UNIVERSITÀ DEGLI STUDI DI MILANO

CORSO DI DOTTORATO IN RICERCA CLINICA

XXXIV CICLO

SSD: MED/03 e MED/39

TUBEROUS SCLEROSIS COMPLEX: IDENTIFICATION OF THE GENETIC CAUSE IN PATIENTS WITH NO MUTATION DETECTED, AND ANALYSIS OF MOSAICISM

Tesi di Dottorato di:
Angela Peron
matricola R12231

Tutor: Prof.ssa Maria Paola Canevini

Coordinatori del Dottorato: Prof. Emilio Berti e Prof. Massimo Del Fabbro

Anno Accademico 2020/2021

"If I have seen further,
it is by standing on the shoulders of giants"
Isaac Newton

Dedicated to all my mentors and to my family

ABSTRACT

Tuberous sclerosis complex (TSC) is an autosomal dominant multisystem disorder characterized by development of hamartomas, intellectual disability, seizures and autism. TSC is caused by inactivating mutations in either the *TSC1* or the *TSC2* genes. A pathogenic variant is not identified in up to 10% of the patients with a clinical diagnosis of TSC despite full molecular assessment. These individuals are referred to as NMI (No Mutation Identified), and it is not clear if they should be monitored and/or treated in similar fashion to those with known etiology of TSC.

To identify the genetic cause of TSC in these patients, we selected ten individuals with a definite clinical diagnosis of TSC and NMI and performed a pilot study. Three different technologies were used and their results were compared: chromosomal microarray, trio whole genome sequencing, and targeted deep-coverage next generation sequencing of *TSC1* and *TSC2*. We identified mosaic variants in *TSC1*/*TSC2* in six patients. No variants in other genes were detected in the remaining individuals.

Based on the results of the pilot study and on recent literature, we then performed targeted deep-coverage *TSC1*/*TSC2* sequencing on 200 affected individuals using peripheral blood DNA and saliva or other tissue samples where available. We identified 24 patients with mosaic pathogenic variants in *TSC1* (n=2) or *TSC2* (n=22), defining a rate of mosaicism of 12%. Mosaic variant allele frequency (VAF) ranged from 2% to 32% in blood and from 2% to 35% in saliva. Most affected individuals had low-level mosaicism (VAF ≤10%). We performed an extensive analysis of the phenotype, and show that individuals with a mosaic variant in *TSC1*/*TSC2* often display normal cognitive functioning, although other TSC-associated neuropsychiatric disorders (TAND) are seen in 62% of the patients. Cortical tubers are invariably present and seizures are diagnosed in 54% of the cohort, but infantile spasms are rare. The number of cutaneous manifestations in these patients is often insufficient to meet diagnostic criteria, except for facial angiofibromas. We observed a high frequency of pulmonary and renal manifestations in our mosaic cohort, which are as severe as those seen in the general TSC population. None of the patients who have reproduced transmitted the variant to their offspring.

In conclusion, our study shows that genome sequencing fails to identify rare variants in new genes related to TSC, and a third gene is therefore unlikely to exist. We demonstrated that at least one out of 10 patients with a clinical diagnosis of TSC carries a mosaic pathogenic variant in *TSC1* or *TSC2*. We also showed that individuals with mosaic variants have a distinctive phenotypic severity, with important implications for surveillance.

RIASSUNTO

La sclerosi tuberosa è una malattia multisistemica autosomica dominante caratterizzata dalla presenza di amartomi, disabilità intellettiva, epilessia e autismo. E' causata da mutazioni inattivanti nel gene *TSC1* o *TSC2*. In circa il 10% dei pazienti con una diagnosi clinica certa non è possibile identificare una variante patogenetica nonostante l'utilizzo dei test genetici appropriati, e non è chiaro se questi pazienti debbano essere monitorati e/o trattati allo stesso modo di coloro nei quali è stata identificata una mutazione.

Al fine di individuare la causa genetica della sclerosi tuberosa in questi pazienti, abbiamo selezionato dieci individui con una diagnosi clinica certa e senza mutazione identificata, e abbiamo eseguito uno studio pilota. Abbiamo applicato tre differenti tecnologie genetiche e ne abbiamo confrontato i risultati: array-CGH, sequenziamento del genoma del trio (probando e genitori), e sequenziamento di nuova generazione ad alta profondità di lettura targettato per i geni *TSC1* e *TSC2*. Le analisi hanno permesso di identificare mutazioni a mosaico in *TSC1*/*TSC2* in sei pazienti. Non sono state identificate varianti in ulteriori geni nei rimanenti pazienti.

Sulla base dei risultati dello studio pilota e dei dati recentemente pubblicati nella letteratura scientifica, abbiamo quindi eseguito il sequenziamento di nuova generazione ad alta profondità di lettura targettato per i geni *TSC1*/*TSC2* in 200 pazienti con sclerosi tuberosa, su DNA estratto da sangue periferico e saliva o campioni di un ulteriore tessuto se disponibili. Abbiamo identificato 24 pazienti con mutazioni a mosaico nel gene *TSC1* (2 pz) o *TSC2* (22 pz), definendo un tasso di mosaicismo pari al 12%. La frequenza allelica della variante a mosaico era del 2-32% nel sangue periferico e del 2-35% nella saliva, e la maggior parte dei pazienti presentava mosaicismo a basso livello (frequenza allelica ≤10%). L'analisi del fenotipo ha permesso di dimostrare che i pazienti con mutazioni a mosaico nei geni *TSC1*/*TSC2* presentavano frequentemente un livello cognitivo nella norma. Tuttavia, altre problematiche neuropsichiatriche sono state riscontrate nel 62% degli affetti. I tuberi corticali erano presenti in quasi tutti i pazienti e l'epilessia nel 54%, ma gli spasmi infantili erano rari. Le manifestazioni cutanee spesso non raggiungevano il numero sufficiente per soddisfare i criteri diagnostici. Al contrario, abbiamo osservato un'elevata frequenza di manifestazioni polmonari e renali gravi nella coorte di pazienti con mutazioni a mosaico. Nessuno dei pazienti con mutazioni a mosaico ha trasmesso la mutazione alla propria prole.

Il presente studio permette di concludere che è improbabile che esista un terzo gene responsabile della sclerosi tuberosa, e dimostra che almeno un paziente su dieci con una diagnosi clinica di questa malattia presenta una mutazione a mosaico in *TSC1* o *TSC2*. Infine, i risultati dimostrano che i pazienti con mutazioni a mosaico presentano un fenotipo caratteristico, con importanti implicazioni per la presa in carico ed il follow-up personalizzati.

INDEX

1 INTRODUCTION

1.1 TUBEROUS SCLEROSIS COMPLEX

Tuberous Sclerosis Complex (TSC, OMIM #191100, #613254) is an autosomal dominant neurocutaneous condition, caused by loss-of-function variants in either the *TSC1* (OMIM #605284) or *TSC2* (OMIM #191092) gene (Northrup et al., 2011). Its incidence is estimated to be 1:6,000-1:10,000 new live births, and the prevalence is thought to be 1:20,000 (O'Callaghan et al., 1998; Henske et al., 2016; Hong et al., 2016).

TSC is a multisystem condition characterized by development of benign tumors (hamartomas) in the central nervous system, eyes, heart, lungs, kidneys, and skin. Additional common manifestations are intellectual disability (ID), seizures, and autism spectrum disorder. There is complete penetrance in TSC, but a great clinical intra- and interfamilial variability is observed.

In 1862 German physician Friedrich Daniel von Recklinghausen recognized for the first time cardiac tumors and brain calcifications in a newborn who died soon after birth (von Recklinghausen, 1862). However, TSC was initially named after French neurologist Désiré-Magloire Bourneville, who identified a girl with infantile spasms, epilepsy, skin findings, and brain tumors (Bourneville, 1880). The term "tuberous sclerosis complex" was coined by Moolten in 1942 to better describe the multisystemic nature of the disease (Moolten, 1942), but the causative genes were discovered only in the 1990's (European Chromosome 16 Tuberous Sclerosis Consortium, 1993; van Slegtenhorst et al., 1997).

1.2 DIAGNOSTIC CRITERIA, PHENOTYPE, AND TREATMENT

Diagnostic criteria for TSC were established in 1998 (Roach et al., 1998) and updated in 2013 during the Tuberous Sclerosis Consensus Conference to include the genetic criteria (Northrup e Krueger, 2013). A new update was published in September 2021 together with recommendations on surveillance and management, and includes minor changes compared to the 2013 criteria (Northrup et al., 2021). The TSC diagnostic criteria include the most specific clinical manifestations of TSC and are summarized in Table 1. A definite diagnosis of TSC is given when 2 major features or 1 major feature with 2 minor features are present. A patient has a possible diagnosis of TSC when she/he has either 1 major feature or ≥2 minor features.

A. Genetic diagnostic criteria

A pathogenic variant in *TSC1* or *TSC2* is diagnostic for TSC.
Most TSC-causing variants are sequence variants that clearly prevent TSC1 or TSC2 protein production. Some variants compatible with protein production (e.g. some missense changes) are well established as disease-causing; other variant types should be considered with caution.

B. Clinical diagnostic criteria

Major criteria

1. Hypomelanotic macules (≥3, at least 5 mm diameter)
2. Angiofibroma (≥3) or fibrous cephalic plaque
3. Ungual fibromas (≥2)
4. Shagreen patch
5. Multiple retinal hamartomas
6. Multiple cortical tubers and/or radial migration lines
7. Subependymal nodule (≥2)
8. Subependymal giant cell astrocytoma (SEGA)
9. Cardiac rhabdomyoma
10. Lymphangioleiomyomatosis (LAM)*
11. Angiomyolipomas (≥2)*

* a combination of LAM and angiomyolipomas without other features does not meet criteria for a definite diagnosis

Minor criteria

1. "Confetti" skin lesions
2. Dental enamel pits (≥3)
3. Intraoral fibromas (≥2)
4. Retinal achromic patch
5. Multiple renal cysts
6. Non-renal hamartomas
7. Sclerotic bone lesions

Table 1: TSC diagnostic criteria. Definite TSC: 2 major features or 1 major feature with 2 minor features. Possible TSC: either 1 major feature or ≥2 minor features. [adapted from Northrup et al., 2021]

Hypomelanotic macules are the most common cutaneous findings. They are seen in >90% of affected individuals, and are usually present at birth or in the first months of life, thus representing one of the first findings that lead to the diagnosis (Northrup and Krueger, 2013). In some patients, hypomelanotic macules are difficult to detect on physical exam and can only be seen using the Wood's lamp (Figure 1).

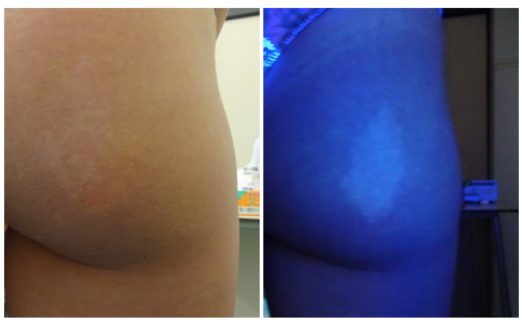


Figure 1: Example of a hypomelanotic macule difficult to detect (left).
The same macule seen with Wood's lamp (right). [re-use not permitted]

Facial angiofibromas are common (75% of patients) and usually occur between age 2 and 5 years (Jozwiak et al., 2000). The other cutaneous and oral manifestations are much less common and highly age-dependent. Figure 2 shows examples of the cutaneous manifestations in individuals with TSC from the TSC Clinic at San Paolo University Hospital in Milan (Italy), and highlights the clinical variability with regard to severity.

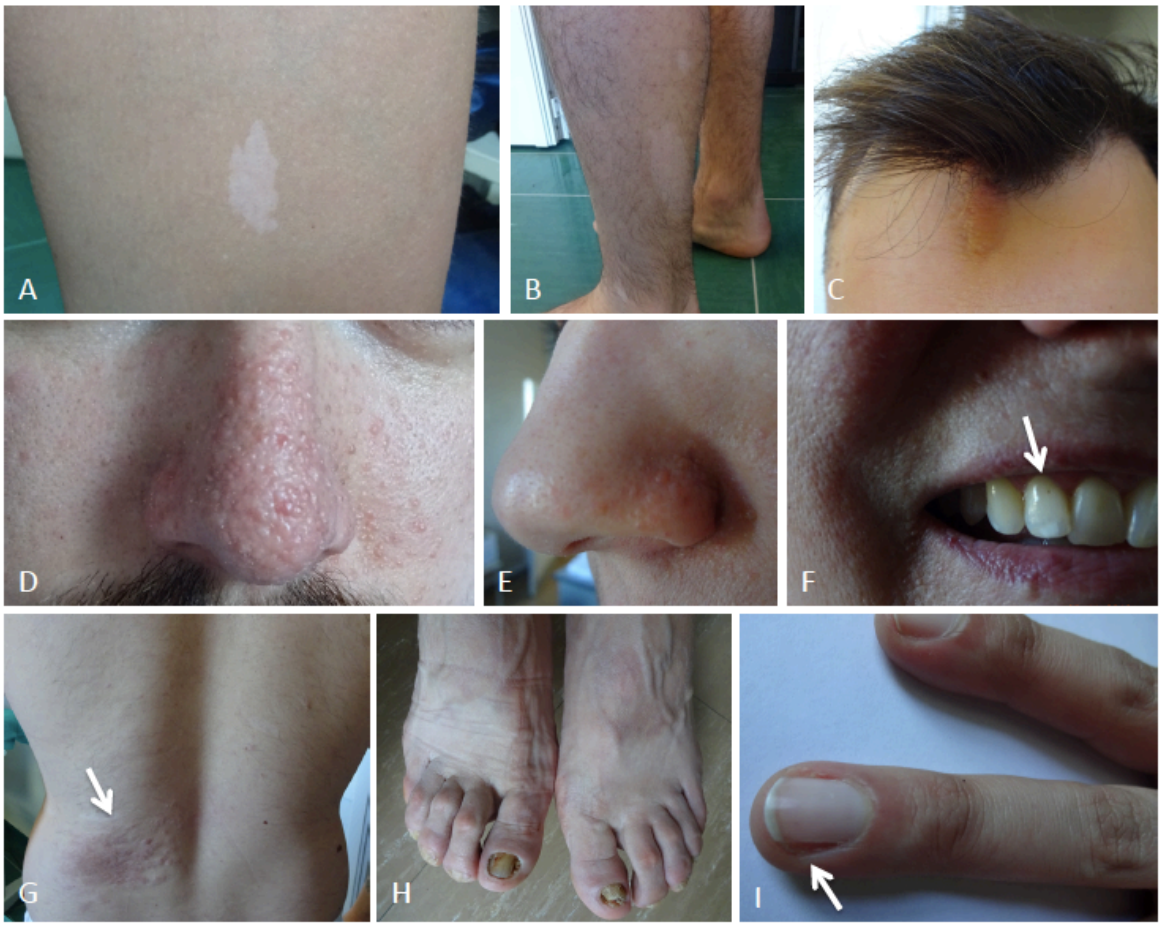


Figure 2: Skin and oral findings in individuals with TSC from our TSC clinic. A and B: hypomelanotic macules. C: fibrous cephalic plaque. D: diffuse facial angiofibromas. E: minimal facial angiofibromas on the nose. F: dental enamel pit (white arrow). G: shagreen patch in the lumbar region (white arrow). H: bilateral ungual fibromas. I: single periungual fibroma (white arrow). [re-use not permitted]

Cortical and sub-cortical tubers are glioneuronal hamartomas located in the grey matter and adjacent white matter. They typically appear as hyperintense lesions on T2-weighted brain MRI scans, and are invariably present in almost all affected individuals. Subependymal nodules (SENs) are asymptomatic areas of heterotopia usually located in the lateral ventricles and present in >80% of patients. When SENs are located near the foramen of Monro, have a size of more than 1 cm in any direction or show serial growth on consecutive imaging regardless of size, they are called

subependymal giant cell astrocytomas (SEGAs) (Roth et al., 2012). SEGAs can lead to hydrocephalus and cause life-threatening consequences.

