Goren et al., HSF Matrices - 9G8, ESE-Finder - SRp40, EIEs from Zhang et al., ESR Sequences from Goren et al.
c.5227C>G (p.R1743G)	43	**Disease causing**, known disease mutation at this position (HGMD CM052391), protein features (might be) affected, splice site changes (prob:0.9999)	**Probably damaging** (pph2_prob = 0.996; pph2_FPR = 0.0222; pph2_TPR = 0.545)	**Deleterious** (score: -4.54)	**Damaging** (score: 0.000) Median sequence conservation: 2.91	**New ESS Site.** Creation of an exonic ESS site. **Potential alteration of splicing.**	Sironi et al. - Motif 2, IIEs from Zhang et al.
c.5227C>T (p.R1743W)	77	**Disease causing**, known disease mutation at this position (HGMD CM052391), protein features (might be) affected, splice site changes (prob:0.9999)	**Probably damaging** (pph2_prob = 1; pph2_FPR = 0.00026; pph2_TPR = 0.00018)	**Deleterious** (score: -7.26)	**Damaging** (score: 0.000) Median sequence conservation: 2.91	**New ESS Site.** Creation of an exonic ESS site. **Potential alteration of splicing.**	IIEs from Zhang et al., ESR Sequences from Goren et al.
c.5228G>A (p.R1743Q)	124	**Disease causing,** known disease mutation at this position (HGMD CM983892 and HGMD CM991217), protein features (might be) affected, splice site changes (prob:0.9999)	**Probably damaging** (pph2_prob = 1; pph2_FPR = 0.00026; pph2_TPR = 0.00018)	**Deleterious** (score: -3.63)	**Damaging** (score: 0.000) Median sequence conservation: 2.91	**New Acceptor Site.** Activation of an exonic cryptic acceptor site, with presence of one or more cryptic branch point(s). **Potential alteration of splicing.**	HSF Matrices

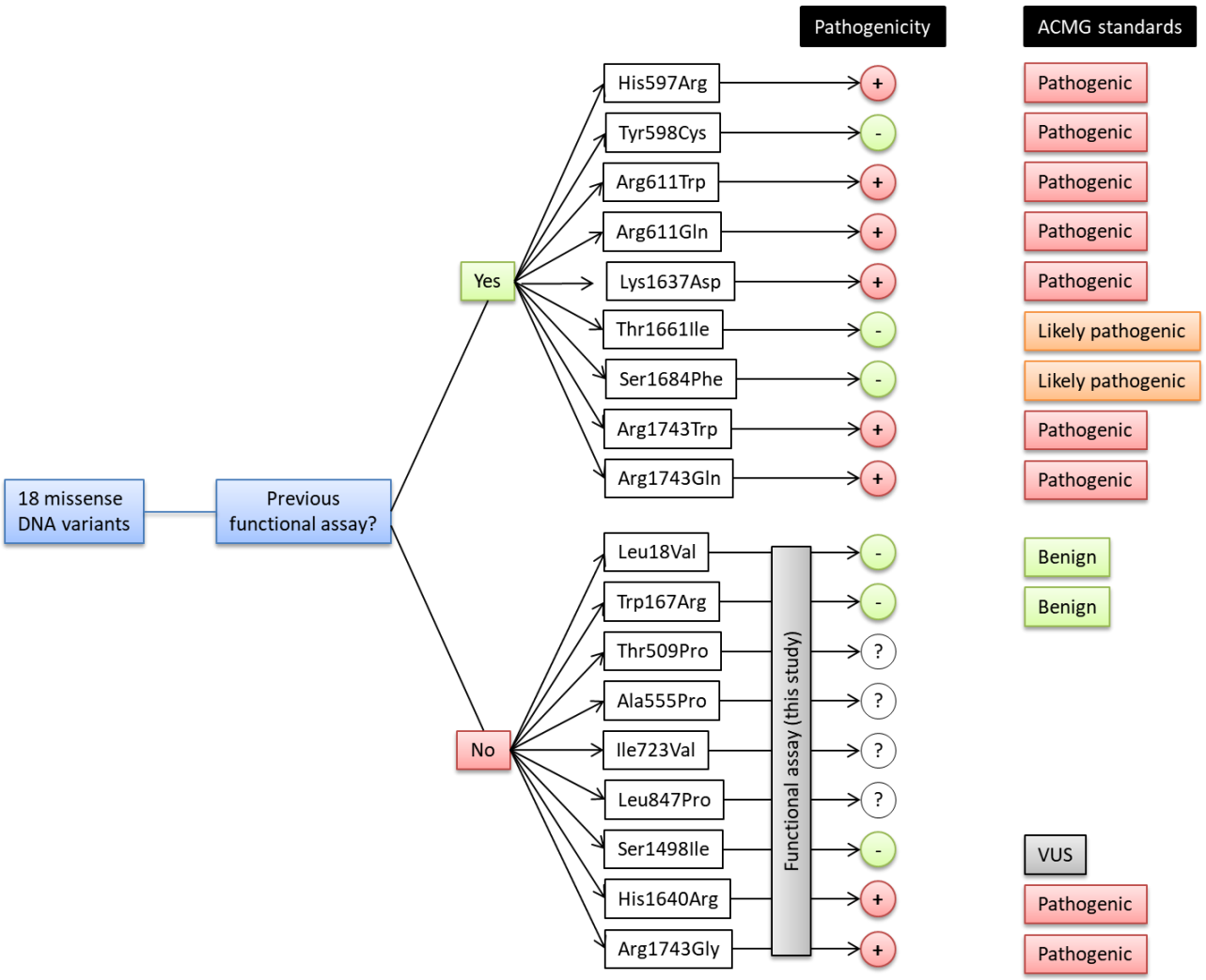


Figure 5: Workflow of the pathogenicity classification of 18 missense DNA variants detected on TSC2. VUS: variant of uncertain significance. ?: Variants under functional investigation.

We adopted the American College of Medical Genetics and Genomics (ACMG, Bethesda, MD, USA) criteria for pathogenicity assessment of missense DNA variants and criteria combination, as illustrated in Table 6 and 6, andSupplementary Table 9 and 11 (Richards, Aziz *et al.* 2015). Besides the six variants that had kept mTOR activated in the functional assay, two others (Tyr598Cys and Arg1743Gly) fulfilled the criteria for pathogenicity (Supplementary Table 9). Moreover, the two variants that inactivated mTOR in the functional assay (Thr1661Ile and Ser1684Phe) classified as likely pathogenic. Among the eight cases of missense DNA variants for which no functional data were available, three were likely benign (Leu18Val, Trp167Arg and Ala555Pro). Two of them (Trp167Arg and Ala555Pro) were from patients with *TSC1* or *TSC2* nonsense variants, one detected in this study and the other from a case available at LOVD. The third likely benign variant Leu18Val was from the same patient with the missense variant Ser1498Ile. Four additional missense DNA variants

classified as likely pathogenic (Thr509Pro, Ile723Val, Leu847Pro and His1640Arg), and Ser1498Ile was considered a variant of uncertain significance, according to ACMG criteria (Table 6 and Figure 5).

Five in-frame deletions identified by Sanger sequencing in the *TSC2* gene were not detected in any databases consulted. All in-frame deletions have been reported before (Table 4). Asn275del and His1746_Arg1751del had been functionally assessed displaying T389/S6K phosphorylation ratios significantly higher than that of wild-type TSC2 (Sancak, Nellist *et al.* 2005, Hoogeveen-Westerveld, Wentink *et al.* 2011). DNA variants c.4842_4844del (p.(Ile1614del)) and c.4823_4825del (p.(Tyr1608del)) have been described as pathogenic (Au, Rodriguez *et al.* 1998, Dabora, Jozwiak *et al.* 2001, Sancak, Nellist *et al.* 2005, Au, Williams *et al.* 2007) but no functional assay had been conducted. These two in-frame deletion variants, Ile1614del and Tyr1608del, were predicted to be disease-causing mutations by Mutation Taster and PROVEAN bioinformatics web tools. The last in frame deletion, c.4527_4529del (p.(Phe1510del)) on exon 35, was not presented in 1000 Genomes Project nether gnomAD databases, but according to ClinVar as benign or likely_benign. On *TSC2*LOVD database, this variant have been reported more than 25 times and classified as probable neutral according a functional assessment which do not find a significant S6K T389 phosphorylation higher than wild type TSC2 (Sancak, Nellist *et al.* 2005, Hoogeveen-Westerveld, Wentink *et al.* 2011).

Table 6: Pathogenicity criteria for classification of TSC2 missense DNA variants with low population frequency (Richards, Aziz et al. 2015)

		***TSC2* DNA variants with significantly low frequency in the general population**	**Missense DNA variants**																**In-frame deletions**				
			c.52C>G (p.L18V)	c.499T>C (p.W167R)	c.1663G>C(p.Ala555Pro))	c.1790A>G (p.(His597Arg))	c.1793A>G (p.(Tyr598Cys))	c.1831C>T (p.(Arg611Trp))	c.1832G>A (p.(Arg611Gln))	c.2167A>G (p.(Ile723Val))	c.2540T>C (p.(Leu847Pro))	c.4493G>T (p.(Ser1498Ile))	c.4909_4911delinsGAC (p.(Lys1637Asp))	c.4919A>G (p.(His1640Arg))	c.4982C>T (p.(Thr1661Ile))	c.5227C>G (p.(Arg1743Gly))	c.5227C>T (p.(Arg1743Trp))	c.5228G>A (p.(Arg1743Gln))	c.824_826del (p.(Asn275del))	c.4527_4529del (p.(Phe1510del))	c.4842_4844del (p.(Ile1614del))	c.4823_4825del (p.(Tyr1608del))	c.5238_5255del (p.(His1746_Arg1751del))
Evidence strength		**Evidence description** **Patient Number:**	30	33	50	68	109	58	38	53	46	30	85	72	113	43	77	124	88	96	37	40	81
Very strong evidence	PVS1	Null variant	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
Strong evidence	PS1	Same amino acid change previously described as pathogenic regarldess of nucleotide change	NO	NO	NO	**YES**	**YES**	**YES**	**YES**	NO	NO	NO	NO	NO	NO	**YES**	**YES**	**YES**	NO	NO	**NO**	**YES**	**YES**
	PS2	*De novo* in patient (both maternity and paternity confirmed)	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	PS3	*In vitro* or *in vivo* functional assay supportive of damaging effect	**NO**	**NO**	NI	**YES**	**NO**	**YES**	**YES**	NI	NI	**NO**	**YES**	**YES**	**NO**	**YES**	**YES**	**YES**	**YES**		**YES**	**YES**	**YES**
	PS4	Prevalence of variant in affected individual significantly increased compared to controls	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Moderate evidence	PM1	Located in a mutational hot spot or critical domain*	NO	NO	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	NO	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	NO	NO	**YES**	**YES**	**YES**
	PM2	Absent from controls**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**NO**	**YES**	**YES**	**YES**

	PM3	For recessive disorders, detected in trans with a pathogenic variant.	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	PM4	Protein length changes as a result of in-frame del/ins in a non repeat	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
	PM5	Novel missense in residue with different missense previously reported as pathogenic.	NO	NO	NO	NO	**YES**	**YES**	**YES**	NO	NO	NO	NO	**YES**	NO	**YES**	**YES**	**YES**	NO	NO	NO	NO	NO
	PM6	Assumed de novo, but without confirmation of paternity and maternity																					
Supporting evidence	PP1	Cosegregation with disease in multiple affected family members	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI	NI
	PP2	Gene has low rate of benign missense variants	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
	PP3	Multiple lines of computational evidence support a deleterious effect on protein	**NO**	**YES**	NO	**YES**	**YES**	**YES**	**YES**	NO	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	YES	NO	**YES**	**YES**	NO
	PP4	Patient phenotype or family history is highly specific for a disease with single gene etiology	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**	**YES**
	PP5	Reported as pathogenic by reputable source but evidence is unavailable	NO	NO	NO	**YES**	**NO**	**YES**	**YES**	NO	NO	NO	**YES**	NO	NO	NO	**YES**	**YES**	NO	NO	NO	NO	NO
Criteria combined before functional assay:			**LB**	**LB**	**LB**	**NA**	**NA**	**NA**	**NA**	**VUS**	**LP**	**VUS**	**NA**	**LP**	**NA**	**P**	**NA**	**NA**	NA	VUS	LP	LP	NA
Criteria combined after functional assay:			**B**	**B**	*******	**P**	**P**	**P**	**P**	**LP**	**LP**	**VUS**	**P**	**P**	**LP**	**P**	**P**	**P**	LP		P	P	P
Functional assay performed by			**DA**	**DA**	**MN**	**FR**	**FR**	**FR**	**FR**	**MN**	**MN**	**DA**	**MN**	**DA**	**MN**	**DA**	**FR**	**FR**	MN	MN	DA	DA	MN

NO: negative criterion; YES: positive criterion; NA: criterion does not apply; NI: criterion not investigated

P: pathogenic; LP: likely pathogenic; B: benign; LB: likely benign; VUS: variant of uncertain significance.

DA: Dufner-Almeida, this study. MN: Mark Nellist, personal communication. FR: former reports as referenced on the text.

*TSC2 domains (approximate amino acid numbers): DUF3384 (55-469); Tuberin domain - hamartin binding (555-903); Rap-GAP domain (1562-1748).

**1000 genomes, Exome Variation and EXAc.

***Waiting for NGS data NGS.

Table 7: Pathogenicity criteria for classification of TSC2 in frame deletions DNA variants with low population frequency (Richards, Aziz et al. 2015)

		***TSC2* DNA variants with significantly low frequency in the general population**	**In-frame deletions**				
			c.824_826del (p.(Asn275del))	c.4527_4529del (p.(Phe1510del))	c.4842_4844del (p.(Ile1614del))	c.4823_4825del (p.(Tyr1608del))	c.5238_5255del (p.(His1746_Arg1751del))
Evidence strength		**Evidence description** **Patient Number:**	88	96	37	40	81
Very strong evidence	PVS1	Null variant	NO	NO	NO	NO	NO
Strong evidence	PS1	Same amino acid change previously described as pathogenic regarldess of nucleotide change	NO	NO	**NO**	**YES**	**YES**
	PS2	*De novo* in patient (both maternity and paternity confirmed)	NI	NI	NI	NI	NI
	PS3	*In vitro* or *in vivo* functional assay supportive of damaging effect	**YES**		**YES**	**YES**	**YES**
	PS4	Prevalence of variant in affected individual significantly increased compared to controls	NA	NA	NA	NA	NA
Moderate evidence	PM1	Located in a mutational hot spot or critical domain*	NO	NO	**YES**	**YES**	**YES**
	PM2	Absent from controls**	**YES**	**NO**	**YES**	**YES**	**YES**

	PM3	For recessive disorders, detected in trans with a pathogenic variant.	NA	NA	NA	NA	NA
	PM4	Protein length changes as a result of in-frame del/ins in a non repeat	NO	NO	NO	NO	NO
	PM5	Novel missense in residue with different missense previously reported as pathogenic.	NO	NO	NO	NO	NO
	PM6	Assumed de novo, but without confirmation of paternity and maternity					
Supporting evidence	PP1	Cosegregation with disease in multiple affected family members	NI	NI	NI	NI	NI
	PP2	Gene has low rate of benign missense variants	NO	NO	NO	NO	NO
	PP3	Multiple lines of computational evidence support a deleterious effect on protein	YES	NO	**YES**	**YES**	NO
	PP4	Patient phenotype or family history is highly specific for a disease with single gene etiology	**YES**	**YES**	**YES**	**YES**	**YES**
	PP5	Reported as pathogenic by reputable source but evidence is unavailable	NO	NO	NO	NO	NO
		Criteria combined before functional assay:	NA	VUS	LP	LP	NA
		Criteria combined after functional assay:	LP		P	P	P
		Functional assay performed by	MN	MN	DA	DA	MN

NO: negative criterion; YES: positive criterion; NA: criterion does not apply; NI: criterion not investigated

P: pathogenic; LP: likely pathogenic; B: benign; LB: likely benign; VUS: variant of uncertain significance.

DA: Dufner-Almeida, this study. MN: Mark Nellist, personal communication. FR: former reports as referenced on the text.

*TSC2 domains (approximate amino acid numbers): DUF3384 (55-469); Tuberin domain - hamartin binding (555-903); Rap-GAP domain (1562-1748).

**1000 genomes, Exome Variation and EXAc.

***pouco sequenciamento de TSC2 por Sanger; TSC1 não sequenciado. Aguardando NGS.

The mTOR activity-based functional assay was employed to evaluate the pathogenicity of five missense variants - c.52C>G (p.L18V), c.499T>C (p.W167R), c.4493G>T (p.S1498I), c.5227C>G (p.R1743G) and c.4919A>G (p.His1640Arg), a splicing variant leading to in-frame deletion of exon 27 - c.3132-3T>G (p.1044del50); and two in-frame deletions, - c.4842_4844del (p.I1614del) and c.4823_4825del (p.Y1608del) identified in the present study. *TSC2* variants S1498I, H1640R and 1608del had TSC2 levels significantly decreased to WT-TSC2 (Figure 6B). No variant impaired TSC1 levels as R611Q did (Figure 6C). For all variants tested, the total S6K levels were not different from those of the WT-TSC2 (Figure 6D). Finally, 044del50, 1614del, H1640R, 1608del and R1734G had relative T389/S6K signal ratio similar to the R611Q and were significantly higher than TSC2-WT (Figure 6E), thereby considered pathogenic variants. On the other hand, the relative T389/S6K signal of S1698Iwas not significantly increased comparatively to R611Q, nor significantly reduced as WT-TSC2, hence remaining a variant of unclear significance (VUS).

In summary, the functional assay confirmed pathogenicity for a pathogenic DNA variant (Arg1723Gly), disclosed pathogenicity for one likely pathogenic missense DNA variant (His1640Arg), and ruled out pathogenicity for a likely pathogenic variant (Ser1498Ile) and two likely benign variants (Leu18Val and Trp167Arg). Thus, half of the missense DNA variants detected were pathogenic (9/18). Among the nine pathogenic missense variants ascertained according to the ACMG pathogenicity criteria, eight were confirmed by the functional assay and one had a borderline result (Figure 5).

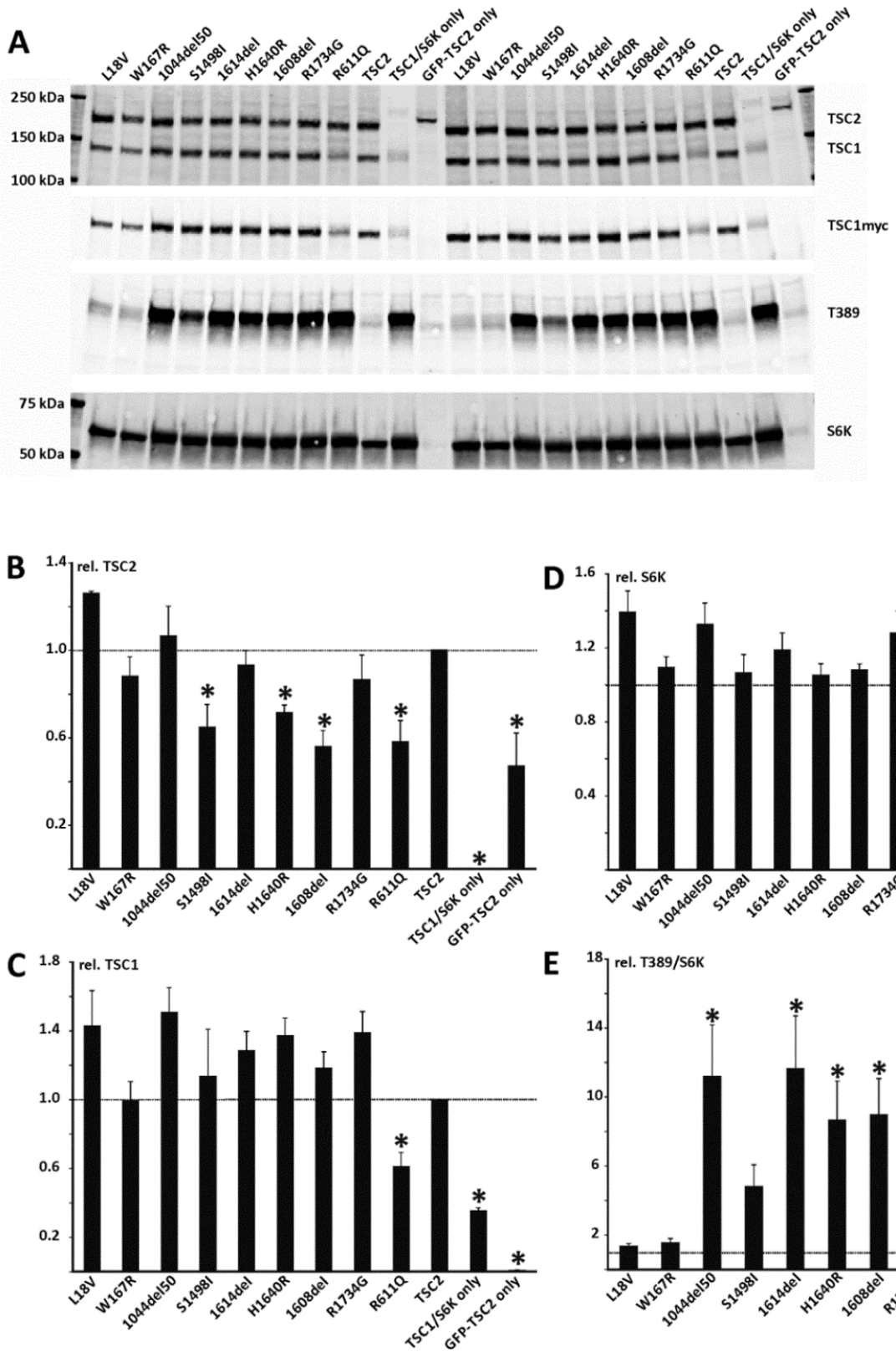


Figure 6: Functional assessment of the TSC2 (NM_000548.3) variants c.52C>G (p.L18V), c.499T>C (p.W167R), c.3132_3282del (p.1044_1094del), c.4493G>T (p.S1498I), c.4840_4842del (p.1614del), c.4919A>G (p.H1640R), c.4823_4825del (p.1680del) and c.5227C>G (p.R1743G). 3H9-1B1 (TSC2/TSC1 double knockout HEK 293T) cells were transfected with the indicated TSC2 variant expression constructs, together with expression constructs for myc-tagged TSC1 and S6K. The known pathogenic TSC2 c.1832G>A (p.R611Q) variant (R611Q) and cells not expressing any TSC2 (TSC1/S6K only) were included as controls. To monitor transfection efficiency, cells were transfected with an expression construct for GFP-TSC2 (GFP-TSC2 only). Twenty-four hours after transfection cells were harvested and the cleared cell lysates analyzed by immunoblotting (A). The signals for TSC2, TSC1, total S6K (S6K) and T389-

phosphorylated S6K (T389) were determined per variant, relative to the wild-type control (TSC2) in two independent, duplicate experiments. The mean TSC2 (B), TSC1 (C) and S6K (D) signals and mean T389/S6K ratio (E) are shown for each variant. In each case the dotted line indicates the signal/ratio for wild-type TSC2 (1.0). Error bars represent the standard error of the mean. TSC2 variants for which the TSC2, TSC1 or S6K signals were significantly reduced compared to wild-type TSC2, or for which the T389/S6K was significantly increased compared to wild-type TSC2, are indicated with an asterisk ($P < 0.05$; Student's paired t-test). Amino acid changes are given according to TSC2 cDNA reference transcript sequence NM_000548.3.

C. Variants that affect splicing

Sixteen *TSC2* pathogenic splice variants were identified in the cohort. Fifteen variants mapped to canonical splice sites: eight at splice donor site and seven at splice acceptor site. The c.482-1G>A and c.482-3C>T belongs to the same patient and were identified in trans. One variant mapped to a deep intronic sequence (Figure 7).

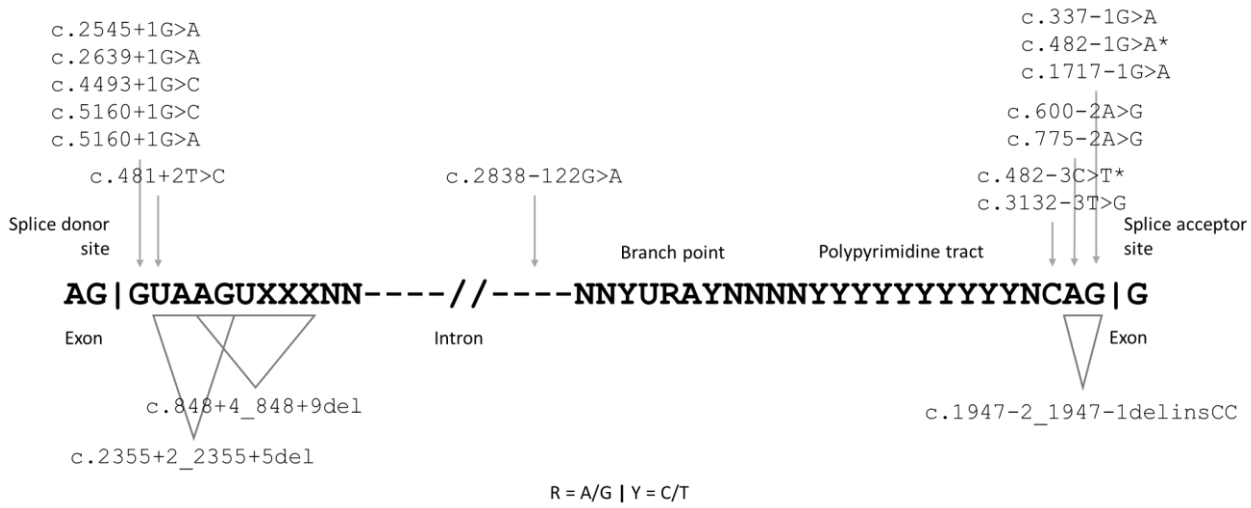


Figure 7: Schematic representation of distribution of pathogenic TSC2 splice variants. Intron canonical splice sites are indicated (donor and acceptor) as well as conserved branch point and downstream polypyrimidine tract. Vertical bar: border between exon and intron. * : Both variants find in the patient 114 in different alleles.

The DNA variant c.4493G>T (p.S1498I) disclosed a non-pathogenic result at the functional mTOR-based assay (Figure 6). On the other hand, splicing-based softwares predicted the c.4493G>T variant instead of impact the protein may prior change the canonical splicing of *TSC2* and cause the skipping of exon 34. Therefore, we interrogated whether c.4493G>T could affect splicing of exon 34 of *TSC2.* The genomic segment of *TSC2* exon 33 through the end of exon 34 from the heterozygous patient DNA was cloned in pCDNA3 vector, and each clone, wild-type or harbouring the nucleotide substitution, was used to transfect the human cell line HEK293. Twenty-four hours after transfection, RT-PCR was performed with vector primers. A 210 bp product was observed for non-recombinant vector cDNA. If the interrogated exon 34 was skipped upon splicing of the mini-gene primary transcript, the expected RT-PCR products would have 408 bp, whereas constitutive splicing products would be expected with 896 bp.

We do not observed products with the length corresponding to the skipped exon. The constitutive splicing product was present for the wild-type. Moreover, an additional RT-PCR reaction was included as an endogenous control, limiting a 275 bp product from *TSC2* exons 8 to 10 mRNA. The endogenous internal control was equally amplified in all samples, including the non-recombinant vector sample (Figure 8). Slow-migrating bands were observed for all samples. All RT-PCR bands were isolated form agarose gel and the Sanger sequence confirmed to be the canonical splicing product of exon 34 mini-gene. We conclude that in HEK293 cells, the *TSC2* variant c.4493T did not effect the splicing. This variant found in patient 14 was seen in patient 50, classified as NMI. However, there is a considerable frequency in the population and two homozygous individuals have already been observed according to the gnomAD database.

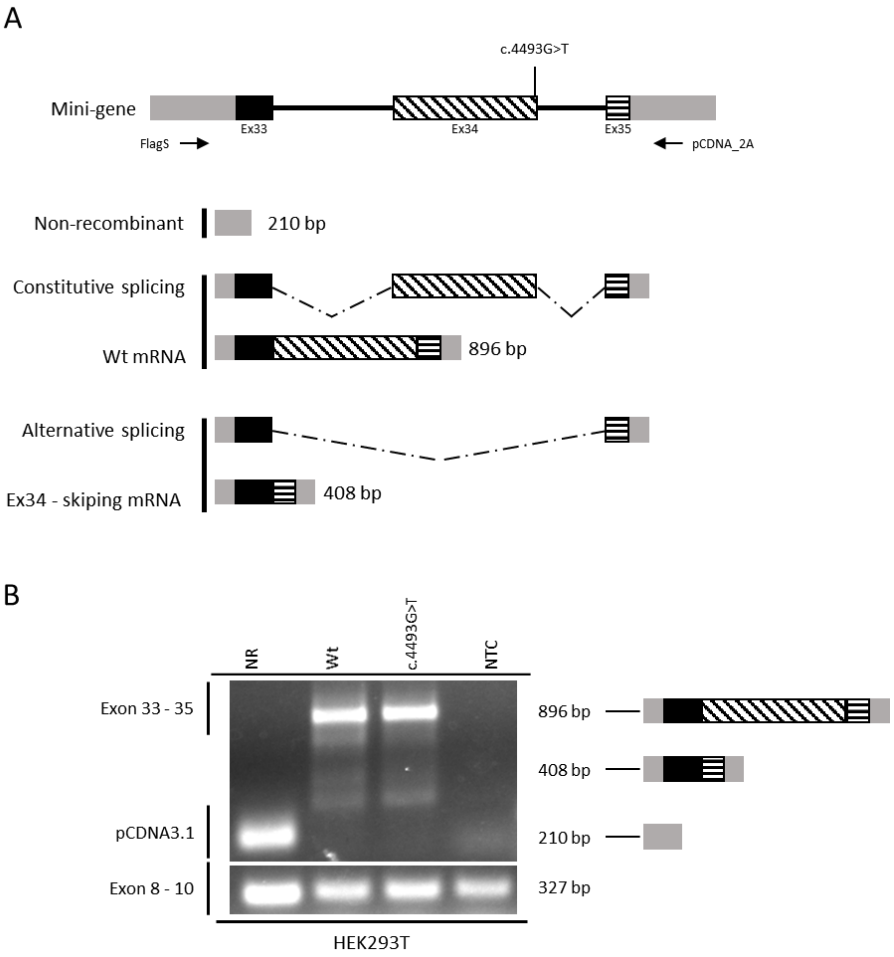


Figure 8: Functional evaluation of a substitution DNA variant in the TSC2 gene (c.4493G>T p.(S1498I) variant). (A) Scheme of predicted constitutive and alternative splicing skipping exon 34 of the mini-gene. The top segment refers to a scheme of the mini-gene: grey blocks are part of vector sequence flanking the mini-gene insert, black blocks are external exons, striped block is the interrogated exon, and black lines are introns. Primer location and identification are indicated by arrows. The location of patient DNA variant that originated the mutant clones, c.4493G>T, is indicated. The expected RT-PCR products of cDNA from cells transfected with either wild-type or deletion clones are indicated, as well as for the non-recombinant clone. (B) Electrophoresis in 1.5% agarose gel of RT-PCR products from HEK293 cell line transfected with non-recombinant vector (NR), wild-type mini-gene (WT), deletion clones and NTC (no-template control). 'Exon 33 – 35' indicates the gel region expected to contain the bands corresponding to the RT-PCR from constitutive and alternative splicing of the mini-gene. Exon 8 – 10 indicate the RT-PCR internal control.

As mentioned, one deep intronic DNA variant (c.2838-122G>A) was considered pathogenic as it had been functionally assessed. It altered splicing of the minigene primary transcript, consequently shifting the reading frame (p.Ser946Argfs*6) (Nellist, Brouwer *et al.* 2015). In an NMI patient, a variant of uncertain clinical significance (c.1361+54_1361+57del) was identified in *TSC2* intron13. This DNA variant has not, to our knowledge, been observed in the general population. In spite of Human Splicing Finder do not predict that this variant probably don´t impact the splicing, both computational tools Acescan2/Rescue and Spliceaid disclosed the 4 bp sequence c.1361+54_1361+57 embedded within a putative intronic splicing enhancer. Of note, Spliceaid recognized it as a potential binding site for the trans-acting protein factor CUG triplet repeat RNA-binding protein 1 / ELAV-like family member 1(CUGBP1/CELF1). The genomic segment of *TSC2* exon 12 through the end of exon 14 from the heterozygous patient DNA was cloned in pCDNA3 vector, and each clone, wild-type or harbouring the deletion, was used to transfect the human cell lines HEK293. If the exon 12 was skipped upon splicing of the mini-gene primary transcript, the expected RT-PCR products would be 430 bp, whereas constitutive splicing products would be expected be 534 bp. We observed no products with the length corresponding to the skipped exon. The constitutive splicing product was present for the wild-type. The endogenous control from *TSC2* exons 8 to 10 mRNA was equally amplified in all samples, (Figure 9). All RT-PCR bands were isolated form agarose gel and the Sanger sequence confirmed to be the constitutive splicing product of exon 12-14 mini-gene. We conclude that in HEK293 cells, the *TSC2* variant c.1361+54_1361+57del allele did not change the splicing output of the mini-gene. Therefore, that variant was not considered pathogenic.

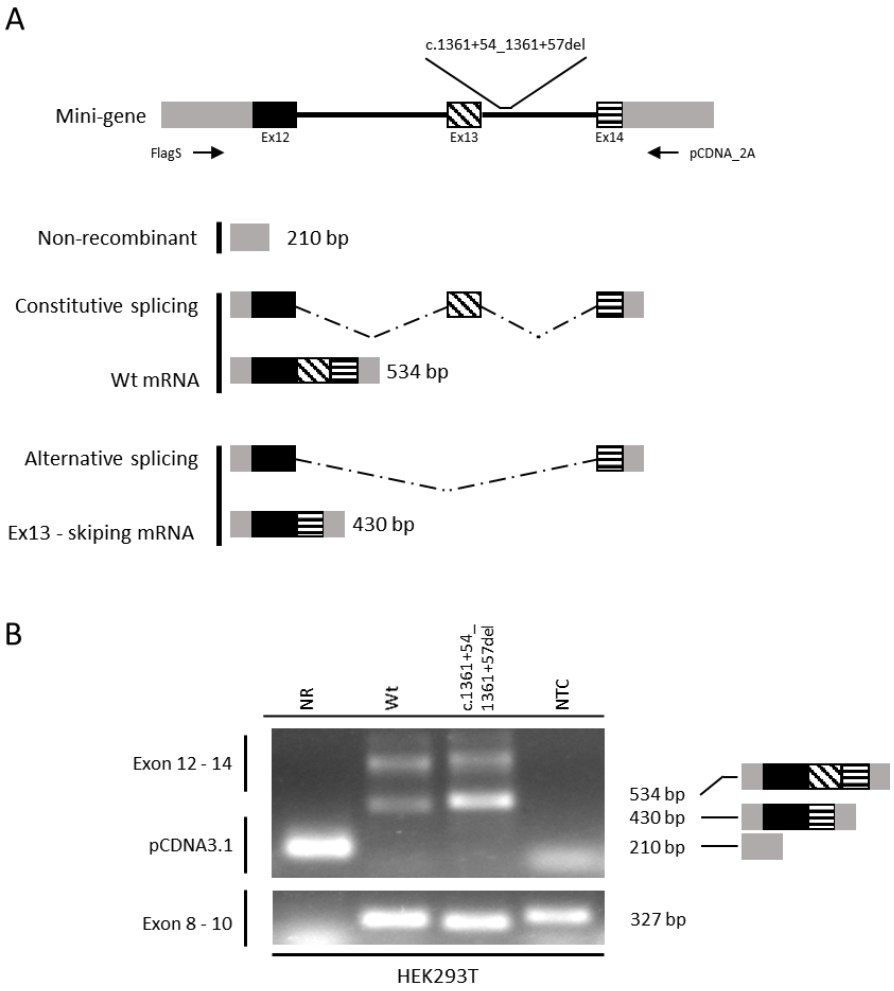


Figure 9: Functional evaluation of a deep intronic DNA variation in the TSC2 gene (c.1361+54_1361+57del variant). (A) Scheme of predicted constitutive and alternative splicing, involving skipping of exon 12, of the c.1361+54_1361+57del variant mini-gene. In The top segment refers to a scheme of the mini-gene is shown: grey blocks are part of vector sequence flanking the mini-gene insert, black blocks are external exons, the striped block is the interrogated exon to be tested, and black lines are introns. Primers location and identification are indicated by arrows. The location of the patient DNA variant that originated the mutant clones, c.1361+54_1361+57del, is indicated. The expected RT-PCR products of cDNA from cells transfected with either wild-type or deletion clones are indicated, as well as for the non-recombinant clone. (B) Electrophoresis in 1.5% agarose gel of RT-PCR products from HEK293 cell line transfected with non-recombinant vector (NR), wild-type mini-gene (WT), deletion clones and NTC (no-template control). 'Exon 12 – 14' indicates the gel region expected to contain the bands corresponding to the RT-PCR from constitutive and alternative splicing of the mini-gene. Exon 8 – 10 indicate the RT-PCR internal control.

Data recently available at gnomAD (Massachusetts Institute of Technology, MIT, Cambridge, MA, USA) revealed the frequency of the c.1361+54_1361+57del allele as 0,2%, found on 467 of 207454 individuals.

D. Large segmental gene deletions and duplications

After Sanger sequencing, DNA from 43 patients with missense variants, in-frame deletions or no mutation identified (NMI) was submitted to MLPA of both *TSC1* and *TSC2* genes. Nine copy number variations (CNV) were identified in NMI patients: two in the *TSC1* and seven in the *TSC2* genes. All CNV were validated by quantitative PCR (qPCR). MLPA and corresponding qPCR results are presented on Table 8 and Figure 10. All

MLPA data were confirmed by qPCR. Both methods inferred somatic mosaicism for the deletions in patients #25 (TSC1 c.-234-?_-144+?del), # 12 (TSC2 c.(975+628)_(1716+41)del), and #51 (TSC2 c.(5068+1_5069-1)_(*102_?)del) (Table 8).

Table 8: MLPA results and qPCR validation of TSC1 and TSC2 CNV.