The eye findings consist of retinal hamartomas and retinal achromic path, which are the ocular equivalent of cortical tubers and hypomelanotic macules, respectively. They are seen in 30-50% of affected individuals and are usually asymptomatic, but are extremely useful for diagnosing TSC (Aronow et al., 2012).

Cardiac rhabdomyomas are benign tumors highly specific to TSC. They can be detectable in utero from the 20th gestational week. In fact, when a cardiac rhabdomyoma is detected on prenatal ultrasound in the general population, the likelihood that the fetus has TSC is 75%–80% (Harding et al., 1990). They are usually asymptomatic and tend to regress spontaneously with age, but in 2% of the patients cardiac rhabdomyomas can cause arrhythmia (Black et al., 1998).

Lymphangioleiomyomatosis (LAM) is a multiple cystic lung disease caused by abnormal growth of smooth muscle cells (Gupta and Henske, 2018). LAM affects almost exclusively adult women (up to 80% by age 40 years). It can be asymptomatic, but can also progress, and represents one of the causes of death in females with TSC. Multifocal micronodular pneumocyte hyperplasia (MMPH) is seen in 40-60% of patients of both sexes and is benign (Gupta and Henske, 2018).

Angiomyolipomas are benign tumors of the kidney composed of vessels, muscle, and fat. They are present in up to 80% of patients and are highly age-dependent. The major risks associated with renal angiomyolipomas is severe bleeding and loss of renal function (Bissler and Kingswood, 2018).

In addition to the clinical features mentioned in the diagnostic criteria, individuals with TSC can exhibit other neurologic and neuropsychiatric manifestations.

Epilepsy is frequently seen (60-80% of patients), and is refractory to treatment in 40-60% of affected individuals. Infantile spams and focal-onset seizures are the most common types of seizures, and are often the first sign leading to further evaluation and diagnosis of TSC in infants (Nabbout et al., 2018). Early-onset seizures, especially if not controlled by anti-seizure medications, are associated with poor neurological outcome. Developmental delay (DD) and ID are seen in half of the patients with TSC. Up to a third of the patients have severe ID, and those with normal cognitive functioning often have a lower IQ than the general population (de Vries et al., 2007). Autism spectrum disorder is more frequent in TSC patients than in the general population (Vignoli et al., 2015; Jeste et al., 2016). Additional neuropsychiatric manifestations are observed in up to 90% of patients and are grouped under the umbrella term Tuberous sclerosis-Associated Neuropsychiatric Disorders (TAND, de Vries et al., 2015).

Sclerotic bone lesions are observed in some patients and should not be misdiagnosed with metastases (Northrup et al., 2021). As a matter of fact, malignancies in TSC are rare, except for renal carcinomas (Peron et al., 2018; Sauter et al., 2021).

For a more detailed description of the clinical manifestations please refer to the articles in the special issue on TSC in the American Journal of Medical Genetics, part C: https://onlinelibrary.wiley.com/toc/15524876/2018/178/3

It must be noted that the clinical manifestations associated with TSC are highly age-dependent, as shown in Figure 3.

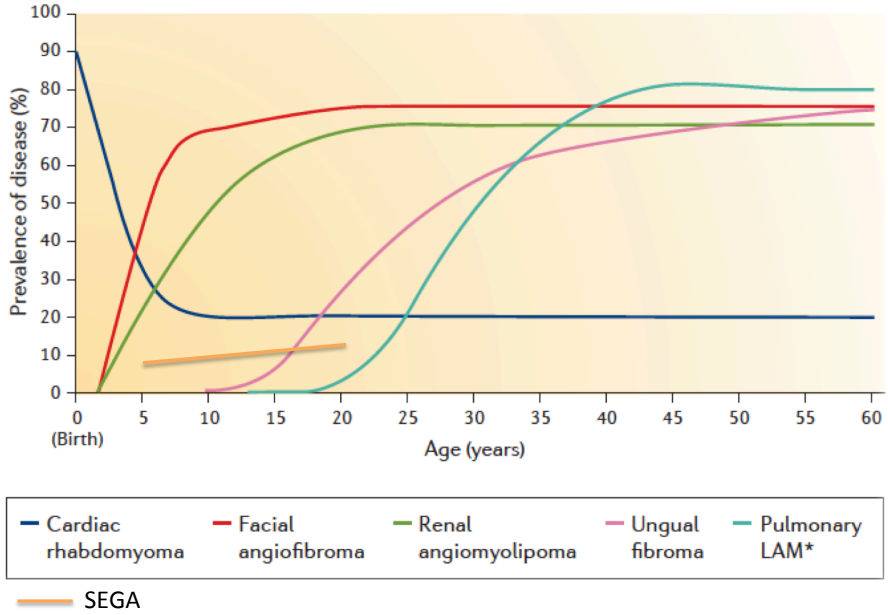


Figure 3: Age-dependency of the main clinical manifestations of TSC. [Modified from Henske et al., 2016]

Surveillance and management should follow the international recommendations (Northrup et al., 2021). With regard to therapies, mTOR inhibitors (Everolimus and Sirolimus) have shown to reduce tumor size of certain manifestations. They have been approved by the Food and Drug Administration (FDA) and by the European Medical Agency (EMA) for the treatment of SEGAs in pediatric and adult patients, LAM and renal angiomyolipomas in adults, and as an adjunctive treatment for patients aged ≥ 2 years with refractory focal-onset seizures. Treatment for the other clinical manifestations is symptomatic.

1.3 GENETIC ASPECTS

1.3.1 THE *TSC1* AND *TSC2* GENES

TSC is caused by heterozygous pathogenic loss-of-function variants in *TSC1* (chr. 9q34.13) or *TSC2* (chr. 16p13.3). *TSC1* and *TSC2* consist of 23 and 42 exons, respectively (Figure 4).

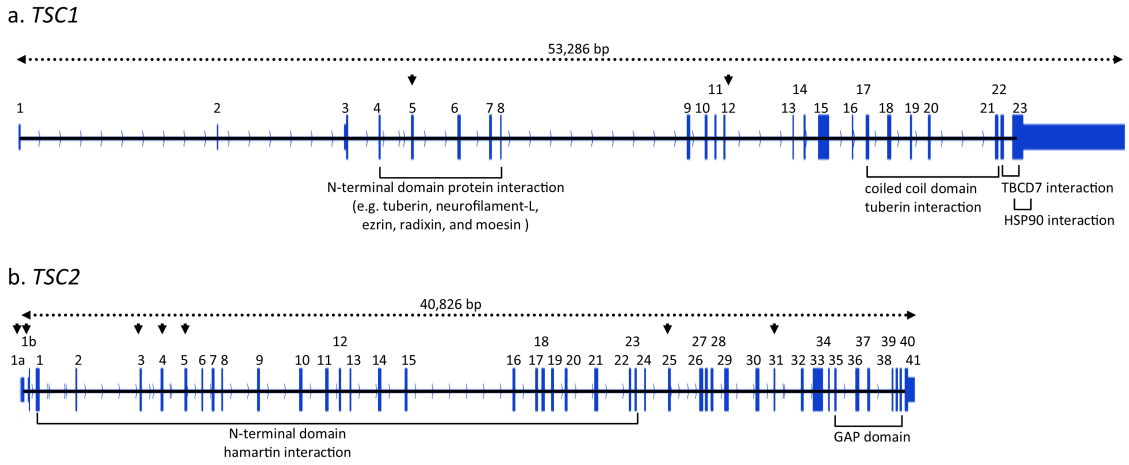


Figure 4: The *TSC1* and *TSC2* genes. [Peron et al., 2018]

At the time of writing, more than 4,000 variants in *TSC1* (of which almost 1,000 unique) and more than 12,000 variants in *TSC2* (of which almost 4,000 unique) have been identified, with only few mutational hotspots as per the Leiden Open Variation Database (http://chromium.lovd.nl/LOVD2/TSC/home.php). Most *TSC1* variants are nonsense and small insertions and deletions (indels), although a few splice and missense variants have been reported (Hoogeveen-Westerveld et al., 2012). The mutational spectrum of *TSC2* comprises all types of variants mentioned above and large genomic deletions, including the *TSC2-PKD1* contiguous gene deletion syndrome (Northrup et al., 2018).

Two-thirds of the patients have pathogenic variants in *TSC2*, whereas 1/3 of affected individuals have pathogenic variants in *TSC1* (Au et al., 2007).

1.3.2 PATHOGENESIS: THE mTOR PATHWAY

TSC1 and *TSC2* encode for hamartin and tuberin, respectively. Hamartin and tuberin interact with each other through their hamartin and tuberin interaction domains and with the protein product of *TBC1D7* to form the TSC complex (Peron et al., 2018). The TSC complex negatively regulates the mechanistic (or mammalian) target of rapamycin (mTOR) signaling pathway (Dibble et al., 2012).

Under physiological conditions, the TSC complex converts RHEB-GTP to the inactive GDP-bound form, and blocks the activation of mTOR. On the contrary, in TSC, loss-of-function pathogenic variants cause RHEB to be bound to GTP (the active form), thus leading to hyperactivation of the mTOR pathway (Figure 5). Hyperactivation of the pathway leads to constitutive deregulation of protein synthesis and consequent cell growth (Henske et al., 2016). The development of hamartomas in TSC requires a second somatic hit (e.g. loss of heterozygosity) and follows Knudson's two-hit hypothesis.

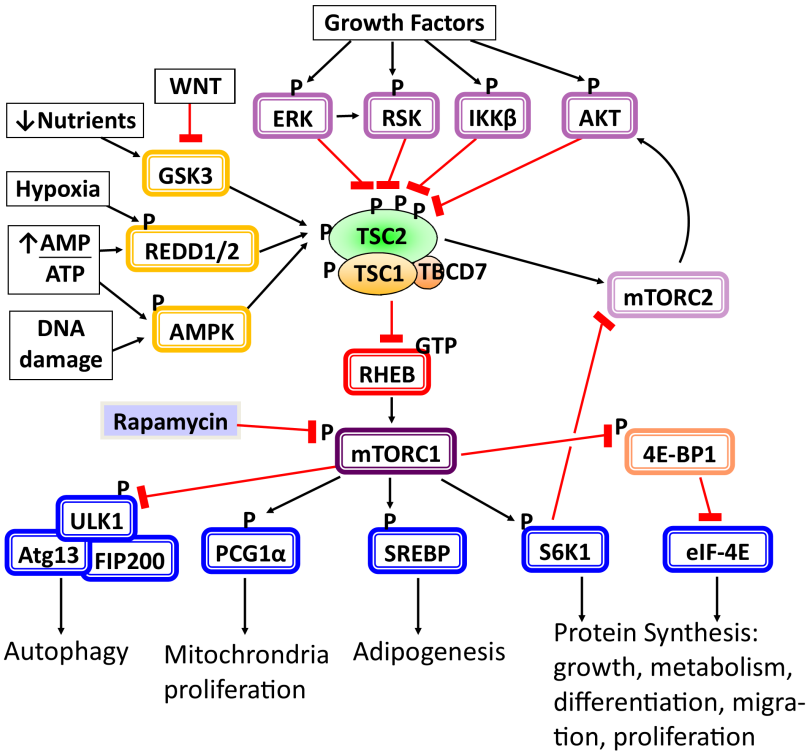


Figure 5: the mTOR pathway. [Peron et al., 2018]

1.3.3 MOLECULAR TESTING

Molecular diagnosis was introduced as an independent diagnostic criterion in 2013 (Northrup and Krueger, 2013), highlighting the importance of a positive genetic test in facilitating the diagnosis especially in individuals who would otherwise remain undiagnosed, such as infants or individuals with insufficient clinical manifestations to meet the clinical criteria. It is important to note that a negative genetic test result does not exclude a TSC diagnosis, and only means that a causative variant could not be identified.

The molecular diagnostic assessment comprises analyses for the detection of point mutations in *TSC1* and *TSC2* and tests for the detection of large deletions/duplications in the same genes. Historically, analysis of the coding regions and intron/exons boundaries of *TSC1* and *TSC2* was

carried out using Denaturing High Performance Liquid Chromatography (DHPLC) and Sanger sequencing, and del/dup analysis was done through multiplex ligation-dependent probe amplification (MLPA) (Peron et al., 2018). The two techniques together had a diagnostic yield of 75-90%, leaving 10-25% of affected individuals with no mutation identified (NMI). The lack of identifiable mutations in *TSC1*/*TSC2* could be due to the presence of somatic mosaicism, mutations in introns or regulatory regions of *TSC1*/*TSC2* (not routinely sequenced), pathogenic variants in a different gene that has not yet been identified, or structural variants.

Concurrently to the starting of the present study, two independent groups demonstrated that targeted *TSC1*/*TSC2* panels through Next Generation Sequencing (NGS) were able to identify pathogenic variants in a higher number of patients, including some individuals with mosaicism and deep intronic variants (Nellist et al., 2015; Tyburczy et al., 2015). As stated in Peron et al. (2018), mosaicism is defined as "the presence of two or more cell populations with different genetic composition within an individual, and it may or may not include the germline (gonadal) cells". Deep intronic variants are single nucleotide variants located in the intronic regions that are not usually covered by conventional testing. Yet, a non-negligible proportion of affected individuals who meet the diagnostic criteria has negative genetic testing even after targeted NGS (Peron et al., 2018). The relatively high number of patients with NMI represents an important challenge and highlights the need of exploring new technologies to try and detect the molecular cause of TSC in these patients.

2 AIMS OF THE STUDY

The aims of the present study were:

1. To investigate the molecular basis of TSC in a systematically phenotyped and homogeneous cohort of TSC patients with no mutation identified on conventional genetic testing.
2. To evaluate the rate of mosaicism in a real-world setting and delineate the phenotype of patients with mosaic *TSC1* or *TSC2* pathogenic variants.

3 MATERIALS AND METHODS

This work consisted of two phases: a pilot study aimed at exploring the molecular cause of TSC in affected individuals with no mutation identified in the *TSC1* or *TSC2* genes, and a full study aimed at better characterizing the results of the pilot study in a large cohort of patients.

3.1 PILOT STUDY

3.1.1 PATIENTS

We included a cohort of individuals with a definite clinical diagnosis of TSC based on the most recent diagnostic criteria at the time of enrollment (Northrup and Krueger, 2013), who had no mutation identified (NMI) on conventional genetic testing. Exclusion criteria were a possible clinical diagnosis and incomplete conventional molecular testing. All the patients had been followed at the TSC clinic of San Paolo Hospital in Milan from January 2001 to February 2015. The TSC clinic is a multidisciplinary center where more than 400 pediatric and adult TSC patients from the Lombardy region and Italy in general have been referred to thus far. It provides both diagnosis and care to these patients, as described in Peron et al. (2018a).

An in-depth analysis of the phenotype of these patients with NMI is reported in Peron et al. (2018b). In brief, we previously found that the phenotype of TSC individuals with NMI is milder than that of patients with a pathogenic variant in *TSC2*, and does not differ from that of patients with a pathogenic variant in *TSC1* except for renal and pulmonary problems, TSC patients with NMI being significantly more affected (Peron et al., 2018b). Conventional genetic tests had been previously performed on DNA extracted from peripheral blood and included the following: DHPLC (Transgenomic, Crewe, UK) and Sanger sequencing to assess single nucleotide variants in the coding regions and intron/exon boundaries of *TSC1* and *TSC2*, and MLPA (MRC-Holland, Probemix TSC2 P046, Probemix TSC1 P124) to investigate large deletions in the two genes (Peron et al., 2018b; Peron et al., 2018c).

From this group of extensively phenotyped patients, we selected the affected individuals for whom trios were available (proband, mother, father) for further clinical evaluation and genetic analyses: the pilot cohort (10 patients). Individuals whose parents were deceased or not available for testing were excluded.

We evaluated all the individuals of the pilot cohort during several encounters, collected their pedigree, family and patients' history, and reviewed all available medical records and imaging data from birth to the time of assessment. The clinical manifestations of each affected individual were recorded during each visit using a standardized examination form, as shown in figure 6. The TSC examination form was adapted from the Neurofibromatosis type I examination form used at the Division of Medical Genetics at the University of Utah in Salt Lake City (US), which was kindly

provided by Prof. David Viskochil. The data collected were then transferred into the database used for TSC research at San Paolo Hospital in Milan (Italy). The clinical manifestations were annotated using the Human Phenotype Ontology (HPO) terms wherever possible, to aid bioinformatic analyses (Kohler et al., 2014).

Physician: _____ **TSC EXAMINATION FORM** Registry #_____

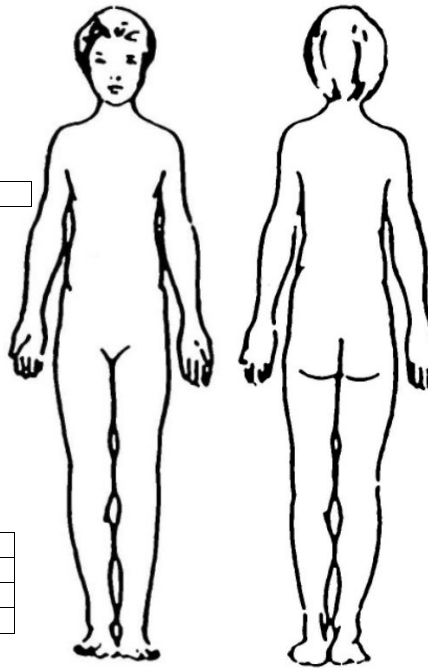
Name (Subject): _____DOB: _____Date of Exam: _____Age at Exam _____

Proband: Y N Name of Proband: _____Relation to Examinee:_____(if other than patient)
Gender: M F Ethnicity: American Indian/Alaska Native Asian Black/African America Native Hawaiian/Pacific Islander White More Than One Race Unknown/ Unreported

TSC Inheritance	Sporadic	Familial: mother father	Unknown Sib(s)_____
Ascertainment	Primary care Dr.	Specialist	Self Referral TSC Support Group
Ht: _____(%)		Wt:_____(%)	
Head Circum: _____(%)		Blood Pressure: /	

Clinical diagnosis of TSC	definite	possible	Age at Dx:

M = Major diagnostic feature (Northrup and Krueger, 2013)
m = minor diagnostic feature
Definite diagnosis: 2 M, or 1 M and ≥2 m, or pathogenic mutation found
Possible diagnosis: 1 M or 2 or more m



DERMATOLOGY

M	Hypomelanotic macules ≥ 5 mm	-	+	
	Number	1 2	3-10	>10
M	Angiofibromas	-	+	
	Number	1-10	11-100	>100
	Fibrous cephalic plaque	-	+	
M	Shagreen patch	-	+	
m	Confetti skin lesions	-	+	
M	Ungual fibromas	-	+	
	Hands, number	Left:	Right:	
	Feet, number	Left:	Right:	

ORAL FINDINGS

m	Dental enamel pits (>3)	-	+
	Number		
m	Intraoral fibromas (≥2)	-	+
	Number		

OCULAR ABNORMALITIES

M	Retinal hamartomas	-	+	
	Number			Bx: Y N
m	Retinal achromic patch	-	+	
	Number			Bx: Y N

BRAIN ABNORMALITIES (age of brain MRI: _____)

M	Cortical tubers	-	+		
	Number			Bx: Y N	
	White matter radial migration lines	-	+		
M	SENs	-	+		
M	SEGA	-	+	Age at onset:	Stable Growing Regress
	Hydrocephalus	-	+		

CARDIAC FINDINGS

M	Cardiac rhabdomyoma	Present	Regressing	No longer present	Never
	Arrythmia	-	+	Treatment:	

RENAL ABNORMALITIES

M	Renal angiomyolipomas	-	+		
	Number			Bx: Y N	
	Size	<3 cm	>3 cm	Stable Growing Regress	
	Hepatic angiomyolipomas	-	+		
m	Renal cysts	-	+	Bx: Y N	
	Renal Cell Carcinoma	-	+	Age at diagnosis:	
	Renal function	normal	insuff		
	Kidney transplant	-	+	Unilateral Bilateral	Age at transplant:
	Embolization	-	+	Age:	

PULMONARY FINDINGS

M	LAM	-	+	Age at onset:
	Symptoms	-	+	Age at onset:
	Pneumothorax	-	+	Number: Age:
	MMPH	-	+	

PSYCHOLOGY/INTELLECTUAL DEVELOPMENT

TAND	- +
Behavior:	ADHD Autism Behavior Problems Comments:
Cognition:	Normal Mild Delay Moderate Delay Severe Delay Unknown IQ (age ___) Full____ Verbal ____ Perform____ Testing: Yes No Type:_____ Edu. Completed:_____
Learning Problems:	None Unknown Present Type: Coordination Prob Y / N Speech Prob Y / N

NEUROLOGY

Seizures:	- +	in the past (stopped at age:_____)		febrile only		
	Spasms: + -					
Age of Seizure Onset:	_____	NA	<1 y	1-3 y	3-18 y	>18 y
Type of seizure at onset:	Spasms	focal seizures	focal+spasms	generalized	unknown	
Current type of seizures:	None	Spasms	focal seizures	focal+spasms	generalized	unknown

ADDITIONAL PROBLEMS

Problem	-	+	unknown	age
GI problems				
Endocrine				
Bone issues				
Neoplasm				
Other:				

MEDICATIONS

Everolimus/Sirolimus	Y/N	dose:	started at age:

GENETICS

Mutation	-	+ TSC1 TSC2	Not tested	unknown
Type of test done	mut	del/dup	NGS	
Type of mutation	deletion	SSV	missense	nonsense
Exact mutation				

IMPRESSIONS/RECOMMENDATIONS:

Figure 6: TSC examination form [download and re-use NOT PERMITTED].

3.1.2 MOLECULAR ANALYSES

For each patient of the pilot cohort and her/his parents, blood and saliva samples were collected by the clinical geneticist during each encounter. In the pilot study we applied the following three different genetic technologies and compared their results and diagnostic yield with respect to the identification of pathogenic variants in patients with TSC and NMI.

Chromosomal microarray (CMA)

Array-based comparative genomic hybridization (array-CGH) analysis on DNA extracted from peripheral blood of each proband was performed at the cytogenetics laboratory of the Istituto Auxologico Italiano (Italy) using the Agilent SurePrint G3 Human CGH Microarray kit 8x60K (Agilent technologies, Santa Clara, CA, USA), with an average resolution of 130 kb. Nucleotide designations were assigned according to the hg19/GRCh37 assembly of the human genome.

Whenever copy number variants (CNVs) were detected, CMA was performed on both parents for segregation analysis of the variant. Detected CNVs were classified based on population frequency (GnomAD-SV, Database of Genomic Variants – DGV), previous reports of overlapping variants, gene content, and segregation analysis. Each CNV was re-evaluated in 2021 using the ACMG/ClinGen recommendations (Riggs et al., 2020).

Whole Genome Sequencing (WGS)

Genomic DNA of the trios (proband, mother, father) was extracted from both peripheral blood and saliva at the genetics laboratory of the Department of Health Sciences at the University of Milan (Italy) and shipped to the University of Utah, Salt Lake City (USA) as part of an international collaboration. Whole-genome libraries were prepared using the KAPA Hyper Prep kit, and trio short-read genome sequencing with a median read depth of 60X was performed on DNA extracted from blood at the NantOmic Facility (Culver City, CA, USA) using the Illumina X Ten platform. We chose this innovative approach, attempted for the first time ever in TSC, because WGS is currently the only method capable of identifying single nucleotide variants and short insertions and deletions (indels) in coding and especially in non-coding regions, and structural variants within the same experiment.

Raw data files were stored in GNomEx (Nix et al., 2010), while variant analysis and filtering was performed using Omicia's (now Fabric Genomics) Opal platform (Coonrod et al., 2013). Using the Variant Annotation Analysis and Search Tool (VAAST, Hu et al., 2013) and a custom filtering strategy (coverage ≥15X, quality score ≥30, MAF ≤5% in 1KG, ExAc and EVS) implemented in the Fabric Genomics Opal framework, we evaluated prioritized genes with genetic damage consistent with both recessive and *de novo* inheritance models in the TSC trios. In addition we used PHEVOR to rerank VAAST gene list according to the HPO terms generated by the clinical team (Kennedy et al., 2014; Kohler et al., 2014). We designed a virtual panel of 55 genes belonging to the mTOR pathway, as we hypothesize that, if a third *TSC* gene exist, it ought to interact with that pathway, and we evaluated impact of *de novo* variants generated by the Rufus reference-free variant calling tool on these genes. Based on our phenotypic assessment of TSC individuals with NMI (Peron et al., 2018b) and two recent studies from Nellist and Tyburczy (2015), we redirected our analyses of WGS data to specifically address mosaicism. We used Freebayes (Garrison&Marth, 2012) for variant calling customized to find *de novo* mosaic variants in the probands (filtering to

variants with at least 3 alternate reads in the proband and none in the parents). We further limited to variants annotated to have a functional impact on protein sequence and those with a maximum allele frequency (AF) in GnomAD of <0.001. Finally, we annotated variants with the Ensembl Variant Effect Predictor (VEP, McLaren et al., 2016). Only variants possibly related to the phenotype were assessed, and secondary findings were neither reviewed nor reported, as per the study protocol.

The same physician who had clinically assessed the patients enrolled in the pilot study evaluated also the WGS data, with the help and assistance of the bioinformatics team of the Utah Center for Genetic Discovery (https://ucgd.genetics.utah.edu), namely Drs. Barry Moore and Brent Pedersen. This permitted a unique analysis of the possible candidate genes, e.g. the physician knows exactly what the phenotypic manifestations of each patient are and does not base the analysis on HPO terms only, allowing for identification and further analysis of genes that may otherwise be not prioritized (Basel-Salmon et al., 2019).

Deep coverage targeted next generation sequencing (NGS) of the coding regions of *TSC1* and *TSC2*

Genomic DNA of each proband was extracted from peripheral blood (QIAmp, DNA mini kit, Qiagen, Germany) and saliva (Oragene Saliva kit, DNA Genotek) following the manufactorer's protocol at the molecular genetics laboratory of San Paolo Hospital in Milan (Italy). DNA concentration was measured by Qubit fluorometer (Thermo Fisher Scientific, USA).

Massively parallel sequencing was performed at the molecular genetics laboratory of San Paolo Hospital in Milan (Italy) using the Miseq platform (Illumina, San Diego, CA, USA). From January 2016 to April 2018 TrueSeq Custom amplicon sequencing TSCA (Illumina, San Diego, CA, USA) was used to amplify all exons and intron-exon boundaries of *TSC1* and *TSC2*. Sequencing of exon 15 of *TSC1* and exons 18, 36, 38 of *TSC2* - which are not covered by the TrueSeq Custom amplicon sequencing TSCA - was performed using the NGS-NEXTERA XT (Illumina, San Diego, CA, USA) with a read depth >5000X, thus covering the entire coding regions of both genes.

After April 2018 libraries were prepared using the Nextera DNA Flex Enrichment custom kit (Illumina, San Diego, CA, USA) and sequenced to generate 151 bp paired-end reads. Sequencing reads were aligned to the hg19 genome reference, and secondary analysis was performed using the standard MiSeq Reporter software (Illumina, San Diego, CA, USA). The BAM files were uploaded onto the Integrative Genome Viewer (IGV), and prioritization of variants was performed using the Illumina Variant Studio and Illumina Variant Interpreter software. Read depth ranged from 2159X to 8659X.

The functional impact of all variants was inspected using different in silico tool (SIFT: http://sift.jcvi.org/; Polyphen-2: http://genetics.bwh.harvard.edu/pph2; Mutation Taster http://www.mutationtaster.org/; fathmm: http://fathmm.biocompute.org.uk/). Exome aggregation

consortium (ExAC, now dismissed), gnomAD, 1000 Genomes database, and the Leiden Open Variation Database (LOVD: http://chromium.lovd.nl/LOVD2/TSC/home.php) were checked to assess the presence/absence of all identified variants. The variants were then classified according to the ACMG guidelines (Richards et al., 2015) and the LOVD.