		MLPA		qPCR	
Patient	**CNV**	**Position**	**ps; p**	**Position**	**rq (sem); p**
		TSC1			
25	c.-234-?_-144+?del	Exon 1	0.80; <0.01	Exon 1	0.79 (0.05); <0.01
29	c.(737+1_738-1)_(*1+_?)del a	Exon 9 - 23	0.60; <0.01	Exon 17	0.40 (0.01); <0.01
		TSC2			
97	c.(774+1_775-1)_(848+1_849-1)del a	Exon 9	0.58; <0.01	Exon 9	0.45 (0.03); <0.01
12	c.(975+628)_(1716+41)del a	Exon 11 - 16	0.74; <0.01	Exon 12	0.62 (0.07); 0.03
59	c.(1599+1_1600-1)_(2545+1_2546-1)del	Exon 16 - 22	0.56; <0.01	Exon 19	0.54 (0.01); <0.01
84	c.(1716+1_1717-1)_(2545+1_2546-1)del	Exon 17 - 22	0.56; <0.01	Exon 19	0.63 (0.02); <0.01
89	c.(2355+1_2356-1)_(*102_?)dup	Exon 22 - 42	1.47; <0.01	Exon 41	2.43 (0.07); <0.01
51	c.(5068+1_5069-1)_(*102_?)del	Exon 39 - PKD1	0.68; <0.01	Exon 41	0.75 (0.02); <0.01
61	c.(?_-30)_(*102_?)del	TSC2	0.54; <0.01	Exon 12	0.50 (0.03); <0.01

MLPA position refers to the extension of the detected deletion or duplication, while qPCR position refers to exon tested to validate MLPA.
ps: mean probe signal; p: Student´s t test p-value; rq: ratio coefficient; sem: standard error of mean.
[a] First description of this mutation

NextGENe analysis of *TSC1* and *TSC2* NGS data from three patients with segmental deletions confirmed the large *TSC1* deletion encompassing exons 9 through 23 (patient #29), and the *TSC2* deletion between exons 11 and 16 (patient #12). The mosaic segmental deletion containing *TSC1* exon 1 (patient #25) was not identified as this exon was not covered by the Nextera probe set. In addition, no other DNA variant was identified by NGS in these patients.

Due to the somatic mosaicism of the *TSC1* exon 1 deletion, the absence of coding region in this exon, and previous reports interrogating its clinical significance (van den Ouweland, Elfferich *et al.* 2011), it was not classified as pathogenic, and patient #29 remained with NMI. The other eight CNV classified as pathogenic.

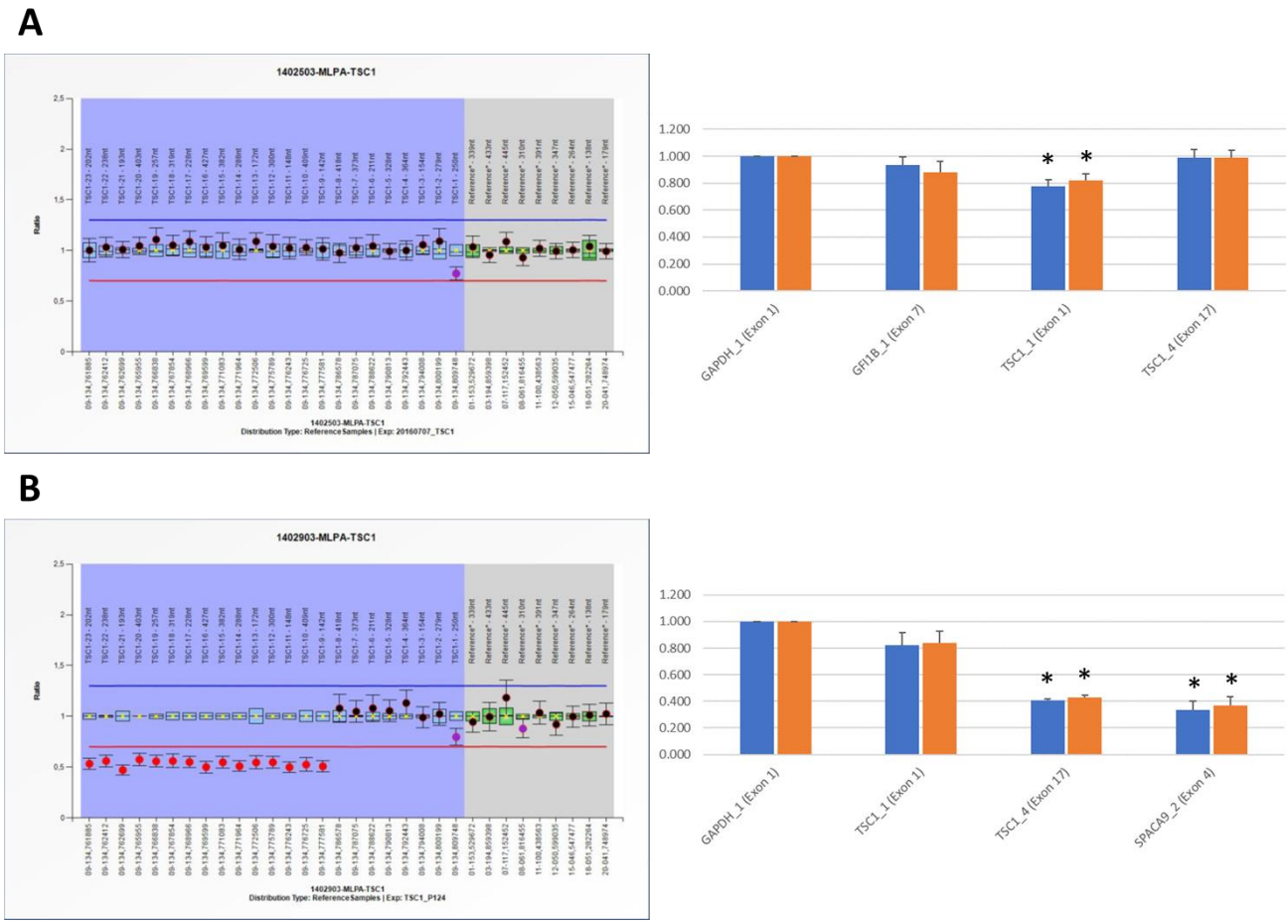
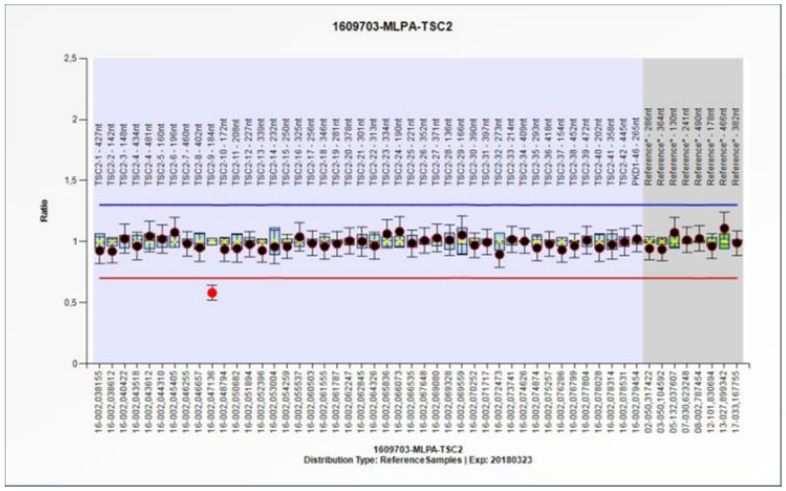
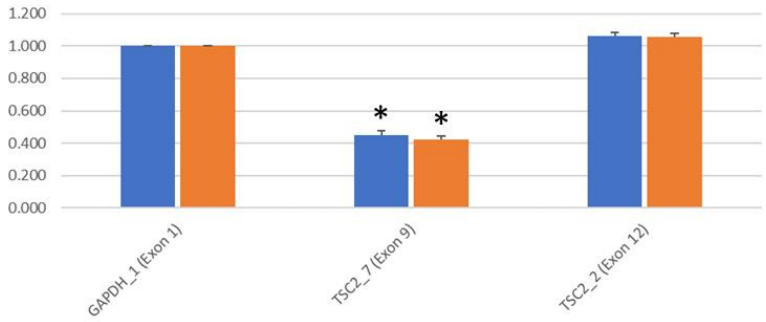


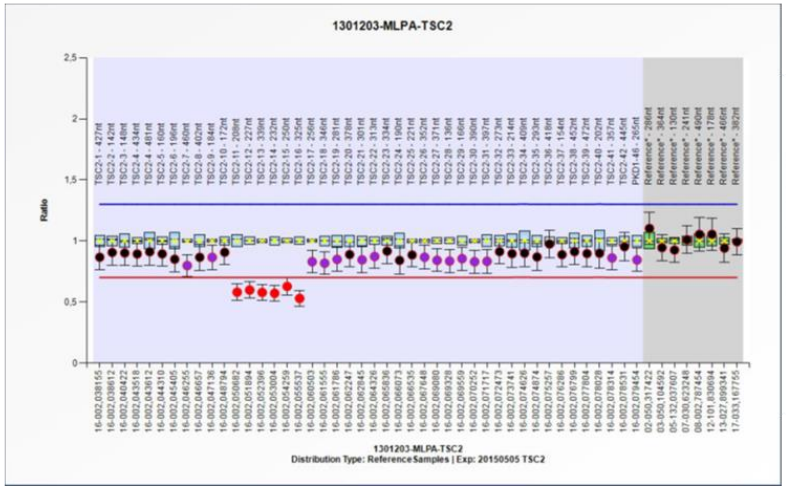
Figure 10: MPLA and qPCR results for *TSC1* or *TSC2* deletions and duplication. Left panel is MLPA report and right panel the qPCR results. **(A)** Patient#25 with mosaic *TSC1* exon 1 deletion. **(B)** Patient#29 with *TSC1* deletion from exon 9 to exon 23, extending to at least exon 4 of *SPACA9* gene. **(C)** Patient#97 with *TSC2* deletion of exon 9. **(D)** Patient#12 with *TSC2* deletion from exon 10 to exon 16. **(E)** Patient#59 with *TSC2* deletion from exon 15 to exon 22. **(F)** Patient#84 with *TSC2* deletion from exon 17 to exon 22. **(G)** Patient#89 with *TSC2* duplication from exon 22 to exon 42. **(H)** Patient#51 with *TSC2* deletion from exon 39 to exon 42, extending to at least exon 46 of downstream *PKD1* gene. **(I)** Patient#61 with whole *TSC2* deletion, extending to at least exon 1 of upstream gene *NTHL1* and o exon 46 of downstream *PKD1* gene (To be continued on next page).

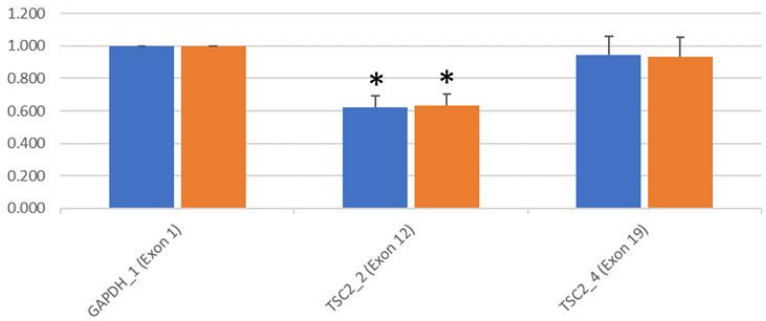
C


1609703-MLPA-TSC2
Ratio
1609703-MLPA-TSC2
Distribution Type: ReferenceSamples | Exp: 20180323

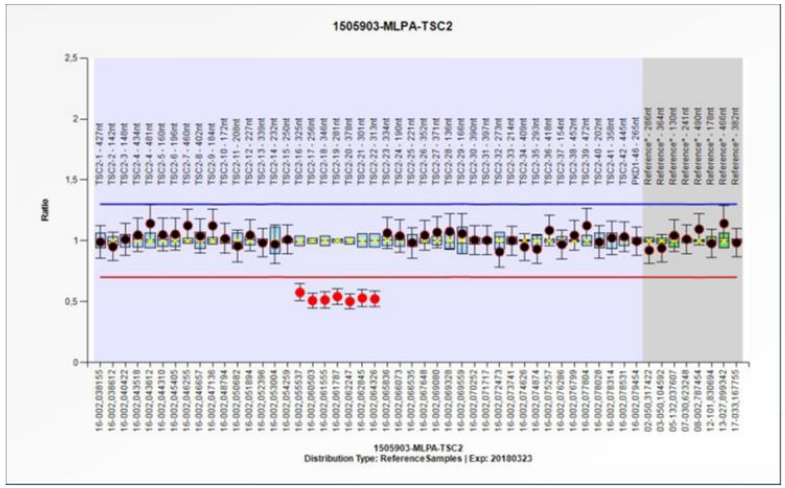

1.200
1.000
0.800
0.600
0.400
0.200
0.000
*
*
GAPDH_1 (Exon 1)
TSC2_7 (Exon 9)
TSC2_2 (Exon 12)

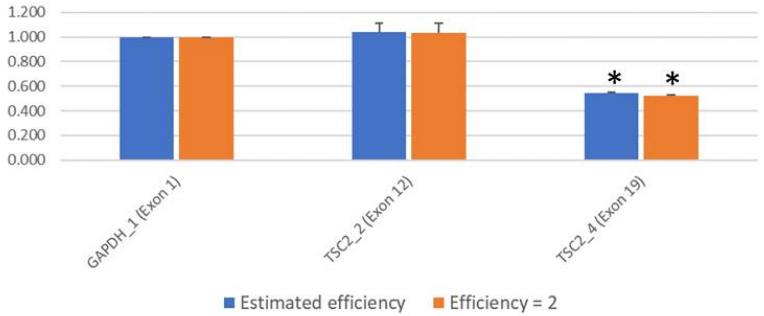
D


1301203-MLPA-TSC2
Ratio
1301203-MLPA-TSC2
Distribution Type: ReferenceSamples | Exp: 20150505 TSC2

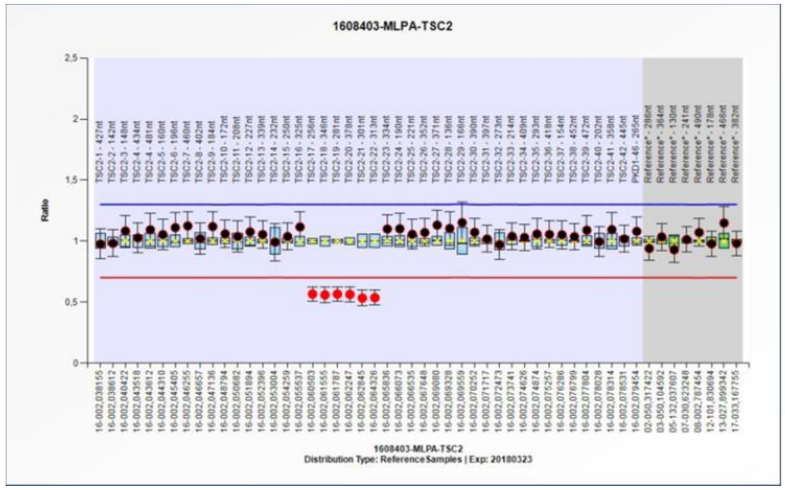

1.200
1.000
0.800
0.600
0.400
0.200
0.000
*
*
GAPDH_1 (Exon 1)
TSC2_2 (Exon 12)
TSC2_4 (Exon 19)

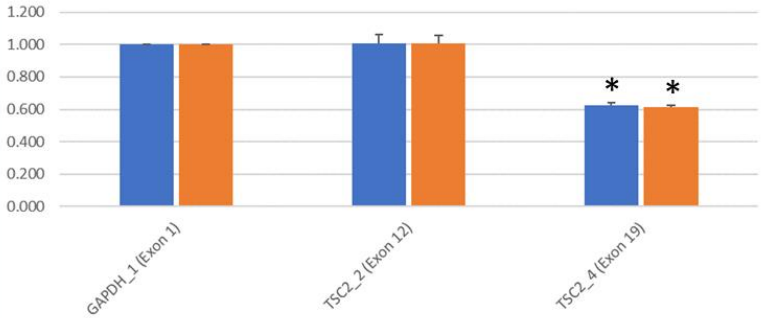
E


1505903-MLPA-TSC2
Ratio
1505903-MLPA-TSC2
Distribution Type: ReferenceSamples | Exp: 20180323

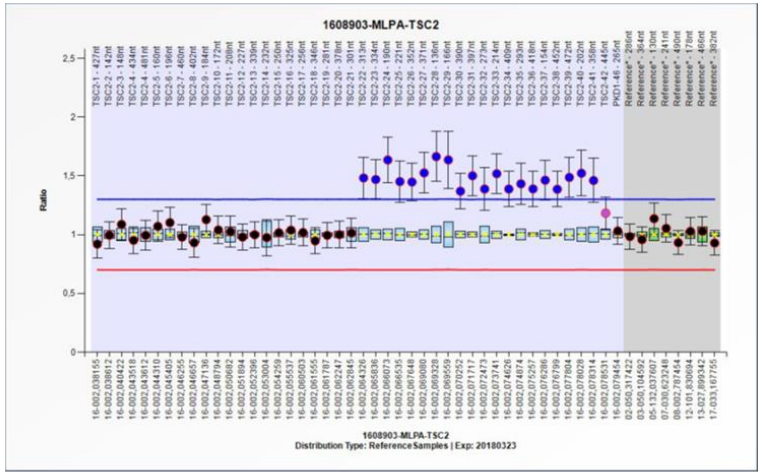

1.200
1.000
0.800
0.600
0.400
0.200
0.000
*
*
GAPDH_1 (Exon 1)
TSC2_2 (Exon 12)
TSC2_4 (Exon 19)
Estimated efficiency
Efficiency = 2

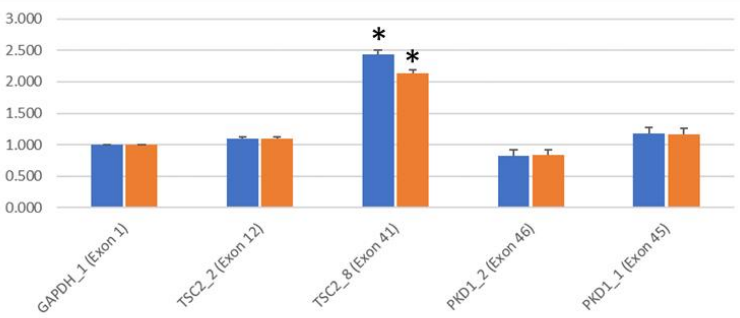
F


1608403-MLPA-TSC2
Ratio
1608403-MLPA-TSC2
Distribution Type: ReferenceSamples | Exp: 20180323

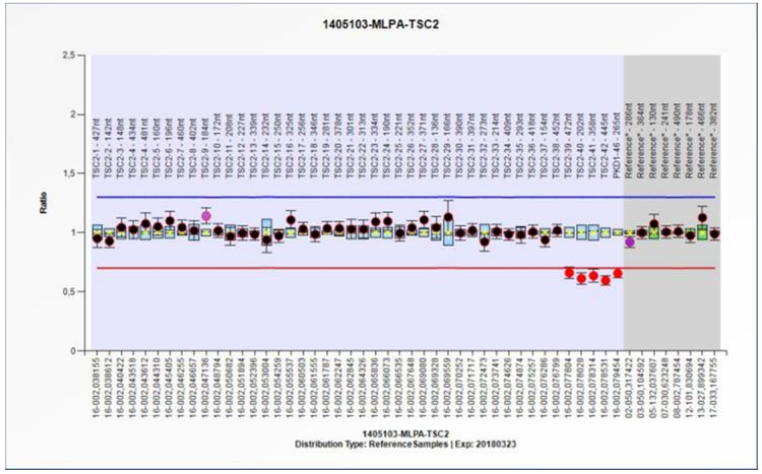

1.200
1.000
0.800
0.600
0.400
0.200
0.000
*
*
GAPDH_1 (Exon 1)
TSC2_2 (Exon 12)
TSC2_4 (Exon 19)

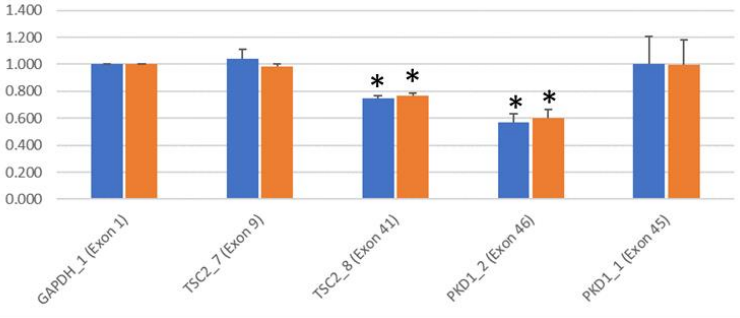
G


1608903-MLPA-TSC2
Ratio
1608903-MLPA-TSC2
Distribution Type: ReferenceSamples | Exp: 20180323

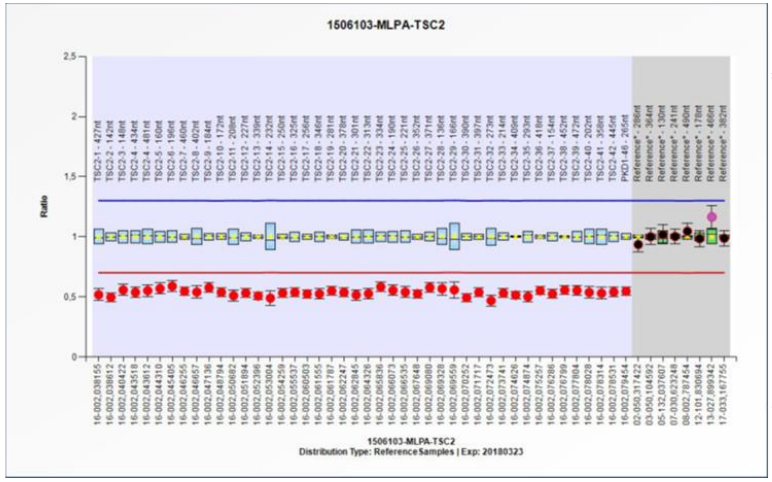

3.000
2.500
2.000
1.500
1.000
0.500
0.000
*
*
GAPDH_1 (Exon 1)
TSC2_2 (Exon 12)
TSC2_8 (Exon 41)
PKD1_2 (Exon 46)
PKD1_1 (Exon 45)

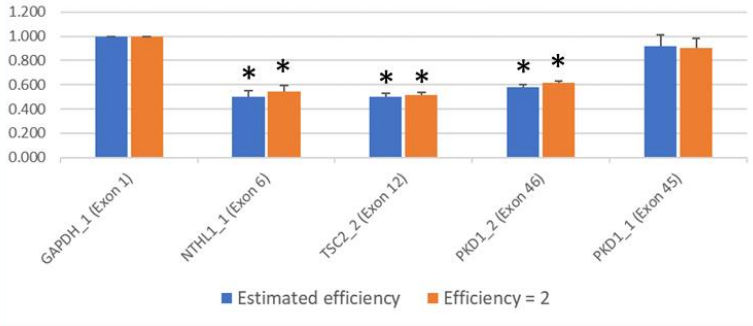
H


1405103-MLPA-TSC2
Ratio
1405103-MLPA-TSC2
Distribution Type: ReferenceSamples | Exp: 20180323


1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
*
*
*
*
GAPDH_1 (Exon 1)
TSC2_7 (Exon 9)
TSC2_8 (Exon 41)
PKD1_2 (Exon 46)
PKD1_1 (Exon 45)

I


1506103-MLPA-TSC2
Ratio
1506103-MLPA-TSC2
Distribution Type: Reference Samples | Exp: 20180323


1.200
1.000
0.800
0.600
0.400
0.200
0.000
GAPDH_1 (Exon 1)
NTHL1_1 (Exon 6)
TSC2_2 (Exon 12)
PKD1_2 (Exon 46)
PKD1_1 (Exon 45)
Estimated efficiency
Efficiency = 2

We attempted to identify the extension of the eight segmental deletions and the *TSC2* duplication. *GFI1B* (Growth Factor-Independent 1B Transcription Repressor - NM_004188.6), a gene positioned in a head-to-head fashion 34 kb upstream *TSC1*, was employed to investigate the boundary of patient #25 mosaic deletion of *TSC1* exon 1. *GFIIB* exon 1 deletion was ruled out by qPCR (rq = 0.93; sem = 0.06; p = 0.34) (Figure 10A). The long extension between *TSC1* exon 1 and upstream *GFI1B* exon 1 (nearly 34 Kb) or downstream *TSC1* exon 2 (*TSC1* intron 1 of 6 Kb) impaired the identification of the breakpoints by PCR.

The extension of *TSC1* deletion c.(737+1_738-1)_(*1+_?)del from patient #29 was analyzed by qPCR. The *SPACA9* (Sperm acrosome-associated 9 - NM_001316898.1) gene is approximately 2 kb downstream of *TSC1* polyadenylation signal. QPCR of *TSC1* 3'-UTR (rq = 0.49; sem = 0.03; p = $4.1x10^{-2}$) and of *SPACA9* exon 4 (rq = 0.38; sem = 0.06; p = $9.3x10^{-3}$) confirmed the deletion comprised both segments (Figure 10B).

Four internal deletions have been detected by MLPA on *TSC2* (Table 9; patients #97, 12, 59, and 84). Table 9 indicates the expected locations of upstream and downstream breakpoints that generated those deletions.

Table 9: *TSC2* internal deletion breakpoint mapping

Patient	CNV	Position	Breakpoint location			
			Expected		Observed	
			Upstream	Downstream	Upstream	Downstream
97	c.(774+1_775-1)_(848+1_849-1)del	Exon 9	Intron 8	Intron 9	Inc	Inc
12	c.(975+628)_(1716+41)del	Exon 11-16	Intron 10	Intron 16	Intron 10	Intron 16
59	c.(1599+1_1600-1)_(2545+1_2546-1)del	Exon 16-22	Intron 15	Intron 22	Inc	Inc
84	c.(1716+1_1717-1)_(2545+1_2546-1)del	Exon 17-22	Intron 16	Intron 22	Inc	Inc

Inc: inconclusive

Various forward and reverse PCR primers were designed to anneal respectively on upstream and downstream sites as an approach to sequence the junction of breakpoints. This strategy was successful to identify the breakpoints of *TSC2* intragenic deletion c.(975+628)_(1716+41)del from patient #12. Sequencing of PCR products amplified with primers annealing to introns 9 and 17 disclosed the deletion spans 6.2 kb, from c.975+627 (intron 10) to c.1716+41 (intron 16). The breakpoint in intron 10 maps to an Alu repeat flanked by other Alu units, whereas the intron 16 breakpoint is 296 bp upstream of a L1 repeat with three Alu repeats downstream. *TSC2* intron 10 and intron 16 sequences, respectively 1,796 bp and 4,820 bp long, have only 11% similarity. However, the 909-bp segment encompassing the short interspersed nuclear element (SINE) sequences within intron 10 and 16 breakpoints are 66% identical (Figure 11). A defined PCR product spanning the breakpoint junction was not obtained from DNA of patients #97, 59 or 84 (Table 9).

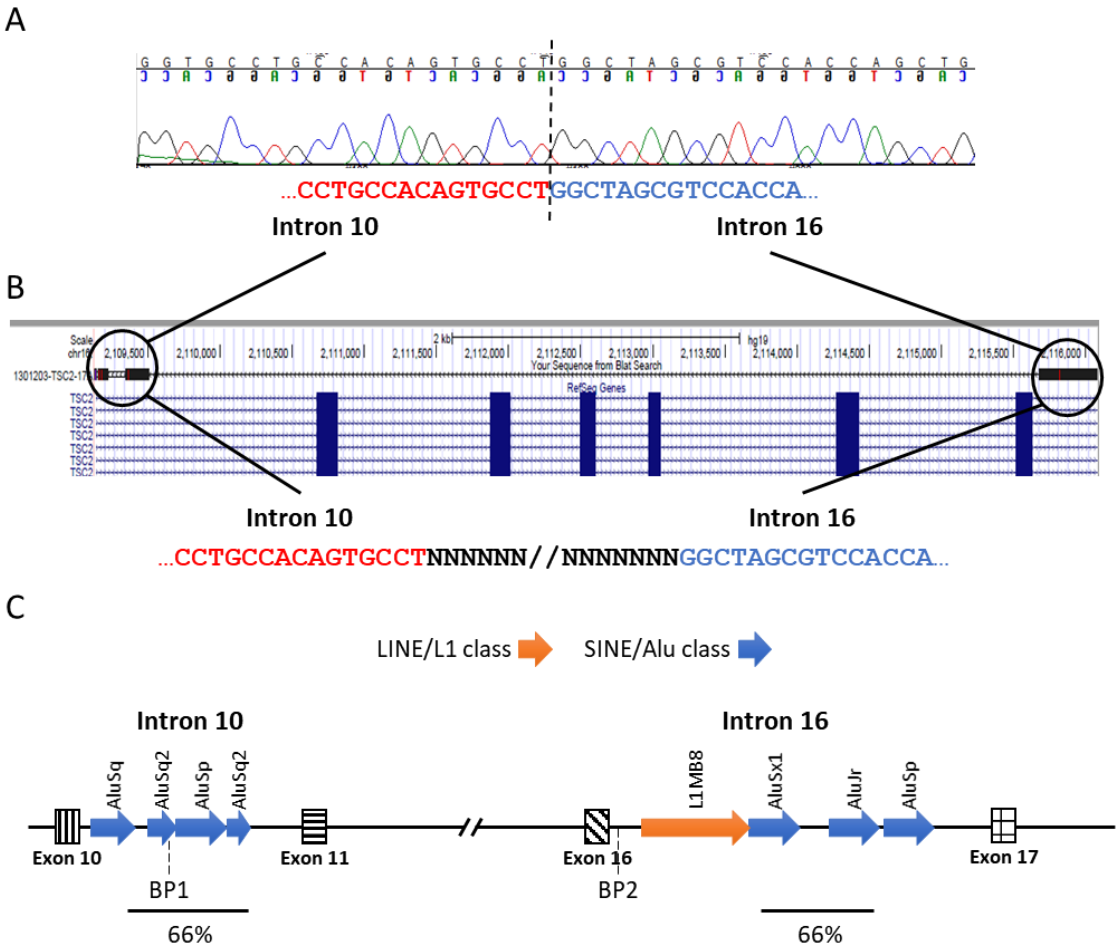


Figure 11: Breakpoint mapping of the *TSC2* intragenic deletion c.(975+_976-)_(1716+_1717-)del . (A) Sanger sequencing of the DNA from the patient with the intragenic *TSC2* deletion c.(975+_976-)_(1716+_1717-)del. (B) UCSC genome browser mapping of the breakpoints in *TSC2* introns 10 and 16. (C) Diagram illustrating the positions of SINE/Alu and LINE/L1 repeat sequences in *TSC2* introns 10 and 16, respectively within and adjacent to each breakpoint. The regions of 66% similarity between introns are indicated.

One segmental duplication and two additional deletions were identified in the *TSC2* gene (Table 8 and Figure 10). *TSC2* is tail-to-tail to *PKD1* (Polycystic Kidney Disease 1 gene, NM_001009944). In other to evaluate if the CNV spanned through *PKD1*, its most 3' exons, 45 and 46, were tested by qPCR in those three cases. In the case of c.(2355+1_2356-1)_(*102_?)dup (Patient #89), neither exon 45 or 46 was duplicated (exon 45: rq = 1.18, sem = 0.1; p = 0.14; exon 46: rq = 0.83; sem = 0.09; p = 0.19) (Figure 10G).

Patient #51 *TSC2* deletion (c.(5068+1_5069-1)_(*102_?)del) of exon 39 extended through *PKD1* exon 46, as validated by qPCR (rq = 0.57; sem = 0.06; p = 0.02), though exon 45 was preserved (rq = 1.0; sem = 0.2; p = 0.99) (Figure 10H). Hence, upstream and downstream breakpoints should be respectively at *TSC2* intron 38 and *PKD1* intron 45. Similarly, the deletion of full *TSC2* gene c.(?_-30)_(*102_?)del (patient #61) encompassed *PKD1* exon 46 as detected by qPCR (rq = 0.58; sem = 0.02; p = $2.2x10^{-3}$), preserving exon 45 (qPCR; rq = 0.92; sem = 0.1; p = 0.42). The *NTHL1* (Endonuclease III-Like 1, MIM#602656) upstream *TSC2* in a tail-to-head fashion. QPCR of *NTHL1* exon 6 (rq = 0.50; sem = 0.05; p = $8.6x10^{-3}$) disclosed that the deletion affected the 3' end of this gene as well.

In summary, Figure 12 and 14 illustrates the eight CNV identified as *TSC1* (1) or *TSC2* (7) pathogenic alterations, as well as the mosaic TSC1 exon 1 deletion, which was considered an alteration of uncertain significance. Their estimated extension is presented according to the gene exon location.

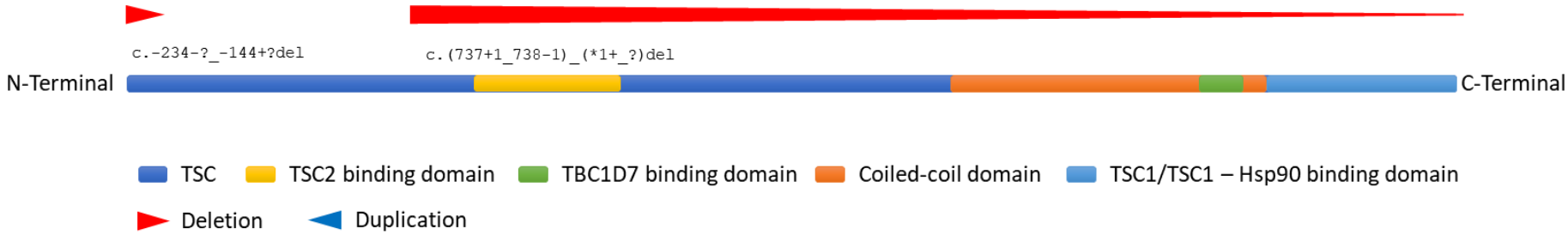


Figure 12: Schematic representation of TSC1 and the segmental deletion c.-234-?_-144+?del of exon 1 on patient#25 c.(737+1_738-1)_(*1+_?)del from exon 9 to exon 23 fond on patient#29.

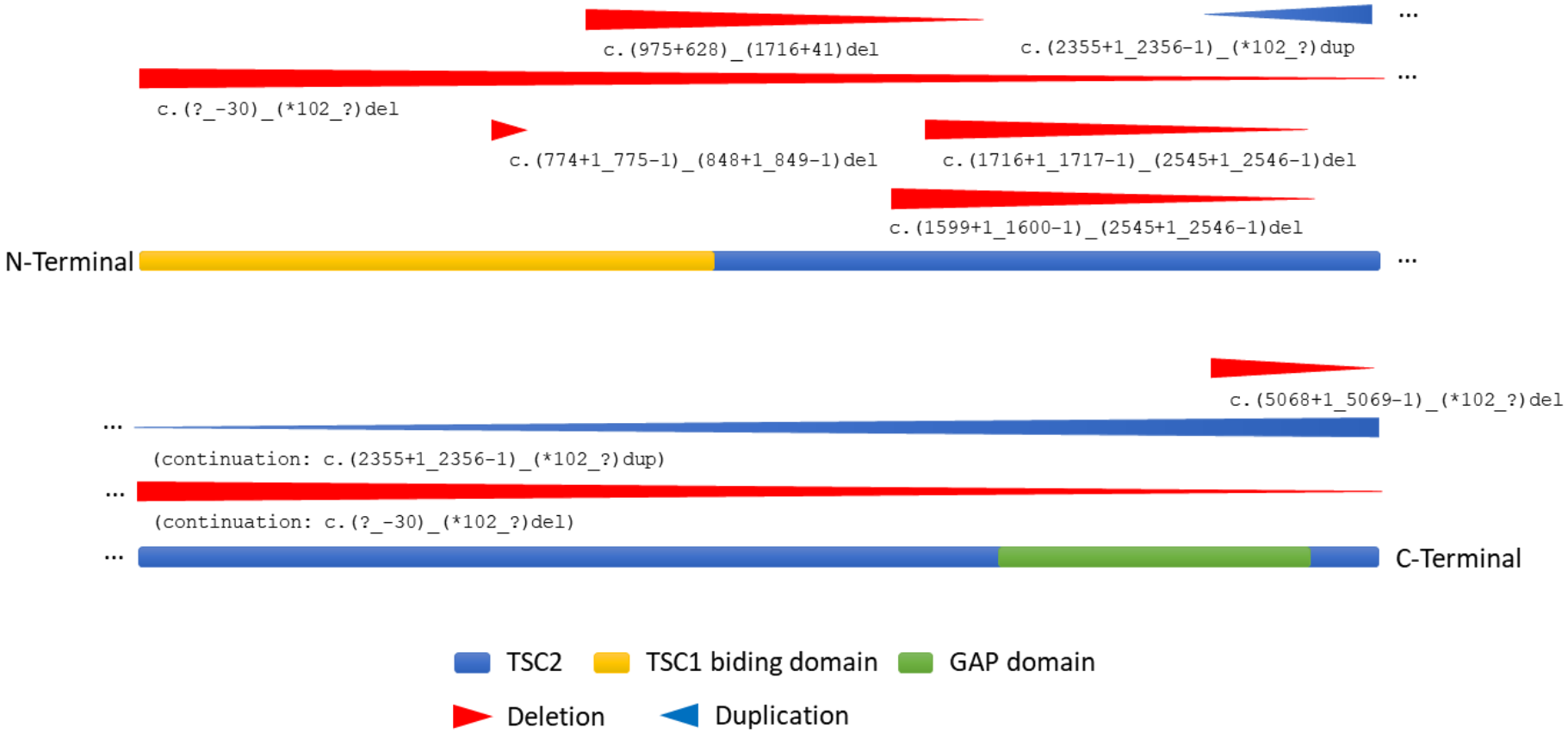


Figure 13: Schematic representation of TSC2 and the segmental deletion c.(975+628)_(1716+41)del from exon 11 to exon 16, c.(5068+1_5069-1)_(*102_?)del from exon 39 to exon 42, c.(1599+1_1600-1)_(2545+1_2546-1)del from exon 16 to exon 22, c.(?_-30)_(*102_?)del whole *TSC2*, c.(1716+1_1717-1)_(2545+1_2546-1)del from exon 17 to exon 22, c.(774+1_775-1)_(848+1_849-1)del from exon 22 to exon 42, and duplication c.(2355+1_2356-1)_(*102_?)dup form exon 22 to 42.