Mosaic variants with variant allele frequency (VAF) <20% were validated with NGS-NEXTERA XT (Illumina, San Diego, CA, USA); the other mosaic variants and all heterozygous variants were confirmed by Sanger sequencing.

3.2 FULL STUDY

3.2.1 PATIENTS AND PHENOTYPING

We included in the full study all individuals with a definite or possible clinical diagnosis of TSC (Northrup and Krueger, 2013) for whom deep-coverage *TSC1* and *TSC2* NGS was requested and performed at the genetics laboratory of San Paolo Hospital in Milan (Italy) between January 1, 2016 and July 8, 2021. This cohort includes: 1. individuals who were referred to the TSC clinic of the same university hospital for the first time (all found to have a definite clinical diagnosis except for <10); 2. TSC patients with a definite clinical diagnosis who had received genetic tests before NGS became available and were thus re-sequenced with the new technology, including the patients from the pilot study; 3. individuals with a definite or possible diagnosis whose testing was requested from physicians from a different institution (<5%).

Based on the results of the genetic tests, we selected those who resulted to have a mosaic variant in either *TSC1* or *TSC2*. Mosaic patients followed at the TSC clinic of San Paolo Hospital were evaluated clinically by a clinical geneticist and a pediatric neurologist with expertise in TSC. For the mosaic patients referred for testing by outside our Institution, referring physicians were asked to provide detailed de-identified medical records. The medical records and imaging data from birth to the time of assessment of each affected individual were retrospectively reviewed, and the clinical manifestations of each affected individual were recorded by the clinical geneticist using the standardized examination form as described above and shown in figure 6, and then transferred in an ad-hoc database.

For each patient we evaluated more than 50 clinical variables, including the following: demographic data (sex, age, age at clinical diagnosis, age at molecular diagnosis, age at last evaluation); presence/absence, exact number and localization of each diagnostic criterion as per Northrup and Krueger (2013) and updated in Northrup et al. (2021); additional manifestations that are not included in the diagnostic criteria (cognitive level, seizures, TAND, bone lesions, malignancies).

The frequencies of each clinical manifestation were calculated using as the denominator the number of individuals for whom clear evidence of presence/absence of each manifestation was available. For clinical features known to be age-dependent, the frequencies were calculated using as the denominator the number of patients older than the lowest age of onset for a specific feature. Qualitative data were described as numbers and percentages. Quantitative data were described as median and ranges. Differences between two groups were calculated using the unpaired T test. A *p* value <0.05 was considered significant.

3.2.2 MOLECULAR ANALYSES

Deep coverage targeted NGS of the coding regions and intron-exon boundaries of *TSC1* and *TSC2* was performed as described above on DNA extracted from peripheral blood and saliva (please see methods of the pilot study for this specific test). DNA extracted from additional tissues (i.e. skin biopsy, SEGA, renal angiomyolipoma) was sequenced whenever needed to confirm mosaicism.

If deep-coverage targeted NGS resulted normal, MLPA of *TSC1* and *TSC2* was carried out using the MRC Holland Probemix *TSC2* P046 and Probemix *TSC1* P124 kits (MCR Holland, Amsterdam, the Netherlands).

The reports of each patient found to have a mosaic variant were reviewed by the clinical geneticist, and each variant was classified based on the LOVD (http://chromium.lovd.nl/LOVD2/TSC/home.php) and according to the ACMG guidelines (Richards et al., 2015). Median read depth and mosaic variant allele frequency (VAF) were calculated for the full study cohort. Mosaicism was defined as follows: 1. low-level mosaicism, when VAF was ≤10%; 2. extremely low-level mosaicism, when VAF was <1%.

The workflow of the present study (pilot study and full study) is presented in Figure 7.

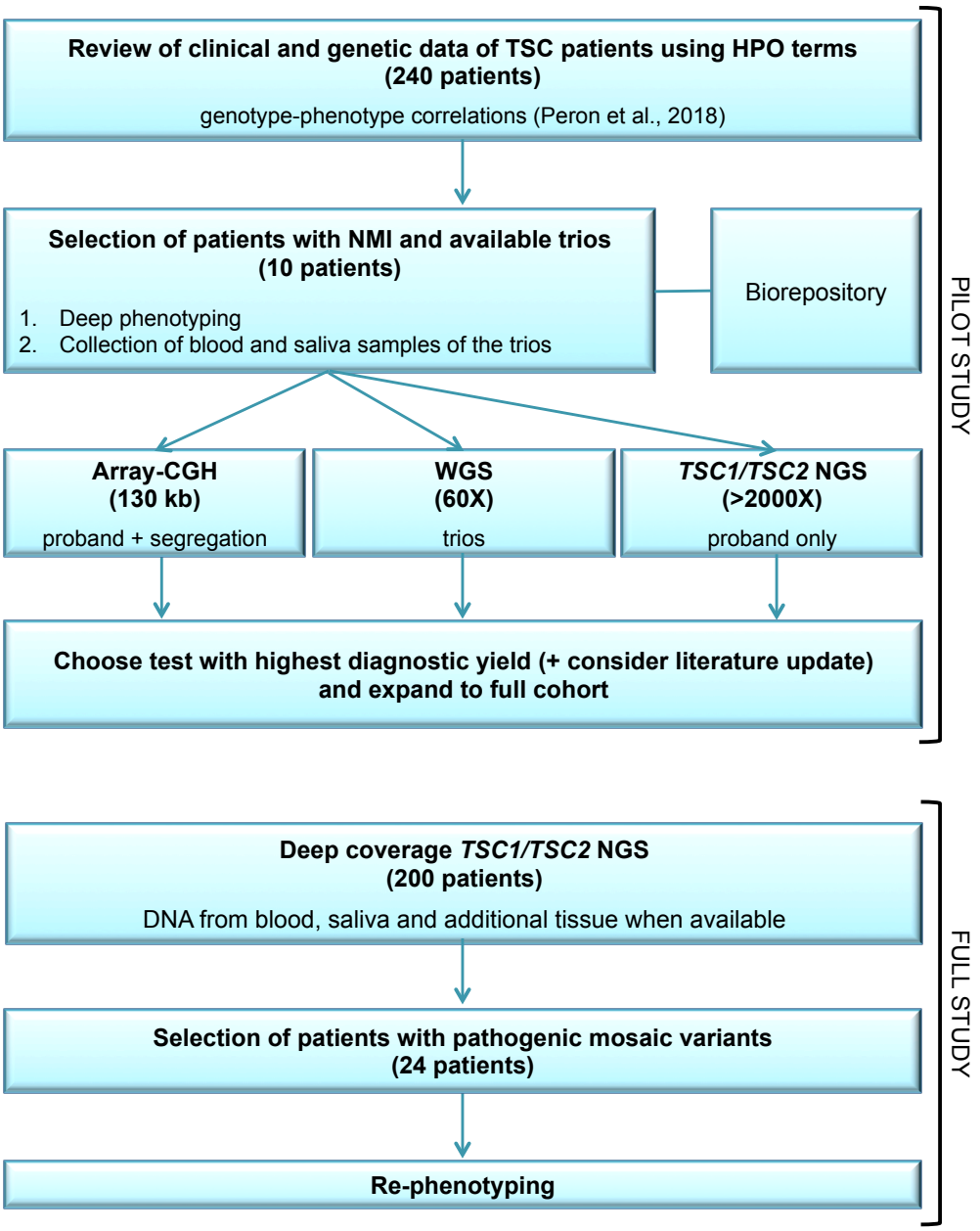


Figure 7: Study workflow. TSC: tuberous sclerosis complex. HPO: Human phenotype ontology. NMI: no mutation identified. CGH: comparative genomic hybridization. WGS: whole genome sequencing. NGS: next generation sequencing.

3.3 ETHICS STATEMENT

For the pilot study, all the patients and/or their legal guardian provided written informed consent to participate in the research. The study was approved by the San Paolo Hospital Ethics Committee

(Comitato Etico Milano Area 1): Protocollo NGS TSC – Identificazione della causa genetica della Sclerosi Tuberosa in pazienti senza mutazione nei geni *TSC1* e *TSC2*, registro sperimentazioni n.2016/ST/179, protocollo n°14156/2017.

For the full study, all genetic analyses were performed in a diagnostic setting based on the new and current clinical practice (Tyburczy et al., 2015; Peron et al., 2018c). All the patients and/or their legal guardians provided written informed consent for genetic analyses. Phenotyping was performed as part of the TSC study approved by the San Paolo Hospital Ethics Committee (Comitato Etico Milano Area 1): protocollo n°9570/2013/rev2020.

4 RESULTS

4.1 PILOT STUDY

Ten affected individuals were included in the pilot study and received CMA, trio WGS and *TSC1*/*TSC2* targeted deep sequencing. The results of the genetic analyses for each patient are summarized in Table 2.

Patient	CMA[1]	Trio WGS (VAF)	Blood NGS (VAF)	Saliva NGS (VAF)
TSC001	NMI	NMI	NMI	NMI
TSC002	NMI*	NMI	*TSC2*: c.1831C>T; p.Arg611Trp (4%)	*TSC2*: c.1831C>T; p.Arg611Trp (4%)
TSC004	NMI	*TSC1*: c.1442_1443ins; p.Ile482LeufsTer4 (het)	NMI	NMI
TSC005	NMI	NMI*	NMI	NMI
TSC007	NMI	NMI	*TSC2*: c.2356-1G>A; p.? (2%)	*TSC2*: c.2356-1G>A; p.? (2%)
TSC008	NMI	*TSC2*: c.1372C>T; p.Arg458Ter (12.7%)	*TSC2*: c.1372C>T; p.Arg458Ter (16%)	*TSC2*: c.1372C>T; p.Arg458Ter (35%)
TSC009	NMI	NMI	NMI	NMI
TSC011	NMI	*TSC2*: c.4129C>T; p.Gln1377Ter (7.4%)	*TSC2*: c.4129C>T; p.Gln1377Ter (7%)	*TSC2*: c.4129C>T; p.Gln1377Ter (5%)
TSC013	NMI	NMI	*TSC2*: c.826_827delAT; p.Met276ValfsTer61 (4%)	*TSC2*: c.826_827delAT; p.Met276ValfsTer61 (4%)
TSC015	NMI	NMI*	*TSC2*: c.2098-1G>A; p.? (5%)	*TSC2*: c.2098-1G>A; p.? (5%)

Table 2: Results of the genetic analyses for each patient. [1] genomic coordinates given according to human assembly hg19/GRCh37. CMA: chromosomal microarray; WGS: whole genome sequencing; VAF: variant allele frequency; NGS: next generation sequencing; NMI: no mutation identified, refers to the absence of pathogenic variants related to the TSC phenotype. het: heterozygous. *variants of unknown significance were identified, see main text.

4.1.1 CHROMOSOMAL MICROARRAY

CMA failed to identify pathogenic microdeletions or microduplications that could explain the patients' phenotype. It did not detect any chromosomal imbalances leading to breakage of any TSC-related genes either. However, a 565 kb microduplication on the long arm of chromosome 19 (19q13.41) was identified in patient TSC002: arr19q13.41 (52,616,637-53,181,882)x3. Segregation analysis showed that the duplication was inherited from the patient's healthy father. The duplication contains several Zinc Finger Protein genes, and the *PPP2R1A* gene (OMIM #605983). *De novo*

missense mutations of *PPP2R1A* – but not whole gene duplications – have been reported in five individuals affected by autosomal dominant intellectual disability (OMIM # 616362) (Deciphering Developmental Disorders Study 2015; Houge et al., 2015). The duplication was inherited from the healthy father and the TSC patient does not exhibit intellectual disability although he presents with severe behavioral problems. Based on the ACMG/ClinGen recommendations on CNV classification, the 19q13.41 duplication identified in patient TSC002 was therefore classified as a variant of unknown clinical significance (VUS) and deemed not to be causative of the TSC phenotype. Of note, this variant is currently being studied as part of a different research project aimed at identifying gene modifiers of the TSC-related neuropsychiatric phenotype.

4.1.2 TRIO WHOLE GENOME SEQUENCING

Through trio WGS we were not able to identify pathogenic variants in genes outside *TSC1* and *TSC2* that could be related to the patients' phenotype. No pathogenic variants were found in any of the AKT/PI3K/mTOR pathway genes. We further reviewed the following candidate genes for each patient on the Integrative Genomic Viewer (IGV) and found no pathogenic variants: *FLCN* (folliculin, known to cause Birt-Hogg-Dubè syndrome, a differential diagnosis of TSC); *G3BP1* and *G3BP2*, as kindly suggested by Prof. Kathrin Thedieck (recently found to be related to mTORC1 signaling in Prentzell et al., 2021), and *TBC1D7*, which encodes for the third subunit of the TSC protein complex (although its pathogenic variants are known to cause a different syndrome characterized by ID, macrocrania, patellar dislocation, and celiac disease (Alfaiz et al., 2014)).

We also manually reviewed the *TSC1* and *TSC2* genes on the IGV to look for variants possibly missed by previous genetic tests or variants in the non-coding regions of *TSC1* and *TSC2* (i.e. deep intronic variants). No pathogenic variants were identified.

In two affected individuals (TSC005 and TSC015) two heterozygous VUSs were detected in the *PHF20* and *FLNA* genes, respectively:

PHF20 [NM_016436.5]: c.870A>G;p.Ile290Met, maternally inherited;

FLNA [NM_001110556.2]: c.1406C>T;p.Pro469Leu, maternally inherited.

The *PHF20* variant was not seen in population databases (1000 Genomes and GnomAD), has a CADD score of 16, and is predicted as deleterious by several *in silico* tools (MutationTaster, Polyphen-2, SIFT). However, *PHF20* has not been associated with disease thus far and the variant was therefore not pursued further. The *FLNA* variant is rare (GnomAD allele frequency: 0.00002520), has a CADD score of 15, and is reported in ClinVar as a VUS. *FLNA* variants are associated with a wide range of heterogeneous diseases including neuronal migration disorders, namely X-linked periventricular nodular heterotopia, which is usually lethal in males (Chen and Walsh, 2015). Moreover, patient TSC005 was later found to have a mosaic pathogenic variant in *TSC2*, and the *FLNA* variant was therefore disregarded.

Interestingly, WGS data analysis of patient TSC004 showed a 62 nucleotide insertion in the *TSC1*

gene causing frameshift and leading to a truncated protein. The variant was inherited from an affected parent. However, multiple assays performed to validate this variant with a different technology did not confirm its presence. The result was not returned to the family as we are currently unable to determine whether the variant detected on WGS is real or is an artifact/false positive. The patient's sample is now being analyzed by Dr. Mark Nellist in the Netherlands with a different technology.

When we performed FREEBAYES-NF analysis of WGS data to specifically look for mosaic variants, we identified two mosaic *TSC2* pathogenic variants in patients TSC008 and TSC011 at an allele frequency of 12.7% and 7.4% respectively, which were confirmed by deep coverage NGS and are discussed below.

4.1.3 DEEP COVERAGE TARGETED *TSC1* AND *TSC2* NGS

Median read depth for targeted sequencing of the *TSC1*/*TSC2* coding regions and intron-exon boundaries was 2903X for DNA extracted from peripheral blood (range 1545-6111X) and 2409X for DNA extracted from saliva (range 1043-4298X).

Targeted sequencing identified mosaic pathogenic variants in *TSC2* in six affected individuals in DNA from both peripheral blood and saliva (table 2). The results were confirmed with NEXTERA XT. All mosaic variants are classified as pathogenic and have been previously reported in the LOVD as pathogenic when heterozygous. Mosaic VAF ranged from 2% to 16% in DNA extracted from peripheral blood and from 2% to 35% in DNA extracted from saliva. No extremely low-level mosaicism (VAF<1%) was identified.

Four affected individuals resulted as having NMI after deep coverage sequencing of the *TSC1*/*TSC2* coding regions and intron-exon boundaries.

We then calculated and compared the mean read depth of deep sequencing for all tissues (blood and saliva) for the patients who were found to have a mosaic pathogenic variant (3248.75X ± 1105.09) and for those who remained with NMI (1330.80X ± 216.97). The mean read depth was significantly different between the two groups: p=0.0018. The clinical manifestations of the patients with mosaic variants and of those remaining with NMI did not differ (data not shown).

4.2 FULL STUDY

We performed targeted deep-coverage *TSC1* and *TSC2* NGS on 200 patients with a definite or possible diagnosis of TSC between 1st January 2016 and 8th July 2021. Affected individuals who received testing on DNA extracted from multiple tissues were counted only once. Breakdown of numbers per year was the following: 2016 (38 tests); 2017 (51 tests); 2018 (35 tests); 2019 (31 tests); 2020 (29 tests); 2021 (16 tests).

4.2.1 RATE OF MOSAICISM IN TSC

Out of the 200 patients who were tested, 25 were found to have a mosaic variant in either *TSC1* or *TSC2*. Of the mosaic variants, 24 were classified as pathogenic (P) or likely pathogenic (LP) according to the ACMG criteria and records found in the LOVD; one was classified as a VUS and excluded. The rate of confirmed mosaicism in our cohort was 12% (24/200). Table 3 shows the P/LP variants identified in each patient and the allele frequency in each tissue that was tested.

Patient ID	Gene[1]	Mosaic variant	Blood VAF (average read depth)	Saliva VAF (average read depth)	Other tissue VAF (average read depth)
P1	*TSC2*	c.521C>A;p.Ser174Ter	10% (na)	NP	NP
P2	*TSC1*	c.2293C>T;p.Gln765Ter	4% (3312X)	NP	NP
P3	*TSC2*	c.1372C>T;p.Arg458Ter	16% (6111X)	35% (3903X)	NP[2]
P4	*TSC2*	c.2098-1G>A;p.?	5% (2377X)	5% (2561X)	NP[3]
P5	*TSC2*	c.1831C>T;p.Arg611Trp	4% (2955X)	4% (2955X)	NP
P6	*TSC2*	c.1600-1G>A;p.?	8.5% (na)	NP	NP
P7	*TSC2*	c.4129C>T;p.Gln1377Ter	7% (2903X)	5% (2257X)	NP
P8	*TSC2*	c.2356-1G>A;p.?	2% (2159X)	2% (3342X)	NP
P9	*TSC2*	c.826_827delAT;p.Met276ValfsTer61	4% (3164X)	4% (4298X)	NP
P10	*TSC2*	c.1769_1776del;p.Leu590ProfsTer24	not detected	2% (921X)	4.5% (1288X)[4]
P11	*TSC2*	c.5044del;p.Leu1682TrpfsTer144	32% (5181X)	NP	NP
P12	*TSC2*	c.4871delT:p.Leu1624ArgfsTer48	NP	2% (1409X)	NP
P13	*TSC2*	c.4620C>G;p.Tyr1540Ter	30% (1873X)	NP	NP
P14	*TCS2*	c.4620C>A;p.Tyr1540Ter	4% (2033X)	4% (1879X)	NP
P15	*TSC2*	c.3094C>T;p. Arg1032Ter	4% (1464X)	NP	NP
P16	*TSC2*	c.770G>A;p.Trp257Ter	3% (1871X)	NP	NP
P17	*TSC2*	c.187C>T;p.Gln63Ter	4% (1756X)	NP	NP[3]
P18	*TSC2*	c.3442C>T;p.Gln1148Ter	5% (1478X)	5% (1478X)	NP
P19	*TSC2*	c.268C>T;p.Gln90Ter	4% (1510X)	3% (514X)	NP
P20	*TSC2*	c.4660C>T;p.Gln1554Ter	3% (na)	NP	NP
P21	*TSC1*	c.2299C>T;p.Gln767Ter	7% (1543X)	8% (21422X)	NP
P22	*TSC2*	c.2251C>T;p.Arg751Ter	9% (2017X)	NP	NP
P23	*TSC2*	c.5228G>A;p.Arg1743Gln	5% (1995X)	NP	NP
P24	*TSC2*	c.5228G>A;p.Arg1743Gln	3% (1843X)	pending	NP

Table 3: Patients with TSC who have a mosaic pathogenic/likely pathogenic variant in *TSC1* or *TSC2*. Patients included in the pilot study and respective corresponding IDs: P3-TSC008; P4-TSC015; P5-TSC002; P7-TSC011; P8-TSC007; P9-TSC013. [1] *TSC1* [NM_000368.4], *TSC2* [NM_000548.3]. [2] Formalin-Fixed Paraffin-Embedded renal tissue was available, but DNA could not be obtained despite attempting extraction from multiple blocks. [3] Formalin-Fixed Paraffin-Embedded cerebral tissue (cortical tuber) was available, but DNA could not be obtained. [4] skin. VAF: variant allele frequency; NP: not performed; na: not available.

Twenty-two individuals (92%) carried a *TSC2* P/LP mosaic variant, and 2 (8%) had a *TSC1* P/LP mosaic variant. With regard to the types of variant, most mutations identified were nonsense (14 patients), whereas a minority were small indel causing frameshift and resulting in premature protein termination (4 patients), variants affecting splicing (3 patients), and missense mutations (3 patients), as illustrated in Figure 8.

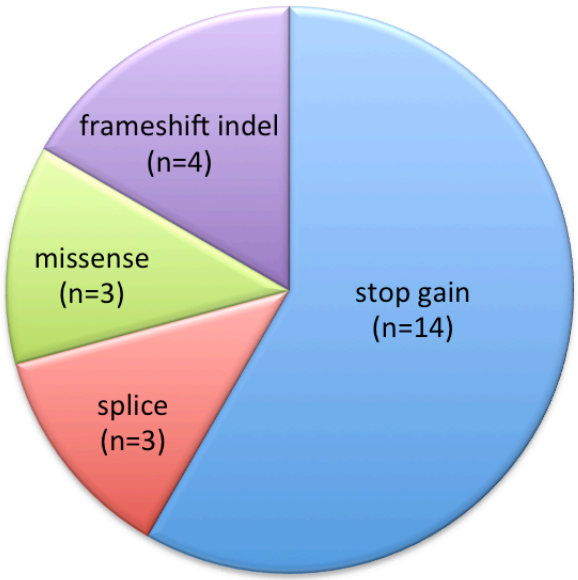


Figure 8: Types of identified pathogenic/likely pathogenic variants.

Median VAF in blood was 4.5% (range 2%-32%). Most patients had low-level mosaicism (VAF ≤10%; n=19). Three individuals had variants with VAF >10%. In one patient the variant seen on other tissues was not detected in blood. In another individual analysis on blood was not performed. Median VAF in saliva was 4% (range 2%-35%). The majority of patients had low-level mosaicism (VAF ≤10%; n=11/12). One individual had a variant with VAF >10%. No extremely low-level mosaicism (VAF<1%) was detected in any of the patients on either peripheral blood DNA or saliva DNA. Testing on both tissues was available for 10 affected individuals: VAF did not differ between the two tissues in nine of them (Table 3). In one patient (P3) the VAF in saliva was much higher than the VAF in blood (35% vs 16%, respectively). Unfortunately, DNA could not be extracted from available Formalin-Fixed Paraffin-Embedded (FFPE) renal tissue of P3 to evaluate the VAF in a different tissue. Testing on three different tissues was available for one affected individual (P10): the molecular diagnosis in this patient was reached thanks to the analysis on saliva, as NGS on blood did not detect any variant. In this patient, the VAF detected on DNA

extracted from two different skin samples (hypomelanotic macule and unaffected skin) was higher than the VAF detected on DNA on saliva (4.5% vs 2%, respectively).

As stated above, an additional patient was found to have a VUS in *TSC2* [NM_000548.3]: c.1027A>C;p.Thr343Pro, with VAF of 17% (median read depth: 1175X). This variant lies in the hamartin binding domain, is absent in the general population (GnomAD VAF=0), has not been reported in the LOVD or in ClinVar, and is predicted as deleterious by *in silico* tools (MutationTaster, SIFT, Polyphen-2). Although these pieces of evidence together with the patient's phenotype suggest pathogenicity, functional studies were not available, and the variant was therefore classified as a VUS according to the ACMG criteria (Richards et al., 2015).

In addition to the patients with mosaic single nucleotide variants or indels, we identified also an affected individual with a mosaic large deletion encompassing exon 41 and the 3' UTR of *TSC2* and part of the *PKD1* gene (contiguous gene deletion) through DHPLC that was previously done at a different laboratory in Turin (Italy). We then reviewed all the MLPA tests performed at the genetics laboratory of San Paolo Hospital (Milan, Italy) between 1st January 2016 and 8th July 2021 (n=69), and found no other patients with mosaic large deletions.

4.2.2 DEEP PHENOTYPING OF MOSAIC PATIENTS

Twenty-four patients had a mosaic P/LP single nucleotide variant in *TSC1* or *TSC2*: 11 males, and 13 females. None of them had a positive family history for TSC. Median age at last follow-up was $24^{+6/12}$ years (range 10 m - $55^{+7/12}$ yrs), with 15 adult individuals and 9 children.

With regard to the clinical diagnosis of TSC, 22 patients had a definite clinical diagnosis, whereas two individuals had a possible diagnosis prior to genetic testing (P21 and P23), according to the diagnostic criteria used at the time of enrollment (Northrup and Krueger, 2013). Median age at clinical diagnosis (or suspicion of diagnosis for P21 and P23) was 18 years (range 5 m - 55 yrs). Median age at molecular diagnosis (= date when the genetic report was issued) was $24^{+6/12}$ years (range $1^{+6/12}$ - $55^{+6/12}$ yrs). The median difference between the age at clinical diagnosis and the age at molecular diagnosis was 7 years (range 1 m - $24^{+4/12}$ yrs).

Figure 9 shows the presence of clinical manifestations for each of the 24 individuals with a P/LP mosaic variant and the sum of all manifestations.

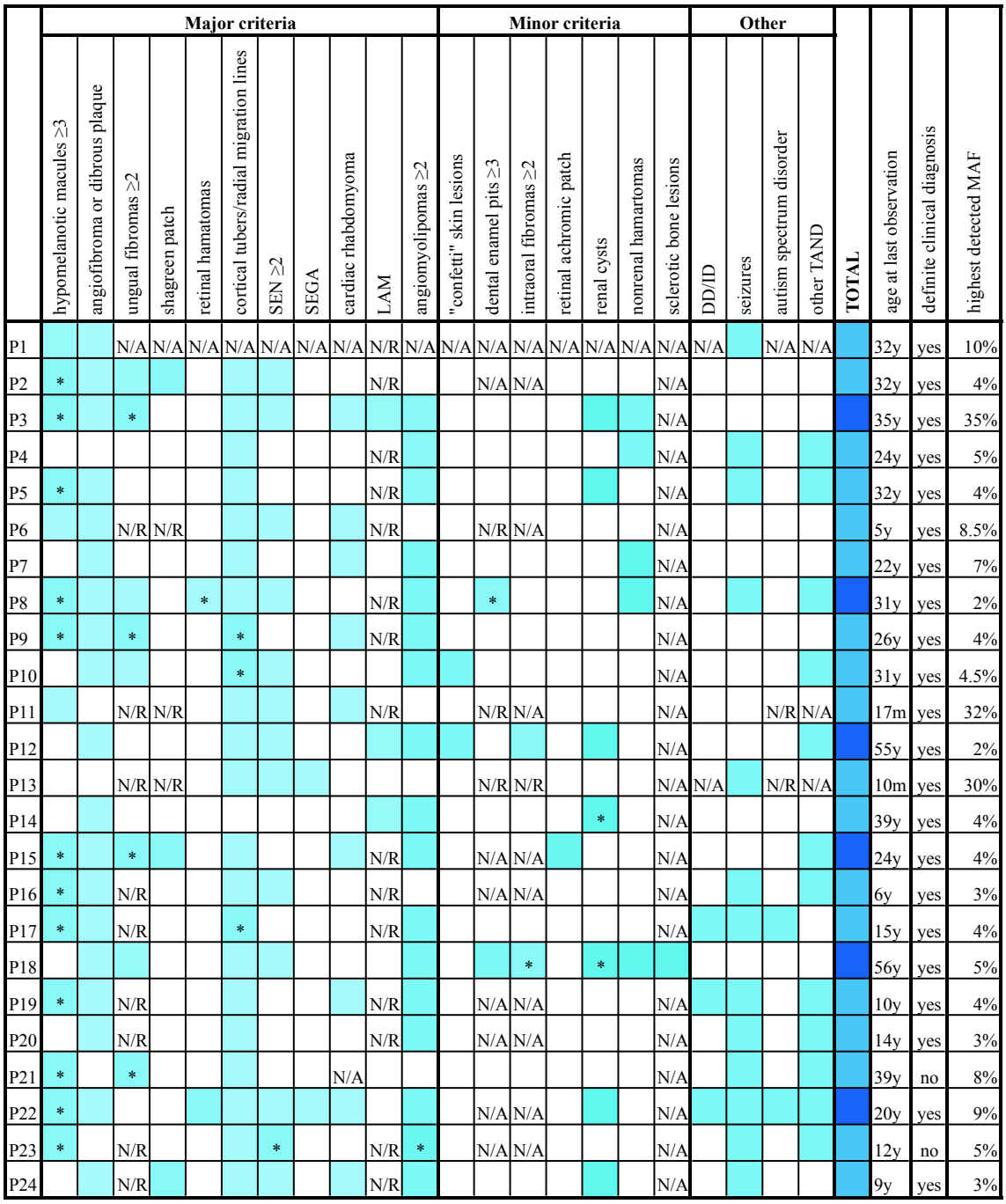


Figure 9: Clinical manifestations of each patient, grouped by major and minor diagnostic criteria (Northrup et al., 2021) and additional common features. White cells: absent; turquoise cells: present; light blue cells: 1-8 manifestations; medium blue cells: 9-15 manifestations; dark blue cells: 16-22 manifestations (no patients in this group). * the manifestation is present, but insufficient in number to meet the specific diagnostic criterion. N/A: information is not available. N/R: information, when absent, is not relevant because the patient was a male (with regard to LAM) or was younger than the age of onset of specific manifestations (16 yrs for LAM, 18 yrs for ungual fibromas, 6 yrs for shagreen patch and dental enamel pits, 2.5 yrs for autism spectrum disorder).