E. Single nucleotides variants

Finally, we identified 108 distinct, single nucleotide variants (SNV; frequency higher than 0.01) without clinical significance, 41 on *TSC1* and 125 on *TSC2*. Of those, 53 variants were identified within exonic segment (Supplementary Table 11) and 113 in introns (Supplementary Table 12). Nineteen variants were not present in 1000 Genomes, dbSNP and COSMIC data banks, and occurred in patients with pathogenic mutations (Supplementary Table 11 and 12).

On NMI patients were find the rare variants *TSC1*:c.738-197T>A (patient#60:) *TSC2*:c.*1G>A (VUS - patient#100), *TSC2*:c.1663G>C (p.A555P) (patient#50) and *TSC2:*c.4005+153T>A (patient#63).

V. Discussion

We analyzed DNA from 115 patients with definite clinical diagnosis of TSC (Northrup, Krueger *et al.* 2013). They were residents from the states of São Paulo (SP), Paraná (PR) or Santa Catarina (SC), situated in the South and Southeast regions of Brazil; and clinically followed up at three large, tertiary referral, SP and PR University hospitals (HSCM-SP, GRAACC-SP, and HC-UFPR). The clinical evaluation of these patients has been the subject of other graduate research studies (Juliana Paula Masters dissertation – Hospital Santa Casa de Misericórdia, 2016, Laís Masulk Masters dissertation – Complexo de Hospital de Clínicas da Universidade Federal do Paraná, Laís Masulk ; and Nasjla Saba, unpublished), and thus will not be discussed here. The vast majority of samples has been first assessed by Sanger sequencing (N=116), whereas 1 samples were initially analysed only by NGS. Pathogenic DNA variants identified by NGS have been confirmed by Sanger sequencing. MLPA was carried out in a subset of patient DNA (N=38), and any altered result further assessed by qPCR.

We report a mutation detection rate of 86,09% (99/115), which is similar to previous reports that sequenced both *TSC1* and *TSC2* coding sequence and intronic boundaries, and assessed exonic copy number variation (Jones, Shyamsundar *et al.* 1999, Dabora, Jozwiak *et al.* 2001, Rendtorff, Bjerregaard *et al.* 2005, Sancak, Nellist *et al.* 2005, Hung, Su *et al.* 2006, Au, Williams *et al.* 2007). Patients with mutation identified, either on *TSC1* or *TSC2* genes, are generally more severely affected than patients with NMI, which had a lower incidence of brain findings on imaging studies, neurological features, renal findings and lower incidence of seizures than those with *TSC2* or *TSC1* mutations (Camposano, Greenberg *et al.* 2009). Our relatively high mutation detection rate by currently standard sequencing approaches are likely due to the origin of the patients. Ninety-eight patients from two tertiary referral hospitals (HC-UFPR and HSCM, SP), and the remaining 19 are from the oncology service of another tertiary referral University hospital (UNIFESP, São Paulo, SP). Tertiary referral hospitals tend to retain the most severe cases.

Despite the high number of distinct TSC-causing mutations already described worldwide in both *TSC1* and *TSC2* genes (LOVD-TSC1, LOVD-TSC2), we report 41 novel pathogenic DNA variants (Table 3 and 3) as well as 19 novel SNV (Supplementary Table 11 and 12).

In regard to sporadic TSC cases, when patients with *TSC2* mutations are compared to those with *TSC1* mutations, they present higher mean number of cerebral cortical tubers and subependymal nodules, and higher mean grade of kidney angiomyolipomas, as well as higher frequency of moderate/severe cases of intellectual disability and retinal hamartomas (Dabora, Franz *et al.* 2011). Accordingly, Sancak, Nellist *et al.* (2005) and Au, Williams *et al.* (2007) reported intellectual disability, angiomyolipoma and retinal phacoma more frequently among TSC patients with *TSC2* than *TSC1* mutations. As mentioned, all patients from our cohort came from tertiary referral hospitals, which tend to assist more severe cases. This could justify the higher numbers of *TSC2* pathogenic variants observed by us. However, if we estimate this ratio specifically for

each hospital, we observe that 7195% (59/82) of all *TSC2* pathogenic variants described here came from HC-UFPR while *TSC1*/*TSC2* ration on HSCM-SP is about 1:1.5. At enrolment in this study, familial cases have been ascribed based solely on parental report. Among the 96 cases with mutation identified. Familial cases are expected to be 30% of all TSC cases (Jones, Shyamsundar *et al.* 1999, Dabora, Jozwiak *et al.* 2001, Au, Williams *et al.* 2007). As we observed lower rates, we interrogate if more familial cases could be part of the HSCM-SP cohort of 23 patients, what could explain the 1:1.5 *TSC1*:*TSC2* partition ratio for pathogenic DNA alterations.

Frameshift and nonsense variants were the mutation types most commonly observed among *TSC1* and *TSC2* pathogenic DNA alterations, respectively. Nonsense and frameshifting mutations corresponded to 94% (16/17) and 51% (41/81), respectively of *TSC1* and *TSC2* pathogenic DNA alterations. In our study, missense, in-frame deletions and splicing DNA variants have been detected only in the *TSC2* gene. Segmental gene deletions and duplications comprehended 8.3% (8/96) of all pathogenic DNA alterations, 6.25% (1/16) and 8.75% (7/80) respectively for *TSC1* and *TSC2*. Of note, 2.5% of TSC patients with mutation identified had *TSC2* deletions extending through *PKD1* gene, the contiguous gene deletion syndrome (Martignoni, Bonetti *et al.* 2002). Our data are similar to another Brazilian study (Rosset, Vairo *et al.* 2017) as well as to reports from many countries This is in agreement with for *TSC1* and *TSC2* mutation classification (Table 10). Because our study does not aim to estimate the frequency of TSC in Brazil, we could not compare previously reports showing no ethnic or geographical bias (Wiederholt, Gomez *et al.* 1985, Sampson, Scahill *et al.* 1989, Osborne, Fryer *et al.* 1991, O'Callaghan, Shiell *et al.* 1998)

Table 10: *TSC1* and *TSC2* pathogenic variants classification and distribution worldwide.

Reference	Country	Patients	Pathogenic Variants	Ratio TSC1/TSC2	TSC1							TSC2						
					Total	Nonsense	Frameshift	Missense	Splicing	Inframe	Duplication/ Deletion	Total	Nonsense	Frameshift	Missense	Splicing	Inframe	Duplication/ Deletion
Jones, A. C. et. al. (1997) *	UK	171	21 (12.3%)	-	21 (100%)	7	12	1	2	-	-	-	-	-	-	-	-	-
Ali, J. B., et. al. (1998) *	UK	83	16 (19.3%)	-	16 (100%)	6	8	-	2	-	-	-	-	-	-	-	-	-
Au, K. S., et. al. (1998) **	USA	90	22 (24.4%)	-	-	-	-	-	-	-	-	22 (100%)	2	8	10	-	2	-
Beauchamp, R. L. et. al. (1998) **	USA	40	21 (52.5%)	-	-	-	-	-	-	-	-	21 (100%)	6	5	6	3	1	-
Young, J. M. et. al. (1998) *	UK	79	27 (34.2%)	-	27 (100%)	13	13	-	1	-	-	-	-	-	-	-	-	-
Jones, A. C. et. al. (1999)	UK	150	120 (80%)	1:4.4	22 (18.3%)	8	12	1	2	-	-	98 (81.7%)	20	21	22	8	5	22
Niida, Y. et. al. (1999)	USA	126	74 (59%)	1:3.6	16 (21.6%)	7	7	-	2	-	-	58 (78.4%)	14	21	13	4	2	4
van Slegtenhorst, M. A. (1999) *	The Netherlands	225	29 (12.8%)	-	29 (100%)	11	17	-	1	-	-	-	-	-	-	-	-	-
Dabora, S. L. et. al. (2001)	USA	224	166 (83%)	1:4.9	28 (15.0%)	11	15	-	2	-	-	138 (85%)	37	43	31	27	-	-
Rendtorff, N. D. et. al. (2005)	Denmark	65	51 (78%)	1:3.6	11 (21.6%)	4	6	-	1	-	-	40 (78.4%)	10	10	10	6	2	2
Hung, C. C. et. al. (2006)	Taiwan	84	64 (76 %)	1:6.1	9 (14%)	5	3	-	-	1	-	55 (86%)	15	21	12	7	-	-
Au, K. S. et. al. (2007)	USA	325	243 (74.8%)	1:3	61 (25%)	21	35	1	4	-	-	182 (75%)	39	57	58	25	-	3
Kozlowski, P. et. al. (2007) ***	USA	261	54 (20.7%)	1:12.5	4 (7%)	-	-	-	-	-	4	50 (93%)	-	-	-	-	-	50
Sasongko, T. H. et. al. (2008)	Japan	21	16 (76.2%)	1.3:1	9 (56.3%)	1	8	-	-	-	-	7 (43.7%)	-	4	1	2	-	-
Chopra, M. et. al. (2011) ****	Australia	45	33 (73%)	1:2.7	9 (27.8%)	-	-	-	-	-	-	24 (72.2%)	-	-	-	-	-	-
Niida, Y. et. al. (2013)	Japan	57	31 (54.4%)	1:1.8	11 (35.5%)	2	5	1	2	-	1	20 (64.5%)	-	6	5[&]	4	4	1
Rosset, C. et. al. (2017)	Brazil	53	47 (88.7%)	1:3.4	11 (23.4%)	6	4	-	1	-	-	37 (76.6%)	12	7	2	8	3	5

* Study only with *TSC1*
** Study only with TSC2
*** Study only with duplication and deletion
**** Only distribution data of *TSC1* and *TSC2* variants
[&] Including one stop codon variant

A. *TSC2* missense and in-frame deletion effect on mTOR

Missense DNA variants cause amino acid substitution in the encoded protein. As *a priori* they are not a loss-of-function variant, the interpretation of their clinical significance is not straight forward, and relies on several bioinformatics tools, although they may be built onto common biological aspects. While PlolyPhen-2 involves comparison of a property of the wild-type (ancestral, normal) allele and the corresponding property of the mutant (derived, disease-causing) allele, Mutation Taster integrates information from different biomedical databases and uses established analysis tools, comprising evolutionary conservation, splice-site changes, loss of protein features and changes that might affect the amount of mRNA. On the other way, SIFT and PROVEAN predict whether a protein sequence variation affects protein function.

Of 18 missense variants and 5 in-frame variants at least two in silico bioinformatic tool predicted to be damage, except to c.2167A>G (p.I723V) and c.5051C>T (p.S1684F) where Mutation Taster was the only one to predict to be deleterious. Moreover, splicing analysis using Human Splicing Finder also predicted for almost all variants submitted to analysis to potential disrupt the splicing. The only exceptions were c.1663G>C (p.A555P), c.1831C>T (p.R611W) and c.2167A>G (p.I723V).

According to previously functional assessments in silico analysis agree with c.1831C>T (p.R611W), c.1832G>A (p.R611Q), c.5227C>T (p.R1743W) and c.5228G>A (p.R1743Q) (Hoogeveen-Westerveld, Ekong *et al.* 2013). However, the *TSC2* variants c.499T>C (p.W167R) and c.1793A>G (p.Y598C), both predicted as pathogenic for all in silico analysis, were found this nonsense variants, therefore disagreeing with predictions and classified as neutral. Moreover, our functional assay disagrees with in silico prediction of c.52C>G (p.L18V) and c.499T>C (p.W167R) which were classified as deleterious but did not activated the mTORC1 pathway (Figure 6). Finally, c.4493G>T (p.S1498I), c.4909_4911delinsGAC (p.K1637D), c.4919A>G (p.H1640R) and c.5227C>G (p.R1743G) were predicted to be deleterious for al in silico analysis, and in these case our functional assessment agree with the pathogenicity (Figure 6). The remaining variants are under investigation.

Data from well-established *in vitro* or *in vivo* functional assays supportive of a damaging effect on the gene product or pathway have been classified as a strong criterion for pathogenicity classification of missense DNA variants (Richards, Aziz *et al.* (2015), Supplementary Table 9).Based on the recognized repressive role of TSC1/2 complex on mTOR kinase activity, a functional assay has been developed (Hoogeveen-Westerveld, Wentink *et al.* 2011, Hoogeveen-Westerveld, Ekong *et al.* 2012, Hoogeveen-Westerveld, Ekong *et al.* 2013, Kwiatkowski, Palmer *et al.* 2015) and further improved (Dufner-Almeida, L. G. *et. al.* submitted). It relies on reporting the phosphorylation status of S6K Thr^{389}, a target of mTOR (Goncharova, Goncharov *et al.* 2002, Tee, Fingar *et al.* 2002). We employed the mTOR activity-based assay to functionally assess five *TSC2* missense variants, two *TSC2* in-frame deletions, and one splicing variant. A genetically modified HEK293T cell line (3H9-1B1) having both *TSC1* and *TSC2* genes fully inactivated by CRISPR/Cas9 (Figure 6; Dufner-Almeida *et al.*,

manuscript submitted) was used in the assay. For each test assay, the 3H9-1B1 cell line was transfected with three different full-length cDNA clones (WT or mutant *TSC2*, WT *TSC1*, and WT *RPS6KB1* – *S6K*). This additional step leads to higher efficiency in reporting S6K Thr^{389} phosphorylation due to more stability of the TSC1/2 complex than assaying endogenous levels (Dufner-Almeida, L. G. *et*. *al.* submitted). TSC1 and S6K protein levels were not affected by the overexpression of any *TSC2* variant tested in the present study (Figure 6C and D). TSC2 levels were reduced by S1498I and 1608del variants (Figure 6B) with no effect on TSC1 and S6K. Because all three proteins, TSC1, TSC2 and S6K, are co-overexpressed, TSC2 levels are unlikely to be solely affected by contaminants from transfection. Hence, variants S1498I and 1608del may render TSC2 unstable, possibly due to incorrect folding and faster degradation than wild-type TSC2. Moreover, as both variants are outside the TSC1-binding domain, the TSC1–TSC2 interaction should not be destabilized (Chong-Kopera, Inoki *et al.* 2006), thus preserving TSC1 levels.

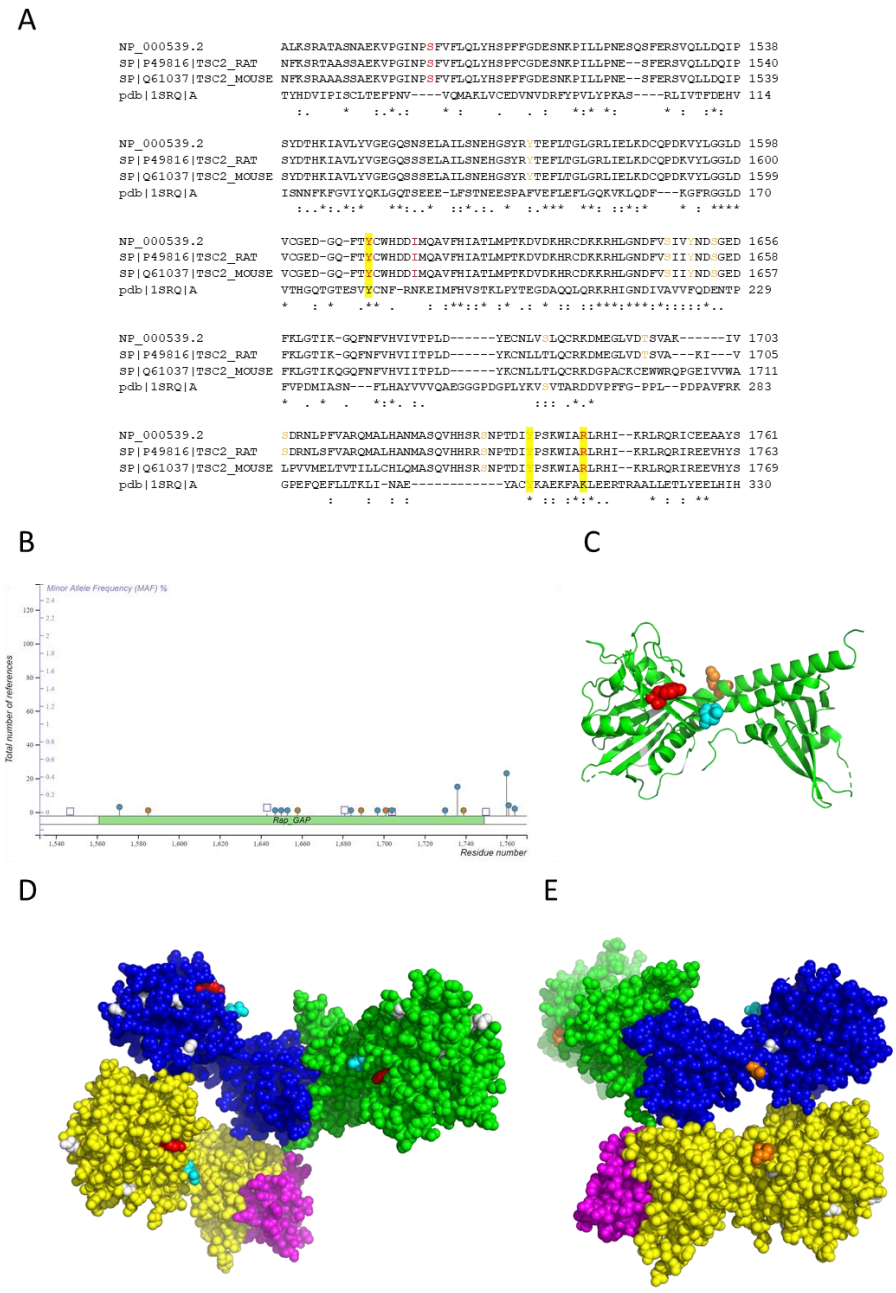


Figure 14: Global alignment of TSC2, phosphorylation site of GAP Domain and positioning of TSC2:p.Y1608del, p.I1614del and p.R1743G variants on predicted *Homo sapiens* GAP domain structure (PDB: 1SRQ). (A) Global alignment of human (NP_000539.2), rat (P49816), mouse (Q61037) TSC2 C-terminal and Rap1GAP catalytic domain (Homo sapiens, pdb|1SRQ|A) (Daumke et al., 2004). Red letters represent amino acids conserved among human, rat and mouse, and red letters highlighted in yellow represent the amino acids conserved between the three organisms and the predicted 1SRQ sequence. Below the alignment, asterisk (*) indicates positions which have a single, fully conserved residue, colon (:) indicates conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet PAM 250 matrix), period (.) indicates conservation between groups of weakly similar properties (scoring =< 0.5 in the Gonnet PAM 250 matrix) and none indicates none conservation between groups. (B) Graphic representation of phosphorylation sites on C-Terminal domain of TSC2. (C) Cartoon representation of tertiary structure of 1SRQ|A generated with PyMOL software. Red sphere represents the TSC2:p.Y1608del variant, cyan sphere represent TSC2:p.I1614del variant, orange sphere represent TSC2:p.R1743G variant and white sphere the phosphorylated residues. (D) and (E), respectively front and back schematic representation of 1SRQ crystal structure of the Rap1GAP catalytic domain (*Homo sapiens*) and the position of TSC2:p.Y1608del, p.I1614del and p.R1743G variants.

We identified five variants that significantly increased T389/S6K ratios, a result from reduced TSC1–TSC2 inhibition of mTORC1 activity (Tee, Fingar *et al.* 2002, Inoki, Li *et al.* 2003, Zhang, Cicchetti *et al.* 2003,

Riemenschneider, Betensky *et al.* 2006). Two missense variants (H1640R and R1743G), one in-frame exon skipping due to splice variant (1044del50), and two in-frame deletions (Y1608del and I1614del) were thus classified as pathogenic (Figure 6). TSC-causing missense variants have been described more often on the coding sequence for *TSC1* or *TSC2* domains, notably TSC2 GAP domain and TSC2/TSC1-interacting domain, such as R611Q on the latter (Nellist, Sancak *et al.* 2005, Hoogeveen-Westerveld, van Unen *et al.* 2012). The missense R1743G variant and the two in-frame deletions (1608del and 1614del) locate within TSC2 GAP domain. Residues I1498, Y1608, I1614 and R1743 are conserved among human, rat and mouse TSC2 (Figure 14A). None of those is close to phosphorylation sites (Figure 14B). However, analysis of TSC2 GAP domain (pdb:1SRQ) by pymol shows that Y1608 and I1614 are close and both are towards the inner face of the structure, while the R1743 locates outside the domain folding and faces a probable catalytic pore (Figure 14D and E). The spatial prediction of each variant and the possible function on activity of TSC2 GAP together with functional assessment (Figure 6) suggest these residues are important for the structural stability of TSC2 GAP domain.

In two other cases (L18V and W167R), the T389/S6K ratios were not significantly different from wild-type TSC2. The variants then classified as non-pathogenic. Variant c.499T>C (p.(W167R)) has been detected in an individual with a nonsense mutation. The functional assay corroborates it is a benign substitution. The missense W167R is conserved among human, rat and mouse TSC2 (Figure 15A) and are not close to phosphorylation site (Figure 15B). Moreover, W167R variant are located within alpha-helix (Figure 15C) but embedded inside the predicted TSC1 binding domain (pdb|5HIU, Figure 15D) of *Chaetomium thermophilum*. The spatial prediction and functional assessment suggests these residues are not important for TSC1/2 complex.

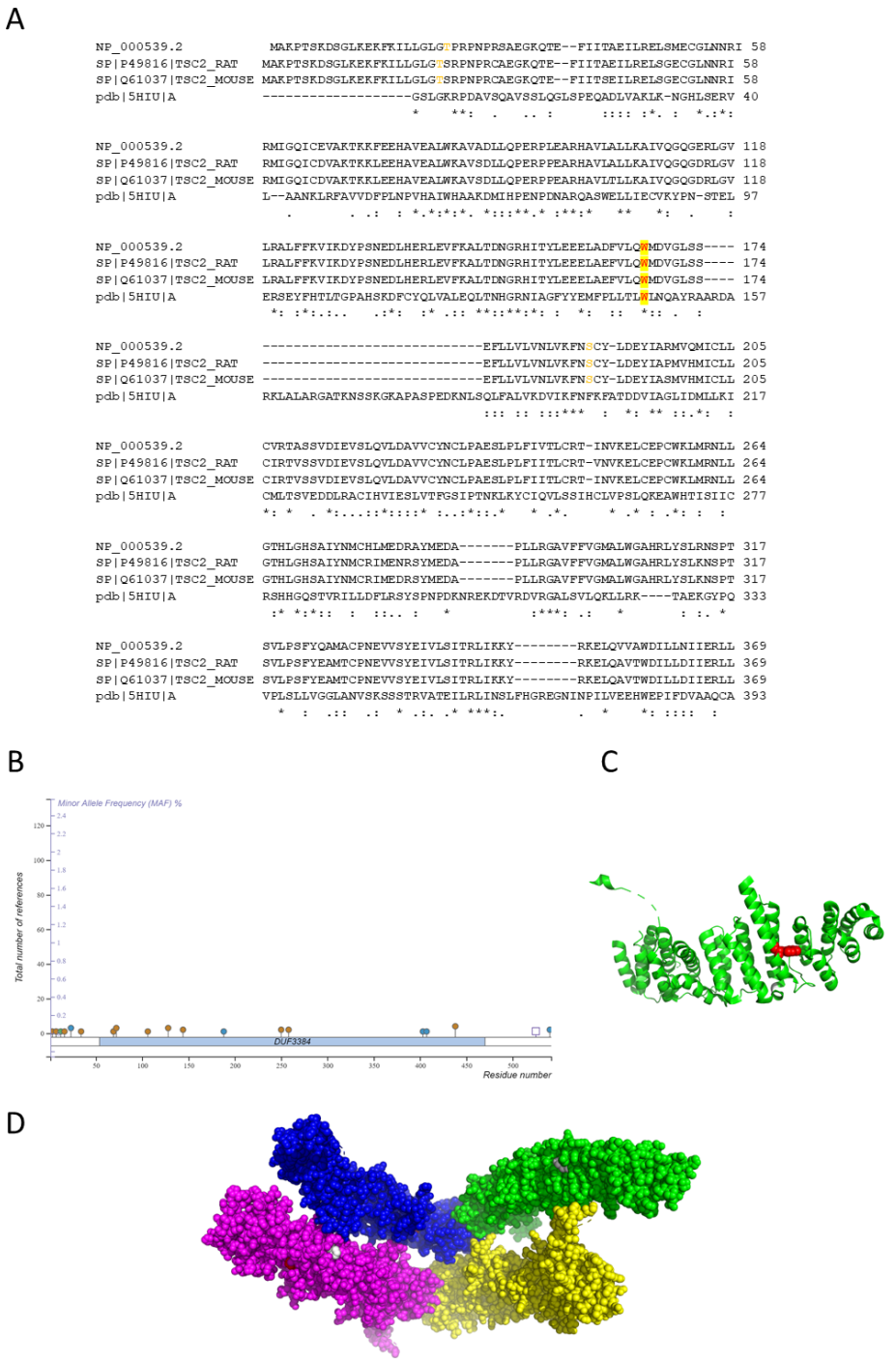


Figure 15: Global alignment of TSC2, phosphorylation site of TSC1 binding domain and positioning of TSC2:W167R variant on predicted *Chaetomium thermophilum* TSC1 binding domain structure (PDB:5HIU). (A) Global alignment of human (NP_000539.2), rat (P49816), mouse (Q61037) TSC2 N-terminal and TSC1 binding domain (*Chaetomium thermophilum*, pdb|5HIU|A) (Zech, R., 2016). Red letters represent amino acids conserved among human, rat and mouse, and red letters highlighted in yellow represent the amino acids conserved between the three organisms and the predicted 5HIU sequence. Below the alignment, asterisk (*) indicates positions which have a single, fully conserved residue, colon (:) indicates conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet PAM 250 matrix), period (.) indicates conservation between groups of weakly similar properties (scoring =< 0.5 in the Gonnet PAM 250 matrix) and none indicates none conservation between groups. (B) Graphic representation of phosphorylation sites on N-Terminal domain of TSC2. (C) Cartoon representation of tertiary structure of 5HIU|A generated with PyMOL software. Red sphere represents the TSC2:p.Y1608del variant, cyan sphere represent TSC2:p.W167R and white sphere the phosphorylated residues. (D) Representation of 5HIU crystal structure of the TSC1 binding domain (*Chaetomium thermophilum*) and the position of TSC2:p.W167R.

B. A 'miscalled' missense DNA variant? Does deep intronic variant disrupt the canonical splicing?

In only one case, c.4493G>T (p.(S1498I)), the T389/S6K ratio was similar to wild-type TSC2 and R611Q. This borderline result led us to consider c.4493G>T with uncertain significance. As they did not significantly affect S6K phosphorylation (Figure 6), we further evaluated the potential of c.4493G>T (p.(S1498I)) to affect splicing. This substitution of unclear significance is on the last nucleotide of exon 34. Nucleotide substitutions, as in putative missense and synonymous variants, may lead to splicing abnormalities if they abolish exonic splicing enhancers (ESE) or silencers (ESS) or activate cryptic exon splice sites (reviewed by Dufner-Almeida *et al.* 2018). Variant c.4493G>T (p.(S1498I)) was predicted by ALAMUT to disrupt splice site (Supplementary Figure 2). ALAMUT employs different softwares to predict effect on pre-mRNA splicing. Acescan2 identifying candidate *cis*-elements in alternative and constitutive splicing in mammalian exons and within the flanking introns. SpliceAid2 searches for motifs versus a database of experimentally assessed target RNA sequences, that can predict the effect of the DNA mutations at the level of the target sequences of the RNA-binding proteins that determine the pattern of mRNA splicing. Finally, HSF combines algorithms based on PWM matrices, Maximum Entropy principle or Motif Comparison to identify and predict mutations' effect on splicing motifs.

To functionally assess the effect of c.4493G>T on splicing of *TSC2* exon 34, the mini-gene strategy was employed by cloning and expressing the genomic DNA segment from exon 33 through exon 35. The internal exon (34) was interrogated to contain an element that regulate its splicing. The expression of the mini-gene in transfected eukaryotic cells produces the processed mRNA. If internal transcribed elements of the mini-gene are recognized by the cell nuclear proteome, differential splicing may take place. Because transcription pattern is dependent of trans elements, splicing variants may have different effect according to tissue. We did not obtain evidence for an effect of the c.4493G>T variant on splicing on HEK 293 cells, however the same variant should have different effect on other tissue, such as on central nervous systems.

A 4-bp deletion (*TSC2*:c.1361+54_1361+57del) from one NMI patient was not identified on 1000 Genomes, but twice on dbSNP (rs137854304) and once on ClinVar (ID 49694) without clinical significance reported. The four bases are embedded within a putative intronic splicing enhancer potentially recognized by a trans-acting protein factor CUGBP1/CELF1 that is widely expressed in over 20 human tissues (Fagerberg, Hallstrom *et al.* 2014). Tissue-specific transcriptional factors could change the splicing pattern of this segment, then the deletion of 4 bp could disrupt the canonical splicing in some tissues, whereas other remain normal.

C. *TSC1* and *TSC2* CNV and genetic counseling implications

Seven segmental deletions and one segmental duplication have been identified in this study as pathogenic alterations (8/101; 7.9%). One deleted segment extended from *TSC1* intron 8 to at least the

downstream gene (*SPACA9*), having a length of more than 23 Kb. Although its breakpoints have not been sequenced, the segment overlaps part of a deletion reported by van den Ouweland, Elfferich et al. (2011) with a breakpoint mapping to nearly 1,400 bases downstream intron 8 and exon 9 junction, and extending downstream for at least 119,000 kb.

The whole deletion of *TSC2* (*TSC2*:c.(?_-30)_(*102_?)del) extended downstream to *PKD1* exons 46 but not exon 45 and upstream to exon 6 of *NTHL1* far from less than 1kb. Homozygosis mutation on this gene are related to an autosomal recessive cancer predisposition syndrome, often with progression to colorectal cancer (Weren, Ligtenberg *et al.* 2015). In this report, the most frequent *NTHL1* mutation has a frequency of 1:250, constituting a high risk for the patient that already has one allele deleted. Therefore, the NTHL1 gene from that patient was sequenced. The deletion was detected, and no further mutation was identified (data not shown). The French Society of Predictive and Personalized Medicine (SFMPP) established guidelines for managing information given on the secondary findings discoveries (SFs) for genes related to cancers (Pujol, Vande Perre *et al.* 2018). Genes were divided into three classes: for class 1 genes, delivering the information on SFs was recommended; for class 2 genes, delivering the information remained questionable; and for class 3 genes, delivering the information on SFs was not recommended because the low level of evidence or data in the literature. Because the *NTHL1* gene belonges to class 3, we decided to not report this incidental finding to family. In one more case the *TSC2* deletions (*TSC2*:c.(5068+1_5069-1)_(*102_?)del) extended to *PKD1* causing a Polycystic Kidney Disease, Infantile Severe, with tuberous sclerosis (PKDTS – MIM# 600273) (O'Callaghan, Edwards *et al.* 1975, Stapleton, Johnson *et al.* 1980). On both cases Q-PCR disclose that deletion extend to exon 46 but not to exon 45 of *PKD1*.

Four intragenic deletion were identified on *TSC2*. In only one case the breakpoint could be sequenced. Of those that break point were not sequenced, the length of flanking introns prevented the sequencing of break point, even in the deletion of exon 9. However, the *TSC2* internal deletion which had both breakpoints sequenced mapped to introns 10 and 16. Each intron breakpoint had SINE nearby (Figure 16). Alu high copy number in the human genome and similarity with members from this SINE family allow for unequal recombination (Gu, Yuan *et al.* 2015). Alu repeats from *TSC2* introns 10 and 16 are 66% identical. The relatively long distance between them (deletion size of 6.2 kb) rules out the possibility of DNA replication-mediated deletion, and reinforces unequal recombination as a likely pathway to have generated the deletion (Abeysinghe, Chuzhanova *et al.* 2003, Abo-Dalo, Kutsche *et al.* 2010, Gu, Yuan *et al.* 2015). Moreover, deletion breakpoints from both cancer and inherited disease are commonly associated to AT-rich sequences (Abeysinghe, Chuzhanova et al. 2003), as found in Alu sequences such as those from *TSC2* introns 10 and 16.

Homozygosis mutation at *NTHL1* cause a familial adenomatous polyposis 3 (MIM#616415), an autosomal recessive cancer predisposition syndrome characterized by the development of multiple colonic adenomas, often with progression to colorectal cancer (Weren, Ligtenberg *et al.* 2015).

The high identity between the SINEs within introns 10 and 16 and the relatively long distance between them suggests unequal recombination as a likely pathway to generate the deletion (Abeysinghe, Chuzhanova et al. 2003, Abo-Dalo, Kutsche et al. 2010, Gu, Yuan et al. 2015).

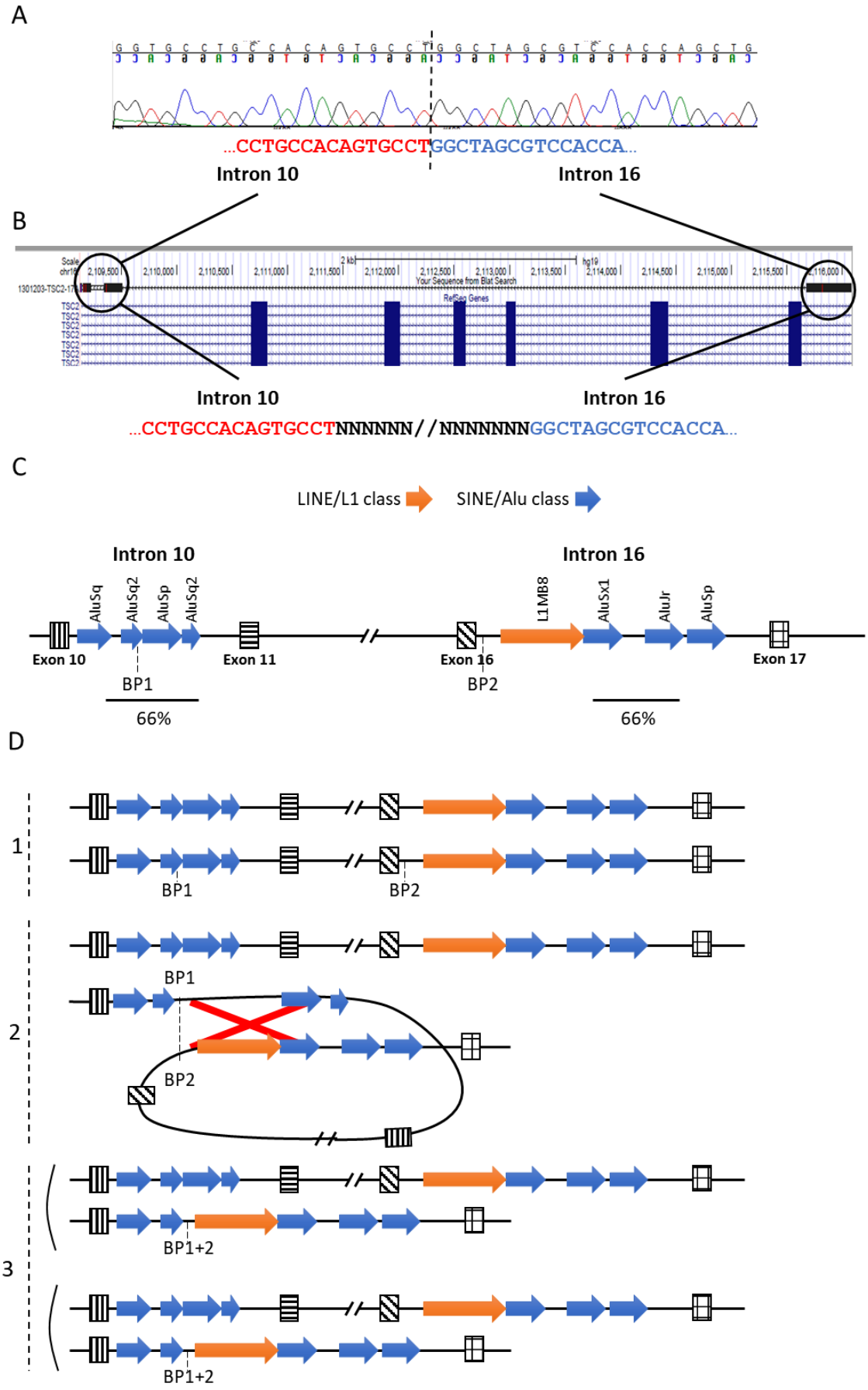

A
...CCTGCCACAGTGCCTGGCTAGCGTCCACCA...
Intron 10
Intron 16
B
Intron 10
Intron 16
...CCTGCCACAGTGCCTNNNNNN//NNNNNNNGGCTAGCGTCCACCA...
C
LINE/L1 class
SINE/Alu class
Intron 10
Intron 16
AluSq
AluSq2
AluSp
AluSq2
L1MB8
AluSx1
AluJr
AluSp
Exon 10
Exon 11
Exon 16
Exon 17
BP1
BP2
66%
66%
D
1
2
3
BP1
BP2
BP1
BP2
BP1+2
BP1+2

Figure 16: Breakpoint mapping of the *TSC2* intragenic deletion c.(975+_976-)_(1716+_1717-)del . (A) Sanger sequencing of the DNA from the patient with the intragenic *TSC2* deletion c.(975+_976-)_(1716+_1717-)del. (B) UCSC genome browser mapping of the breakpoints in *TSC2* introns 10 and 16. (C) Diagram illustrating the positions of SINE/Alu and LINE/L1 repeat sequences in *TSC2* introns 10 and 16, respectively within and adjacent to each breakpoint. The regions of 66% similarity between introns are indicated. (D) Steps representing the predicted intra-chromosomal crossing-over leading to deletion. 1: normal chromosomal pairing; 2: intra-chromosomal pairing. Red cross represents the crossing over. 3: Daughter cells with normal and TSC2 intragenic deleted chromosome.