Table 4 reports the frequencies of each clinical manifestation in the present cohort.

Phenotype	Frequency in mosaic patients (%)	Frequency in the general TSC population[1]
Skin		
Hypomelanotic macules >5mm	15/24 (63%)	
Hypomelanotic macules ≥3	3/24 (12.5%)	77%
Hypomelanotic macules <3	12	
Facial angiofibromas (age >2 yrs)	19/22 (86%)	82%
Unilateral angiofibromas	3	
Fibrous plaque	7/23 (30%)	
Fibrous cephalic plaque	5/23 (22%)	
Shagreen patch (age >6 yrs)	3/20 (15%)	48%
Ungual fibromas (age >18 yrs)	8/14 (57%)	
Ungual fibromas ≥2	4/14 (29%)	23%
Ungual fibromas <2	4	
"Confetti" skin lesions	2/23 (9%)	
Dental enamel pits (age >6 yrs)	2/13 (15%)	
Dental enamel pits <3	1	
Intraoral fibromas	2/13 (15%)	
Intraoral fibromas <2	1	
Neuroradiology		
Cortical tubers	22/23 (96%)	
Multiple cortical tubers	19/23 (83%)	95%
Single cortical tuber	3	
Radial migration lines	6/23 (26%)	
Subependymal nodules	13/23 (57%)	
Subependymal nodules ≥2	12/23 (52%)	94%
Subependymal nodules <2	1	
SEGA	2/23 (9%)	40%
Ophthalmology		
Retinal hamartomas	2/12 (9%)	
Multiple retinal hamartomas	1/12 (8%)	18%
Single retinal hamartoma	1	
Retinal achromic patch	1/23 (4%)	
Cardiology		
Cardiac rhabdomyoma	9/22 (41%)	40%
Cardiac arrhythmia	0	
Pulmonology		
LAM (women, age ≥16 yrs)	3/7 (43%)	37%
MMPH (age ≥16 yrs)	2/8 (25%)	
Nephrology		
Renal angiomyolipomas	17/23 (74%)	

Bilateral renal angiomyolipomas	14/17 (82%)	
Renal angiomyolipomas ≥2	16/23 (70%)	68%
Renal angiomyolipomas <2	1	
Renal cysts	7/23 (30%)	
Bilateral renal cysts	4/7 (57%)	
Multiple renal cysts	5/23 (22%)	50%
Single renal cyst	2	
Renal embolization	0	
Nephrectomy	2	
Hepatic angiomyolipomas	5/23 (22%)	
Neuropsychiatric manifestations		
Normal cognitive level	18/22 (82%)	
Borderline intellectual functioning	1/22 (4%)	
DD/ID	3/22 (14%)	50%[2]
Autism spectrum disorder	2/21 (10%)	19%
Other TAND	13/21 (62%)	
Epilepsy	13/24 (54%)	88%
Infantile spams	3/24 (13%)	
Seizure-free at last observation	7/12 (58%)	
Other manifestations		
Sclerotic bone lesions	1	
Malignancies	1/23 (4%)	6%[3]

Table 4: Frequencies of the clinical manifestations in the patients with mosaic pathogenic variants and frequencies in the general TSC population (assumed to be non-mosaic). [1] from a cohort of 240 adult patients from the Natural History Database, in Giannikou et al. (2019). [2] from the TOSCA study in Kingswood et al. (2017). [3] from Peron et al. (2016). SEGA: subependymal giant cell astrocytoma; LAM: lymphangioleiomyomatosis; MMPH: Multifocal micronodular pneumocyte hyperplasia; DD: developmental delay; ID: intellectual disability; TAND: tuberous sclerosis associated neuropsychiatric disorders.

Cutaneous and oral findings

Hypomelanotic macules larger than 5 mm were seen in more than half of the patients. However, 12 individuals had less than three macules, thus not reaching the cut-off number to be counted has a diagnostic criterion. They all had one or two hypomelanotic macules. In one individual the lesion was seen only after performing skin examination with a Wood's lamp. Facial angiofibromas were observed in the majority of affected individuals. Three of them presented with unilateral angiofibromas. When present, fibrous plaques were located mainly in the cephalic area, but were observed also on the trunk, and on the upper and lower limbs in two affected women. Ungual fibromas were present in eight adult patients, but four of these individuals exhibited only a single fibroma, thus not reaching the cut-off number to be counted has a diagnostic criterion. The same was observed for dental enamel pits and for intraoral fibromas, where one patient had only one pit

and another patient had only one fibroma. Three individuals exhibited also café-au-lait macules.

Neuroradiological manifestations

All patient except for one presented with cortical tubers. However, three individuals had one single tuber. With regard to subependymal nodules, one of the patients had one single SEN. Two patients presented with a SEGA, of neonatal onset and at the age of 17 years, respectively. Close follow-up in the first one showed that the astrocytoma was stable over time, thus not requiring surgery or the use of mTOR inhibitors for the time being. The second patient received surgery at the same time of epilepsy surgery.

Eye findings

Three affected individuals showed ocular manifestations on eye exam. As seen for skin and brain findings, there was a patient with a single retinal hamartoma, thus not reaching the cut-off number to be counted has a diagnostic criterion.

Cardiac manifestations

Less than half affected individuals had cardiac rhabdomyoma/s. In five patients a single tumor was detected, whereas in four patients multiple cardiac tumors were seen. In four patients a cardiac rhabdomyoma was observed in the past and was no longer present at the time of last evaluation. Five individuals (aged 5, 22, $1^{+5/12}$, 24, and 9 years) had a rhabdomyoma on last cardiac ultrasound, and in two of them the tumor was regressing.

Pulmonary findings

Three women exhibited LAM. They were diagnosed with LAM at the age of 26 and 35 years; in one patient the exact age at diagnosis was unknown. Two of them were asymptomatic. Of these, one patient had normal pulmonary function tests (PFTs), normal diffusing capacity for carbon monoxide (DLCO) testing, and oxygen saturation >99% at last evaluation; she was on Everolimus. The other patient had PFTs and DLCO testing within the normal limits, though slightly worse than the previous exam; she was not on mTOR inhibitors. The third woman with LAM has emphysema; her DLCO testing worsened over time, whereas PFTs remained normal; high resolution CT scan (HRTC) showed progression of the disease from age 45 to 50 years, with stable radiographic findings thereafter; she was a smoker, and was not on mTOR inhibitors. No patient with mosaic variants was known to have had a pneumothorax.

Nephrological manifestations

Kidney manifestations varied widely among the affected individuals. One adult male patient had a single small renal angiomyolipoma, whereas three women exhibited large, bilateral and growing

angiomyolipomas that resulted in nephrectomy in two of them, at the age of 19 and 27 years. One of these women has later been started on an mTOR inhibitor (Everolimus) following evidence of progressively increasing size of the angiomyolipomas in the remaining kidney. Although renal cysts were often multiple, two patients had one single cyst.

Neuropsychiatric manifestations

The majority of individuals had normal cognitive functioning. Where available, total IQ ranged from 85 to 119. One patient presented with borderline functioning, and three showed ID (1 mild, 2 moderate). None of the affected individuals exhibited severe ID. More than half of the patients presented with a TAND other than ID or autism spectrum disorder. They consisted of: anxiety disorder (5 patients), sleep problems (4 patients), eating disorder (3 patients), dyscalculia and/or dysgraphia (2 patients), attention deficit hyperactivity disorder (1 patient), and depressed mood (1 patient).

Half of the patients had epilepsy, with a median age of onset of 26 months (1 m – 28 yrs). Three had infantile spasms. Focal seizures were the predominant seizure type in 10 individuals. Of the patients with epilepsy, 58% were seizure-free at the time of last assessment. Four affected individuals received epilepsy surgery, and three have been seizure-free since then (Vannicola et al., 2021). Of the affected individuals without epilepsy, two patients exhibited a single febrile seizure with no further episodes, and one patient had a single seizure and was not started on anti-seizure medications.

Other findings

Sclerotic bone lesions of the spine were detected in an adult woman on abdominal CT scan. We are not aware of any other patient with sclerotic bone lesions after reviewing the available imaging.

One patient was diagnosed with cancer, namely carcinoma in situ of the uterine cervix, and was reported in Peron et al. (2016). All the individuals with mosaic pathogenic variants were alive at the time of writing.

Table 5 shows the frequencies of the clinical manifestations in the patients with mosaic pathogenic variants in different studies.

Phenotype of mosaic patients	Present cohort, age 1-55 yrs (%), n=24	Giannikou et al 2019, age >18 yrs (%), n=24	Ogòrek et al 2020, age ≤2 yrs (%), n=9
Hypomelanotic macules ≥3	3/24 (12.5%)	8/24 (33%)	3/8 (38%)
Facial angiofibromas	19/22 (86%)	22/24 (92%)	1/8 (13%)
Shagreen patch	3/20 (15%)	4/23 (17%)	1/8 (13%)
Ungual fibromas ≥2	4/14 (29%)	3/24 (13%)	N/A
"Confetti" skin lesions	2/23 (9%)	N/A	0/8 (0%)
Multiple cortical tubers	19/23 (83%)	14/24 (58%)	7/8 (88%)
Radial migration lines	6/23 (26%)	N/A	3/8 (38%)
Subependymal nodules ≥2	12/23 (52%)	14/24 (58%)	8/8 (100%)
SEGA	2/23 (9%)	0/24 (0%)	2/8 (25%)
Multiple retinal hamartomas	1/12 (8%)	4/21 (19%)	0/5 (0%)
Retinal achromic patch	1/23 (4%)	N/A	1/6 (17%)
Cardiac rhabdomyoma	9/22 (41%)	2/17 (12%)	7/8 (88%)
LAM (women, age ≥16 yrs)	3/7 (43%)	1/22 (5%)	N/A
Renal angiomyolipomas ≥2	16/23 (70%)	20/24 (83%)	1/8 (13%)
Multiple renal cysts	5/23 (22%)	9/24 (38%)	2/8 (25%)
DD/ID	3/22 (14%)	N/A	1/8 (13%)
Autism spectrum disorder	2/21 (10%)	4/24 (17%)	0/8 (0%)
Epilepsy	13/24 (54%)	7/24 (29%)	3/8 (38%)

Table 5: Frequencies of the clinical manifestations in the patients with mosaic pathogenic variants in different studies. SEGA: subependymal giant cell astrocytoma; LAM: lymphangioleiomyomatosis; DD: developmental delay; ID: intellectual disability; N/A: not available.

Number of clinical manifestations and asymmetry of manifestations

By analyzing each affected individual, a trend regarding number and location of the clinical manifestations emerged. If we look at each clinical manifestation, there are 14 patients who presented with a single lesion instead of multiple lesions, as opposed to what is seen in the majority of patients with heterozygous variants. In other words: although individuals with mosaic variants exhibit several clinical features related to TSC, the manifestations are often single in number or unilateral and not bilateral. This was seen mostly with regard to cutaneous manifestations, especially hypomelanotic macules and ungual fibromas (12 patients), but was noted also for cortical tubers (3 patients), renal cysts (2 patients), and retinal hamartomas (1 patient).

An example of this is represented by patient P15. He presented with multiple cortical tubers exclusively on the left side, a single hypomelanotic macule on the left arm, angiofibromas on the

left side of his face (Figure 10A), a shagreen patch, a single ungual fibroma on his left hand, a retinal achromic patch on the right eye, a cardiac rhabdomyoma, bilateral renal angiomyolipomas that were more numerous and larger on the left kidney. He was found to carry the c.3094C>T;p.Arg1032Ter mosaic variant in *TSC2*, with VAF of 4% in DNA extracted from blood.

However, it must be noted that other patients with mosaic variants of the same type (i.e. truncating variants) and detected at a similar VAF in blood (i.e. ≈4%) presented with a more generalized and/or bilateral phenotype, as seen in patients P9 and P18 for instance (Figure 10B and 10C).

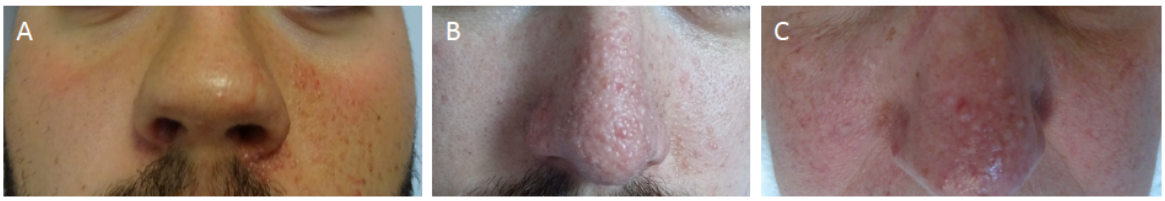


Figure 10: Photographs of selected individuals with mosaic truncating variants in *TSC2* with VAF 4-5% in DNA extracted from blood, showing a different pattern of facial angiofibromas with regard to number, size, and location (unilateral vs. bilateral). A: P15, showing small unilateral angiofibromas on the left cheek; B: P9, showing bilateral angiofibromas of the nose and cheeks; C: P18, showing bilateral angiofibromas of the nose.

As mentioned above, we identified an additional patient with a mosaic large deletion in addition to the 24 individuals with mosaic single nucleotide variants. Since the large deletion encompasses both the *TSC2* and the *PKD1* genes, which is known to cause a contiguous gene syndrome with clinical manifestations of both TSC and polycystic kidney disease, this patient is discussed separately. This patient is an adult male with normal cognitive functioning and no seizures. He was diagnosed with TSC at the age of 12 months, and received the molecular diagnosis at the age of $6^{+4/12}$ years. His phenotype consists of: bilateral facial angiofibromas, bilateral cortical tubers, a subependymal nodule of 10 mm detected at the age of 17 years and stable in size over time, a cardiac rhabdomyoma that spontaneously regressed, unilateral renal angiomyolipoma on the left kidney that required embolization, and bilateral renal cysts.

4.2.3 TRANSMISSION OF MOSAIC VARIANTS TO THE OFFSPRING

TSC patients with heterozygous variants have a 50% risk of passing the mutation to their offspring at each pregnancy and regardless of the fetus' sex. For TSC patients with a mosaic pathogenic variant, a recurrence risk of up to 50% is given, since it is impossible to calculate the exact recurrence risk because the VAF in the gonads is unknown (Peron et al., 2018).

We therefore sought to evaluate the risk of disease transmission of individuals with a mosaic variant. Four patients from the present cohort had pregnancies. They were all women. One patient had a single pregnancy, which resulted in the birth of an unaffected baby; prenatal diagnosis through chorionic villi sampling was performed following the recent molecular diagnosis in the

woman. Another patient had two pregnancies, which both resulted in early miscarriages, prior to the molecular diagnosis in the woman. Another patient had a son, who was tested after the molecular diagnosis in the mother and has not inherited the pathogenic variant. Another woman had two children, who both tested negative for the variant after the molecular diagnosis was reached in the mother. No male patients from our cohort are known to have yet reproduced.

5 DISCUSSION

A significant proportion of individuals with a clinical diagnosis of TSC have NMI despite full diagnostic workup. It is not clear if these patients should be monitored and/or treated in similar fashion to those with known etiology of TSC. This highlights the need of exploring new technologies to try and detect the molecular cause of TSC in these patients and, once identified, to better delineate their phenotype.

In the pilot study we compared three different genetic technologies to answer this question. Chromosomal microarray was not useful in identifying the genetic cause of TSC. Whole genome sequencing was able to detect a mutation in 2 out of 10 affected individuals. No pathogenic variants were identified in genes other than *TSC1* and *TSC2*. This supports the hypothesis that a third gene causing TSC (i.e. *TSC3*) is extremely unlikely to exist. If pathogenic variants were discovered in other genes by future reanalysis of genome data, a careful re-evaluation of the phenotype would need to be carried out, as the patients may have a different though overlapping disease. To the best of our knowledge, there are no other research groups that have been able to identify a third gene thus far, and all attempts have resulted in the identification of pathogenic variants in *TSC1* or *TSC2* that had been previously missed by conventional testing (Qin et al., 2010; Tyburczy et al., 2015). A recently published study claimed to have identified pathogenic variants in genes other than *TSC1*/*TSC2* in three patients with TSC by applying exome sequencing (Kovesdi et al., 2021). However, a critical review of the study reveals that the variants in two of those genes are more likely to be risk factors for certain manifestations exhibited by the patients and that the biallelic pathogenic variants in the third gene explain only part of the patient's phenotype.

Going back to WGS, the only two patients with a molecular diagnosis obtained through WGS were found to have mosaic pathogenic variants in *TSC2*, which were confirmed using targeted deep-sequencing. We were able to identify these two mosaic variants only when applying FREEBAYES-NT, whereas VAAST failed to detect them as it is not designed for mosaicism. WGS - performed at a read depth of 60X instead of the conventional 30X - was able to identify mosaic variants with VAF of at least 7% and missed the mosaic variants with a lower VAF (5% and lower) detected on deep-coverage *TSC1*/*TSC2* sequencing. We did not identify any deep intronic pathogenic variant in either *TSC1*/*TSC2* or other genes.

The pilot study showed that deep-coverage *TSC1*/*TSC2* sequencing had the highest diagnostic yield, and confirmed that mosaic variants in *TSC1* or *TSC2* explain the majority of patients with NMI on conventional genetic testing, as found in the pivotal studies by the Kwiatkowski and Nellist groups (Nellist et al., 2015; Tyburczy et al., 2015). We can therefore conclude that targeted *TSC1* and *TSC2* deep sequencing is the preferable genetic test for TSC, and WGS - or exome sequencing - should not be used to diagnose TSC patients in clinical practice, as stated in current recommendations (Peron et al., 2018c; Ogòrek et al., 2020; Northrup et al., 2021).

As nicely explained by Prof. Leslie Bieseker during his lecture at the latest European Society of Human Genetics Conference, when aiming at reaching a molecular diagnosis and a high diagnostic yield, one should balance breadth (number of genes or regions analyzed, i.e. with exome or genome sequencing) versus depth (i.e. deep sequencing), depending on the most likely mechanism of disease in a certain condition. It is now clear that clinicians and laboratories should prefer depth when testing individuals with TSC. In this regard, it must be noted that the four patients who remained with NMI in the pilot study received *TSC1*/*TSC2* deep sequencing prior to the introduction of the improved enrichments kits at our laboratory, and that the read depth for their sequencing was statistically significantly lower than the read depth obtained for the other patients. Therefore, it is possible that mosaic variants in these patients may have been missed if they were in a region with poor coverage. Looking at this subject from a different perspective, our results suggest that not all deep sequencing panels are suitable for TSC testing. They need to be specifically designed and have an adequate read depth to detect low-level mosaicism, and the requesting physician should pay attention to this aspect when ordering *TSC1* and *TSC2* testing, especially if the phenotype suggests that the affected individual may be a mosaic. A limitation of our targeted panel is that it only covers coding regions and intron-exon boundaries of *TSC1* and *TSC2*, thus missing deep intronic pathogenic variants. Deep intronic variants have been reported (Tyburczy et al., 2015). However, based on the previously published studies, this occurrence seems to be rare. Furthermore, the specific analysis of genome data to look for *TSC1* and *TSC2* non-coding variants in the 10 trios from the pilot study did not identify any deep intronic pathogenic variant.

Based on the results of the pilot study and of the literature, we proceeded with the full study, aimed at characterizing the molecular and clinical characteristics and risk of transmission in patients with mosaic *TSC1* or *TSC2* pathogenic variants.

In the full study, we used deep-coverage sequencing of the coding regions and intron-exon boundaries of *TSC1* and *TSC2*, and identified 24 patients with mosaicism. The rate of mosaicism is 12%, which is surprisingly high for a rare Mendelian multisystem disease. Clinicians should be aware of the fact that at least 1 in 10 patients with TSC have mosaic pathogenic variants and should request the most appropriate test for these patients, ensuring that sequencing at a high read depth is available in the laboratory. Regarding the mosaicism rate, we want to disclose that there may be a possible ascertainment bias in our study since the patients included in our cohort were evaluated in a reference center for TSC, with a specific expertise in individuals whose previous genetic testing had resulted as negative. Nevertheless, we retrospectively reviewed the reasons of referral, and only a minority of patients were referred specifically because of NMI while the majority were patients already known and followed on a regular basis at our TSC clinic, thus ensuring that the rate of mosaicism is actually not biased.

Mosaicism is an emerging topic in TSC and, to the best of our knowledge, there are only three research groups who are currently studying mosaicism in TSC: the group led by Prof. David Kwiatkowski in Boston (USA), the group led by Dr. Mark Nellist in the Netherlands, and our group.

Most patients from our cohort were mosaic for a *TSC2* variant, but mosaic variants in *TSC1* were found as well. The mutational spectrum included all types of variants, with a prevalence for nonsense mutations, as seen for heterozygous *TSC1*/*TSC2* pathogenic variants. The majority of mosaic variants that we detected in our cohort are among those most commonly found in TSC in a heterozygous state as per the LOVD (http://chromium.lovd.nl/LOVD2/TSC/home.php). Interestingly, ten variants (P3, P5, P6, P13, P14, P15, P19, P22, P23 and P24) were found in a mosaic state also in the three previously published studies that evaluated TSC mosaicism (Tyburczy et al., 2015; Giannikou et al., 2019; Ogòrek et al., 2020). This could reflect an enrichment of these recurrent variants in mosaic patients, thus suggesting that particular attention be paid to these regions when performing *TSC1*/*TSC2* deep sequencing.

Mosaic VAF varied widely, from 2% to 35%. The majority of the patients in our cohort had low-level mosaicism (VAF ≤10%.), and three had variants with higher VAF (>10%). We did not identify any individuals with extremely low-level mosaicism (<1%), as opposed to the previous studies, despite applying deep sequencing with a similar read depth. We could not perform a comparison between the groups with higher and lower VAF due to the small sample size, but we would like to note that a lower VAF does not necessarily reflects the presence of a milder phenotype, as described by Byers and colleagues (2018) and seen in several patients from our cohort (P12, P18, and P22).

Since the cohort was heterogeneous with regard to age and the clinical manifestations of TSC are known to be highly age-dependent, we could not evaluate possible phenotypic differences between the two groups. As noted by Giannikou et al. (2019), also in our cohort the mosaic VAF did not differ among different normal tissues (blood and saliva DNA), except for one patient. In P10 analysis on saliva and skin was necessary to identify a mosaic variant that was not seen on blood, and the VAF in the skin tissue sample was higher than the VAF in saliva. So did Giannikou et al. (2019) in six individuals. These observations argue in favor of choosing saliva as the preferred sample to analyze in affected individuals with suspected mosaicism, provided the laboratory has validated the test on saliva. In fact, buccal swabs are easy to obtain, not invasive, and could be applied broadly in any genetics clinic. It must be noted that Giannikou et al. (2019) found a significantly higher VAF in skin DNA than in blood and saliva DNA, whereas we did not have the opportunity to test this difference except for P10. Since the diagnostic yield of deep sequencing in our study is in line with that of Giannikou et al. (2019), we posit that skin biopsy can reasonably be reserved to unsolved cases that should be referred to centers with great expertise in assessing these type of individuals. Nevertheless, we highlight the need of proceeding with testing on different

tissues when sequencing on blood/saliva is negative and large deletions have been ruled out by MLPA. In the future, improved or different technologies could be applied to identify the missing causative variants in the patients who remain with NMI, and the raw data generated with genome sequencing could be re-analyzed to search for structural variants or in cis-regulatory elements outside the coding genome that interfere with the mTOR pathway gene regulation. Another option could be to analyze cell-free DNA using droplet digital polymerase chain reaction (ddPCR), as recently applied to vascular malformations (Zenner et al., 2021).

We performed an extensive analysis of the phenotype of the patients with a mosaic variant, and evaluated all clinical manifestations in detail for the first time. We found that the majority of them have a mild neurodevelopmental phenotype despite having a high frequency of multiple cortical tubers, and often have cardiac, pulmonary and renal findings. Since our cohort includes both children and adults, our results likely represent the true spectrum of clinical variability of TSC patients with mosaic variants. For the same reason however, it was impossible to compare statistically our cohort with the previously published studies evaluating selected manifestations, which include only adults (Giannikou et al., 2019; Treichel et al., 2019) or only children younger than 2 years of age (Ogòrek et al., 2020). The differences we observed (Table 5) in the three studies reflect the age-dependency of the clinical manifestations. Our results show that individuals with a mosaic variant in *TSC1* or *TSC2* have often normal cognitive functioning, although other TANDs – especially anxiety and behavioral problems – are seen in more than half of the patients and should therefore be investigated and treated. Seizures are diagnosed in about half of the cohort. They infrequently involve infantile spams and are more commonly focal-onset seizures that can have an onset in adulthood. Although cortical tubers are invariably present, SENs were detected in only half of the patients in our cohort. Interestingly, cutaneous manifestations - especially hypomelanotic macules - are present in almost all the affected individuals, but their number is often insufficient to meet each diagnostic criterion, except for facial angiofibromas. This was observed also for retinal hamartomas and renal cysts. We hypothesize that, for certain clinical manifestations, the difference between the phenotype of individuals with mosaic and germline mutations does not lie in the presence/absence of clinical features but in the number of single lesions that are present for each feature. Mosaicism depends on the timing of the mutational event in embryonic life (Nathan et al., 2017). It is generally accepted that the earlier the event the more generalized the phenotype (i.e. several different tissues affected), and the later the event the more confined the phenotype (only one or two tissues affected). As discussed by Treichel and colleagues (2019), a lower proportion of cells with the mutation in mosaic patients intuitively suggests that fewer cells are potentially subject to a second hit event following Knudson's hypothesis. This could explain the lower frequency of ID, autism and epilepsy in our cohort of individuals with mosaic variants. However, we observed a high frequency of LAM and renal manifestations in our mosaic cohort, which is at least comparable to and as severe as that seen in the general TSC population, leading to life-

threatening consequences and nephrectomy in two adult women. Similarly, a SEGA was diagnosed in two individuals with mosaic variants. Therefore, TSC patients with a mosaic pathogenic variant should continue surveillance throughout their life. This is especially important for those individuals without neurodevelopmental manifestations who may be inclined to loosen the frequency of their medical appointments in adolescence and adulthood, when pulmonary and renal findings manifest and whose outcome can be dramatically changed by the use of FDA- and EMA-approved mTOR inhibitors (Everolimus and Sirolimus). Of note, all previous studies that evaluated the phenotype of mosaic TSC patients were published by the most experienced research group on this subject (Kwiatkowski lab) with or without the collaboration of the National Institutes of Health. This is the first independent study originating from a different research group. As more individuals with mosaic variants are identified and more research is done on this subject, it is likely that we will gain detailed information about genotype-phenotype correlations to formulate anticipatory guidelines, especially for patients diagnosed prenatally or in the first years of life. This will ideally result in a tailored surveillance and more personalized and cost-effective medicine.

With regard to risk of disease transmission to the offspring, none of the patients in our cohort had affected children or pregnancies with affected fetuses. Other studies have documented the presence of affected children born to mosaic TSC individuals (Giannikou et al., 2019; Treichel et al., 2019), though significantly less frequently than patients with germline TSC (Treichel et al., 2019). As of today and based on the current knowledge about transmission of mosaic variants to the offspring, identifying a patient with a mosaic pathogenic variant in *TSC1* or *TSC2* would not change their recurrence risk (Peron et al., 2018), but would certainly open up a series of different options for future pregnancies for the couple. For these reasons, genetic counselling should always be offered to patients with mosaic variants, and preimplantation/prenatal options should be discussed.

6 CONCLUSIONS

In conclusion, our study shows that genome sequencing fails to identify rare variants in new genes that cause TSC, and a third gene is therefore unlikely to exist. We confirm that the gold standard molecular analysis for TSC is deep-coverage targeted sequencing of the *TSC1* and *TSC2* genes, and tissues different from blood may be needed to find a pathogenic variant.