D. NMI patients

An NMI patient presented a mosaic *TSC1* exon 1 deletion. The mosaic deletion was confirmed by Q-PCR, and deletion of upstream gene *GFI1B* has been ruled out. MLPA false-positive deletion of only one exon can be due to SNV impairing probe hybridization (van Dijk, Rombout et al. 2005, Kim, Cho et al. 2016). However, no DNA variant has been detected by Sanger sequencing of the patient *TSC1* exon 1. Due to the long extent of *TSC1* intron 1 (9,447 bp) and intergenic segment upstream (nearly 34 kb), we could not map the breakpoints of *TSC1* exon 1 deletion from our patient by a PCR-based approach with different primer pairs (data not shown). A study by van den Ouweland, Elfferich et al. (2011) identified five out of 327 TSC patients with deletions affecting *TSC1* exon 1, preserving exon 2. Upstream breakpoints from four deletions mapped differently between *TSC1* and *GFI1B* genes. Likewise, downstream breakpoints from the four cases mapped on *TSC1* intron 1 between 300 and more than 4 kb away the exon-intron junction. In one case of non-mosaic *TSC1* exon 1 deletion, only the wild-type allele was expressed, suggesting the deletion inactivated transcription of the gene (van den Ouweland, Elfferich *et al.* 2011). As we could not evaluate the *TSC1* mRNA levels from our patient, the patient classified as NMI and the mosaic *TSC1* exon 1 deletion of uncertain significance. If this is definitely not the TSC-causing mutation, it remains a possibility that the patient may have a pathogenic alteration in mosaicism that could not have been detected by sequencing due to limited sensitivity.

The second NMI patient has the c.1361+54_1361+57del variant in *TSC2* gene intron 13, which did not affect the splicing of the minigene primary transcript in HEK293 cell line (Figure 9). Since the patient mother has facial angiofibromas, we ruled out the possibility that the disease in the affected child is due to a novel mutation in mosaicism. As the patient had no segmental alteration on *TSC1* or *TSC2* detected by Sanger sequencing, NGS-Panel and MLPA, it is more likely that a DNA variant belongs to deep intronic sequence do not cover by our primer designed or to sequence that potentially altering *TSC1*/*2* gene expression regulation causes the disease.

NMI patients tend to correspond to less severe cases (Camposano, Greenberg *et al.* 2009). This class of patient should be submitted to whole gene-targeted NGS to access the deep intronic regions and possible mosaics (Nellist, Brouwer *et al.* 2015, Tyburczy, Dies *et al.* 2015).

Of note, one splicing mutation is a deep intronic variant (c.2838-122G>A) previously reported as pathogenic by a functional assay (Nellist, Brouwer *et al.* 2015) and appear in two unrelated patients. As demonstrated by recent reports (Nellist, Brouwer *et al.* 2015, Tyburczy, Dies *et al.* 2015), deep intronic DNA variants affecting splicing along with somatic mosaicism explains part of the NMI cases. Identification of deep

intronic variants was possible because in our study many primers designed for Sanger sequencing annealed on intronic sequences at least 200 bp away the exon/intron junctions.

Somatic mosaicism has been described in many individuals with TSC. The likely overestimate of the percent of mosaicism in some studies because are caused by the sensitivity of the pathogenic variant screening methods and the difficulty of detecting low-level mosaicism. It has been estimated that levels of mosaicism greater than 20% in lymphocyte DNA can be detected with confidence by Sanger sequencing and next-generation sequencing methods (Qin, Kozlowski *et al.* 2010). Using targeted deep next-generation sequencing of TSC1 and TSC2, somatic mosaic pathogenic variants at 9.5% in an affected individual who did not have a pathogenic variant identified by exon sequence analysis or gene-targeted deletion/duplication analysis was reported (Nellist, Brouwer *et al.* 2015).However, it has not yet been determined whether levels of somatic mosaicism below 9.5% can be detected by deep next-generation sequencing. Although detection of somatic mosaicism between 1% and 9.5% has been reported (Tyburczy, Dies *et al.* 2015).

VI. References

Abeysinghe, S. S., N. Chuzhanova, M. Krawczak, E. V. Ball and D. N. Cooper (2003). "Translocation and gross deletion breakpoints in human inherited disease and cancer I: Nucleotide composition and recombination-associated motifs." Hum Mutat **22**(3): 229-244.

Abo-Dalo, B., K. Kutsche, V. Mautner and L. Kluwe (2010). "Large intragenic deletions of the NF2 gene: breakpoints and associated phenotypes." Genes Chromosomes Cancer **49**(2): 171-175.

Adzhubei, I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova, P. Bork, A. S. Kondrashov and S. R. Sunyaev (2010). "A method and server for predicting damaging missense mutations." Nat Methods **7**(4): 248-249.

Ambar, G. and S. Chiavegatto (2009). "Anabolic-androgenic steroid treatment induces behavioral disinhibition and downregulation of serotonin receptor messenger RNA in the prefrontal cortex and amygdala of male mice." Genes Brain Behav **8**(2): 161-173.

Au, K. S., J. A. Rodriguez, J. L. Finch, K. A. Volcik, E. S. Roach, M. R. Delgado, E. Rodriguez, Jr. and H. Northrup (1998). "Germ-line mutational analysis of the TSC2 gene in 90 tuberous-sclerosis patients." Am J Hum Genet **62**(2): 286-294.

Au, K. S., A. T. Williams, E. S. Roach, L. Batchelor, S. P. Sparagana, M. R. Delgado, J. W. Wheless, J. E. Baumgartner, B. B. Roa, C. M. Wilson, T. K. Smith-Knuppel, M. Y. Cheung, V. H. Whittemore, T. M. King and H. Northrup (2007). "Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States." Genet Med **9**(2): 88-100.

Beauchamp, R. L., A. Banwell, P. McNamara, M. Jacobsen, E. Higgins, H. Northrup, P. Short, K. Sims, L. Ozelius and V. Ramesh (1998). "Exon scanning of the entire TSC2 gene for germline mutations in 40 unrelated patients with tuberous sclerosis." Hum Mutat **12**(6): 408-416.

Camposano, S. E., E. Greenberg, D. J. Kwiatkowski and E. A. Thiele (2009). "Distinct clinical characteristics of tuberous sclerosis complex patients with no mutation identified." Ann Hum Genet **73**(2): 141-146.

Choi, Y. and A. P. Chan (2015). "PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels." Bioinformatics **31**(16): 2745-2747.

Chong-Kopera, H., K. Inoki, Y. Li, T. Zhu, F. R. Garcia-Gonzalo, J. L. Rosa and K. L. Guan (2006). "TSC1 stabilizes TSC2 by inhibiting the interaction between TSC2 and the HERC1 ubiquitin ligase." J Biol Chem **281**(13): 8313-8316.

Consortium, T. E. C. T. S. (1993). "Identification and characterization of the tuberous sclerosis gene on chromosome 16." Cell **75**(7): 1305-1315.

Dabora, S. L., D. N. Franz, S. Ashwal, A. Sagalowsky, F. J. DiMario, Jr., D. Miles, D. Cutler, D. Krueger, R. N. Uppot, R. Rabenou, S. Camposano, J. Paolini, F. Fennessy, N. Lee, C. Woodrum, J. Manola, J. Garber and E. A. Thiele (2011). "Multicenter phase 2 trial of sirolimus for tuberous sclerosis: kidney angiomyolipomas and other tumors regress and VEGF- D levels decrease." PLoS One **6**(9): e23379.

Dabora, S. L., S. Jozwiak, D. N. Franz, P. S. Roberts, A. Nieto, J. Chung, Y. S. Choy, M. P. Reeve, E. Thiele, J. C. Egelhoff, J. Kasprzyk-Obara, D. Domanska-Pakiela and D. J. Kwiatkowski (2001). "Mutational analysis in a cohort of 224 tuberous sclerosis patients indicates increased severity of TSC2, compared with TSC1, disease in multiple organs." Am J Hum Genet **68**(1): 64-80.

Desmet, F. O., D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres and C. Beroud (2009). "Human Splicing Finder: an online bioinformatics tool to predict splicing signals." Nucleic Acids Res **37**(9): e67.

Dufner-Almeida, L. (2014). Estudo mutacional em pacientes com o complexo da esclerose tuberosa. Dissertação de Mesteado Dissertação de Mestrado, Universidade de São Paulo.

Fagerberg, L., B. M. Hallstrom, P. Oksvold, C. Kampf, D. Djureinovic, J. Odeberg, M. Habuka, S. Tahmasebpoor, A. Danielsson, K. Edlund, A. Asplund, E. Sjostedt, E. Lundberg, C. A. Szigyarto, M. Skogs, J. O. Takanen, H. Berling, H. Tegel, J. Mulder, P. Nilsson, J. M. Schwenk, C.

Lindskog, F. Danielsson, A. Mardinoglu, A. Sivertsson, K. von Feilitzen, M. Forsberg, M. Zwahlen, I. Olsson, S. Navani, M. Huss, J. Nielsen, F. Ponten and M. Uhlen (2014). "Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics." Mol Cell Proteomics **13**(2): 397-406.

Freeman, P. J., R. K. Hart, L. J. Gretton, A. J. Brookes and R. Dalgleish (2018). "VariantValidator: Accurate validation, mapping, and formatting of sequence variation descriptions." Hum Mutat **39**(1): 61-68.

Frey, U. H., H. S. Bachmann, J. Peters and W. Siffert (2008). "PCR-amplification of GC-rich regions: 'slowdown PCR'." Nat Protoc **3**(8): 1312-1317.

Garami, A., F. J. T. Zwartkruis, T. Nobukuni, M. Joaquin, M. Roccio, H. Stocker, S. C. Kozma, E. Hafen, J. L. Bos and G. Thomas (2003). "Insulin Activation of Rheb, a Mediator of mTOR/S6K/4E-BP Signaling, Is Inhibited by TSC1 and 2." Molecular Cell **11**(6): 1457-1466.

Gerstein, M. B., C. Bruce, J. S. Rozowsky, D. Zheng, J. Du, J. O. Korbel, O. Emanuelsson, Z. D. Zhang, S. Weissman and M. Snyder (2007). "What is a gene, post-ENCODE? History and updated definition." Genome Res **17**(6): 669-681.

Goncharova, E. A., D. A. Goncharov, A. Eszterhas, D. S. Hunter, M. K. Glassberg, R. S. Yeung, C. L. Walker, D. Noonan, D. J. Kwiatkowski, M. M. Chou, R. A. Panettieri, Jr. and V. P. Krymskaya (2002). "Tuberin regulates p70 S6 kinase activation and ribosomal protein S6 phosphorylation. A role for the TSC2 tumor suppressor gene in pulmonary lymphangioleiomyomatosis (LAM)." J Biol Chem **277**(34): 30958-30967.

Gu, S., B. Yuan, I. M. Campbell, C. R. Beck, C. M. Carvalho, S. C. Nagamani, A. Erez, A. Patel, C. A. Bacino, C. A. Shaw, P. Stankiewicz, S. W. Cheung, W. Bi and J. R. Lupski (2015). "Alu-mediated diverse and complex pathogenic copy-number variants within human chromosome 17 at p13.3." Hum Mol Genet **24**(14): 4061-4077.

Hoogeveen-Westerveld, M., R. Ekong, S. Povey, I. Karbassi, S. D. Batish, J. T. den Dunnen, A. van Eeghen, E. Thiele, K. Mayer, K. Dies, L. Wen, C. Thompson, S. P. Sparagana, P. Davies, C. Aalfs, A. van den Ouweland, D. Halley and M. Nellist (2012). "Functional assessment of TSC1 missense variants identified in individuals with tuberous sclerosis complex." Hum Mutat **33**(3): 476-479.

Hoogeveen-Westerveld, M., R. Ekong, S. Povey, K. Mayer, N. Lannoy, F. Elmslie, M. Bebin, K. Dies, C. Thompson, S. P. Sparagana, P. Davies, A. M. van Eeghen, E. A. Thiele, A. van den Ouweland, D. Halley and M. Nellist (2013). "Functional assessment of TSC2 variants identified in individuals with tuberous sclerosis complex." Hum Mutat **34**(1): 167-175.

Hoogeveen-Westerveld, M., L. van Unen, A. van den Ouweland, D. Halley, A. Hoogeveen and M. Nellist (2012). "The TSC1-TSC2 complex consists of multiple TSC1 and TSC2 subunits." BMC Biochem **13**: 18.

Hoogeveen-Westerveld, M., M. Wentink, D. van den Heuvel, M. Mozaffari, R. Ekong, S. Povey, J. T. den Dunnen, K. Metcalfe, S. Vallee, S. Krueger, J. Bergoffen, V. Shashi, F. Elmslie, D. Kwiatkowski, J. Sampson, C. Vidales, J. Dzarir, J. Garcia-Planells, K. Dies, A. Maat-Kievit, A. van den Ouweland, D. Halley and M. Nellist (2011). "Functional assessment of variants in the TSC1 and TSC2 genes identified in individuals with Tuberous Sclerosis Complex." Hum Mutat **32**(4): 424-435.

Hornbeck, P. V., B. Zhang, B. Murray, J. M. Kornhauser, V. Latham and E. Skrzypek (2015). "PhosphoSitePlus, 2014: mutations, PTMs and recalibrations." Nucleic Acids Res **43**(Database issue): D512-520.

Hung, C. C., Y. N. Su, S. C. Chien, H. H. Liou, C. C. Chen, P. C. Chen, C. J. Hsieh, C. P. Chen, W. T. Lee, W. L. Lin and C. N. Lee (2006). "Molecular and clinical analyses of 84 patients with tuberous sclerosis complex." BMC Med Genet **7**: 72.

Inoki, K., Y. Li, T. Xu and K. L. Guan (2003). "Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling." Genes Dev **17**(15): 1829-1834.

Jones, A. C., M. M. Shyamsundar, M. W. Thomas, J. Maynard, S. Idziaszczyk, S. Tomkins, J. R. Sampson and J. P. Cheadle (1999). "Comprehensive mutation analysis of TSC1 and TSC2-and phenotypic correlations in 150 families with tuberous sclerosis." Am J Hum Genet **64**(5): 1305-1315.

Kim, M. J., S. I. Cho, J. H. Chae, B. C. Lim, J. S. Lee, S. J. Lee, S. H. Seo, H. Park, A. Cho, S. Y. Kim, J. Y. Kim, S. S. Park and M. W. Seong (2016). "Pitfalls of Multiple Ligation-Dependent Probe Amplifications in Detecting DMD Exon Deletions or Duplications." J Mol Diagn **18**(2): 253-259.

Knudson, A. G., Jr. (1971). "Mutation and cancer: statistical study of retinoblastoma." Proc Natl Acad Sci U S A **68**(4): 820-823.

Kozlowski, P., P. Roberts, S. Dabora, D. Franz, J. Bissler, H. Northrup, K. S. Au, R. Lazarus, D. Domanska-Pakiela, K. Kotulska, S. Jozwiak and D. J. Kwiatkowski (2007). "Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations." Hum Genet **121**(3-4): 389-400.

Kumar, P., S. Henikoff and P. C. Ng (2009). "Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm." Nat Protoc **4**(7): 1073-1081.

Kwiatkowski, D. J., M. R. Palmer, S. Jozwiak, J. Bissler, D. Franz, S. Segal, D. Chen and J. R. Sampson (2015). "Response to everolimus is seen in TSC-associated SEGAs and angiomyolipomas independent of mutation type and site in TSC1 and TSC2." Eur J Hum Genet **23**(12): 1665-1672.

Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics." Bioinformatics **25**: 7.

Li, W., L. H. Zhou, B. D. Gao, L. Y. Li, C. G. Zhong, F. Gong, H. M. Xiao, T. Song and G. X. Lu (2011). "[Mutation screening and prenatal diagnosis of tuberous sclerosis complex]." Zhonghua Yi Xue Yi Chuan Xue Za Zhi **28**(4): 361-366.

Martignoni, G., F. Bonetti, M. Pea, R. Tardanico, M. Brunelli and J. N. Eble (2002). "Renal disease in adults with TSC2/PKD1 contiguous gene syndrome." Am J Surg Pathol **26**(2): 198-205.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly and M. A. DePristo (2010). "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data." Genome Res **20**(9): 1297-1303.

Moavero, R., M. Pinci, R. Bombardieri and P. Curatolo (2011). "The management of subependymal giant cell tumors in tuberous sclerosis: a clinician's perspective." Childs Nerv Syst **27**(8): 1203-1210.

Nellist, M., R. W. Brouwer, C. E. Kockx, M. van Veghel-Plandsoen, C. Withagen-Hermans, L. Prins-Bakker, M. Hoogeveen-Westerveld, A. Mrsic, M. M. van den Berg, A. E. Koopmans, M. C. de Wit, F. E. Jansen, A. J. Maat-Kievit, A. van den Ouweland, D. Halley, A. de Klein and I. W. F. van (2015). "Targeted Next Generation Sequencing reveals previously unidentified TSC1 and TSC2 mutations." BMC Med Genet **16**: 10.

Nellist, M., O. Sancak, M. A. Goedbloed, C. Rohe, D. van Netten, K. Mayer, A. Tucker-Williams, A. M. van den Ouweland and D. J. Halley (2005). "Distinct effects of single amino-acid changes to tuberin on the function of the tuberin-hamartin complex." Eur J Hum Genet **13**(1): 59-68.

Ng, P. C. and S. Henikoff (2006). "Predicting the effects of amino acid substitutions on protein function." Annu Rev Genomics Hum Genet **7**: 61-80.

Niida, Y., N. Lawrence-Smith, A. Banwell, E. Hammer, J. Lewis, R. L. Beauchamp, K. Sims, V. Ramesh and L. Ozelius (1999). "Analysis of bothTSC1 andTSC2 for germline mutations in 126 unrelated patients with tuberous sclerosis." Human Mutation **14**(5): 412-422.

Niida, Y., A. O. Stemmer-Rachamimov, M. Logrip, D. Tapon, R. Perez, D. J. Kwiatkowski, K. Sims, M. MacCollin, D. N. Louis and V. Ramesh (2001). "Survey of somatic mutations in tuberous sclerosis complex (TSC) hamartomas suggests different genetic mechanisms for pathogenesis of TSC lesions." Am J Hum Genet **69**(3): 493-503.

Northrup, H., D. A. Krueger and G. International Tuberous Sclerosis Complex Consensus (2013). "Tuberous sclerosis complex diagnostic criteria update: recommendations of the 2012 Iinternational Tuberous Sclerosis Complex Consensus Conference." Pediatr Neurol **49**(4): 243-254.

O'Callaghan, F. J., A. W. Shiell, J. P. Osborne and C. N. Martyn (1998). "Prevalence of tuberous sclerosis estimated by capture-recapture analysis." Lancet **351**(9114): 1490.

O'Callaghan, T. J., J. A. Edwards, M. Tobin and B. K. Mookerjee (1975). "Tuberous sclerosis with striking renal involvement in a family." Archives of Internal Medicine **135**(8): 1082-1087.

Osborne, J. P., A. Fryer and D. Webb (1991). "Epidemiology of tuberous sclerosis." Ann N Y Acad Sci **615**: 125-127.

Padma Priya, T. and A. B. Dalal (2012). "Tuberous sclerosis: diagnosis and prenatal diagnosis by MLPA." Indian J Pediatr **79**(10): 1366-1369.

Peron, A., K. S. Au and H. Northrup (2018). "Genetics, genomics, and genotype-phenotype correlations of TSC: Insights for clinical practice." Am J Med Genet C Semin Med Genet **178**(3): 281-290.

Piva, F., M. Giulietti, A. B. Burini and G. Principato (2012). "SpliceAid 2: a database of human splicing factors expression data and RNA target motifs." Hum Mutat **33**(1): 81-85.

Pujol, P., P. Vande Perre, L. Faivre, D. Sanlaville, C. Corsini, B. Baertschi, M. Anahory, D. Vaur, S. Olschwang, N. Soufir, N. Bastide, S. Amar, M. Vintraud, O. Ingster, S. Richard, P. Le Coz, J. P. Spano, O. Caron, P. Hammel, E. Luporsi, A. Toledano, X. Rebillard, A. Cambon-Thomsen, O. Putois, J. M. Rey, C. Herve, C. Zorn, K. Baudry, V. Galibert, J. Gligorov, D. Azria, B. Bressac-de Paillerets, N. Burnichon, M. Spielmann, D. Zarca, I. Coupier, O. Cussenot, A. P. Gimenez-Roqueplo, S. Giraud, A. S. Lapointe, P. Niccoli, I. Raingeard, M. Le Bidan, T. Frebourg, A. Rafii and D. Genevieve (2018). "Guidelines for reporting secondary findings of genome sequencing in cancer genes: the SFMPP recommendations." Eur J Hum Genet.

Qin, W., J. A. Chan, H. V. Vinters, G. W. Mathern, D. N. Franz, B. E. Taillon, P. Bouffard and D. J. Kwiatkowski (2010). "Analysis of TSC cortical tubers by deep sequencing of TSC1, TSC2 and KRAS demonstrates that small second-hit mutations in these genes are rare events." Brain Pathol **20**(6): 1096-1105.

Qin, W., P. Kozlowski, B. E. Taillon, P. Bouffard, A. J. Holmes, P. Janne, S. Camposano, E. Thiele, D. Franz and D. J. Kwiatkowski (2010). "Ultra deep sequencing detects a low rate of mosaic mutations in tuberous sclerosis complex." Hum Genet **127**(5): 573-582.

Rakowski, S. K., E. B. Winterkorn, E. Paul, D. J. Steele, E. F. Halpern and E. A. Thiele (2006). "Renal manifestations of tuberous sclerosis complex: Incidence, prognosis, and predictive factors." Kidney Int **70**(10): 1777-1782.

Rendtorff, N. D., B. Bjerregaard, M. Frodin, S. Kjaergaard, H. Hove, F. Skovby, K. Brondum-Nielsen, M. Schwartz and G. Danish Tuberous Sclerosis (2005). "Analysis of 65 tuberous sclerosis complex (TSC) patients by TSC2 DGGE, TSC1/TSC2 MLPA, and TSC1 long-range PCR sequencing, and report of 28 novel mutations." Hum Mutat **26**(4): 374-383.

Richards, S., N. Aziz, S. Bale, D. Bick, S. Das, J. Gastier-Foster, W. W. Grody, M. Hegde, E. Lyon, E. Spector, K. Voelkerding, H. L. Rehm and A. L. Q. A. Committee (2015). "Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology." Genet Med **17**(5): 405-424.

Riemenschneider, M. J., R. A. Betensky, S. M. Pasedag and D. N. Louis (2006). "AKT activation in human glioblastomas enhances proliferation via TSC2 and S6 kinase signaling." Cancer Res **66**(11): 5618-5623.

Roach, E. S., M. R. Gomez and H. Northrup (1998). "Tuberous sclerosis complex consensus conference: revised clinical diagnostic criteria." J Child Neurol **13**(12): 624-628.

Rosset, C., F. Vairo, I. C. Bandeira, R. L. Correia, F. V. de Goes, R. T. B. da Silva, L. S. M. Bueno, M. C. S. de Miranda Gomes, H. C. R. Galvao, J. Neri, M. I. Achatz, C. B. O. Netto and P. Ashton-Prolla (2017). "Molecular analysis of TSC1 and TSC2 genes and phenotypic correlations in Brazilian families with tuberous sclerosis." PLoS One **12**(10): e0185713.

Rozen, S. and H. J. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology." Humana Press: 2.

Sampson, J. R., S. J. Scahill, J. B. Stephenson, L. Mann and J. M. Connor (1989). "Genetic aspects of tuberous sclerosis in the west of Scotland." J Med Genet **26**(1): 28-31.

Sancak, O., M. Nellist, M. Goedbloed, P. Elfferich, C. Wouters, A. Maat-Kievit, B. Zonnenberg, S. Verhoef, D. Halley and A. van den Ouweland (2005). "Mutational analysis of the TSC1 and TSC2 genes in a diagnostic setting: genotype--phenotype correlations and comparison of diagnostic DNA techniques in Tuberous Sclerosis Complex." Eur J Hum Genet **13**(6): 731-741.

Schwarz, J. M., C. Rodelsperger, M. Schuelke and D. Seelow (2010). "MutationTaster evaluates disease-causing potential of sequence alterations." Nat Methods **7**(8): 575-576.

Stapleton, F. B., D. Johnson, G. W. Kaplan and W. Griswold (1980). "The cystic renal lesion in tuberous sclerosis." The Journal of Pediatrics **97**(4): 574-579.

Tee, A. R., D. C. Fingar, B. D. Manning, D. J. Kwiatkowski, L. C. Cantley and J. Blenis (2002). "Tuberous sclerosis complex-1 and -2 gene products function together to inhibit mammalian target of rapamycin (mTOR)-mediated downstream signaling." Proc Natl Acad Sci U S A **99**(21): 13571-13576.

Tee, A. R., B. D. Manning, P. P. Roux, L. C. Cantley and J. Blenis (2003). "Tuberous Sclerosis Complex Gene Products, Tuberin and Hamartin, Control mTOR Signaling by Acting as a GTPase-Activating Protein Complex toward Rheb." Current Biology **13**(15): 1259-1268.

Tyburczy, M. E., K. A. Dies, J. Glass, S. Camposano, Y. Chekaluk, A. R. Thorner, L. Lin, D. Krueger, D. N. Franz, E. A. Thiele, M. Sahin and D. J. Kwiatkowski (2015). "Mosaic and Intronic Mutations in TSC1/TSC2 Explain the Majority of TSC Patients with No Mutation Identified by Conventional Testing." PLoS Genet **11**(11): e1005637.

Tyburczy, M. E., J. A. Wang, S. Li, R. Thangapazham, Y. Chekaluk, J. Moss, D. J. Kwiatkowski and T. N. Darling (2013). "Sun exposure causes somatic second-hit mutations and angiofibroma development in tuberous sclerosis complex." Human Molecular Genetics **23**(8): 20232029.

van den Ouweland, A. M., P. Elfferich, B. A. Zonnenberg, W. F. Arts, T. Kleefstra, M. D. Nellist, J. M. Millan, C. Withagen-Hermans, A. J. Maat-Kievit and D. J. Halley (2011). "Characterisation of TSC1 promoter deletions in tuberous sclerosis complex patients." Eur J Hum Genet **19**(2): 157-163.

van Dijk, M. C., P. D. Rombout, S. H. Boots-Sprenger, H. Straatman, M. R. Bernsen, D. J. Ruiter and J. W. Jeuken (2005). "Multiplex ligation-dependent probe amplification for the detection of chromosomal gains and losses in formalin-fixed tissue." Diagn Mol Pathol **14**(1): 9-16.

van Slegtenhorst, M., R. de Hoogt, C. Hermans, M. Nellist, B. Janssen, S. Verhoef, D. Lindhout, A. van den Ouweland, D. Halley, J. Young, M. Burley, S. Jeremiah, K. Woodward, J. Nahmias, M. Fox, R. Ekong, J. Osborne, J. Wolfe, S. Povey, R. G. Snell, J. P. Cheadle, A. C. Jones, M. Tachataki, D. Ravine, J. R. Sampson, M. P. Reeve, P. Richardson, F. Wilmer, C. Munro, T. L. Hawkins, T. Sepp, J. B. Ali, S. Ward, A. J. Green, J. R. Yates, J. Kwiatkowska, E. P. Henske, M. P. Short, J. H. Haines, S. Jozwiak and D. J. Kwiatkowski (1997). "Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34." Science **277**(5327): 805-808.

van Slegtenhorst, M., M. Nellist, B. Nagelkerken, J. Cheadle, R. Snell, A. van den Ouweland, A. Reuser, J. Sampson, D. Halley and P. van der Sluijs (1998). "Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products." Hum Mol Genet **7**(6): 7.

van Slegtenhorst, M. A., A. Verhoef, A. Tempelaars, L. Bakker, Q. Wang, M. Wessels, R. Bakker, M. Nellist, D. Lindhout, D. Halley and A. van den Ouweland (1999). "Mutational spectrum of the TSC1 gene in a cohort of 225 tuberous sclerosis complex patients: no evidence for genotype-phenotype correlation." J Med Genet **36**(4): 5.

Wang, K., M. Li and H. Hakonarson (2010). "ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data." Nucleic Acids Research **38**(16): e164-e164.

Weren, R. D., M. J. Ligtenberg, C. M. Kets, R. M. de Voer, E. T. Verwiel, L. Spruijt, W. A. van Zelst-Stams, M. C. Jongmans, C. Gilissen, J. Y. Hehir-Kwa, A. Hoischen, J. Shendure, E. A. Boyle, E. J. Kamping, I. D. Nagtegaal, B. B. Tops, F. M. Nagengast, A. Geurts van Kessel, J. H. van Krieken, R. P. Kuiper and N. Hoogerbrugge (2015). "A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer." Nat Genet **47**(6): 668-671.

Wiederholt, W. C., M. R. Gomez and L. T. Kurland (1985). "Incidence and prevalence of tuberous sclerosis in Rochester, Minnesota, 1950 through 1982." Neurology **35**(4): 3.

Wildeman, M., E. van Ophuizen, J. T. den Dunnen and P. E. Taschner (2008). "Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker." Hum Mutat **29**(1): 6-13.

Wilson, P. J., V. Ramesh, A. Kristiansen, C. Bove, S. Jozwiak, D. J. Kwiatkowski, M. P. Short and J. L. Haines (1996). "Novel Mutations Detected in the TSC2 Hene From Both Sporadic and Familial TSC Patients." Human Molecular Genetics **5**(2): 249-256.

Zerbino, D. R., P. Achuthan, W. Akanni, M. R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, C. G. Giron, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T. Hourlier, O. G. Izuogu, S. H. Janacek, T. Juettemann, J. K. To, M. R. Laird, I. Lavidas, Z. Liu, J. E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D. N. Murphy, V. Newman, M. Nuhn, D. Ogeh, C. K. Ong, A. Parker, M. Patricio, H. S. Riat, H. Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, A. Zadissa, A. Frankish, S. E. Hunt, M. Kostadima, N. Langridge, F. J. Martin, M. Muffato, E. Perry, M. Ruffier, D. M. Staines, S. J. Trevanion, B. L. Aken, F. Cunningham, A. Yates and P. Flicek (2018). "Ensembl 2018." Nucleic Acids Res **46**(D1): D754-D761.

Zhang, H., G. Cicchetti, H. Onda, H. B. Koon, K. Asrican, N. Bajraszewski, F. Vazquez, C. L. Carpenter and D. J. Kwiatkowski (2003). "Loss of Tsc1/Tsc2 activates mTOR and disrupts PI3K-Akt signaling through downregulation of PDGFR." J Clin Invest **112**(8): 1223-1233.

VII. Supplementary Tables

Supplementary Table 1: PCR and Sanger sequencing primers for the *TSC1* (NG_012386.1).

Targeted segment	Name	Oligonucleotide sequence (5' – 3')
Promoter + exon 1	gTSC1_1S	AAATGTTTAGCCCAGGAAGGA
	gTSC1_1A	GCCGGAGATAGCGTGTAATAA
	gTSC1_2A	CATCTTGGACGTACAGCACCT
	gTSC1_2S	CCGTCTATCCTTCCTTTCGAG
Exon 2	gTSC1_3S	TTGGATTTTAACCCGGAACTC
	gTSC1_3A	TCAGGCACTGAATACAAGCAA
Exon 3	gTSC1_4A	GGGGTTCACTGCATGATTCT
	gTSC1_4S	CCTCTTCATAAACTCGCCAAAG
Exon 4	gTSC1_5S	CAGAACTGTAATGCTGCACAAA
	gTSC1_5A	TTCAAGAATCATGGGTCCTACA
Exon 5	gTSC1_6S	TTTTATCTGCATGACCCTTGC
	gTSC1_6A	CCATACTTGCATGGACAAGGT
Exon 6	gTSC1_7S	CAGTAGAGTTGGGGCTCAGTG
	gTSC1_7A	GCACCCAAGATATTCCCTCA
Exons 7+8	gTSC1_8S	CTGAAGAGGAGGGCAGAAGTT
	gTSC1_8A	ATTAGTCCTCCGCCTGTGAA
	gTSC1_9A	AATTTCCCTGTCTGCCGTTA
	gTSC1_29S	CAATCCCTAGGCAGCCACTA
Exon 9	gTSC1_10S	TTTCCATTTTGAGGCTACACC
	gTSC1_10A	TTCCAGAGACAAAGTTGCAAAA
Exon 10	gTSC1_28S	ACCTAAAACCACACACTAACCC
	gTSC1_11A	GGAATGCTAAGTCATCCACGA
Exons 11+12	gTSC1_12S	GGGAAAATTTCACACTGCTCA
	gTSC1_12A	CACACCTTGAGAGCAGCTTGT
	gTSC1_13A	CCCAGGGATTTGCAATAAGT
	gTSC1_13S	CGGCAGTTTTTCTAATAGTTGG
Exons 13+14	gTSC1_14S	CATCCCAACAATTTGAGAATCA
	gTSC1_14A	GGCATCACTTTACCTGGCATA
	gTSC1_15A	TCCCAGAATTTCCTTGTTTCC
	gTSC1_15S	CCATGTCCAGCCTTCTCTGT
Exon 15	gTSC1_16S	GGATGCCACTTTTTCTCCTCT
	gTSC1_16A	TCCCAATTTAGGTGCACAGAG
	gTSC1_17A	GATGACAAAATGATGGGCTGT
	gTSC1_17S	CACACCAAAGCAAGCCTTTAC
Exons 16+17	gTSC1_18S	TTATGCCATTGCAGATTTTGAC
	gTSC1_18A	GGAAGGACTGGGAACTCTGAC
	gTSC1_19A	ACTTGGCAACACTTGAGATCCT
	gTSC1_19S	AAGCTAACAACACATGGGAAGG
Exon 18	gTSC1_20S	GCAAACTGATCCCTGAGAAGA
	gTSC1_20A	AGTTGGGGAACCTCTGTCCTA
Exon 19	gTSC1_21S	CAGAATCTTTCTGCAGCATCC
	gTSC1_21A	CAGCACCAAAAACATGAACCT
Exon 20	gTSC1_22S	CCATTATGTCAGGGACTGTGAA
	gTSC1_22A	TAGCTGGACCACGGAGTAGTG
Exons 21+22	gTSC1_23S	GCTTGGGGATAGATTTCAAGG
	gTSC1_23A	ACACGGAGTGAGCTGAGTGTT
	gTSC1_24A	TGCAGCTGTCCTCTGAAAGAT
	gTSC1_24S	GTCAAACTCCAGGCAAGGTAA
Exon 23	gTSC1_25S	CATATGGCCACAGGAAGTGTT
	gTSC1_28A	CCGTCCCATTTCCACACATG
	gTSC1_26A	CAGAAAGGCTACTGGTCATGC
	gTSC1_26S	GGGAGACGACTATGGGAGAAG

Supplementary Table 2: PCR and Sanger sequencing primers for the *TSC2* (NG_005895.1).

Targeted segment	Name	Oligonucleotide sequence (5' – 3')
Promoter + exon 1	gTSC2_1S	CGAGGACAGCAAGTTCACTG
	gTSC2_1A	GTTTGCCGTCTCTCCTCTACC
	gTSC2_2A	GAGCTTGCTGGGAGTTGTAGTT
	gTSC2_2S	CTACCTGCTGCAGCCTCTCT
Exon 2	gTSC2_3S	GGTAGAGGAGAGACGGCAAAC
	gTSC2_3A	AAGTGTGCCTGAACCAGGTC
Exon 3	gTSC2_4S	CGGCTCGTCAAGTGAATCTT
	gTSC2_4A	GTCAGCTGTCAACCATGTTCC
Exon 4	gTSC2_5S	TGAGACTGTCCCATGACTTCC
	gTSC2_5A	AGGGCAAAACAACACCGTAG
Exon 5	gTSC2_43S	CCTGCCCTGTACAATGCTGA
	gTSC2_43A	CAAGCCCCAGAGACTCACAG
Exon 6	gTSC2_44S	GATCCTAGTGTCCGTGCGTAG
	gTSC2_44A	CGGAGCTGAACTTAGGACCAT
Exon 7	gTSC2_8S	GCTCTCATCTGATGTCTTGGTT
	gTSC2_8A	GTCATTGATGCTGTCATCCAC
Exon 8	gTSC2_9S	GTCCCCCATGTAAGTCAGGAT
	gTSC2_9A	ATCTCCTCCCAAAGACAGAGG
Exon 9	gTSC2_10S	CTGTCTCCCATGAATGGTTGT
	gTSC2_10A	GGCTAAGTAGTTGGGGAGCAC
Exon 10	gTSC2_45S	GTGTTACTGCTGGCCTCTGT
	gTSC2_11A	CAGCTCACTGCACACAGAAAC
Exon 11	gTSC2_12S	GGATTCAGTTGCTGGTCTGTC
	gTSC2_12A	ACTAATGCGGTCCTCCAAAGT
Exon 12	gTSC2_46S	CCTCTGGTGCCAAGTCCATG
	gTSC2_45A	CCCTAAGCTGAGTGTTCCTGG
Exon 13	gTSC2_14S	CAGTTTCCTCCCACCTGTGT
	gTSC2_14A	GGAGCATCTCTCCAGACGAC
Exon 14	gTSC2_15S	GTGCTAGCTTGCTTTCCAGTC
	gTSC2_15A	AGACTGGCTGAAACGAACTCA
Exon 15	gTSC2_16S	GCTGCTCCTTGTGAGTTGTG
	gTSC2_16A	ACTGTGCAGAAACCAAAATGC
Exon 16	gTSC2_17S	CTCAGAACCATGAGCCTGTGT
	gTSC2_17A	AGCGTGTGCTACTGGTATGCT
Exon 17	gTSC2_18S	GTTGATGACTGCCCTGATGAT
	gTSC2_18A	TTAGAGCGACAAGCCACAGAT
Exon 18	gTSC2_19S	CAGAGTCCTGTTCAGCCTGTC
	gTSC2_19A	GAAGCAAGAGAAGCAGCTGAG
Exon 19	gTSC2_20S	CTACATGTACGCGGGACCTC
	gTSC2_20A	GCCTTCTGGACCCTAGAGACA
Exon 20	gTSC2_21S	GTGCCCTACTCCCTGCTCTT
	gTSC2_46A	GCTCGCAGTCTTTTGGGGAA
Exon 21	gTSC2_47S	GTGTGTTACTTGGCAGGCAC
	gTSC2_22A	GTGGACAGGGAACACTGGAT
Exon 22	gTSC2_23S	GAGTCTGCTCGGGTAGCTCA
	gTSC2_23A	ACCTGAGCTCCTGAAGTCACA
Exon 23	gTSC2_24S	TCACGGATCACACAAATGGTA
	gTSC2_24A	GAGCCCACCTTAGTGATGAAA
Exon 24	gTSC2_25S	CGCACCTCTACAGGAACTTTG
	gTSC2_25A	GAGTGAGCACACCCAGACAGT
Exon 25	gTSC2_26S	TCATCACTAAGGTGGGCTCAG
	gTSC2_26A	AACCCCCAATTCCACAAGTAG
Exon 26	gTSC2_27S	ACCCACACACGTTTAATTTGC
	gTSC2_27A	GAATACGAAAAGGCCAAAACC
Exons 27 + 28	gTSC2_28S	AATGTGGTCCACGTGATTCTC
	gTSC2_28A	GACTTAGTCCCCAGGCTGGTA
Exon 29	gTSC2_29S	CGCTCCCTGTCTTCTAGGTCT
	gTSC2_29A	CAGAGAAGGGCTCCAGGACT
Exon 30	gTSC2_48S	CTTGAGGCTGGTGGTTTTGC
	gTSC2_47A	AGAGGGCCAAGTCTGCAATC
Exon 31	gTSC2_31S	TGAGGGGTGCAAAGAGTAGG
	gTSC2_31A	GGAGAACAATGGTGCTGAGG

Exon 32	gTSC2_32S	GACGTCTATTCACGGGAGGA
	gTSC2_32A	CTAAACAGCTGCCACCCATC
Exon 33	gTSC2_33S	GTTACGAGGGCTGGTTTCAG
	gTSC2_33A	ACACTGCGTGAGCAGAGGTAT
Exon 34	gTSC2_34S	ATACCTCTGCTCACGCAGTGT
	gTSC2_34A	AGCTGCAGGAACACGAAACT
	gTSC2_35A	CTCTTTAAGGCGTCCCTCTCT
	gTSC2_35S	CTGTGGACCTCTCCTTCCAG
Exon 35	gTSC2_36S	AGCCTCCAATGCAGAGAAAGT
	gTSC2_36A	GTGTCGTATGATGGGATCTGG
Exon 36	gTSC2_37S	TGTTCCTGCAGCTCTACCATT
	gTSC2_37A	TGTCAGCTCACTGACCAACAG
Exon 37	gTSC2_38S	GAGGGAAGAGAGGGAGTCAAG
	gTSC2_38A	GGCACCTCCTGATTACTCCA
Exon 38	gTSC2_39S	CTCCCATCCAGTCCTGCTAC
	gTSC2_39A	TCTGCACTTGCCAGTTACTCC
Exons 39 + 40	gTSC2_49S	CAGAGGGGAAAGTTCAGGGG
	gTSC2_40A	GTAGATATCGGTGGGGTTGGA
Exon 41	gTSC2_41S	AAGTCTCCCCAGACATGGAG
	gTSC2_41A	CACAAACTCGGTGAAGTCCTC
Exon 42	gTSC2_42S	CCGATATCTACCCCTCCAAGT
	gTSC2_48A	CTTCTAGAGCCTCGACACCC

Supplementary Table 3: *TSC1* exon 1 slowdown PCR cycle program.

Cycles	Time	Temperature	Ramp
1X	5min	95°C	2,5°/s
53X	30sec.	95°C	2,5°/s
	30sec.	70°C (-0,3°/cycle)	1,5°/s
	40sec.	72°C	2,5°/s
15X	30sec.	95°C	2,5°/s
	30sec.	58°C	1,5°/s
	40sec.	72°C	2,5°/s
1X	7min	72°C	-
1X	∞	10°C	-

Supplementary Table 4: Standard PCR cycle program.

Cycles	Time	Temperature
1X	5 min	95°C
30X	1 min	95°C
	1 min	60°C
	1 min	72°C
1X	8 min	72°C
1X	∞	4°C

Supplementary Table 5: qPCR primers for *TSC1* (NM_000368.4), *TSC2* (NM_000548.3), *GAPDH* (NM_002046), *GFI1B* (NM_004188.6), *SPACA9* (NM_001316898.1), *NTHL1* (NM_002528) and *PKD1* (NM_001009944).

Segment	Name	Oligonucleotide sequence (5' – 3')	Product size (bp)	DNA concentration (ng/µL)	Efficiency
		TSC1			
Exon 1	qPCRgTSC1_1S qPCRgTSC1_1A	AGGGACTGTGAGGTAAACAGC AGGAAGCCCCCATAAAAAGGAG	85	3.125	2.0111
Exon 17	qPCRgTSC1_4S qPCRgTSC1_4A	CAGATGAGATCCGCACCCTC AGCTGCTGCTTTGATCACCT	131	3.125	2.0324
		TSC2			
Exon 9	qPCRgTSC2_7S qPCRgTSC2_7A	CCACAGCGCCATCTACAACA AAGTCATGGGCCTCGCTCTA	195	3.125	1.9019
Exon 12	qPCRgTSC2_2S qPCRgTSC2_2A	TCCATGACCTGTTGACCACG CGCACATCTCTCCACCAGTT	95	3.125	2.0527
Exon 19	qPCRgTSC2_4S qPCRgTSC2_4A	CTGAGAAGAAGACCAGCGGC GAAGAGCAGGGAGTAGGGCA	107	6.25	1.9124
Exon 41	qPCRgTSC2_8S qPCRgTSC2_8A	TCACAGGTGCATCATAGCCG CCCATATTCCCTACCCGCTG	107	3.125	2.2134
		GADPH			
Intron 1	gGAPDH_1S gGAPDH_1A	GCTCCCACCTTTCTCATCC CTGCAGCGTACTCCCCAC	147	6.25	2.0545
		GFI1B			
Exon 1	qPCRgGFI1B_1S qPCRgGFI1B_1A	TATTGCGAGGCTGAGTATGATGG TTCATATCCCACAGTTCGACTTCA	96	3.125	1.9485
		SPACA9			
Exon 4	qPCRgSPACA9_2S qPCRgSPACA9_2A	GAGGCCCGGAACCACTAC GAGTGGGCGATCCATTCTTTC	111	3.125	2.1000
		NTHL1			
Exon 6	qPCRgNTHL1_1S qPCRgNTHL1_1A	TCAATGGACTCTTGGTGGGC AAAGCCACTTCACAGACGGT	163	6.25	2.1350
		PKD1			
Exon 45	qPCRgPKD1_1S qPCRgPKD1_1A	GGCTCTCTACCCTGTGTCCT CGGAGAATAACAGCCCCCAG	112	3.125	2.1407
Exon 46	qPCRgPKD1_2S qPCRgPKD1_2A	TAGGTGTGGTGGCGTTATGG CTCTGGGGTGATGAGAGTGC	107	6.25	2.1537

Supplementary Table 6: qPCR cycle program.

Cycling		Time	Temperature	Ramp
Holding	1X	2 min	50°C	100%
	1X	10 min	95°C	100%
Cycling	35X	15 seg	95°C	100%
		1 min	60°C	100%
Melting	1X	15 seg	95°C	100%
	1X	1 min	60°C	100%
	1X	30 seg	95°C	1%
	1X	15 seg	60°C	100%

Supplementary Table 7: List of softwares and websites consulted.

Program	Web site
Primer3	http://primer3.ut.ee/
Primer-BLAST	https://www.ncbi.nlm.nih.gov/tools/primer-blast/

BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Beacon Designer Free Edition	http://www.premierbiosoft.com/qOligo/Oligo.jsp?PID=1
BLAT	http://genome.ucsc.edu/cgi-bin/hgBlat?command=start
Human Genome Variation Society	http://varnomen.hgvs.org/
Variant Validator	https://variantvalidator.org/
Mutalyzer	http://www.lovd.nl/mutalyzer/
Leiden Open Variation Database TSC1	http://chromium.lovd.nl/LOVD2/TSC/home.php?select_db=TSC1
Leiden Open Variation Database TSC2	http://chromium.lovd.nl/LOVD2/TSC/home.php?select_db=TSC2
Ensembl	http://www.ensembl.org/index.html
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/
Mutation Taster	http://www.mutationtaster.org/
SIFT and PROVEAN	http://provean.jcvi.org/genome_submit_2.php
PhosphositePlus	https://www.phosphosite.org/homeAction.action?jsessionid=D937EFB1B37CAF55C7289FE52C813AA1
Acescan2	http://genes.mit.edu/acescan2/index.html
SpliceAid2	http://193.206.120.249/splicing_tissue.html
Human Splicing Finder 3.1	http://www.umd.be/HSF3/
RepeatMasker	http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker
UniProtKB BLAST	https://www.uniprot.org/help/uniprotkb

Supplementary Table 8: Primer pairs for function assessment of splicing variant.

Segment	Name	Oligonucleotide sequence (5' – 3')
Mini-gene Exon 12-14	gTSC2_51S	GCCCATGGGGATCCACCTTGGACAGCCCGGAGC
	gTSC2_50A	TCTAGAACTCGAGCTCATAGAACTGCCTGTTGATGAGC
Mini-gene Exon 13-35	gTSC2_53S	GCCCATGGGGATCCACTCCGCCGTGGTCATGGAG
	gTSC2_52A	TCTAGAAGCGGCCGCCTCATTGGGCAGCAGGATTGGC
Endogenous Exons 8-10	cTSC2_2S	TGGTCTGCTACAACTGCCTG
	cTSC2_1A	CGGCGAGTTCCTGAGAGAAT

BamHI restriction site: GGATTC
XhoI restriction site: CTCGAG
NotI restriction site: GCGGCCGC

Supplementary Table 9: Rules for combining criteria to classify sequence variants according to ACMG (Richards, Aziz *et al.* 2015).

Rules for combining criteria to classify sequence variants		
Pathogenic	(i) 1 Very strong (PVS1) AND	(a) ≥1 Strong (PS1–PS4) OR (b) ≥2 Moderate (PM1–PM6) OR (c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) OR (d) ≥2 Supporting (PP1–PP5) OR
	(ii) ≥2 Strong (PS1–PS4) OR	
	(iii) 1 Strong (PS1–PS4) AND	(a)≥3 Moderate (PM1–PM6) OR (b)2 Moderate (PM1–PM6) AND ≥2 Supporting (PP1–PP5) OR (c)1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5).
Likely pathogenic	(i) 1 Very strong (PVS1) AND 1 moderate (PM1– PM6) OR (ii) 1 Strong (PS1–PS4) AND 1–2 moderate (PM1–PM6) OR (iii) 1 Strong (PS1–PS4) AND ≥2 supporting (PP1–PP5) OR (iv) ≥3 Moderate (PM1–PM6) OR (v) 2 Moderate (PM1–PM6) AND ≥2 supporting (PP1–PP5) OR (vi) 1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5)	
Benign	(i) 1 Stand-alone (BA1) OR (ii) ≥2 Strong (BS1–BS4).	
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1– BP7) OR (ii) ≥2 Supporting (BP1–BP7).	
Uncertain significance	(i) Other criteria shown above are not met OR (ii) the criteria for benign and pathogenic are contradictory.	

Supplementary Table 10: Evidence of benign impact according to ACMG (Richards, Aziz *et al.* 2015).

	Evidence of benign impact
Stan-alone	BA1: Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium
Strong	BS1: Allele frequency is greater than expected for disorder (see Table 6) BS2: Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age BS3: Well-established *in vitro* or *in vivo* functional studies show no damaging effect on protein function or splicing BS4: Lack of segregation in affected members of a family
Supporting	BP1: Missense variant in a gene for which primarily truncating variants are known to cause disease BP2: Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern BP3: In-frame deletions/insertions in a repetitive region without a known function BP4: Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.) BP5: Variant found in a case with an alternate molecular basis for disease BP6 Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation BP7 A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved

Supplementary Table 11: *TSC1* and *TSC2* exonic single nucleotide variants.

Variant	Exon	Existing variation	1000 Genomes								gnomAD								
			AF	AFR	AMR	EAS	EUR	SAS	AA	EA	AF	AFR	AMR	ASJ	EAS	FIN	NFE	OTH	SAS
TSC1																			
c.-99C>T	02/23	rs114755636	0.004	0.015	0.001	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.-129A>T	02/23	rs116951280	0.005	0.000	0.013	0.000	0.014	0.001	-	-	-	-	-	-	-	-	-	-	-
c.840A>G (p.=)	09/23	rs1171852730	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.965T>C (p.Met322Thr)	10/23	rs1073123	0.135	0.234	0.094	0.074	0.131	0.098	0.220	0.138	0.125	0.217	0.086	0.141	0.088	0.102	0.138	0.126	0.104
c.1342C>T (p.Pro448Ser)	14/23	rs118203518	0.004	0.015	0.001	0.000	0.000	0.000	0.011	0.000	0.001	0.011	0.001	0.000	0.000	0.000	0.000	0.000	0.000
c.1335A>G (p.=)	14/23	rs7862221	0.139	0.233	0.097	0.074	0.144	0.101	0.219	0.146	0.132	0.218	0.088	0.150	0.088	0.115	0.148	0.135	0.107
c.1760A>G (p.Lys587Arg)	15/23	rs118203576	0.018	0.000	0.128	0.000	0.002	0.000	0.002	0.000	0.026	0.002	0.175	0.000	0.000	0.012	0.000	0.026	0.001
c.2829C>T (p.=)	22/23	rs4962081	0.062	0.095	0.097	0.023	0.064	0.032	0.091	0.082	0.079	0.089	0.154	0.053	0.028	0.046	0.081	0.076	0.051
c.*4437G>A	23/23	rs11553763	0.103	0.214	0.123	0.040	0.068	0.043	-	-	-	-	-	-	-	-	-	-	-
c.*3679G>A	23/23	rs1050700	0.726	0.595	0.648	0.825	0.749	0.832	-	-	-	-	-	-	-	-	-	-	-
c.*1507G>A	23/23	rs739441	-	0.660	0.752	0.821	0.830	0.834	-	-	-	-	-	-	-	-	-	-	-
c.*1488C>T	23/23	rs739442	0.451	0.056	0.496	0.581	0.612	0.651	-	-	-	-	-	-	-	-	-	-	-
c.*1362G>T	23/23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.*1275T>G	23/23	rs2809244	-	0.778	0.738	0.682	0.753	0.780	-	-	-	-	-	-	-	-	-	-	-
c.*289delT	23/23	rs11323835	0.439	0.137	0.473	0.566	0.515	0.616	-	-	-	-	-	-	-	-	-	-	-
c.*191dupC	23/23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.*107T>C	23/23	rs116917669	0.017	0.002	0.020	0.002	0.044	0.024	-	-	-	-	-	-	-	-	-	-	-
c.*1G>A	23/23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.3324C>T (p.=)	23/23	rs35593170	0.016	0.003	0.095	0.006	0.002	0.001	0.001	0.001	0.018	0.002	0.125	0.000	0.003	0.002	0.001	0.013	0.001
c.3282G>A (p.=)	23/23	rs116747861	0.005	0.017	0.001	0.000	0.000	0.000	0.013	0.001	0.002	0.014	0.002	0.005	0.000	0.000	0.001	0.002	0.000

c.3127_3129del (p.Ser1043del)	23/23	rs397514812	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TSC2																			
c.52C>G (p.Leu18Val)	02/42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.499T>C (p.Trp167Arg)	06/42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.881G>A (p.Gly294Glu)	10/42	rs45487497	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1100G>A (p.Arg367Gln)	11/42	rs1800725	0.009	0.002	0.016	0.000	0.014	0.017	0.005	0.020	0.014	0.003	0.011	0.027	0.000	0.007	0.018	0.015	0.017
c.1443G>T (p.Glu481Asp)	14/42	CS091000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1578C>T (p.=)	15/42	rs34012042,	0.035	0.002	0.046	0.000	0.073	0.070	0.015	0.070	0.057	0.011	0.033	0.100	0.000	0.086	0.067	0.064	0.070
c.1731C>T (p.=)	17/42	rs144122318	0.000	0.000	0.000	0.000	0.001	0.000	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.2073C>T (p.=)	19/42	rs45512398	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.011	0.000	0.000	0.000	0.001	0.001
c.2580T>C (p.=)	23/42	rs13337626	0.063	0.093	0.042	0.002	0.088	0.074	0.087	0.086	0.070	0.088	0.035	0.077	0.001	0.111	0.082	0.069	0.063
c.3126G>C (p.=)	27/42	rs36078782	0.013	0.047	0.003	0.000	0.000	0.000	0.037	0.000	0.003	0.043	0.003	0.000	0.000	0.000	0.000	0.003	0.000
c.3431T>A (p.Val1144Glu)	30/42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.3557A>G (p.Tyr1186Cys)	30/42	rs137854421	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
c.3768G>A (p.=)	31/42	rs201599540	0.000	0.000	0.000	0.001	0.000	0.000	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.3889G>A (p.Ala1297Thr)	33/42	rs45517319	0.001	0.000	0.000	0.000	0.003	0.002	0.002	0.004	0.003	0.001	0.001	0.004	0.000	0.005	0.004	0.002	0.002
c.3914C>T (p.Pro1305Leu)	33/42	rs45517320	0.009	0.033	0.003	0.000	0.000	0.000	0.021	0.000	0.002	0.026	0.002	0.000	0.000	0.000	0.000	0.000	0.000
c.3915G>A (p.=)	33/42	rs11551373	0.022	0.077	0.006	0.001	0.002	0.000	0.066	0.000	0.005	0.068	0.004	0.000	0.000	0.000	0.000	0.002	0.000
c.3986G>A (p.Arg1329His)	33/42	rs45517323,	0.020	0.070	0.007	0.001	0.002	0.000	0.048	0.001	0.005	0.054	0.002	0.001	0.001	0.011	0.001	0.004	0.000
c.4316G>A (p.Gly1439Asp)	34/42	rs150397923	0.005	0.020	0.000	0.000	0.000	0.000	0.013	0.000	0.001	0.017	0.001	0.000	0.000	0.000	0.000	0.000	0.000
c.4536C>T (p.=)	35/42	rs35986575	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.005	0.003	0.000	0.001	0.021	0.000	0.002	0.004	0.005	0.002
c.4716G>T	37/42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

(p.=)																			
c.4908C>T (p.=)	38/42	rs115200071	0.002	0.008	0.000	0.000	0.000	0.000	0.007	0.001	0.001	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.4959C>T (p.=)	38/42	rs45517384	0.003	0.002	0.009	0.000	0.008	0.000	0.003	0.015	0.010	0.003	0.007	0.001	0.000	0.020	0.014	0.010	0.001
c.4982C>T (p.Thr1661Ile)	38/42	rs776541842	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.5010C>T (p.=)	39/42	rs376306544	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.5025G>A (p.=)	39/42	rs35118875	0.004	0.014	0.003	0.000	0.000	0.000	0.016	0.000	0.001	0.014	0.001	0.000	0.000	0.000	0.000	0.001	0.000
c.5051C>T (p.Ser1684Phe)	39/42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.5202T>C (p.=)	41/42	rs1748	0.276	0.748	0.199	0.000	0.184	0.074	0.634	0.178	0.178	0.662	0.112	0.210	0.000	0.207	0.177	0.179	0.081
c.5312C>T (p.Pro1771Leu)	42/42	rs137854214	0.001	0.004	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.5397G>C (p.=)	42/42	rs1051771	0.041	0.002	0.055	0.000	0.101	0.064	0.019	0.099	0.073	0.015	0.033	0.094	0.000	0.144	0.096	0.075	0.058
c.*26G>A	42/42	rs13332015	0.039	0.138	0.016	0.000	0.000	0.000	0.110	0.000	0.009	0.122	0.006	0.000	0.000	0.000	0.000	0.005	0.000
c.*35G>A	42/42	rs200025534	0.002	0.006	0.000	0.000	0.000	0.000	0.008	0.000	0.001	0.008	0.000	0.000	0.000	0.000	0.000	0.001	0.000
c.*61_*62delAA	42/42	rs36032671	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Supplementary Table 12: *TSC1* and *TSC2* intronic single nucleotide variants.

Variant	Intron	Existing variation	1000 Genomes						gnomAD										
			AF	AFR	AMR	EAS	EUR	SAS	AA	EA	AF	AFR	AMR	ASJ	EAS	FIN	NFE	OTH	SAS
TSC1																			
c.2976-43G>C	22/22	rs7853849	0.025	0.084	0.016	0.000	0.003	0.000	0.066	0.001	0.007	0.077	0.009	0.000	0.000	0.000	0.001	0.008	0.000
c.2976-54G>A	22/22	rs118203740	0.007	0.019	0.007	0.000	0.003	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2625+68G>A	20/22	rs1076160	0.412	0.055	0.450	0.559	0.511	0.618	-	-	-	-	-	-	-	-	-	-	-
c.2502+131C>T	19/22	rs7020175	0.082	0.107	0.059	0.072	0.086	0.073	-	-	-	-	-	-	-	-	-	-	-
c.2502+51A>G	19/22	rs75802666	0.080	0.004	0.042	0.198	0.067	0.099	0.009	0.050	0.073	0.009	0.079	0.024	0.197	0.124	0.054	0.067	0.080
c.2392-35T>C	18/22	rs11243931	0.175	0.363	0.108	0.074	0.146	0.101	0.346	0.148	0.142	0.348	0.094	0.150	0.088	0.117	0.149	0.138	0.107
c.2392-222C>T	18/22	rs141011030	0.005	0.000	0.013	0.000	0.014	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2391+59G>C	18/22	rs117301441	0.003	0.000	0.004	0.000	0.012	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2391+34G>A	18/22	rs116518821	0.010	0.027	0.013	0.001	0.004	0.000	0.022	0.003	0.004	0.025	0.008	0.017	0.000	0.000	0.002	0.008	0.000
c.2041+268A>G	16/22	rs12345576	0.175	0.364	0.108	0.074	0.146	0.101	-	-	-	-	-	-	-	-	-	-	-
c.1439-37C>T	14/22	rs10901220	0.139	0.234	0.097	0.074	0.144	0.101	0.204	0.132	0.133	0.224	0.090	0.152	0.090	0.118	0.153	0.135	0.107
c.1334-55C>G	13/22	rs7872606	0.082	0.106	0.059	0.072	0.086	0.073	-	-	-	-	-	-	-	-	-	-	-
c.1334-174T>C	13/22	rs117147322	0.017	0.002	0.020	0.001	0.042	0.024	-	-	-	-	-	-	-	-	-	-	-
c.1333+209C>T	13/22	rs7872860	0.139	0.235	0.097	0.074	0.144	0.101	-	-	-	-	-	-	-	-	-	-	-
c.1333+187A>G	13/22	rs114683952	0.010	0.027	0.013	0.000	0.004	0.000	-	-	-	-	-	-	-	-	-	-	-
c.1142-26_1142-25del	11/22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1142-33A>G	11/22	rs6597586	0.135	0.232	0.094	0.074	0.131	0.098	0.218	0.138	0.124	0.217	0.086	0.142	0.088	0.102	0.138	0.126	0.103
c.1141+63G>A	11/22	rs77891974	0.011	0.038	0.006	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.738-197T>A	08/22	rs1336366611	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.-80-71T>A	02/22	rs45607538	0.001	0.000	0.004	0.000	0.002	0.001	-	-	-	-	-	-	-	-	-	-	-
TSC2																			
c.226-101C>T	03/41	rs144408387	0.008	0.002	0.027	0.000	0.017	0.001	-	-	-	-	-	-	-	-	-	-	-

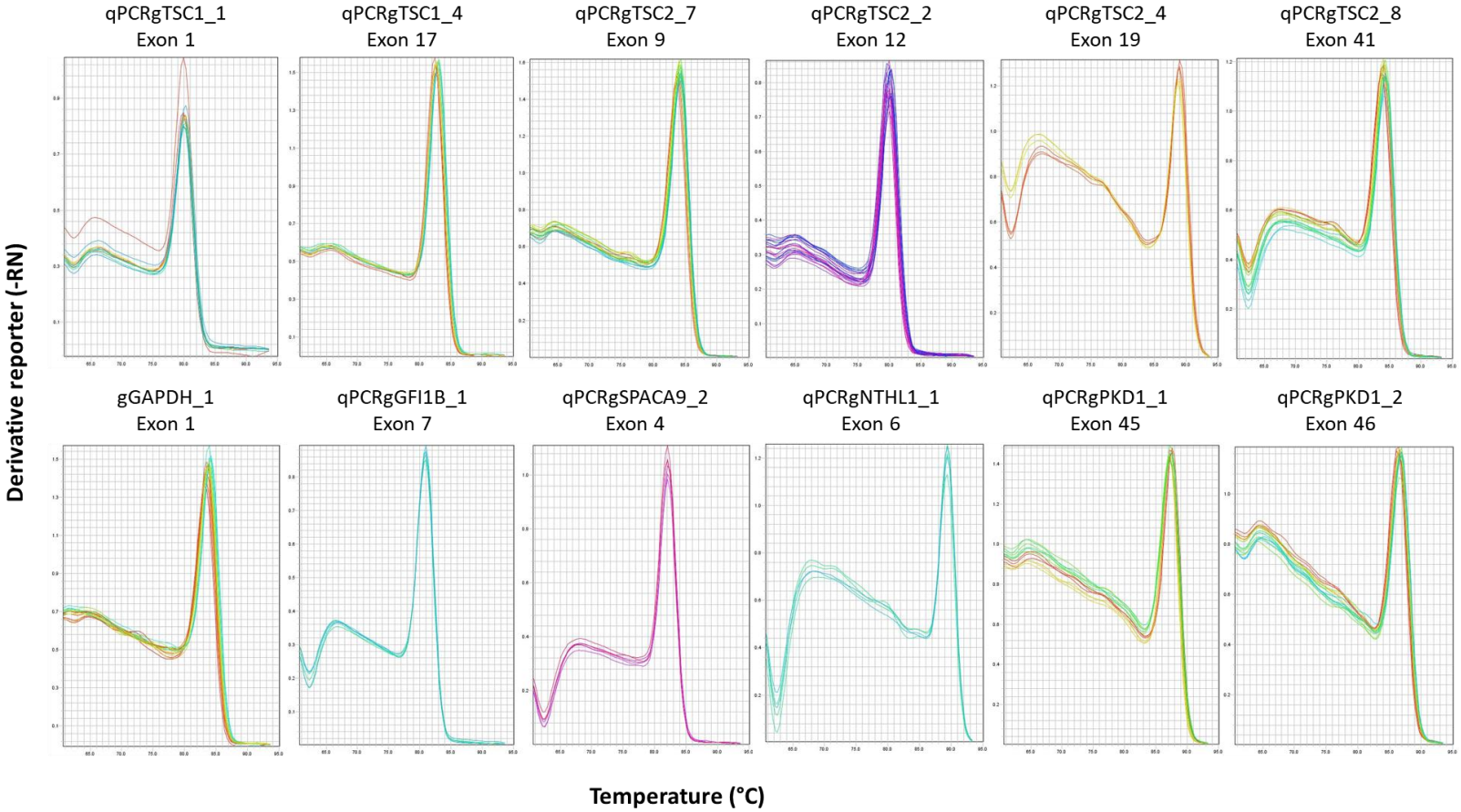
c.226-82C>T	03/41	rs7185742	0.003	0.000	0.003	0.001	0.011	0.003	-	-	-	-	-	-	-	-	-	-	-
c.336+33G>T	04/41	rs45517104	0.009	0.034	0.001	0.000	0.000	0.000	0.033	0.000	0.002	0.040	0.001	0.000	0.000	0.000	0.000	0.002	0.000
c.336+66G>C	04/41	rs920782009	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.482-114T>C	05/41	rs77037371	0.012	0.041	0.009	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.482-68C>G	05/41	rs2516734	0.153	0.374	0.151	0.000	0.136	0.032	-	-	-	-	-	-	-	-	-	-	-
c.482-59G>A	05/41	rs200833855	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.600-84C>T	06/41	rs35364892	0.007	0.008	0.001	0.000	0.003	0.020	-	-	0.009	0.007	0.003	0.036	0.000	0.006	0.008	0.010	0.019
c.648+12C>A	07/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.848+113A>G	09/41	rs56077830	0.011	0.038	0.004	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.848+143G>T	09/41	rs78160478	0.046	0.042	0.076	0.000	0.099	0.022	-	-	-	-	-	-	-	-	-	-	-
c.976-100C>G	10/41	rs2074968	-	0.398	0.499	0.714	0.547	0.665	-	-	-	-	-	-	-	-	-	-	-
c.976-63G>A	10/41	rs12927333	0.035	0.002	0.046	0.000	0.075	0.064	-	-	-	-	-	-	-	-	-	-	-
c.1257+109dupG	12/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1258-223G>A	12/41	rs924724434	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1361+54_1361+57del	13/41	rs137854304	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1362-151G>T	13/41	rs545796760	0.000	0.000	0.000	0.000	0.002	0.000	-	-	-	-	-	-	-	-	-	-	-
c.1443+65T>G	14/41	rs115599333	0.007	0.027	0.000	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.1444-201G>A	14/41	rs60763712	0.051	0.181	0.017	0.000	0.002	0.000	-	-	-	-	-	-	-	-	-	-	-
c.1599+85_1599+87del	15/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1599+162C>T	15/41	rs74819416	0.009	0.030	0.006	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.1599+216T>G	15/41	rs17654678	0.038	0.003	0.042	0.001	0.112	0.044	-	-	-	-	-	-	-	-	-	-	-
c.1599+282_1599+283insCTGGGG	15/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1716+183A>G	16/41	rs8063461	0.615	0.350	0.716	0.839	0.590	0.698	-	-	-	-	-	-	-	-	-	-	-
c.1717-121G>A	16/41	rs1275330311	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.1717-55T>C	16/41	rs7187438	0.393	0.714	0.297	0.221	0.371	0.229	-	-	-	-	-	-	-	-	-	-	-
c.1840-172G>A	17/41	rs532002248	0.000	0.000	0.001	0.001	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.1840-136T>C	17/41	rs113749674	0.013	0.046	0.006	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2097+43A>G	19/41	rs186681035	0.003	0.001	0.001	0.000	0.010	0.001	0.001	0.010	0.009	0.003	0.003	0.012	0.000	0.026	0.013	0.009	0.005
c.2098-114A>G	19/41	rs6600185	0.110	0.393	0.040	0.000	0.002	0.000	-	-	-	-	-	-	-	-	-	-	-

c.2221-28A>G	20/41	rs7196184	0.109	0.392	0.040	0.000	0.002	0.000	0.330	0.001	0.026	0.354	0.019	0.005	0.000	0.000	0.001	0.017	0.001
c.2356-63A>G	21/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.2545+26G>A	22/41	rs45517242	0.000	0.000	0.000	0.000	0.002	0.000	0.002	0.005	0.003	0.001	0.002	0.001	0.000	0.004	0.004	0.005	0.000
c.2545+31C>T	22/41	rs45517243	0.005	0.014	0.007	0.000	0.000	0.000	0.015	0.000	0.001	0.017	0.001	0.000	0.001	0.000	0.000	0.000	0.000
c.2545+95dupG	22/41	rs397515216	0.063	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.2545+135C>T	22/41	rs139579913	0.004	0.012	0.003	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2545+157G>C	22/41	rs9941195	0.056	0.201	0.020	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2545+194A>G	22/41	rs7185909	0.035	0.123	0.013	0.000	0.002	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2545+240C>T	22/41	rs9921591	0.056	0.201	0.020	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2545+244C>G	22/41	rs143553665	-	0.012	0.001	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2546-12C>T	22/41	rs13331451	0.206	0.620	0.107	0.000	0.085	0.052	0.521	0.074	0.095	0.552	0.063	0.074	0.000	0.098	0.072	0.083	0.048
c.2639+44C>G	23/41	rs1800715	0.215	0.624	0.115	0.000	0.097	0.074	0.526	0.087	0.107	0.557	0.069	0.102	0.000	0.111	0.085	0.099	0.063
c.2639+44C>T	23/41	rs1800715	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.2640-74G>A	23/41	rs143128054	0.004	0.014	0.003	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2640-65C>T	23/41	rs138967036	0.004	0.014	0.003	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.2640-26G>T	23/41	rs45483797	0.003	0.012	0.001	0.000	0.000	0.000	0.013	0.000	0.001	0.011	0.001	0.000	0.000	0.000	0.000	0.000	0.000
c.2743-40A>G	24/41	rs45517265	0.004	0.000	0.006	0.000	0.015	0.002	0.002	0.010	0.007	0.002	0.007	0.004	0.000	0.001	0.010	0.010	0.007
c.2837+93G>A	25/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.2838-4A>G	25/41	rs45517272	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.001	0.000	0.000	0.003	0.000	0.000	0.001	0.001	0.001
c.2966+92_2966+94dup	26/41	rs397515128	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.2966+217G>T	26/41	rs118124428	0.008	0.002	0.009	0.000	0.030	0.001	-	-	-	-	-	-	-	-	-	-	-
c.2967-41C>T	26/41	rs972958530	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.3131+34C>T	27/41	rs45487103	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.001	0.000	0.001	0.002	0.000	0.000	0.001	0.001	0.000
c.3132-30A>G	27/41	rs45517283	0.003	0.012	0.001	0.000	0.000	0.000	0.013	0.000	0.001	0.012	0.001	0.000	0.000	0.000	0.000	0.000	0.000
c.3397+126C>T	29/41	rs115243225	0.051	0.184	0.019	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.3610+46dupC	30/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.3611-136C>T	30/41	rs376053649	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.3611-100T>C	30/41	rs370512781	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.3611-28C>T	30/41	rs45517303	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

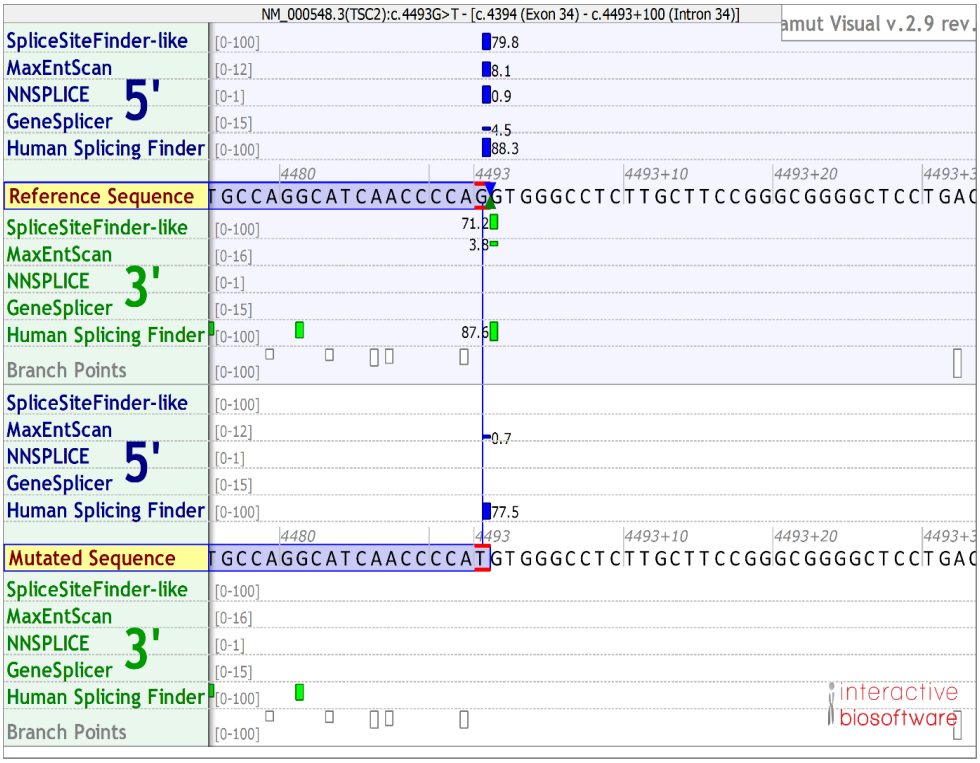
c.3815-124C>T	31/41	rs143001593	0.003	0.009	0.001	0.000	0.001	0.000	-	-	-	-	-	-	-	-	-	-	-
c.3815-38C>T	31/41	rs371592734	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000
c.3883+8C>G	32/41	rs45517316	0.012	0.043	0.004	0.000	0.000	0.000	0.033	0.000	0.003	0.039	0.003	0.000	0.000	0.000	0.000	0.002	0.000
c.3883+66G>A	32/41	rs74002774	0.011	0.042	0.003	0.000	0.000	0.000	0.035	0.000	-	-	-	-	-	-	-	-	-
c.3883+78G>A	32/41	rs1800705	0.011	0.003	0.017	0.000	0.027	0.011	-	-	-	-	-	-	-	-	-	-	-
c.3883+127G>A	32/41	rs562317274	0.001	0.004	0.000	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.3883+226G>A	32/41	rs9928737	0.188	0.409	0.180	0.000	0.188	0.088	-	-	-	-	-	-	-	-	-	-	-
c.3884-68C>T	32/41	rs114209035	0.004	0.014	0.003	0.000	0.000	0.000	-	-	0.001	0.012	0.001	0.000	0.000	0.000	0.000	0.001	0.000
c.3884-56C>G	32/41	rs1800724	0.036	0.001	0.038	0.000	0.078	0.079	-	-	0.060	0.009	0.029	0.092	0.000	0.110	0.071	0.059	0.066
c.4005+84C>T	33/41	rs30259	0.040	0.002	0.053	0.046	0.074	0.041	-	-	-	-	-	-	-	-	-	-	-
c.4005+153T>A	33/41	rs28535326	0.019	0.070	0.006	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.4006-60C>A	33/41	rs376519405	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.4006-8C>T	33/41	rs45517325	0.002	0.000	0.004	0.000	0.003	0.002	0.000	0.004	0.002	0.001	0.004	0.000	0.000	0.000	0.004	0.004	0.001
c.4493+110G>A	34/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.4569+44C>T	35/41	rs142227748	0.010	0.039	0.000	0.000	0.000	0.000	0.025	0.000	0.002	0.032	0.001	0.000	0.000	0.000	0.000	0.001	0.000
c.4569+46C>T	35/41	rs45482793	0.004	0.001	0.003	0.000	0.018	0.001	0.003	0.018	0.010	0.002	0.005	0.010	0.000	0.008	0.017	0.011	0.003
c.4569+47G>A	35/41	rs45517351	0.019	0.065	0.013	0.000	0.000	0.001	0.059	0.001	0.005	0.059	0.003	0.006	0.000	0.000	0.000	0.004	0.000
c.4570-38G>A	35/41	rs770778907	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
c.4662+199G>A	36/41	rs1379483764	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.4663-88C>T	36/41	rs146465188	0.006	0.017	0.009	0.000	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.4663-56C>G	36/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.4849+75C>T	37/41	rs76029733	0.028	0.011	0.026	0.043	0.031	0.035	-	-	-	-	-	-	-	-	-	-	-
c.4849+76G>A	37/41	rs144647393	0.015	0.057	0.001	0.000	0.001	0.000	-	-	-	-	-	-	-	-	-	-	-
c.4849+97delA	37/41	rs142421783	0.056	0.201	0.020	0.001	0.000	0.000	-	-	-	-	-	-	-	-	-	-	-
c.4850-133C>T	37/41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.4850-109T>C	37/41	rs13335638	0.241	0.595	0.195	0.000	0.198	0.090	-	-	-	-	-	-	-	-	-	-	-
c.4990-7C>T	38/41	rs45457095	0.009	0.027	0.012	0.000	0.000	0.000	0.028	0.001	0.002	0.025	0.002	0.006	0.000	0.000	0.000	0.002	0.001
c.5160+37C>T	40/41	rs45517400	0.001	0.000	0.001	0.000	0.003	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.000
c.5160+43C>T	40/41	rs45517402	-	-	-	-	-	-	-	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

c.5161-28_5161-25delTGAG	40/41	rs1799758	0.005	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
c.5161-10A>C	40/41	rs1800718	0.257	0.616	0.213	0.000	0.236	0.090	0.543	0.219	0.200	0.563	0.128	0.250	0.000	0.258	0.214	0.217	0.097
c.5260-49C>T	41/41	rs13332221	0.168	0.458	0.140	0.000	0.116	0.020	0.397	0.102	0.106	0.410	0.083	0.142	0.000	0.114	0.103	0.125	0.028
c.5260-25C>G	41/41	rs13332222	0.201	0.576	0.153	0.000	0.118	0.020	0.478	0.105	0.114	0.507	0.089	0.142	0.000	0.115	0.106	0.130	0.029
c.5260-15C>T	41/41	rs45517416	0.002	0.004	0.000	0.000	0.001	0.002	0.002	0.001	0.002	0.002	0.001	0.000	0.000	0.001	0.002	0.002	0.002

VIII. Supplementary Figures



Supplementary Figure 1: Melt curve of pair of primers used to in qPCR. The first row there are primers for *TSC1* and *TSC2* sorted according to exon position on gene. On second row, *GAPDH* was the reference for qPCR, *GFI1B* and *SPACA9* respectively upstream and downstream genes for *TSC1,* and *NTHL1* and *PKD1* respectively upstream and downstream genes for *TSC2*. They are sorted according to exon position on gene



Supplementary Figure 2: ALAMUT splicing site prediction analysis for TSC2 (NM_000548.3) c.4493G>T substitution.

CHAPTER 3

List of Acronyms

3´-UTR: 3’- untranslated region

5’-UTR: 5’-untranslated regions

DMEMplus: DMEM supplemented with 5% calf serum and penicillin-streptomycin

DMEMminus: DMEM with no supplementation

DPBS: Dulbecco’s Phosphate Buffered Saline

DMEM: Dulbecco's Modified Eagle’s Medium

EDTA: Ethylenediamine tetra-acetic acid

FDR: False discovery rate

GO: Gene ontology

HSP90: Heat shock protein 90-alpha

O/N: Overnight

TPM: Transcript per million

UTR: Untranslated region

List of Tables

List of Figures

Summary

I. Introduction

Human *TSC1* (NG_012386.1) exons 1, 2 and the first 80 bp of exon 3 comprise the 5´ untranslated region (UTR) of its mRNA. *TSC1* 3'-UTR corresponds to the last 4,887 bp of exon 23 (Figure 1A). There are 68% of identity between human and mouse *TSC1* 3'-UTR, suggesting it lays as a platform for potential binding sites for translation and stability regulatory factors. Full-length *TSC1* mature mRNA (NM_000368.4) consists of 8,626 nucleotides, and is expressed in nearly all tissues. RNA-Seq expression data from the Genotype-Tissue Expression (GTEx) project reveals the highest median expression of transcript per million (TPM) in the cerebellum (Figure 1B).

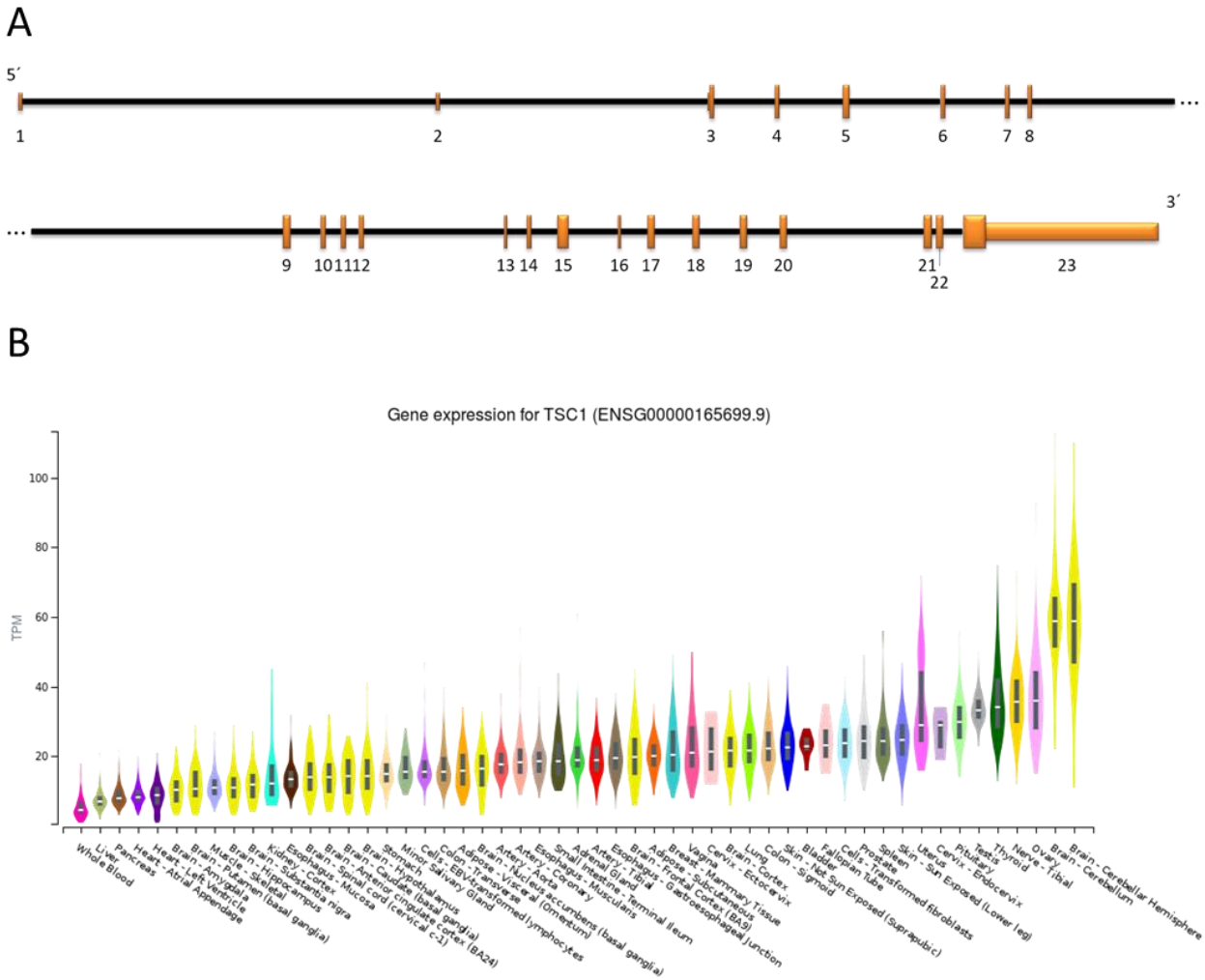


Figure 1: *TSC1* locus and mRNA expression. (A) Schematic diagram of the *TSC1* gene. Black lines represent introns, broad and narrow orange bars highlight respectively coding exons and the UTR regions of the gene. (B) Plot of *TSC1* RNA-Seq expression data from the GTEx project (https://www.gtexportal.org/home/gene/TSC1). TPM: transcripts per million.

TSC2 (NG_005895.1) exon 1 and the first 29 bp of exon 2 encompass the 5`-UTR of its mRNA. The 3'-UTR comprises the last 102 bp of exon 42. The full-length *TSC2* transcript (NM_000548.3) consists of 5,675

nucleotides and, similar to *TSC1*, is expressed in most tissues, with the highest mRNA amount in the cerebellum (Figure 2B). Overall *TSC2* mRNA levels are nearly twice that observed for *TSC1* (Figure 1B and Figure 2B).

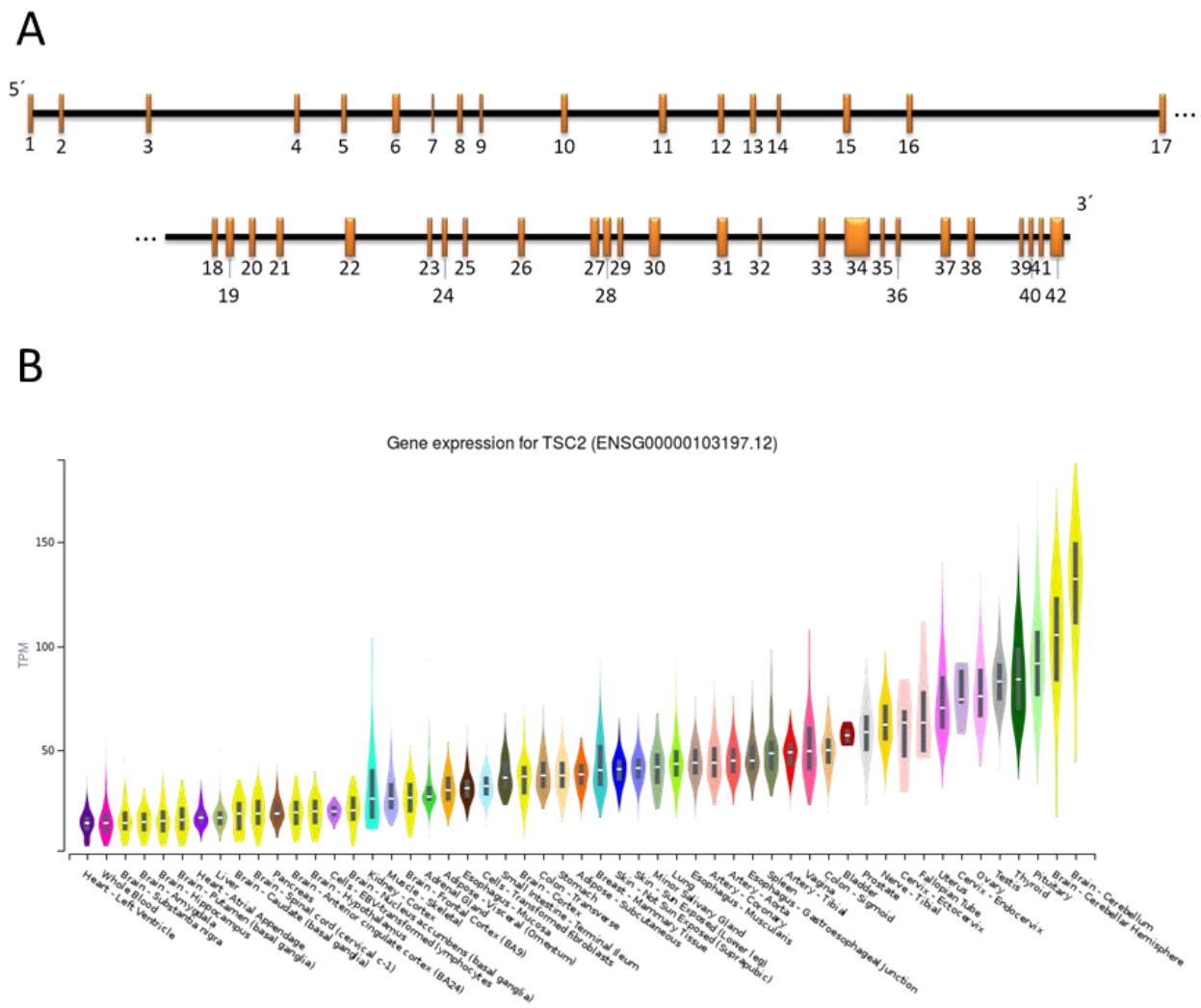


Figure 2: *TSC2* locus and mRNA expression. (A) Schematic diagram of the *TSC2* gene exons and introns. Black lines represent introns, broad and narrow orange bars highlight coding exons and the UTR regions, respectively. (B) *TSC2* RNA-Seq expression data from the GTEx project (https://www.gtexportal.org/home/gene/TSC2). TPM: transcripts per million.

TSC1 protein (NP_000359.1) is a 130-kDa (1,164 amino acids) hydrophilic protein displaying a C-terminal domain required for interaction with TSC2 (residues 762-848) (van Slegtenhorst, Nellist *et al.* 1998), predicted as a coiled-coil region. The 200-kDa TSC2 protein (NP_000539.2) has 1,807 amino acids, an N-terminal domain (residues 555-903) required for interaction with TSC1, and a C-terminal 163-residue GTPase activating protein GAP domain (GAP-Rheb residues 1562 to 1764) (Plank, Yeung *et al.* 1998, van Slegtenhorst, Nellist *et al.* 1998).The 200 kDa TSC2 protein (NP_000539.2) has 1,807 amino acids, an N-terminal domain (residues 555-903) required for interaction with TSC1, and a C-terminal 186-residue GTPase activating protein GAP domain (GAP-Rheb residues 1562-1748) (Plank, Yeung *et al.* 1998, van Slegtenhorst, Nellist *et al.* 1998).

Under conditions of amino acid depletion, Rag GTPase recruits TSC2 to the lysosome, interacting with Rheb. mTORC1 is thus released from the lysosome and becomes inactivated. However, upon amino acid withdrawal, cells with pathogenic variants on TSC2 cannot completely release TORC1 from the lysosome, hence failing to adjust to physiologically conditions (Demetriades, Doumpas *et al.* 2014).

In cells lacking the TSC2 protein, TSC1 homodimers are prone to self-aggregation in detergent-resistant complexes, presenting a coarsely dotted cytoplasmic distribution by immunofluorescence analysis (Nellist, Sancak *et al.* 2005). If both genes are co-overexpressed, TSC1 protein becomes solubilized with a more homogeneous cytoplasmic distribution. Thus, it has been proposed that TSC2 may act as TSC1 chaperone (Nellist, van Slegtenhorst *et al.* 1999). On the other hand, Benvenuto, Li *et al.* (2000) suggested that TSC1 stabilizes TSC2.

TBC1D7, which has been reported as a protein from the TSC1/2 complex, stabilizes TSC1 dimers (Gai, Chu *et al.* 2016). The crystal structure of TBC1D7 in complex with the C-terminal part of TSC1 (TSC1-CC, residues 939–992, encoded by exon 22) revealed that two units of TSC1-CC form a parallel homodimer - two symmetric surfaces for interaction with TBC1D7 (Santiago Lima, Hoogeveen-Westerveld *et al.* 2014, Gai, Chu *et al.* 2016, Qin, Wang *et al.* 2016), stabilizing the TSC complex.

The molecular chaperone heat-shock protein 90-alpha (HSP90, NP_001017963.2), an essential component of the cellular homeostatic machinery in eukaryotes, may interact with TSC1 in the absence or presence of TSC2. The association between TSC1 and TSC2 decreases under heat shock, which does not affect the association between TSC1 and HSP90 (Inoue, Uyama *et al.* 2010). TSC1 has been proposed as HSP90 co-chaperone (Woodford, 2017; Sager 2018a, 2018b), as HSP90 binds *in vitro* and *in vivo* to denatured proteins and displays ATP-dependent anti-aggregation properties (Panaretou, Prodromou *et al.* 1998).

TSC2 had been identified as a target of AKT kinase regulation (Potter, Huang *et al.* 2001, Dan, Sun *et al.* 2002, Manning, Tee *et al.* 2002, Inoki, Li *et al.* 2003). The pathway of phosphatidylinositol-3 and AKT kinases was characterized as the main physiological regulation of TSC complex turnover, leading to a decrease in the affinity between TSC1 and TSC2, and degradation of both proteins mediated by their phosphorylation (Potter, Huang *et al.* 2001, Dan, Sun *et al.* 2002, Inoki, Li *et al.* 2003). The observed response to heat shock appears to be related to the phosphorylation of TSC2 by AKT, stimulating its degradation by the proteasome (Inoue, Uyama *et al.* 2010). The interaction of TSC1 with HSP90 is required for its localization in the outer membrane of the mitochondria, with or without TSC2, and in a manner dependent on the phosphorylation of TSC1 Thr^{417} (Inoue, Uyama *et al.* 2010). These data are corroborated by other observations that the knockout of *Tsc2* maintains Tsc1 at reasonable levels (Zhang, Gao *et al.* 2003, Pollizzi, Malinowska-Kolodziej *et al.* 2009). On the other hand, Western blotting of mouse $Tsc1^{-/-}$ embryonic fibroblast protein has shown reduced levels of the Tsc2 protein (Astrinidis, Senapedis *et al.* 2006), suggesting that lack of expression of both *Tsc1* alleles may

subject Tsc2 to poly-ubiquitination and degradation by the proteasome. Overexpression of *TSC1* in Rat-2 cells or rat embryonic fibroblasts elevates the level of TSC2, and binding of TSC1 to TSC2 inhibits TSC2 ubiquitination. Finally, inhibition of the proteasome represses TSC2 degradation (Benvenuto, Li *et al.* 2000).

The stability of TSC1-TSC2 protein complex is central for the control of mTOR signaling. Further understanding of the complexity of *TSC1* and *TSC2* gene regulation is important as heterologous gene replacement can be a potential future therapy alternative for specific TSC lesions (Prabhakar, Zhang *et al.* 2015). However, it is as yet unknown whether in the absence of the TSC1 or TSC2 protein there may be transcriptional or post-transcriptional feedback control of *TSC1* and *TSC2* gene expression. Although it is clear that there is inter-regulation of TSC1 and TSC2 protein stability, it is necessary to understand the levels and mechanisms of regulation of their specific endogenous gene expression.

II. Aim and objectives

A. Aim

To assess *TSC1* and *TSC2* gene product levels in search for possible transcriptional or translational feedback regulatory loops in cells lacking TSC1 or TSC2 proteins.

B. Objectives

1. To compare TSC1 and TSC2 protein levels between the reference human cell line and human cell lines with bi-allelic inactivation of *TSC1* and or *TSC2* genes;
2. To assess the activity of the mTORC1 pathway reported by the phosphorylation of S6 protein Ser^{235} and Ser^{236} residues; having nutrient starvation, insulin treatment and rapamycin supplementation as proof-of-concept controls;
3. To compare *TSC1* and *TSC2* mRNA levels between the reference human cell line and each derived cell line with bi-allelic inactivation of *TSC1* and or *TSC2* genes.

III. Material and Methods

Experiments were conducted in the Netherlands, mostly at Erasmus Medical Centre, Rotterdam, Amsterdam Univesrity or at the GenomeScan (Leiden, the Netherland). Data analises was performed at Instituto de Biociências, Universidade de São Paulo (São Paulo, Brazil) and Hospital Sírio-Libanês, São Paulo, Brazil.

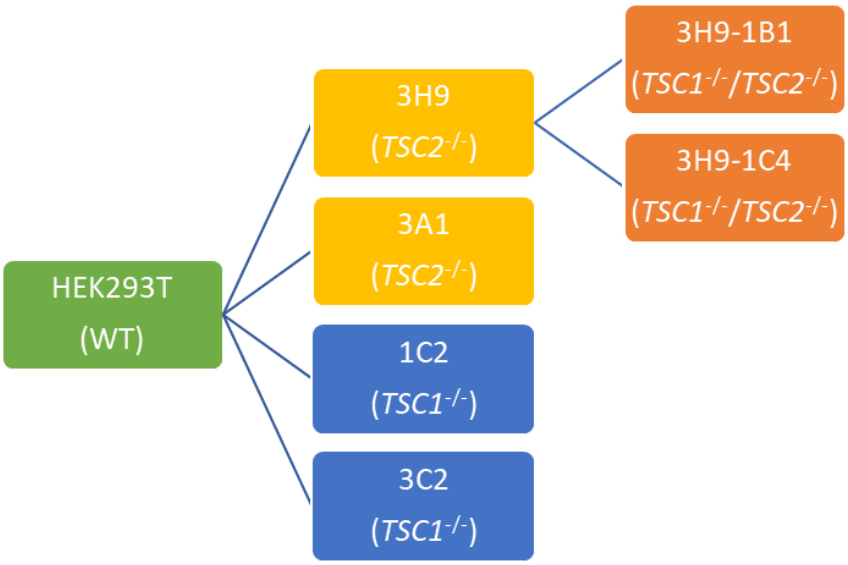
A. Cell lines

Seven cell lines employed had been generated before: HEK293T cell line, and six cell lines derived from HEK293T in which *TSC1* or *TSC2* gene genomic segments have been genetically edited by the CRISPR-Cas9 strategy (Dufner-Almeida et al., submitted). Figure 3 depicts the outflow of genetic modifications, and Table 1 presents cell line names and *TSC1* and *TSC2* genotypes.

B. Cell culture

Cell lines were seeded in round tissue culture dish (6 cm; Sigma-Aldrich, St. Louis, Missouri, USA) in Dulbecco's Modified Eagle's Medium (DMEM, BioWhittaker, Radnor, Pennsylvania, USA), supplemented with 5% fetal calf serum and penicillin-streptomycin, constituting the complete medium or DMEMplus, and maintained at 37C and 10% CO_2. After splitting to ten round dishes, cells were kept under the same culturing conditions until 80% confluence. Cells from three plates (group DMEMplus) were harvested and submitted to DNA, RNA and protein isolation. After washing the cells three times in PBS (Dulbecco's Phosphate Buffer Saline (0.0095M) – BioWhittaker - Radnor, Pennsylvania, USA), medium from the seven remaining plates was switched to DMEM minimum medium (without supplementation), considered as DMEMminus. The starvation period lasted for 16 hours (overnight), after which RNA and protein were isolated from two plates (group DMEMminus). Rapamycin (Cell Signaling Technology, Danvers, Massachusetts, USA) suspended in DMSO was added to cells from the five remainder dishes to final concentration of 10 nM. After one hour, RNA and protein were isolated from two plates (group DMEMminus - RAPA-1h) while culture continued for further 15 hours in three plates. On the fourth day, RNA and protein were isolated from two plates (DMEMminus - RAPA O/N) and cells in the final plate were treated with insulin at 200 nM suspended in HEPES buffer (Sigma-Aldrich, St. Louis, Missouri, USA) for 10 minutes (group DMEMminus – RAPA O/N – INS), washed and harvested for protein isolation (Figure 4). All seven cell lines were simultaneously submitted to the same protocol, and the experiment was repeated at least three times.

A



B

Cell Line	TSC1 genotype	TSC2 genotype
HEK293T (WT)	c.965T>C (p.M322T) heterozygote	wild-type
1C2 ($TSC1^{-/-}$)	c.-13_5del homozygote c.965T>C (p.M322T) heterozygote c.3194_3206del (p.T1065Kfs*22) homozygote	wild-type
3C2 ($TSC1^{-/-}$)	c.989_3199del (p.S331_M1067del) heterozygote c.1002_3204del (p.T335Kfs*22) heterozygote c.1002dup (p.T335Dfs*6) homozygote c.3196insT (p.T1066Ifs*33) homozygote	wild-type
3A1 ($TSC2^{-/-}$)	c.965T>C (p.M322T) heterozygote	c.121dup (p.T41Nfs*26) homozygote c.116_4989+358del (p.I39Rfs*42) heterozygote
3H9 ($TSC2^{-/-}$)	c.965T>C (p.M322T) heterozygote	c.122_4990-284del38842 (p.T41Rfs*42) homozygote
3H9-1B1 ($TSC1^{-/-}$/$TSC2^{-/-}$)	c.-8_3196del (p.?) homozygote	c.122_4990-284del38842 (p.T41Rfs*42) homozygote
3H9-1C4 ($TSC1^{-/-}$/$TSC2^{-/-}$)	c.-8_3196del (p.?) homozygote	c.122_4990-284del38842 (p.T41Rfs*42) homozygote

Figure 3: Cladogram and genotype of wildtype cell line and six edited cell lines by the CRISPR-Cas9. (A) Schematic representation of parental wild type HEK293T cell line and derived cell line. (B) Genotype of each seven cell lines.

C. Protein isolation and Western Blotting

Cells were lysed in 50 mM Tris-HCl (pH 7.6), 100 mM NaCl, 50 mM NaF, 1% Triton-X-100, 1 mM ethylenediamine tetra-acetic acid (EDTA) and complete protease inhibitor cocktail (Roche, Basileia,

Switzerland). Proteins lysate were submitted to Criterion™ TGX™ Precast Gel (4 -15%; Bio-Rad, Hercules, California, USA) electrophoresis at 100V for 10-15 minutes followed by one hour at 120V; and transferred to nitrocellulose filters in the Trans-blot® Turbo™ (Bio-Rad, Hercules, California, USA) using the pre-set program for high molecular weight (2.5 Amp, 25 V, 10min.). Finally, the membrane proteins were blocked in 5% powder milk in DPBS (Dulbecco's Phosphate Buffer Saline (0.0095M) – BioWhittaker - Radnor, Pennsylvania, USA) for 30 minutes, followed by specific primary and secondary antibody incubations, and fluorescence captured under the conditions described in chapter 2. The experiment was repeated at least three times and Student's *t* test was used for pairwise comparison.

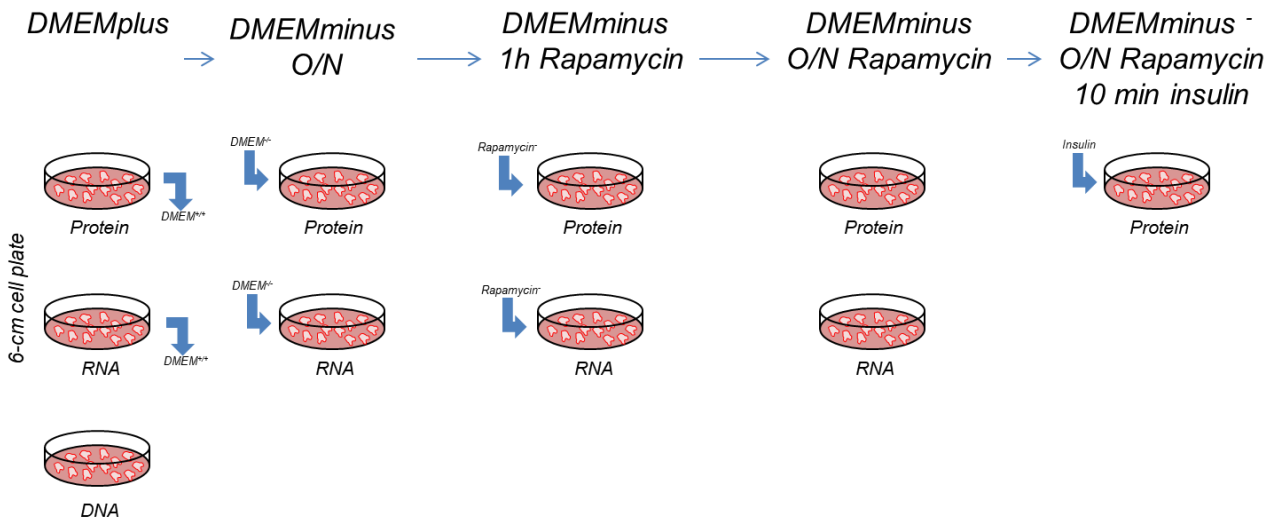


Figure 4: Workflow of cell culture, DNA isolation, RNA Isolation and protein isolation from all seven cell lines. O/N: overnight.

D. Total RNA isolation and RNA sequencing

Total RNA was isolated with RNeasy® Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer conditions. Briefly, cells were washed in DPBS, and then lysis buffer - 1% β-mercaptoethanol in RTL buffer (QIAGEN, Hilden, Germany) - was added. Cells were detached and transferred to a 1.5-mL tube. One volume of 70% ethanol was added, homogenized, and the solution was transferred to an RNeasy® mini-spin column, and centrifuged for 15 seconds at 8,000 X *g*. Column was washed with buffer RW1 and centrifuge for 15 sec at 12000g., and contents incubated with DNase I for 15 minutes at room temperature. Finally, the column was washed three times and RNA eluted in 50 µL of RNase-free water. RNA concentration and quality were measured using the NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) and BioAnalyzer (Thermo Fisher Scientific, Waltham, Massachusetts, USA) (Supplementary Table 1).

A total of 30 RNA samples (Table 1) were used for library construction and Illumina HiSeq 4000 sequencing at GenomeScan (Leiden, The Netherlands). To assess the quality of the samples, concentration was determined using the Fragment Analyzer (Supplementary Figure 1). Sample preparation was performed

according to NEBNext Ultra Directional RNA Library Prep Kit for Illumina protocol (NEB #E7420S/L; New England Biolabs, Ipswich, Massachusetts, USA). Briefly, mRNA was isolated from total RNA using oligo-dT magnetic beads. After mRNA fragmentation on broad peak between 300-500 bp, cDNA was synthesized following kit specifications. Ligation with sequencing adapters followed, and the resulting product was submitted to PCR bridge amplification. Product quality and yield after sample preparation was measured with the Fragment Analyzer (Supplementary Figure 2). The size of the resulting products was consistent with the expected size distribution (a broad peak between 300-500 nucleotides). A cDNA concentration of 3.0 nM was used (Supplementary Table 2). Clustering and cDNA sequencing using the Illumina cBot and HiSeq 4000 were performed according to the manufacturer's protocols (Supplementary Table 3).

Table 1: Total of samples per cell line for each treatment - DMEMplus or DMEMminus.

Cell Line	Treatment	
	DMEMplus	DMEMminus
HEK293T (WT)	3	3
1C2 ($TSC1^{-/-}$)	2	2
3C2 ($TSC1^{-/-}$)	2	2
3A1 ($TSC2^{-/-}$)	2	2
3H9 ($TSC2^{-/-}$)	2	2
3H9-1B1 ($TSC1^{-/-}$/$TSC2^{-/-}$)	2	2
3H9-1C4 ($TSC1^{-/-}$/$TSC2^{-/-}$)	2	2

The mRNA-Seq pipeline starts with quality filtering and trimming of the sequence of optimized standard thresholds: Illumina sequencing adapters were removed, bases with read Q-score below 22 trimmed off, base reads with at least 36 nucleotide-long kept in the data set. Alignment was based on Burrows–Wheeler Transform to ensure maximum sensitivity for mapping reads to exons of genes. All the genes found in the reference annotation were included, even if they do not have any associated expression. Each expressed gene is listed separately per Ensembl Gene ID.

Statistical analysis and comparison between treatment and between samples were performed using DEseq2 script (Love, Huber *et al.* 2014) on R program (Version 3.5.1,). Briefly, average of the normalized count values divided by size factors, taken over all samples (baseMean); expression changed due to treatment or cell lines (log2FoldChange), standard error estimate for the log2 fold change (lfcSE), *p*value and *p*adj (Benjamini-Hochberg BH-adjusted *p*values for false positive results) were obtained. All *p*values in results' table assigned as NA were excluded from analysis because it contained an extreme count outlier, or the expression was not detected. Only genes with *p*adj<0.1 were considered differentially expressed.

Comparisons between gene sets identified from RNA-Seq analysis between 1C2 and HEK293T cell lines from DMEMplus or DMEMminus groups or from 1C2 (DMEMplus) and 1C2 (DMEMminus) groups yielded differentially expressed genes under significance of adjusted p value <0.1. Differentially up- or down-regulated

genes were submitted to gene ontology (GO) grouping for non-redundant biological processes, employing the over-representation enrichment analysis at WEB-based GEne SeT AnaLysis Toolkit (2019 Version; http://webgestalt.org), a set of computational tools developed and maintained at Dr. Bing Zhang laboratory at the Baylor College of Medicine (Houston, TX, USA). Gene ID was gene symbol, and the reference set for enrichment analysis was the human genome. Additionally, the following cut-off values were considered: at least five genes for a GO category, a maximum of 2,000 genes for a GO category, ten GO categories expected from set cover, and significance level for false discovery rate (FDR) < 0.05 among the top 10 GO categories (Benjamini statistical test corrected by Hochberg).

IV. Results

A. TSC1 and TSC2 protein analyses

In the present study, we screened six HEK293T-derived cell lines with bi-allelic inactivation of the *TSC1* and or *TSC2* genes (Figure 3B) for the expression of these genes at both mRNA and protein levels. As presented in Figure 5 (panels A-C), under DMEMplus conditions, the protein encoded by each CRISPR-Cas9-mediated knocked-out gene is not detected, as expected. Knocking-out *TSC1* or *TSC2* genes in homozygosity does not appear to affect the expression of AKT or S6 proteins (Figure 5, D and E), but to relate to at least 50% reduction of the partnering protein, e.g., TSC2 when *TSC1* is inactivated, or TSC1 when *TSC2* is knocked out (Figure 5, B and C).

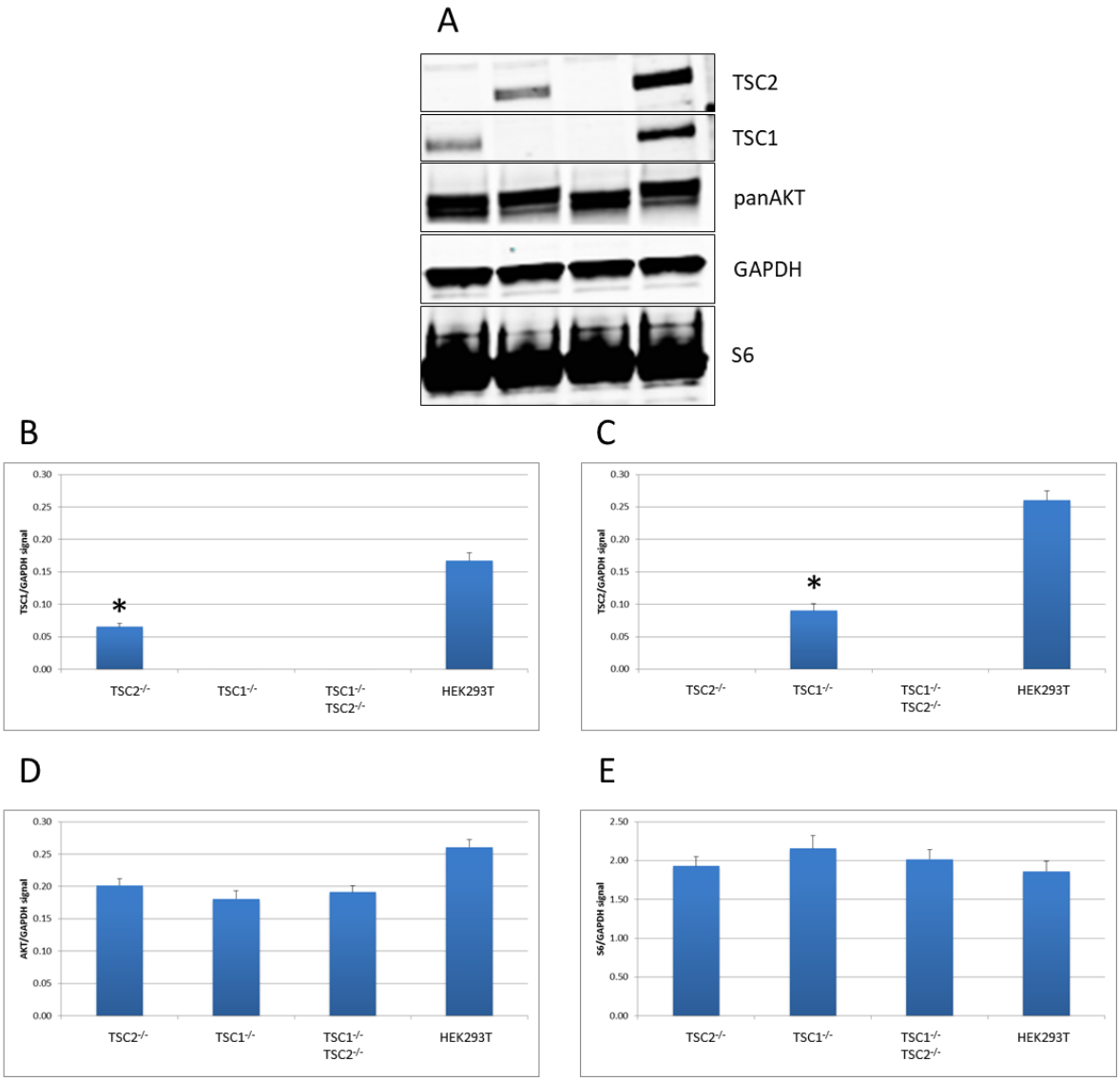


Figure 5: TSC1, TSC2, AKT and S6 protein levels assessed by Western blotting. (A) Western blots showing protein bands for TSC2, TSC1, panAKT, GAPDH and S6 protein, as shown on the right side, respectively from HEK293T-derived cell lines *TSC2*$^{-/-}$, *TSC1*$^{-/-}$, *TSC2*$^{-/-}$/*TSC1*$^{-/-}$ and HEK293T reference. (B-E) Signal relative to GAPDH signal respectively for TSC1 (B), TSC2 (C), AKT (D) and S6 (E) for HEK293T-derived cell lines *TSC2*$^{-/-}$, *TSC1*$^{-/-}$, *TSC2*$^{-/-}$/*TSC1*$^{-/-}$ and HEK293T reference. All values were normalized to GAPDH. Asterisk represents a statistically significant difference between *TSC1*$^{-/-}$ or *TSC2*$^{-/-}$ under the same culturing condition to HEK293T (p-value < 0.05).

TSC1 and TSC2 proteins were reassessed on blots in all seven cell lines in a single electrophoresis for each culturing condition. Lack of TSC1 and TSC2 proteins was reconfirmed for the respective knock-out genotypes (Figure 6). As expected, lack of TSC1 or TSC2 protein directly related to the levels of S6/Ser^{235}/Ser^{236} phosphorylation, reporting mTORC1 hyperactivity (Figure 6). For instance, all cells from DMEMplus group and the knocked-out cells from DMEMminus group had S6/Ser^{235}/Ser^{236} hyperphosphorylated (Figure 6, first and second panels from second bottom row). On the contrary, DMEMminus reference HEK293T cells did not (Figure 6, second panel from second bottom row). Addition of rapamycin, for one or 16 hours, down-regulated mTORC1 activity as shown by hypophosphorylation of the reporter S6/Ser^{235}/Ser^{236} (Figure 6, third to fifth panels from second bottom row).

The PI3 kinase -Akt pathway is activated under growth conditions. Akt activity depends on two phosphorylated residues; one located in the catalytic domain (threonine 308 - Akt Thr^{308}), and the second, serine 473 (Akt Ser^{473}), in the C-terminal regulatory domain (Nicholson and Anderson 2002). Thus, Akt residues Thr^{308} and Ser^{473} are phosphorylated in response to insulin-dependent PI3K activity. As observed in figure 6, complete medium (left panel) and ten-minute insulin treatment (far right panel) of cells associate with positive blot signals for both Akt phosphorylated residues. Serum withdrawal did not change Akt activity, as reported by the blot signals observed for antibodies specific for the phosphorylated residues (Figure 6, second panel). Signals for both phosphorylated residues reporting AKT activity decreased when cells were treated overnight with rapamycin for 16 hours (Figure 6, second right panel), though they appeared unchanged when rapamycin treatment lasted for only one hour (Figure 6, middle panel). As Akt is upstream to the TSC1/2 complex, its residues Thr^{308} and Ser^{473} were expected to be phosphorylated in all cell lines. However, 3H9-1B1 cell line showed unphosphorylated Akt under all conditions, except for insulin treatment.

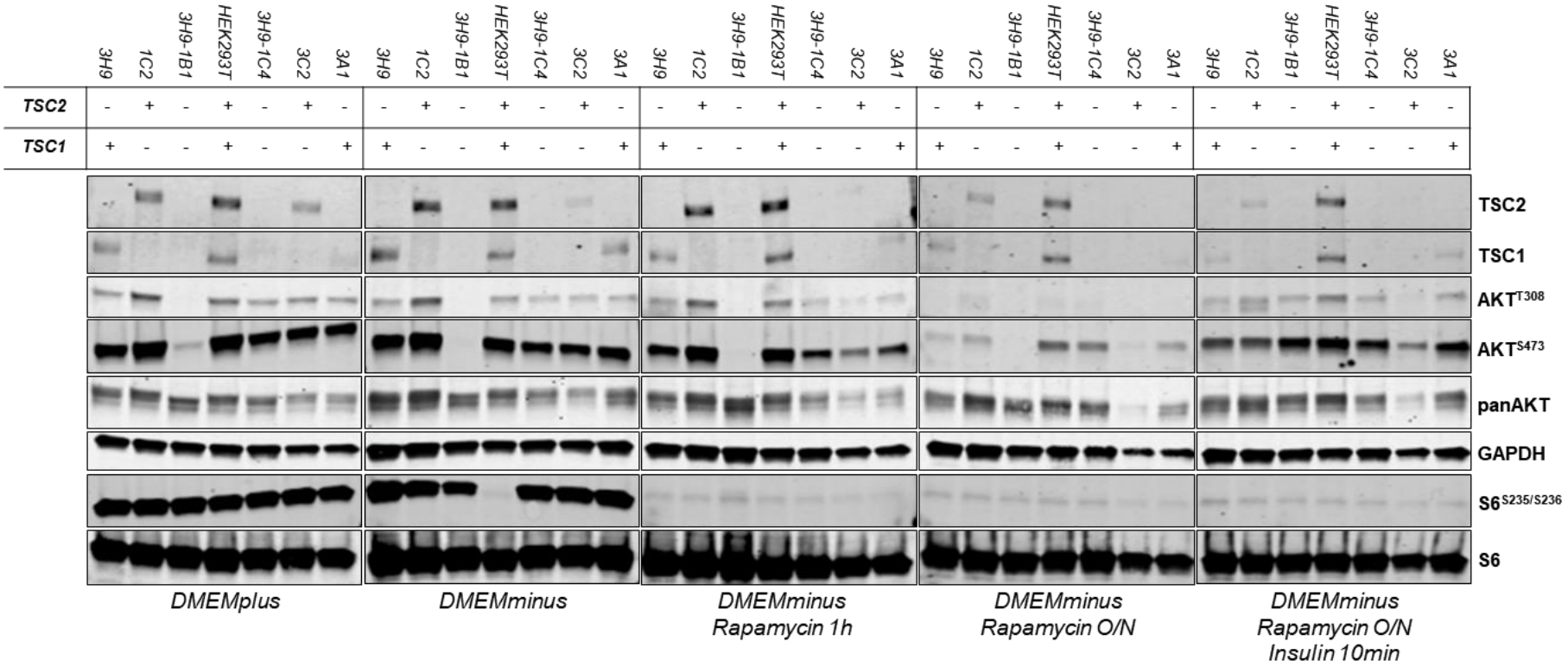


Figure 6: Specific protein profile from genetically modified cell lines. (A) Table of cell lines showing if there is expression (+) or knock-out (-) of HEK293T *TSC1* and *TSC2* genes. (B) Western Blot of proteins isolated from cell lines according to table (A) expressing (+) *TSC1* and *TSC2* or not (-). Antibodies employed on blots detect proteins identified on the right end side (TSC2, TSC1, phosphorylated Thr^{308} from AKT, phosphorylated Ser^{473} from AKT, panAKT, GAPDH, both phosphorylated Ser^{235} and Ser^{236} from S6; and S6. The five blot panels represent distinct cell growth media and treatments, as labeled from left to right: DMEMplus, DMEMminus, DMEMminus and rapamycin for 1 hour, DMEMminus and rapamycin for 16 hours (O/N), and DMEMminus, rapamycin for 16 hours (O/N) and 10-minute treatment with insulin.

An apparent decrease was observed for TSC1 protein amount in serum-deprived cells treated with rapamycin, in the lack of TSC2 expression (Figure 6, cell lines 3H9 and 3A1, three right panels). Moreover, treatment with rapamycin for 16 hours also apparently reduced TSC2 protein levels in the two cell lines without TSC1 protein (1C2 and 3C2). However, gel lanes seem to have different amount of lysate protein, causing an apparent reduction of total protein levels.

B. *TSC1* and *TSC2* mRNA analyses

Reference HEK293T cells, under complete (DMEMplus) or minimum (DMEMminus) medium, disclosed *TSC2* mRNA RPKM values respectively 2.13 and 2.70 times higher than *TSC1* transcript RPKM (Figure 7). These data show that independently on the presence of calf serum *TSC2* mRNA levels are at least twice that of *TSC1* mRNA in the reference HEK293T cell line, as previously reported in databases for other cell lines and tissue (Figure 1 and Figure 2).

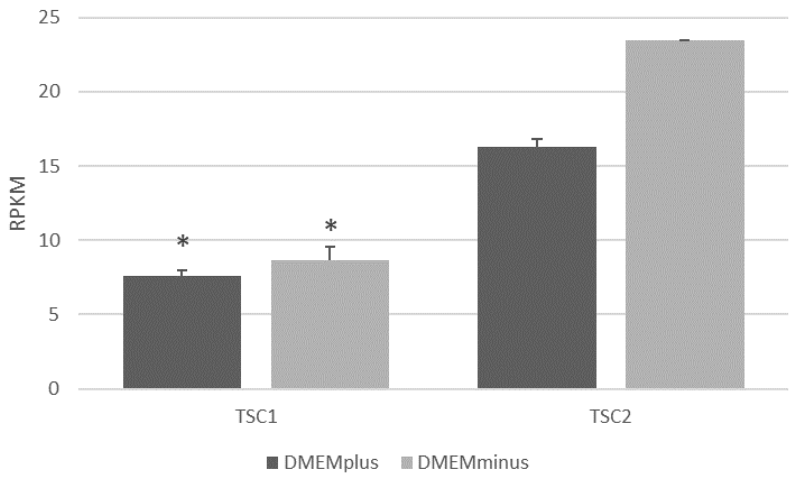


Figure 7: *TSC1* and *TSC2* mRNA RPKM upon DMEMplus and DMEMminus treatments. Student *t*-test was calculated, and asterisk represents a statistically significant difference between *TSC1* and *TSC2* RPKM, under the same culturing condition (*p*-value < 0.05).

As *TSC2*-to-*TSC1* mRNA ratio appeared similar to RNA-Seq reports on other tissues and cell lines, we subsequently analyzed their mRNA amount in HEK293T-derived cell lines for two groups, DMEMplus and DMEMminus. For that, we employed DESeq2 pairwise comparisons to the reference HEK293T cell line, a test based on binary logarithmic fold change of the gene total mRNA read count. In DMEMplus group, in the lack of the TSC1 protein (cell lines 1C2 and 3C2, and 3H9-1B1 and 3H9-1C4), there is a significant reduction of *TSC1* mRNA (Figure 8A). The fold change of those reductions was in average 39% (34% to 44%) as the binary logarithm varied between -0.60 and -0.83 (Figure 8A).

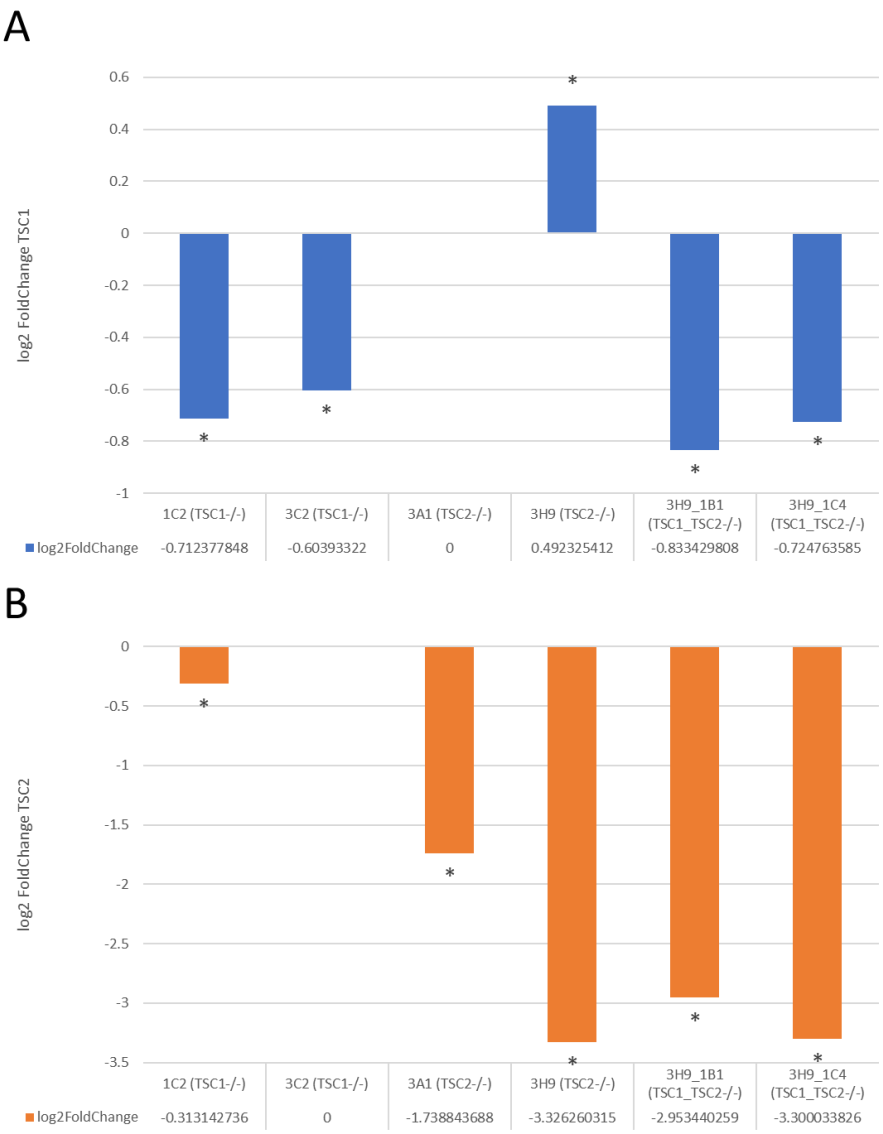


Figure 8: Log2FoldChange of *TSC1* (A) and *TSC2* (B) transcript amount upon DMEMplus treatment. Asterisk = *p*-value<0.05.

The absence of the TSC2 protein (cell lines 3A1, 3H9, 3H9-1B1 and 3H9-1C4) related to a significant decrease in *TSC2* mRNA amount (Figure 8B). Notably, the fold change of these *TSC2* mRNA reductions were considerably higher, in average 84% (70% to 90%) as the binary logarithm varied between -1.74 and -3.33 (Figure 8B). The limited length of the *TSC2* mRNA from the cell lines 3A1, 3H9, 3H9-1B1 and 3H9-1C4 (85% the size of the full-length wild-type *TSC2* mRNA) generates a relatively lower number of reads. Thus, when adjusted to the wild-type transcript length and data size (RKPM) a drastic reduction is expected. Remarkably, 3A1 cell line, which is heterozygous for the gross *TSC2* deletion, presented a less severe reduction of *TSC2* mRNA (70%; Figure 8B).

When the effect of one protein absence was evaluated on the mRNA levels from the partnering gene, cell lines with the same knock-out genotype were discordant. Cells without the TSC1 protein had no effect or

discrete reduction of *TSC2* mRNA (1C2 and 3C2; Figure 8B). On the other hand, cells lacking the TSC2 protein had equal amounts or significantly more *TSC1* mRNA than the reference HEK293T cell line (Figure 8A).

When the same mRNA analysis was conducted for the DMEMminus group, we observed fold-change reductions of *TSC2* mRNA in the absence of TSC2 protein similar to the comparisons from the DMEMplus group, in average 88% reduction (78% to 93%). The lowest reduction (78%) was from the cell line with one allele with *TSC2* gross deletion. We additionally disclosed a concordant, significant, more discrete reduction of *TSC2* mRNA in both cell lines (1C2 and 3C2) that have no TSC1 protein (Figure 9B). In the DMEMminus group, four cell lines had undetectable change in *TSC1* mRNA levels in the binary logarithmic scale. One cell line without TSC1 protein (1C2) had a significant reduction of *TSC1* mRNA (39% decrease), as observed for the DMEMplus group (Figure 9A). In another cell line, which lacks TSC2 protein (3H9), *TSC1* mRNA increased nearly 70% when compared to the reference HEK293T cell line (Figure 9A).

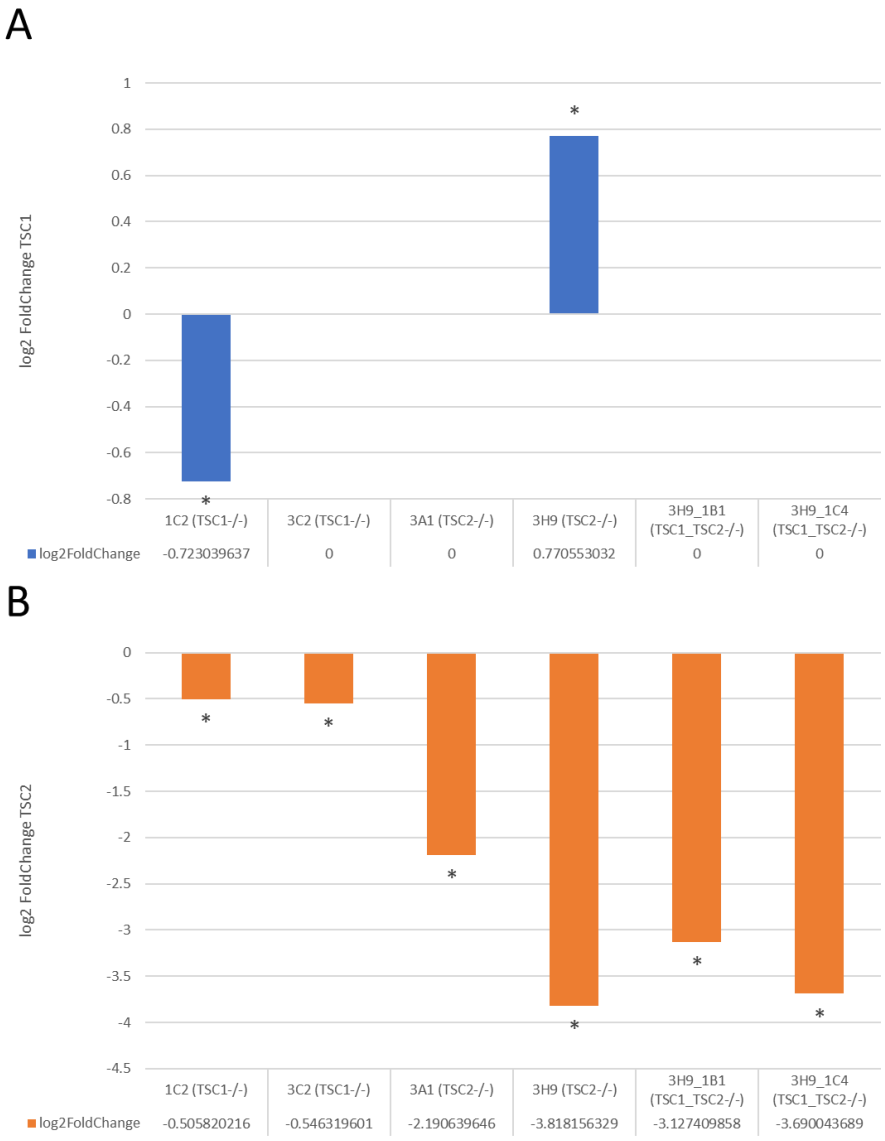


Figure 9: Log2FoldChange of *TSC1* (A) and *TSC2* (B) transcript amount upon DMEMminus treatment. Asterisk = *p*-value<0.05.

Table 2: Observed *TSC1* and *TSC2* mRNA and protein levels for each knock-out cell line comparatively to HEK293T.

Knocked-out gene(s)		*TSC1*		*TSC2*		*TSC1 and TSC2*	
Cell line		1C2	3C2	3A1	3H9	3H9-1B1	3H9-1C4
Protein	TSC1	-	-	↓	↓	-	-
	TSC2	↓	↓	-	-	-	-
mRNA (DMEMplus)	*TSC1*	↓	↓	-	↑	↓	↓
	TSC2	↓	-	↓	↓	↓	↓
mRNA (DMEMminus)	*TSC1*	↓	-	-	↑	-	-
	TSC2	↓	↓	↓	↓	↓	↓

(-) unchanged or not observed; (↓) reduced; (↑) increased.

As observed in Table 2, 1C2 was the only cell line that independently on the medium consistently decreased *TSC1* and *TSC2* mRNA levels. This cell line expresses no TSC1 protein. It presented decreased TSC2 protein as well as *TSC1* and *TSC2* mRNA amounts. The two CRISPR/Cas9-generated deletions in 1C2 are fairly small, in homozygosity. One deletion abolishes the translational start codon, and the other generates a frameshift, with no significant impact on the *TSC1* mRNA length. Therefore, we selected the 1C2 cell line for further experiments, aiming at global analyses of the effects of the lack of TSC1 protein on the cell transcriptome.

C. Global RNA analysis in the 1C2 cell line

The number of differentially expressed genes in the 1C2 cell line compared to HEK293T (reference) gene set are indicated in Table 3 for both DMEMminus and DMEMplus groups. 1C2 genes with less mRNA or more mRNA than the reference cell line were considered respectively down- or up-regulated. In all comparisons, Benjamini-Hochberg test was employed with FDR <0.05.

Table 3: Number of genes down- or up-regulated in 1C2 cell line in regard to HEK293T in DMEMplus or DMEMminus groups.

Groups compared	Number of down-regulated genes and labeled group ID*	Number of up-regulated genes and labeled group ID*
1C2 X HEK293T (DMEMminus)	645 (B2)	474 (C2)
1C2 X HEK293T (DMEMplus)	1,411 (B3)	1,746 (C3)

*Group ID in parenthesis

Gene ontology (GO) analysis of 645 genes down-regulated in 1C2 in minimum medium (DMEMminus) significantly grouped them in ten sets, as presented in Table 4. The gene sets at this table is similar to the gene sets identified for down-regulated genes from DMEMplus group (N=1,411; Table 5), and relate to post-transcription steps of gene expression.

Table 4: Gene sets significantly enriched in down-regulated genes in 1C2 cell line in regard to HEK293T, in DMEMminus group.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0006401	RNA catabolic process	341	10.66	43.15	0	0
GO:0090150	Establishment of protein localization to membrane	313	97.86	48.03	0	0
GO:0006413	Translational initiation	192	60.03	84.96	0	0
GO:0070972	Protein localization to endoplasmic reticulum	137	42.83	95.72	0	0
GO:0006605	Protein targeting	412	12.88	38.04	5,55E-12	9,44E-10
GO:0002181	Cytoplasmic translation	92	28.76	73.01	5,84E-09	8,28E-07
GO:0022613	Ribonucleoprotein complex biogenesis	440	13.76	28.35	4,54E-05	5,51E-03
GO:0071826	Ribonucleoprotein complex subunit organization	245	76.60	32.64	2,10E-03	2,22E-5
GO:0034248	Regulation of cellular amide metabolic process	385	12.04	21.60	1,99E-4	1,88E-2
GO:0006284	Base-excision repair	43	13.44	52.07	3,39E-4	2,88E-2

Table 5: Gene sets significantly enriched in down-regulated genes in 1C2 cell line in regard to HEK293T, in DMEMplus group.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0022613	Ribonucleoprotein complex biogenesis	440	40.71	27.51	0	0
GO:0006401	RNA catabolic process	341	31.55	29.80	0	0
GO:0006413	Translational initiation	192	17.76	37.72	0	0
GO:0070972	Protein localization to endoplasmic reticulum	137	12.67	37.08	4,44E-12	9,44E-10
GO:0071826	Ribonucleoprotein complex subunit organization	245	22.66	26.91	3,63E-09	6,00E-07
GO:0002181	Cytoplasmic translation	92	85.11	39.95	4,23E-09	6,00E-07
GO:0016072	rRNA metabolic process	236	21.83	24.73	2,58E-06	3,13E-04
GO:0090150	Establishment of protein localization to membrane	313	28.96	21.76	2,25E-05	2,39E-03
GO:0008380	RNA splicing	417	38.58	19.70	5,59E-05	5,28E-03
GO:0006605	Protein targeting	412	38.11	19.68	7,44E-05	6,33E-03

Up-regulated genes were significantly enriched in two gene sets for the DMEMminus group (Table 6), and in at least ten GO sets for the DMEMplus group (Table 7). Those gene sets relate to response to topologically incorrect protein and autophagic mechanism in DMEMminus group. The DMEMplus group of up-regulated genes was enriched in gene sets that relate to cell growth and morphogenesis.

Table 6 Gene sets significantly enriched in up-regulated genes in 1C2 cell line in regard to HEK293T, in DMEMminus group.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0035966	Response to topologically incorrect protein	188	43.417	36.85	7,63E-6	6,48E-3
GO:0061919	Process utilizing autophagic mechanism	473	10.923	22.89	1,07E-4	4,56E-2

Table 7: Gene sets significantly enriched in up-regulated genes in 1C2 cell line in regard to HEK293T, in DMEMplus group.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO: 0034330	Cell junction organization	285	19.73	23.32	5,77E-04	4,90E-5
GO: 0022604	Regulation of cell morphogenesis	473	32.74	18.94	7,86E-03	3,34E-4
GO: 2000147	Positive regulation of cell motility	493	34.12	18.17	3,17E-6	8,45E-4
GO: 0010975	Regulation of neuron projection development	475	32.88	18.25	3,97E-6	8,45E-4
GO: 0060560	Developmental growth involved in morphogenesis	225	15.57	22.47	5,16E-6	8,78E-4
GO: 0090066	Regulation of anatomical structure size	487	33.71	17.80	8,74E-6	1,12E-4
GO: 0045785	Positive regulation of cell adhesion	392	27.132	18.797	9,25E-6	1,12E-3
GO: 0048638	Regulation of developmental growth	314	21.733	19.325	2,91E-5	3,09E-3
GO: 0048732	Gland development	434	30.039	17.644	3,70E-5	3,49E-3
GO: 0045926	Negative regulation of growth	250	17.304	20.227	5,17E-5	4,39E-3

As shown in Table 3, we labeled as B2, B3, C2, and C3 the sets of differentially expressed genes from comparisons between 1C2 and HEK293T cell lines, if they were respectively down-regulated in DMEMminus or DMEMplus groups, or up-regulated in DMEMminus or DMEMplus groups. These gene group identifications were used to show in Table 8 the intersections that were generated between gene sets from two experimental groups. The number of retrieved genes is indicated in the same table.

Table 8: Number of genes observed when groups assigned in Table 3 were intersected.

Intersected groups	Number of genes from intersection or intersection subtracted
B2 ∩ B3	513
C2 ∩ C3	326
B2 - (B2 ∩ B3)	133
B3 - (B2 ∩ B3)	1,234
C2 - (C2 ∩ C3)	149
C3 - (C2 ∩ C3)	1,086

B2: down-regulated genes in DMEMminus group from 1C2 X HEK293T comparison
B3: down-regulated genes in DMEMplus group from 1C2 X HEK293T comparison
C2: up-regulated genes in DMEMminus group from 1C2 X HEK293T comparison
C3: up-regulated genes in DMEMplus group from 1C2 X HEK293T comparison

The intersections between down-regulated gene sets (B2 ∩ B3) represent significantly down-regulated genes independently on the presence of serum. As shown in Table 9, the twelve gene sets enriched in this intersection group (N = 513) relates to post-transcritptional steps of gene expression, notably mRNA translation, and ATP generation. Thus, these biological processes appear down-regulated in the lack of the TSC1 protein, independently on culture serum (trophic factors). Differentially expressed genes uniquely detected in the presence of serum were enriched in gene sets also directly related to gene expression,

however, mostly to the nuclear steps of gene expression (Table 10). Down-regulated genes detected in DMEMminus but not in DMEMplus samples (B2 - (B2 ∩ B3), N = 133) were not significantly enriched in gene sets.

Table 9: Gene sets significantly enriched in down-regulated genes in 1C2 cell line in regard to HEK293T, independently on the serum presence.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0090150	Establishment of protein localization to membrane	313	79.68	55.22	0	0
GO:0006413	Translational initiation	192	48.88	98.20	0	0
GO:0070972	Protein localization to endoplasmic reticulum	137	34.88	10.90	0	0
GO:0006401	RNA catabolic process	341	86.81	47.23	1,11E-12	2,36E-10
GO:0006605	Protein targeting	412	10.49	41.95	5,55E-12	9,44E-10
GO:0002181	Cytoplasmic translation	92	23.42	85.39	1,24E-09	1,76E-07
GO:0022613	Ribonucleoprotein complex biogenesis	440	11.20	32.14	6,45E-06	7,84E-04
GO:0071826	Ribonucleoprotein complex subunit organization	245	62.37	36.88	7,53E-04	8,00E-6
GO:0009141	Nucleoside triphosphate metabolic process	304	77.39	24.55	3,02E-4	2,85E-2
GO:0009123	Nucleoside monophosphate metabolic process	321	81.72	23.25	5,91E-4	5,02E-2
GO:0006091	Generation of precursor metabolites and energy	469	11.940	20.101	9,64E-4	7,44E-2
GO:0016072	rRNA metabolic process	236	60.080	24.967	1,07E-3	7,63E-2

Table 10: Gene sets significantly enriched in down-regulated genes in 1C2 cell line in regard to HEK293T, detected in DMEMplus but not in DMEMminus samples

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0022613	Ribonucleoprotein complex biogenesis	440	29.50	25.76	1,33E-10	1,13E-07
GO:0006397	mRNA processing	487	32.65	21.74	3,64E-06	1,55E-03
GO:0008380	RNA splicing	417	27.96	22.53	8,74E-06	2,48E-03
GO:0006403	RNA localization	228	15.29	27.47	1,70E-05	3,61E-03
GO:0006401	RNA catabolic process	341	22.86	23.18	7,26E-05	1,23E-6
GO:1903311	Regulation of mRNA metabolic process	266	17.84	24.11	6,12E-04	8,66E-6
GO:0016072	rRNA metabolic process	236	15.82	24.64	1,39E-03	1,68E-5
GO:0010608	Posttranscriptional regulation of gene expression	486	32.59	19.33	3,15E-03	3,34E-5
GO:0051169	Nuclear transport	356	23.87	20.95	4,94E-03	4,66E-5
GO:0071826	Ribonucleoprotein complex subunit organization	245	16.43	23.13	1,03E-6	8,76E-5

There was no significant enrichment in gene sets for up-regulated genes in 1C2 cell line when compared to HEK293, independently on the serum presence (C2 ∩ C3, N = 326). Gene sets such as response to topologically incorrect protein and regulation of protein stability were enriched in up-regulated genes in DMEMminus but not in DMEMplus (Table 11). Enriched gene sets for genes up-regulated in DMEMplus but not DMEMminus samples relate to cell morphogenesis (Table 12).

Table 11: Gene sets significantly enriched in up-regulated genes in 1C2 cell line in regard to HEK293T, detected in DMEMminus but not in DMEMplus samples.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0035966	Response to topologically incorrect protein	188	14.47	62.19	1,45E-5	1,23E-2
GO:0031647	Regulation of protein stability	267	20.55	48.65	4,02E-5	1,70E-2

Table 12: Gene sets significantly enriched in up-regulated genes in 1C2 cell line in regard to HEK293T, detected in DMEMplus but not in DMEMminus samples

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0034330	Cell junction organization	285	15.34	26.08	2,43E-04	2,06E-5
GO:0022604	Regulation of cell morphogenesis	473	25.46	19.64	3,70E-6	1,57E-3
GO:0060560	Developmental growth involved in morphogenesis	225	12.11	23.95	1,12E-5	3,18E-3
GO:0048638	Regulation of developmental growth	314	16.90	21.30	1,50E-5	3,18E-3
GO:0043087	Regulation of GTPase activity	472	25.40	18.50	3,40E-5	5,79E-3
GO:2000147	Positive regulation of cell motility	493	26.53	18.09	4,97E-5	7,04E-3
GO:0045785	Positive regulation of cell adhesion	392	21.10	18.96	7,60E-5	8,59E-3
GO:0010975	Regulation of neuron projection development	475	25.56	17.99	8,08E-5	8,59E-3
GO:0031589	Cell-substrate adhesion	332	17.87	19.59	1,11E-4	1,04E-2
GO:0090066	Regulation of anatomical structure size	487	26.21	17.55	1,45E-4	1,23E-2

When only the 1C2 cell line was considered for global RNA analyses, and the groups DMEMminus and DMEMplus (reference) were compared, we retrieved 72 and 51 genes that were respectively down - and up-regulated in 1C2 DMEMminus group. Gene sets that revealed a significant enrichment are presented in Table 13 and Table 14. We observed that response to topologically incorrect protein is both down- and up-regulated. Response to endoplasmic reticulum stress is down-regulated, whereas protein folding and response to temperature stress are up-regulated, consistently with the role of TSC1 protein as heat-shock protein co-chaperone.

Table 13: Gene sets significantly enriched in down-regulated genes in DMEMminus group in regard to DMEMplus, in the 1C2 cell line.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0034976	Response to endoplasmic reticulum stress	268	10.47	76.22	9,07E-6	7,71E-3
GO:0006730	One-carbon metabolic process	25	0.10	30.64	1,23E-4	3,49E-2
GO:0035966	Response to topologically incorrect protein	188	0.73	81.49	9,04E-5	3,49E-2
GO:0070972	Protein localization to endoplasmic reticulum	137	0.54	93.18	1,94E-4	4,13E-2

Table 14: Gene sets significantly enriched in up-regulated genes in DMEMminus group in regard to DMEMplus, in the 1C2 cell line.

Gene Set	Description	Size	Expect	Ratio	pValue	FDR
GO:0035966	Response to topologically incorrect protein	188	0.49510	18.178	1,08E-05	6,22E-03
GO:0051131	Chaperone-mediated protein complex assembly	20	0.052671	94.929	1,46E-05	6,22E-03
GO:0006457	Protein folding	210	0.55304	12.657	1,09E-6	3,09E-4
GO:0009266	Response to temperature stimulus	201	0.52934	11.335	1,30E-5	2,76E-3
GO:0070841	Inclusion body assembly	22	0.057938	51.780	2,51E-5	4,27E-3
GO:0046677	Response to antibiotic	316	0.83220	72.098	1,62E-4	2,29E-2
GO:0044839	Cell cycle G2/M phase transition	213	0.56094	89.136	2,27E-4	2,75E-2
GO:0071559	Response to transforming growth factor beta	238	0.62678	79.773	3,78E-4	4,02E-2

V. Discussion

A. Cell line characterization

In this study, we employed six cell lines derived from HEK293T cell line by CRISPR-Cas9 editing of the *TSC1* (chromosome 9q34.1) and or *TSC2* (chromosome 16p13.3) genes. HEK293 cell line has been first obtained after immortalization by adenovirus transduction of primarily cultured human embryonic kidney (HEK) cells (Graham, Smiley *et al.* 1977). Although the identity of the cell that originally derived the HEK293 line is unknown, later analyses disclosed that HEK293 cells exhibit some adrenal cell properties, including neuronal-like (Lin, Boone *et al.* 2014). Therefore, the original HEK293 line should not be considered as a kidney-derived cell line, but a complex, heterogeneous cell line, likely derived from cells of the adrenal glands, which are anatomically related to the kidneys (Lin, Boone *et al.* 2014). Among further genomic modifications these cells have undergone, HEK293T cell line, employed in our study, originated upon SV40 transfection of the HEK293 cell line (Rio, Clark *et al.* 1985). HEK293T presents a combination of karyotypes with some chromosomes in three copies, leading to the classification of the cell line as hypotriploid or pseudotriploid (Lin, Boone *et al.* 2014). In 19 HEK293T chromosome metaphases examined by (Lin, Boone *et al.* 2014), six cells presented t(9,17) translocations, five had duplicated chromosome 9 short arm, two cells had deleted 9q, one cell had a 9q duplication, and two cells had a translocation involving chromosome 16 (t(16,21) and t(16,22). Although Lin, Boone *et al.* (2014) did not demonstrate the translocation breakpoints nor 9q34.1 and 16p13.3 trisomy in HEK293T cells, their data tend to suggest a situation of eventual di- or monosomy for the *TSC1* and *TSC2* genes. As *TSC1* and *TSC2* genes are considered haplosufficient in non-neuronal cells, we believe that HEK293T chromosome 9 and 16 ploidies are not a concern for the present study.

CRISPR-Cas9-edited HEK293T cell lines analyzed in the present study have been previously produced in the Nellist laboratory (Erasmus Medical Centre, Rotterdam, The Netherlands), and have *TSC1* and *TSC2* genotypes indicated in Figure 3. Our Western blot analyses have shown that all cell lines that had either gene inactivated did not express the respective protein (Figures 5 and 6), and had mTORC1 hyperactivated, as presented in the functional assay by S6-phosphorylation reporter (Figure 6). Hyperactivated mTOR was observed under both DMEMplus and DMEMminus conditions, similarly to reports from the literature (Tee, Fingar *et al.* 2002, Garami, Zwartkruis *et al.* 2003). Amino acids, such as leucine and arginine, activate mTORC1 independently on growth factors from serum in cells where *TSC1* or *TSC2* is knocked out (Byfield, Murray *et al.* 2005, Nobukuni, Joaquin *et al.* 2005, Smith, Finn *et al.* 2005). As DMEM minimum medium contains 15 out of the 20 amino acids, it is believed that the growth factor-independent, amino-acid-dependent mTORC1 activation pathway should be undertaken. This pathway is regulated by the Rag GTPases, which allow mTORC1 to interact with the lysosome membrane. Therefore, in DMEMminus conditions, cells negative for TSC1 or TSC2 present hyperphosphorylated S6 Ser^{235} and Ser^{236} (Figure 6, second panel). Under conditions of amino

acid depletion, Rag GTPases recruit TSC2 to the lysosome, interacting with Rheb. mTORC1 is thus released from the lysosome and becomes inactivated (Bar-Peled, Schweitzer *et al.* 2012). As seen in Figure 6 (third and fourth panels), mTORC1 activity was inhibited upon rapamycin treatment for one or 16 hours, as shown by hypophosphorylation of the reporter S6/Ser^{235}/Ser^{236}.

Upon insulin or insulin-like growth factor (IGF) binding, the insulin/IGF receptor becomes activated by auto-phosphorylation, starting a signalling cascade that activates the PI3 kinase (PI3K)-Akt pathway. PI3K activity on membrane lipids is essential for Akt and PDK1 (phosphoinositide-dependent kinase 1) translocation to the cell membrane. PDK1 has been demonstrated to phosphorylate Akt Thr^{308}, thus inhibiting the TSC1/2 complex (Kwiatkowski and Manning 2005). On the other hand, considerable efforts have been made to understand the roles of different kinases shown to phosphorylate Akt Ser^{473}. For instance, mTORC2 (mechanistic target of rapamycin complex 2) has been reported to phosphorylate Akt Ser^{473} under a variety of physiological conditions, and to facilitate phosphorylation of AKT/Thr^{308} by PDK1 (Sarbassov, Guertin *et al.* 2005) . Similarly to mTORC1, the large molecular complex mTORC2 contains mTOR and mLST8 (mammalian lethal with Sec13 protein 8), but differs from the former complex as they present different additional partnering proteins, subcellular localization, functions and sensitivity to rapalogs. MTORC2 phosphorylates several members of protein kinase families A, B (Akt), C and G, affecting cytoskeleton remodelling, ion transport and cell survival, growth, proliferation (Reviewed by Menon, 2018).

We observed that serum starvation for 16 hours did not apparently affect the phosphorylation status of Akt residues Thr^{308} and Ser^{473} not even when followed by one-hour treatment with rapamycin (Figure 6). It is as yet difficult to explain the lack of effect of serum deprivation on Akt dephosphorylation. However, as there should be a reciprocal relation between Akt and the two mTOR complexes, different possibilities could be raised, such as inhibition of specific phosphatases, a delayed response to repress mTORC2 or susceptibility of Akt Thr^{308} and Ser^{473} to additional kinases.

The *TSC1*$^{-/-}$, *TSC2*$^{-/-}$ 3H9-1B1 cell line had AKT hypophosphorylated in all treatments, a situation restored upon insulin treatment. A marked reduction in Akt activation was reported in *Tsc2*$^{-/-}$/*Tp53*$^{-/-}$ and *Tsc1*$^{-/-}$ cells in response to serum (Zhang, Cicchetti *et al.* 2003). The combination of double knock-out 3H9-1B1 cell line and the intrinsic inhibition of p53 in HEK293T (Lilyestrom, Klein *et al.* 2006) could explain why this cell line did not respond to Akt kinases (Figure 6). However, 3H9-1B1 cell line did not considerably differ from 3H9-1C4, except for Akt phosphorylation status.

Knocking-out *TSC1* or *TSC2* genes by CRISPR-Cas9 genome edition did not affect the overall protein levels of upstream AKT protein or downstream S6 protein (Figure 6). The TSC2 levels in *TSC1*$^{-/-}$ cells were reduced relative to the wild-type cell line HEK293T; likewise, the protein levels of TSC1 in *TSC2*$^{-/-}$ cells (Table 2). Our data do not corroborate observations that the knockout of *Tsc2* maintains levels of Tsc1 (Zhang, Gao

et al. 2003, Pollizzi, Malinowska-Kolodziej *et al.* 2009). However, embryonic fibroblasts obtained from murine *Tsc1* knockout model (MEF *Tsc1*$^{-/-}$) showed reduced levels of Tsc2 (Astrinidis, Senapedis *et al.* 2006). *TSC1* transient transfection increased endogenous levels of TSC2 in wild-type cells, and transient co-transfection of *TSC1* with *TSC2* in the same cell type increased TSC2 levels (Benvenuto, Li *et al.* 2000). These observations can be attributed to the Tsc1 property as a co-chaperone of Hsp90 and properly folding of Tsc2 (Woodford, Sager *et al.* 2017). The absence of *TSC1* in human cell line might have the same effects as in mouse cells, what could inhibit the folding of TSC2, thus leading to its ubiquitination and degradation. In the same way, low levels of TSC2 in knocked-out cell lines should lead to ubiquitination and degradation of TSC1. Part of non-degraded TSC1 could accumulate in the cell as partially unfolded, self-aggregated protein, as suggested by Nellist, van Slegtenhorst *et al.* (1999). As TSC1 protein has been considered a chaperone-allied protein, this scenario considering both protein degradation and aggregation could be closely examined to answer as well if other protein species are also unfolded and present in the TSC1 aggregates.

B. *TSC1* and *TSC2* gene expression regulation

The expression of *TSC2* has been observed to be higher than that of the *TSC1*gene at the mRNA level (Figure 1 and 2), independently whether cells are growing under complete or minimum medium (Figure 7). However, the half-life of mRNA depends on transcription activity as well as decay mechanisms. Thus it is recommended to refer to as alteration in mRNA amount or levels, as the responsible mechanisms cannot be inferred without further experiments. Furthermore, one cannot infer a quantitative correlation between mRNA and the respective protein taking into account the mRNA amount, as there are various post-transcriptional steps regulating gene expression.

Regardless of medium conditions, the reduction of *TSC2* mRNA expression has been more pronounced than those in *TSC1* mRNA levels. This higher decrease in expression of *TSC2* may be due to the genomic edition, which subtracted nearly 4.8 Kb of *TSC2* coding sequence (Figure 3B). Consequently, the *TSC2*$^{-/-}$ cell lines should provide significantly lower RNA-Seq *TSC2* read count, and thus they do not constitute an appropriate model for *TSC1/2* gene regulation studies. Furthermore, the unaffected *TSC1* mRNA levels in the DMEMminus group (similar to HEK293T *TSC1* mRNA levels) suggest that the *TSC1* mRNA down-regulation should be sensitive to the presence of growth factors, TSC1 and TSC2 proteins. As mTORC1 has been functionally involved in transcriptional control as well translational regulation, it should be interesting to study if rapamycin treatment affects *TSC1* mRNA in the 1C2 cell line. We also disclosed a discrete reduction of *TSC2* mRNA in both cell lines that have no TSC1 protein (Figure 9B) under minimum medium condition.

The extension of the mRNA 5'-untranslated regions (5'-UTR) is similar for human *TSC1* (234 bp) and *TSC2* (200 bp) mRNA. The 230-bp 5'UTR of mouse *Tsc1* is 88% (202/230) similar to the human orthologous 5'UTR. The 119-bp 5'UTR of mouse *Tsc2* is 70% (83/119) similar to the 5'UTR of the human orthologue.

However, human *TSC1* (4,887 bp) and *TSC2* (109 bp) 3'- untranslated region (3´-UTR) lengths are significantly different. Human and mouse *Tsc1* 3'-UTR length and sequence are conserved, the latter containing 3,999 bp. On the other hand, although human and mouse *Tsc2* 3'-UTRs have similar lengths, the latter with 131 bp, the sequences display lower similarity (51%, 56/109). The longer *TSC1* 3´-UTR may infer that this mRNA is more responsive to translational regulation and more stable than TSC2 mRNA, as it may harbor *cis* regulatory motifs in that transcribed segment. Furthermore, changes in mRNA amount may build a feedback control loop on TSC gene transcription. Studies based on quantitative RT-PCR of the *TSC1* and TSC2 cDNA, combined with the use of chemical transcriptional, translational or proteasomal inhibitors are likely to clarify the type of regulation to which TSC mRNA and proteins should be submitted. The data presented here, along with the anatomy of the *TSC1* gene, suggest that *cis* elements in the mRNA 3'-UTR region could regulate the stability of the transcript. Finally, if RNA decay and translation appear to regulate the steady state of *TSC1* protein and mRNA, it would be worth screening *TSC1* mRNA for *cis*-regulating motifs.

C. Protein stress response in TSC1-negative cell line

TSC1$^{-/-}$ 1C2 cell line has one deletion encompassing the translational start codon and another causing frameshift, basically with no impact on *TSC1* RNA length. Considering that more than 80% of TSC patients have point or small indel pathogenic variants (Niida, Lawrence-Smith *et al.* 1999, van Slegtenhorst, Verhoef *et al.* 1999, Dabora, Jozwiak *et al.* 2001, Au, Williams *et al.* 2007), the 1C2 cell line should be a good model study loss-of-function *TSC1* mutations. Although it was not an objective of the present work, we were provided with RNA-Seq from the seven cell lines under different growth conditions. As the 1C2 cell line appeared to consistently affect *TSC1* and *TSC2* mRNA and protein, its RNA-Seq data were analyzed under complete or minimum medium in comparison to HEK293T reference cell line.

Under growth conditions, 1C2 down-regulated mRNAs related to posttranscriptional steps of gene expression, notably RNA processes and translation (Table 10), and was not considerably different from gene sets enriched in down-regulated genes for starvation conditions (Table 4). However, the latter group was more represented by translation-associated GOs. It is important to note that the differentially expressed genes have not been identified or studied on an individual basis. Therefore, it is possible that translation-associated GOs be enriched in translation repression genes, which are expected to be down-regulated when mTOR is hyperactivated (Goncharova, Goncharov *et al.* 2002).

Likewise, it is important to put in evidence that the GO response to unfolded proteins has been identified in both down-regulated and up-regulated gene sets, inferring that, if truly positive evidences, a detailed analysis of each gene function is advised, as positive and negative regulatory functions could be represented. These data, as well as the gene sets for the ER stress response are consistent with the role of

TSC1 as HSP90 co-chaperone (Woodford, Sager *et al.* 2017, Sager, Woodford *et al.* 2018, Sager, Woodford *et al.* 2018).

The DMEMplus group presented up-regulated genes significantly enriched in cell growth and morphogenesis sets while, in DMEMminus set, topologically incorrect protein and autophagic mechanisms were enriched. Cell size phenotypes have been observed in *Tsc1*$^{-/-}$ cells from *Drosophila melanogaster* that dramatically increased in size, yet differentiated normally (Potter, Huang *et al.* 2001). Autophagic mechanisms are related to the TSC1/2 proteins, which up-regulate this pathway, which becomes inhibited upon mTORC1 hyperactivation (Martina, Chen *et al.* 2012, Efeyan, Zoncu *et al.* 2013).

We did not find significant enrichment in gene sets for up-regulated genes independently on the presence of serum. However, up-regulated gene sets in DMEMplus group did relate to cell morphogenesis. This transcriptional trait for cell morphology of 1C2 *TSC1*$^{-/-}$ are in agreement with previously reports, in which overexpression of the *TSC1* gene product hamartin results in changes of cell morphology (Benvenuto, Li *et al.* 2000).

Finally, when only the 1C2 cell line was considered for global RNA analyses between DMEMminus and DMEMplus groups, more than 50 down-regulated genes have been detected. Similarly to up-regulated genes on DMEMminus group, response to topologically incorrect protein is both down- and up-regulated, which means that incorrect protein folding is an intrinsic characteristic of the 1C2 *TSC1*$^{-/-}$ cell line. Interestingly, response to endoplasmic reticulum (ER) stress is down-regulated. The ER is the cellular compartment where synthesis of proteins for the secretory pathway takes place. The ER stress response has been observed in mTORC1 activation in TSC2-null Elt3 cells via c-MYC translation (Babcock, Nguyen *et al.* 2013). mTORC1 hyperactivation increases the basal level of ER stress via an accumulation of unfolded proteins (Clarke, Chambers *et al.* 2014, Johnson, Dunlop *et al.* 2018). Taking together, a stress on this cellular compartment due to mTORC1 hyperactivation may cause an accumulation of unfolded proteins leading to a negative feedback and down regulate genes related to post-transcritptional steps, as observed for 1C2 cell line. It is worth to note that ER stress can develop in TSC-associated cells due to mTORC1-driven protein translation. mTORC1 inhibition of ER stress by RAD001 suppressed angiomyolipoma cell proliferation in a cytostatic manner (Siroky, Yin *et al.* 2012). Finally, protein folding and response to temperature stress are up-regulated in the 1C2 cell line, consistent with the role of TSC1 protein as a heat-shock co-chaperone protein (Woodford, Sager *et al.* 2017).

VI. References

Astrinidis, A., W. Senapedis and E. P. Henske (2006). "Hamartin, the tuberous sclerosis complex 1 gene product, interacts with polo-like kinase 1 in a phosphorylation-dependent manner." Hum Mol Genet **15**(2): 287-297.

Au, K. S., A. T. Williams, E. S. Roach, L. Batchelor, S. P. Sparagana, M. R. Delgado, J. W. Wheless, J. E. Baumgartner, B. B. Roa, C. M. Wilson, T. K. Smith-Knuppel, M. Y. Cheung, V. H. Whittemore, T. M. King and H. Northrup (2007). "Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States." Genet Med **9**(2): 88-100.

Babcock, J. T., H. B. Nguyen, Y. He, J. W. Hendricks, R. C. Wek and L. A. Quilliam (2013). "Mammalian target of rapamycin complex 1 (mTORC1) enhances bortezomib-induced death in tuberous sclerosis complex (TSC)-null cells by a c-MYC-dependent induction of the unfolded protein response." J Biol Chem **288**(22): 15687-15698.

Bar-Peled, L., L. D. Schweitzer, R. Zoncu and D. M. Sabatini (2012). "Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1." Cell **150**(6): 1196-1208.

Benvenuto, G., S. Li, S. J. Brown, R. Braverman, W. C. Vass, J. P. Cheadle, D. J. Halley, J. R. Sampson, R. Wienecke and J. E. DeClue (2000). "The tuberous sclerosis-1 (TSC1) gene product hamartin suppresses cell growth and augments the expression of the TSC2 product tuberin by inhibiting its ubiquitination." Oncogene **19**(54): 6306-6316.

Byfield, M. P., J. T. Murray and J. M. Backer (2005). "hVps34 is a nutrient-regulated lipid kinase required for activation of p70 S6 kinase." J Biol Chem **280**(38): 33076-33082.

Clarke, H. J., J. E. Chambers, E. Liniker and S. J. Marciniak (2014). "Endoplasmic reticulum stress in malignancy." Cancer Cell **25**(5): 563-573.

Dabora, S. L., S. Jozwiak, D. N. Franz, P. S. Roberts, A. Nieto, J. Chung, Y. S. Choy, M. P. Reeve, E. Thiele, J. C. Egelhoff, J. Kasprzyk-Obara, D. Domanska-Pakiela and D. J. Kwiatkowski (2001). "Mutational analysis in a cohort of 224 tuberous sclerosis patients indicates increased severity of TSC2, compared with TSC1, disease in multiple organs." Am J Hum Genet **68**(1): 64-80.

Dan, H. C., M. Sun, L. Yang, R. I. Feldman, X. M. Sui, C. C. Ou, M. Nellist, R. S. Yeung, D. J. Halley, S. V. Nicosia, W. J. Pledger and J. Q. Cheng (2002). "Phosphatidylinositol 3-kinase/Akt pathway regulates tuberous sclerosis tumor suppressor complex by phosphorylation of tuberin." J Biol Chem **277**(38): 35364-35370.

Demetriades, C., N. Doumpas and A. A. Teleman (2014). "Regulation of TORC1 in response to amino acid starvation via lysosomal recruitment of TSC2." Cell **156**(4): 786-799.

Efeyan, A., R. Zoncu, S. Chang, I. Gumper, H. Snitkin, R. L. Wolfson, O. Kirak, D. D. Sabatini and D. M. Sabatini (2013). "Regulation of mTORC1 by the Rag GTPases is necessary for neonatal autophagy and survival." Nature **493**(7434): 679-683.

Gai, Z., W. Chu, W. Deng, W. Li, H. Li, A. He, M. Nellist and G. Wu (2016). "Structure of the TBC1D7-TSC1 complex reveals that TBC1D7 stabilizes dimerization of the TSC1 C-terminal coiled coil region." J Mol Cell Biol.

Garami, A., F. J. T. Zwartkruis, T. Nobukuni, M. Joaquin, M. Roccio, H. Stocker, S. C. Kozma, E. Hafen, J. L. Bos and G. Thomas (2003). "Insulin Activation of Rheb, a Mediator of mTOR/S6K/4E-BP Signaling, Is Inhibited by TSC1 and 2." Molecular Cell **11**(6): 1457-1466.

Goncharova, E. A., D. A. Goncharov, A. Eszterhas, D. S. Hunter, M. K. Glassberg, R. S. Yeung, C. L. Walker, D. Noonan, D. J. Kwiatkowski, M. M. Chou, R. A. Panettieri, Jr. and V. P. Krymskaya (2002). "Tuberin regulates p70 S6 kinase activation and ribosomal protein S6 phosphorylation. A role for the TSC2 tumor suppressor gene in pulmonary lymphangioleiomyomatosis (LAM)." J Biol Chem **277**(34): 30958-30967.

Graham, F. L., J. Smiley, W. C. Russell and R. Nairn (1977). "Characteristics of a human cell line transformed by DNA from human adenovirus type 5." J Gen Virol **36**(1): 59-74.

Inoki, K., Y. Li, T. Xu and K. L. Guan (2003). "Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling." Genes Dev **17**(15): 1829-1834.

Inoue, H., T. Uyama, T. Suzuki, M. Kazami, O. Hino, T. Kobayashi, K. Kobayashi, T. Tadokoro and Y. Yamamoto (2010). "Phosphorylated hamartin-Hsp70 complex regulates apoptosis via mitochondrial localization." Biochem Biophys Res Commun **391**(1): 1148-1153.

Johnson, C. E., E. A. Dunlop, S. Seifan, H. D. McCann, T. Hay, G. J. Parfitt, A. T. Jones, P. J. Giles, M. H. Shen, J. R. Sampson, R. J. Errington, D. M. Davies and A. R. Tee (2018). "Loss of tuberous sclerosis complex 2 sensitizes tumors to nelfinavir-bortezomib therapy to intensify endoplasmic reticulum stress-induced cell death." Oncogene **37**(45): 5913-5925.

Kwiatkowski, D. J. and B. D. Manning (2005). "Tuberous sclerosis: a GAP at the crossroads of multiple signaling pathways." Hum Mol Genet **14 Spec No. 2**: R251-258.

Lilyestrom, W., M. G. Klein, R. G. Zhang, A. Joachimiak and X. J. S. Chen (2006). "Crystal structure of SV40 large T-antigen bound to p53: interplay between a viral oncoprotein and a cellular tumor suppressor." Genes & Development **20**(17): 2373-2382.

Lin, Y. C., M. Boone, L. Meuris, I. Lemmens, N. Van Roy, A. Soete, J. Reumers, M. Moisse, S. Plaisance, R. Drmanac, J. Chen, F. Speleman, D. Lambrechts, Y. Van de Peer, J. Tavernier and N. Callewaert (2014). "Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations." Nat Commun **5**: 4767.

Love, M. I., W. Huber and S. Anders (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biol **15**(12): 550.

Manning, B. D., A. R. Tee, M. N. Logsdon, J. Blenis and L. C. Cantley (2002). "Identification of the Tuberous Sclerosis Complex-2 Tumor Suppressor Gene Product Tuberin as a Target of the Phosphoinositide 3-Kinase/Akt Pathway." Molecular Cell **10**(1): 151-162.

Martina, J. A., Y. Chen, M. Gucek and R. Puertollano (2012). "MTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB." Autophagy **8**(6): 903-914.

Nellist, M., O. Sancak, M. A. Goedbloed, C. Rohe, D. van Netten, K. Mayer, A. Tucker-Williams, A. M. van den Ouweland and D. J. Halley (2005). "Distinct effects of single amino-acid changes to tuberin on the function of the tuberin-hamartin complex." Eur J Hum Genet **13**(1): 59-68.

Nellist, M., M. A. van Slegtenhorst, M. Goedbloed, A. M. van den Ouweland, D. J. Halley and P. van der Sluijs (1999). "Characterization of the cytosolic tuberin-hamartin complex. Tuberin is a cytosolic chaperone for hamartin." J Biol Chem **274**(50): 35647-35652.

Nicholson, K. M. and N. G. Anderson (2002). "The protein kinase B/Akt signalling pathway in human malignancy." Cellular Signalling **14**(5): 381-395.

Niida, Y., N. Lawrence-Smith, A. Banwell, E. Hammer, J. Lewis, R. L. Beauchamp, K. Sims, V. Ramesh and L. Ozelius (1999). "Analysis of bothTSC1 andTSC2 for germline mutations in 126 unrelated patients with tuberous sclerosis." Human Mutation **14**(5): 412-422.

Nobukuni, T., M. Joaquin, M. Roccio, S. G. Dann, S. Y. Kim, P. Gulati, M. P. Byfield, J. M. Backer, F. Natt, J. L. Bos, F. J. Zwartkruis and G. Thomas (2005). "Amino acids mediate mTOR/raptor signaling through activation of class 3 phosphatidylinositol 3OH-kinase." Proc Natl Acad Sci U S A **102**(40): 14238-14243.

Panaretou, B., C. Prodromou, S. M. Roe, R. O'Brien, J. E. Ladbury, P. W. Piper and L. H. Pearl (1998). "ATP binding and hydrolysis are essential to the function of the Hsp90 molecular chaperone in vivo." EMBO J **17**(16): 4829-4836.

Plank, T. L., R. S. Yeung and E. P. Henske (1998). "Hamartin, the product of the tuberous sclerosis 1 (TSC1) gene, interacts with tuberin and appears to be localized to cytoplasmic vesicles." Cancer Research **58**(21): 4766-4770.

Pollizzi, K., I. Malinowska-Kolodziej, C. Doughty, C. Betz, J. Ma, J. Goto and D. J. Kwiatkowski (2009). "A hypomorphic allele of Tsc2 highlights the role of TSC1/TSC2 in signaling to AKT and models mild human TSC2 alleles." Hum Mol Genet **18**(13): 2378-2387.

Potter, C. J., H. Huang and T. Xu (2001). "Drosophila Tsc1 functions with Tsc2 to antagonize insulin signaling in regulating cell growth, cell proliferation, and organ size." Cell **105**(3): 357-368.

Prabhakar, S., X. Zhang, J. Goto, S. Han, C. Lai, R. Bronson, M. Sena-Esteves, V. Ramesh, A. Stemmer-Rachamimov, D. J. Kwiatkowski and X. O. Breakefield (2015). "Survival benefit and phenotypic improvement by hamartin gene therapy in a tuberous sclerosis mouse brain model." Neurobiol Dis **82**: 22-31.

Qin, J., Z. Wang, M. Hoogeveen-Westerveld, G. Shen, W. Gong, M. Nellist and W. Xu (2016). "Structural Basis of the Interaction between Tuberous Sclerosis Complex 1 (TSC1) and Tre2-Bub2-Cdc16 Domain Family Member 7 (TBC1D7)." J Biol Chem **291**(16): 8591-8601.

Rio, D., S. Clark and R. Tjian (1985). "A mammalian host-vector system that regulates expression and amplification of transfected genes by temperature induction." Science **227**(4682): 23-28.

Sager, R. A., M. R. Woodford and M. Mollapour (2018). "The mTOR Independent Function of Tsc1 and FNIPs." Trends Biochem Sci.

Sager, R. A., M. R. Woodford, O. Shapiro, M. Mollapour and G. Bratslavsky (2018). "Sporadic renal angiomyolipoma in a patient with Birt-Hogg-Dube: chaperones in pathogenesis." Oncotarget **9**(31): 22220-22229.

Santiago Lima, A. J., M. Hoogeveen-Westerveld, A. Nakashima, A. Maat-Kievit, A. van den Ouweland, D. Halley, U. Kikkawa and M. Nellist (2014). "Identification of regions critical for the integrity of the TSC1-TSC2-TBC1D7 complex." PLoS One **9**(4): e93940.

Sarbassov, D. D., D. A. Guertin, S. M. Ali and D. M. Sabatini (2005). "Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex." Science **307**(5712): 1098-1101.

Siroky, B. J., H. Yin, J. T. Babcock, L. Lu, A. R. Hellmann, B. P. Dixon, L. A. Quilliam and J. J. Bissler (2012). "Human TSC-associated renal angiomyolipoma cells are hypersensitive to ER stress." Am J Physiol Renal Physiol **303**(6): F831-844.

Smith, E. M., S. G. Finn, A. R. Tee, G. J. Browne and C. G. Proud (2005). "The tuberous sclerosis protein TSC2 is not required for the regulation of the mammalian target of rapamycin by amino acids and certain cellular stresses." J Biol Chem **280**(19): 18717-18727.

Tee, A. R., D. C. Fingar, B. D. Manning, D. J. Kwiatkowski, L. C. Cantley and J. Blenis (2002). "Tuberous sclerosis complex-1 and -2 gene products function together to inhibit mammalian target of rapamycin (mTOR)-mediated downstream signaling." Proc Natl Acad Sci U S A **99**(21): 13571-13576.

van Slegtenhorst, M., M. Nellist, B. Nagelkerken, J. Cheadle, R. Snell, A. van den Ouweland, A. Reuser, J. Sampson, D. Halley and P. van der Sluijs (1998). "Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products." Human molecular genetics **7**(6): 1053-1057.

van Slegtenhorst, M., M. Nellist, B. Nagelkerken, J. Cheadle, R. Snell, A. van den Ouweland, A. Reuser, J. Sampson, D. Halley and P. van der Sluijs (1998). "Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products." Hum Mol Genet **7**(6): 7.

van Slegtenhorst, M. A., A. Verhoef, A. Tempelaars, L. Bakker, Q. Wang, M. Wessels, R. Bakker, M. Nellist, D. Lindhout, D. Halley and A. van den Ouweland (1999). "Mutational spectrum of the TSC1 gene in a cohort of 225 tuberous sclerosis complex patients: no evidence for genotype-phenotype correlation." J Med Genet **36**(4): 5.

Woodford, M. R., R. A. Sager, E. Marris, D. M. Dunn, A. R. Blanden, R. L. Murphy, N. Rensing, O. Shapiro, B. Panaretou, C. Prodromou, S. N. Loh, D. H. Gutmann, D. Bourboulia, G. Bratslavsky, M. Wong and M. Mollapour (2017). "Tumor suppressor Tsc1 is a new Hsp90 co-chaperone that facilitates folding of kinase and non-kinase clients." EMBO J **36**(24): 3650-3665.

Zhang, H., G. Cicchetti, H. Onda, H. B. Koon, K. Asrican, N. Bajraszewski, F. Vazquez, C. L. Carpenter and D. J. Kwiatkowski (2003). "Loss of Tsc1/Tsc2 activates mTOR and disrupts PI3K-Akt signaling through downregulation of PDGFR." J Clin Invest **112**(8): 1223-1233.

Zhang, Y., X. Gao, L. J. Saucedo, B. Ru, B. A. Edgar and D. Pan (2003). "Rheb is a direct target of the tuberous sclerosis tumour suppressor proteins." Nat Cell Biol **5**(6): 578-581.

VII. Supplementary Tables

Supplementary Table 1: RNA sample information. RQN: RNA quality number. QC: Quality control. All samples passed on RQN and QC.

Customer ID	Material Type	RQN	Entry QC (ng/ul)
293T (DMEM +/+)	total RNA	10	1551.86
1C2 (DMEM +/+)	total RNA	10	1262.16
3H9 (DMEM +/+)	total RNA	10	1582.19
3H9-1B1 (DMEM +/+)	total RNA	10	1197.36
293T (DMEM -/-)	total RNA	10	1138.53
1C2 (DMEM -/-)	total RNA	10	885.78
3H9 (DMEM -/-)	total RNA	10	501.64
3H9-1B1 (DMEM -/-)	total RNA	10	754.1
293T (DMEM +/+)	total RNA	10	981.95
3C2 (DMEM +/+)	total RNA	10	911.37
3A1 (DMEM +/+)	total RNA	10	1016.13
3H9-1C4 (DMEM +/+)	total RNA	10	773.92
293T (DMEM -/-)	total RNA	10	858.66
3C2 (DMEM -/-)	total RNA	10	621.08
3A1 (DMEM -/-)	total RNA	10	816.27
3H9-1C4 (DMEM -/-)	total RNA	10	960.41
293T (DMEM +/+)	total RNA	10	867.13
3C2 (DMEM +/+)	total RNA	10	768.35
1C2 (DMEM +/+)	total RNA	10	1023.06
3A1 (DMEM +/+)	total RNA	10	547.62
3H9 (DMEM +/+)	total RNA	10	1357.19
3H9-1C4 (DMEM +/+)	total RNA	10	466.22
3H9-1B1 (DMEM +/+)	total RNA	10	774.4
293T (DMEM -/-)	total RNA	10	911.17
3C2 (DMEM -/-)	total RNA	10	887.26
1C2 (DMEM -/-)	total RNA	10	871.46
3A1 (DMEM -/-)	total RNA	10	1108.38
3H9 (DMEM -/-)	total RNA	10	1198.26
3H9-1C4 (DMEM -/-)	total RNA	10	1134.05
3H9-1B1 (DMEM -/-)	total RNA	10	825.42

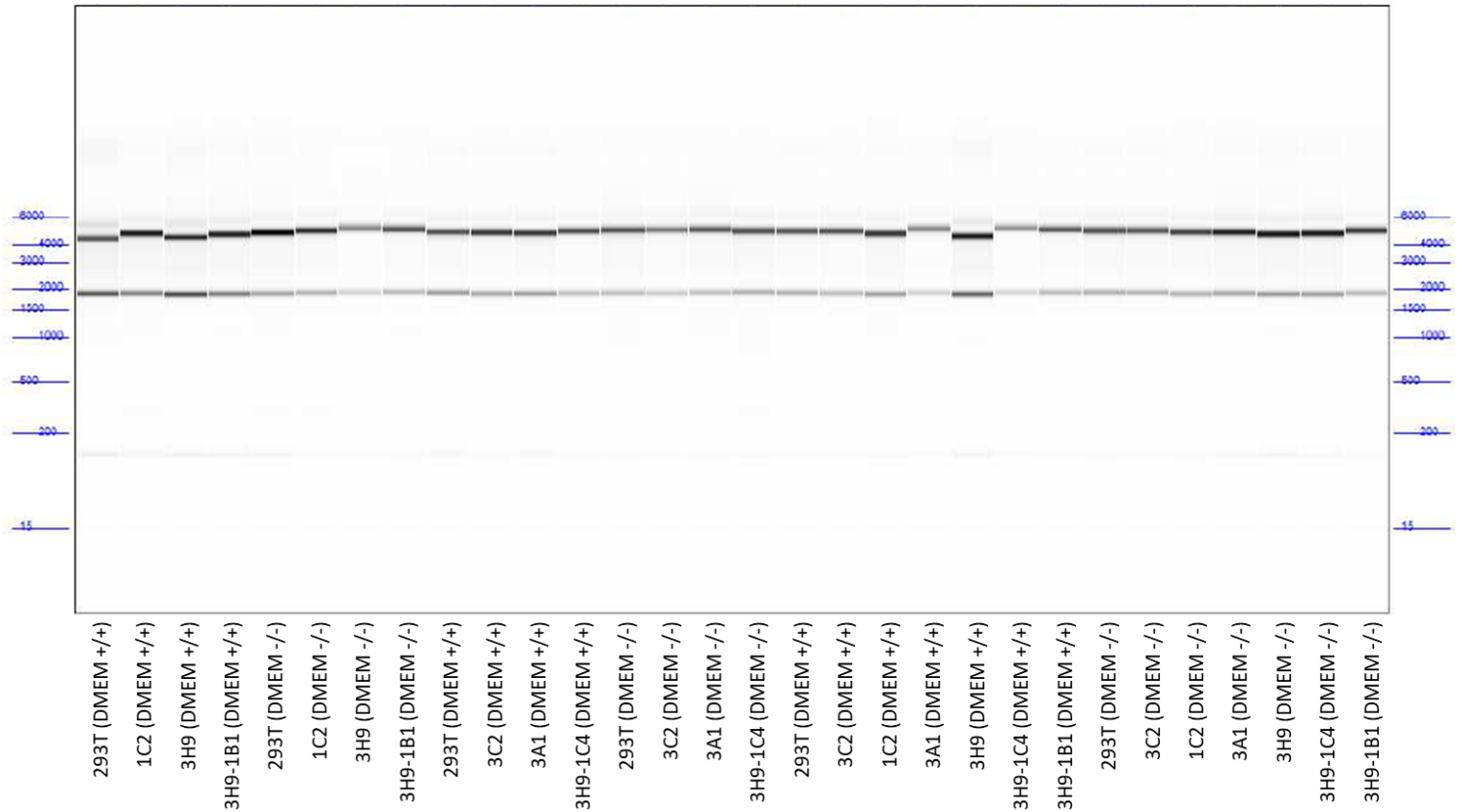
Supplementary Table 2: Library construction and Molarity.

Customer ID	Molarity (nM)
293T (DMEM +/+)	5.33
1C2 (DMEM +/+)	3.75
3H9 (DMEM +/+)	4.40
3H9-1B1 (DMEM +/+)	5.00
293T (DMEM -/-)	7.30
1C2 (DMEM -/-)	7.57
3H9 (DMEM -/-)	8.10
3H9-1B1 (DMEM -/-)	4.28
293T (DMEM +/+)	6.82
3C2 (DMEM +/+)	7.05
3A1 (DMEM +/+)	10.40
3H9-1C4 (DMEM +/+)	4.44
293T (DMEM -/-)	6.89
3C2 (DMEM -/-)	6.56
3A1 (DMEM -/-)	7.13
3H9-1C4 (DMEM -/-)	9.10
293T (DMEM +/+)	6.62
3C2 (DMEM +/+)	5.05
1C2 (DMEM +/+)	7.48
3A1 (DMEM +/+)	6.44
3H9 (DMEM +/+)	7.08
3H9-1C4 (DMEM +/+)	9.79
3H9-1B1 (DMEM +/+)	4.23
293T (DMEM -/-)	5.32
3C2 (DMEM -/-)	4.71
1C2 (DMEM -/-)	5.01
3A1 (DMEM -/-)	6.72
3H9 (DMEM -/-)	3.68
3H9-1C4 (DMEM -/-)	3.56
3H9-1B1 (DMEM -/-)	3.13

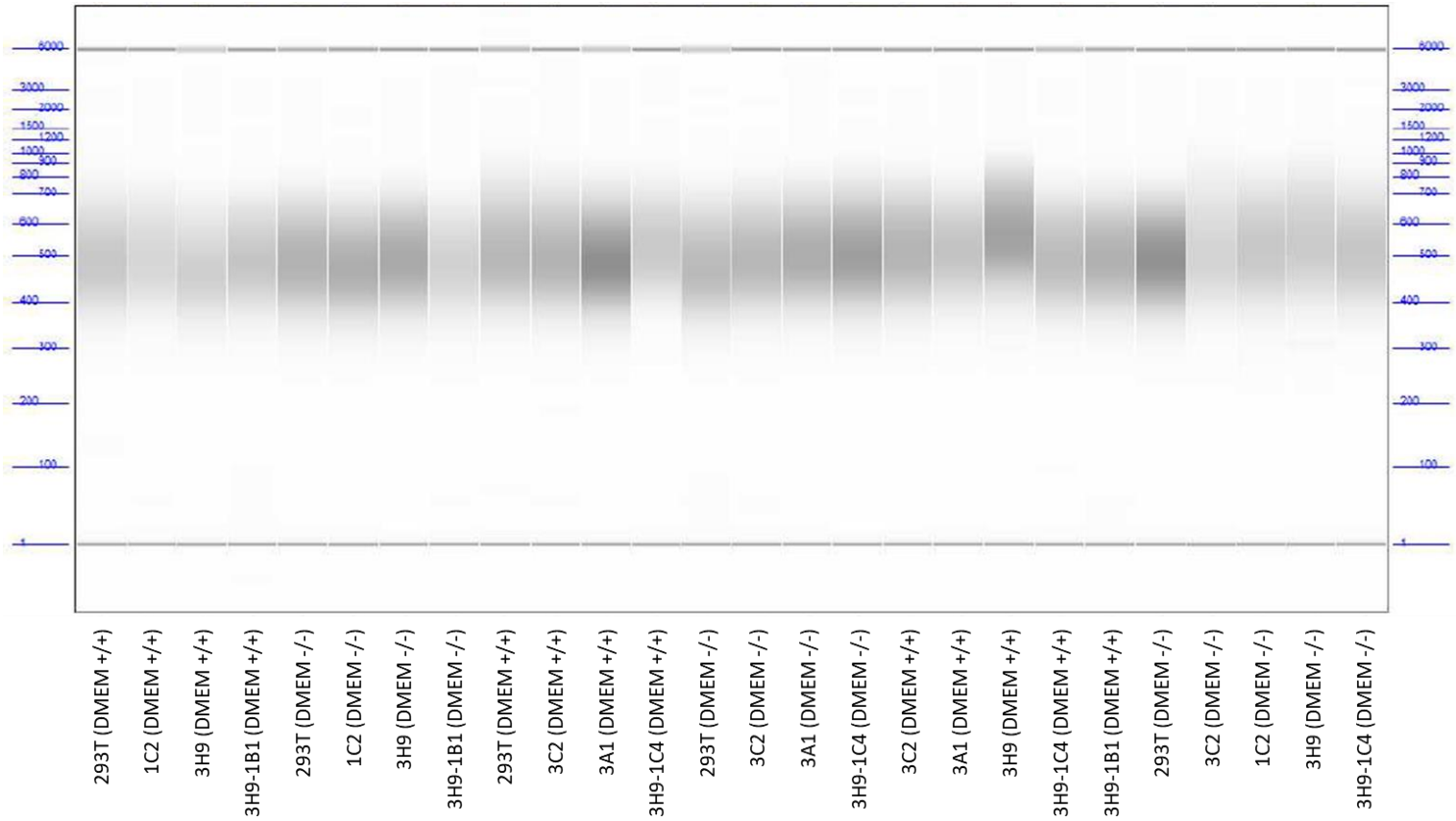
Supplementary Table 3: Run yields, clusters and percentage of Q30 per sample.

Sample ID	Yield (Mb)	Clusters	% >=Q30
293T (DMEM +/+)	4,359	14,435,335	90.68
1C2 (DMEM +/+)	3,543	11,731,379	90.14
3H9 (DMEM +/+)	4,518	14,957,866	91.35
3H9-1B1 (DMEM +/+)	4,192	13,881,889	90.68
293T (DMEM -/-)	4,357	14,428,154	90.83
1C2 (DMEM -/-)	4,204	13,918,424	90.96
3H9 (DMEM -/-)	4,281	14,175,446	90.64
3H9-1B1 (DMEM -/-)	3,578	11,847,647	90.72
293T (DMEM +/+)	3,835	12,698,726	90.25
3C2 (DMEM +/+)	3,611	11,954,942	90.54
3A1 (DMEM +/+)	3,984	13,191,862	90.69
3H9-1C4 (DMEM +/+)	3,970	13,146,674	89.93
293T (DMEM -/-)	4,052	13,416,176	90.31
3C2 (DMEM -/-)	3,840	12,715,334	90.37
3A1 (DMEM -/-)	3,554	11,770,421	90.06
3H9-1C4 (DMEM -/-)	5,331	17,651,436	90.11
293T (DMEM +/+)	6,269	20,757,887	91.08
3C2 (DMEM +/+)	4,114	13,620,545	90.06
1C2 (DMEM +/+)	6,498	21,514,943	90.54
3A1 (DMEM +/+)	4,255	14,088,555	90.05
3H9 (DMEM +/+)	3,561	11,791,372	89.18
3H9-1C4 (DMEM +/+)	3,894	12,895,789	90.23
3H9-1B1 (DMEM +/+)	3,529	11,685,430	89.36
293T (DMEM -/-)	3,711	12,287,753	90.18
3C2 (DMEM -/-)	7,429	24,598,652	89.78
1C2 (DMEM -/-)	3,477	11,513,170	89.65
3A1 (DMEM -/-)	5,368	17,775,044	89.78
3H9 (DMEM -/-)	4,043	13,384,615	88.99
3H9-1C4 (DMEM -/-)	3,959	13,110,984	89.72
3H9-1B1 (DMEM -/-)	3,940	13,045,184	89.55

VIII. Supplementary Figures



Supplementary Figure 1: Gel image of total RNA fragment analysis. Separation at 8.0 kV, for 40.0 min.



Supplementary Figure 2: Gel image of NGS Fragment 1-6000 bp analysis. Separation at 6.0 kV, for 50.0 min.