We demonstrated that mosaicism for *TSC1* or *TSC2* accounts for 12% of the molecular diagnoses of TSC. This is the highest rate of mosaicism seen in a genetic condition thus far, where at least one patient out of ten is a mosaic. We also showed that individuals with mosaic variants in *TSC1* or *TSC2* have a distinctive phenotypic severity, with important implications for surveillance. The identification of larger numbers of patients with mosaicism will likely help elucidate the numerous nuances of mosaicism.

7 REFERENCES

Alfaiz AA, Micale L, Mandriani B, Augello B, Pellico MT, Chrast J, Xenarios I, Zelante L, Merla G, Reymond, A. TBC1D7 mutations are associated with intellectual disability, macrocrania, patellar dislocation, and celiac disease. Human Mutation 2014; 35(4), 447-451. doi: 10.1002/humu.22529

Aronow M, Nakagawa J, Gupta A, Traboulsi E, Singh A. Tuberous sclerosis complex: genotype/phenotype correlation of retinal findings. Ophthalmology 2012; 199: 1917-1923. doi: 10.1016/j.ophtha.2012.03.020

Au KS, Williams AT, Roach ES, Batchelor L, Sparagana SP, Delgado MR, Wheless JW, Baumgartner JE, Roa BB, Wilson CM, Smith-Knuppel TK, Cheung MC, Whittemore WH, King TM, Northrup H. Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States. Genet Med 2007; 9: 88-100. doi: 10.1097/gim.0b013e31803068c7

Basel-Salmon L, Orenstein N, Markus-Bustani K, Ruhrman-Shahar N, Kilim Y, Magal N, Hubshman MW, Bazak L. Improved diagnostics by exome sequencing following raw data reevaluation by clinical geneticists involved in the medical care of the individuals tested. Genet Med 2019; 21(6): 1443-1451. doi: 10.1038/s41436-018-0343-7

Bissler JJ, Kingswood CJ. Renal manifestation of tuberous sclerosis complex. Am J Med Genet C Semin Med Genet 2018; 178(3): 338-347. doi: 10.1002/ajmg.c.31654

Black M, Kadlez M, Smallhorn J, Freedom R. Cardiac rhabdomyomas and obstructive left heart disease: histologically but not functionally benign. Ann Thorac Surg 1998; 65: 1388-1390. doi: 10.1016/s0003-4975(98)00180-5

Bourneville DM. Sclerose tubereuse des circonvultions cerebrales: idiotie et epilepsie hemiplegique. Arch Neurol 1880; 1: 81-91.

Byers HM, Jensen DM, Glass IA, Bennett JT. Minimal mosaicism, maximal phenotype: Discordance between clinical and molecular findings in two patients with tuberous sclerosis. Am J Med Genet C Semin Med Genet 2018; 178(3): 374-378. doi: 10.1002/ajmg.c.31656

Chen MH, Walsh CA. FLNA-Related Periventricular Nodular Heterotopia. 2002 Oct 8 [updated 2015 Sep 17]. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Mirzaa G, Amemiya A, editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993–2021.

Coonrod EM, Margraf RL, Russell A, Voelkerding KV, Reese MG. Clinical analysis of genome next-generation sequencing data using the Omicia platform. Expert Rev Mol Diagn 2013; 13: 529-40. doi: 10.1586/14737159.2013.811907

Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. Nature 2015; 223-228. doi: 10.1038/nature14135

de Vries PJ, Prather PA. The tuberous sclerosis complex. N Engl J Med 2007; 356: 392. doi: 10.1056/NEJMc062928.

de Vries PJ, Whittemore VH, Leclezio L, Byars AW, Dunn D, Ess KC, Hook D, King BH, Sahin M, Jansen A. Tuberous Sclerosis Associated Neuropsychiatric Disorders (TAND) and the TAND Checklist. Pediatric Neurology 2015; 52: 25-35. doi: 10.1016/j.pediatrneurol.2014.10.004.

Dibble CC, Elis W, Menon S, Qin W, Klekota J, Asara JM, Finan PM, Kwiatkowski DJ, Murphy LO, Manning BD. TBC1D7 is a third subunit of the TSC1-TSC2 complex upstream of mTORC1. Mol Cell 2012; 47(4): 535-546. doi: 10.1016/j.molcel.2012.06.009.

European Chromosome 16 Tuberous Sclerosis Consortium. Identification and characterization of the tuberous sclerosis gene on chromosome 16. Cell 1993; 75(7): 1305-1315. doi: 10.1016/0092-8674(93)90618-z

Garrison E and Marth G. "Haplotype-based variant detection from short-read sequencing." arXiv preprint arXiv:1207.3907 (2012). https://arxiv.org/abs/1207.3907

Giannikou K, Lasseter KD, Grevelink JM, Tyburczy ME, Dies KA, Zhu Z, Hamieh L, Wollison BM, Thorner AR, Ruoss SJ, Thiele EA, Sahin M, Kwiatkowski DJ. Low-level mosaicism in tuberous sclerosis complex: prevalence, clinical features, and risk of disease transmission. Genet Med 2019; 21(11): 2639-2643. doi: 10.1038/s41436-019-0562-6

Gupta N, Henske EP. Pulmonary manifestations in tuberous sclerosis complex. Am J Med Genet C Semin Med Genet 2018; 178(3): 326-337. doi: 10.1002/ajmg.c.31638

Harding C, Pagon R. Incidence of tuberous sclerosis in patients with cardiac rhabdomyoma. Am J Med Genet 1990; 37: 443-446. doi: 10.1002/ajmg.1320370402

Hoogeveen-Westerveld M, Ekong R, Povey S, et al. Functional assessment of TSC1 missense variants identified in individuals with tuberous sclerosis complex. Hum Mutat 2012; 33: 476-479. doi: 10.1002/humu.22007

Houge G, Haesen D, Vissers LE, Mehta S, Parker MJ, Wright M, Vogt J, McKee S, Tolmie JL, Cordeiro N, Kleefstra T, Willemsen MH, Reijnders MR, Berland S, Hayman E, Lahat E, Brilstra EH, van Gassen KL, Zonneveld-Huijssoon E, de Bie CI, Hoischen A, Eichler EE, Holdhus R, Steen VM, Døskeland SO, Hurles ME, FitzPatrick DR, Janssens V. B56-delta-related protein phosphatase 2A dysfunction identified in patients with intellectual disability. J Clin Invest 2015; 125(8): 3051-3062. doi: 10.1172/JCI79860

Hu H, Huff CD, Moore B, Flygare S, Reese MG, Yandell M. VAAST 2.0: improved variant classification and disease-gene identification using a conservation-controlled amino acid substitution matrix. Genet Epidemiol 2013; 37(6): 622-634. doi: 10.1002/gepi.21743

Jozwiak S, Schwartz RA, Janniger CK, Bielicka-Cymerman J. Usefulness of diagnostic criteria of tuberous sclerosis complex in pediatric patients. J Child Neurol 2000; 15: 652-659. doi: 10.1177/088307380001501003

Kennedy B, Kronenberg Z, Hu H, Moore B, Flygare S, Reese MG, Jorde LB, Yandell M, Huff C. Using VAAST to Identify Disease-Associated Variants in Next-Generation Sequencing Data. Curr Protoc Hum Genet 2014; 81: 1-25. doi: 10.1002/0471142905.hg0614s81

Kingswood JC, d'Augères GB, Belousova E, Ferreira JC, Carter T, Castellana R, Cottin V, Curatolo P, Dahlin M, de Vries PJ, Feucht M, Fladrowski C, Gislimberti G, Hertzberg C, Jozwiak S, Lawson JA, Macaya A, Nabbout R, O'Callaghan F, Benedik MP, Qin J, Marques R, Sander V, Sauter M, Takahashi Y, Touraine R, Youroukos S, Zonnenberg B, Jansen AC; TOSCA consortium and TOSCA investigators. TuberOus SClerosis registry to increase disease Awareness (TOSCA) - baseline data on 2093 patients. Orphanet J Rare Dis 2017; 12(1): 2. doi: 10.1186/s13023-016-0553-5

Kohler S, Doelken SC, Mungall CJ, Bauer S, Firth HV, Bailleul-Forestier I, Black GC, Brown DL, Brudno M, Campbell J, FitzPatrick DR, Eppig JT, Jackson AP, Freson K, Girdea M, Helbig I, Hurst JA, Jähn J, Jackson LG, Kelly AM, Ledbetter DH, Mansour S, Martin CL, Moss C, Mumford A, Ouwehand WH, Park SM, Riggs ER, Scott RH, Sisodiya S, Van Vooren S, Wapner RJ, Wilkie AO, Wright CF, Vulto-van Silfhout AT, de Leeuw N, de Vries BB, Washingthon NL, Smith CL, Westerfield M, Schofield P, Ruef BJ, Gkoutos GV, Haendel M, Smedley D, Lewis SE, Robinson PN. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res 2014; 42: D966-974. doi: 10.1093/nar/gkt1026

Kovesdi E, Ripszam R, Postyeni E, Horvath EB, Kelemen A, Fabos B, Farkas V, Hadzsiev K, Sumegi K, Magyari L, Moreno PG, Bauer P, Melegh B. Whole Exome Sequencing in a Series of

Patients with a Clinical Diagnosis of Tuberous Sclerosis Not Confirmed by Targeted *TSC1/TSC2* Sequencing. Genes 2021; 12(9): 1401. doi: 10.3390/genes12091401

Henske EP, Jóźwiak S, Kingswood JC, Sampson JR, Thiele EA. Tuberous sclerosis complex. Nat Rev Dis Primers 2016; 2: 16035. doi: 10.1038/nrdp.2016.35

Hong CH, Tu HP, Lin JR, Lee CH. An estimation of the incidence of tuberous sclerosis complex in a nationwide retrospective cohort study (1997-2010). Br J Dermatol 2016; 174(6): 1282-1289. doi: 10.1111/bjd.14415

Jeste SS, Varcin KJ, Hellemann GS, Gulsrud AC, Bhatt R, Kasari C, Wu JY, Sahin M, Nelson CA 3rd. Symptom profiles of autism spectrum disorder in tuberous sclerosis complex. Neurology 2016; 87(8): 766-72. doi: 10.1212/WNL.0000000000003002

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The Ensembl Variant Effect Predictor. Genome Biol 2016; 17(1): 122. doi: 10.1186/s13059-016-0974-4

Moolten SE. Hamartial nature of the tuberous sclerosis complex and its bearing on the tumor problem: report of one case with tumor anomalies of the kidney and adenoma sebaceum. Arch Intern Med 1942; 69: 589-623.

Nabbout R, Belousova E, Benedik MP, Carter T, Cottin V, Curatolo P, Dahlin M, D Amato L, d'Augères GB, de Vries PJ, Ferreira JC, Feucht M, Fladrowski C, Hertzberg C, Jozwiak S, Lawson JA, Macaya A, Marques R, O'Callaghan F, Qin J, Sander V, Sauter M, Shah S, Takahashi Y, Touraine R, Youroukos S, Zonnenberg B, Jansen A, Kingswood JC; TOSCA Consortium and TOSCA Investigators. Epilepsy in tuberous sclerosis complex: Findings from the TOSCA Study. Epilepsia Open 2018; 4(1): 73-84. doi: 10.1002/epi4.12286

Nathan N, Keppler-Noreuil KM, Biesecker LG, Moss J, Darling TN. Mosaic Disorders of the PI3K/PTEN/AKT/TSC/mTORC1 Signaling Pathway. Dermatol Clin. 2017; 35(1): 51-60. doi: 10.1016/j.det.2016.07.001

Nellist M, Brouwer RW, Kockx CE, van Veghel-Plandsoen M, Withagen-Hermans C, Prins-Bakker L, Hoogeveen-Westerveld M, Mrsic A, van den Berg MM, Koopmans AE, de Wit MC, Jansen FE, Maat-Kievit AJ, van den Ouweland A, Halley D, de Klein A, van IJcken WF. Targeted Next Generation Sequencing reveals previously unidentified TSC1 and TSC2 mutations. BMC Med Genet 2015; 16(1): 10. doi: 10.1186/s12881-015-0155-4

Nix DA, Di Sera TL, Dalley BK, Milash BA, Cundick RM, Quinn KS, Courdy SJ. Next generation tools for genomic data generation, distribution, and visualization. BMC Bioinformatics 2010; 11: 455-466. doi: 10.1186/1471-2105-11-455

Northrup H, Koenig MK, Au KS. Tuberous Sclerosis Complex. 1999 [updated 2020]. In: Pagon et al. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2014.

Northrup H, Krueger DA; International Tuberous Sclerosis Complex Consensus Group. Tuberous sclerosis complex diagnostic criteria update: recommendations of the 2012 International Tuberous Sclerosis Complex Consensus Conference. Pediatr Neurol 2013: 49: 243-254. doi: 10.1016/j.pediatrneurol.2013.08.001

Northrup H, Aronow ME, Bebin EM, Bissler J, Darling TN, de Vries PJ, Frost MD, Fuchs Z, Gosnell ES, Gupta N, Jansen AC, Jóźwiak S, Kingswood JC, Knilans TK, McCormack FX, Pounders A, Roberds SL, Rodriguez-Buritica DF, Roth J, Sampson JR, Sparagana S, Thiele EA, Weiner HL, Wheless JW, Towbin AJ, Krueger DA; International Tuberous Sclerosis Complex Consensus Group. Updated International Tuberous Sclerosis Complex Diagnostic Criteria and Surveillance and Management Recommendations. Pediatr Neurol 2021; 123: 50-66. doi: 10.1016/j.pediatrneurol.2021.07.011

O'Callaghan FJ, Shiell AW, Osborne JP, Martyn CN. Prevalence of tuberous sclerosis estimated by capture-recapture analysis. Lancet 1998; 351(9114): 1490. doi: 10.1016/S0140-6736(05)78872-3.

Ogórek B, Hamieh L, Lasseter K, Bagwe S, Machado T, Herranz-Ors C, Thorner AR, Nag A, Gulleman P, Giannikou K, Young LR, Pujana MÀ, Darling TN, El-Chemaly S, Moss J, Henske EP, Kwiatkowski DJ. Generalised mosaicism for TSC2 mutation in isolated lymphangioleiomyomatosis. Eur Respir J 2019; 54(4): 1900938. doi: 10.1183/13993003.00938-2019

Ogórek B, Hamieh L, Hulshof HM, Lasseter K, Klonowska K, Kuijf H, Moavero R, Hertzberg C, Weschke B, Riney K, Feucht M, Scholl T, Krsek P, Nabbout R, Jansen AC, Benova B, Aronica E, Lagae L, Curatolo P, Borkowska J, Sadowski K, Domańska-Pakieła D, Janson S, Kozlowski P, Urbanska M, Jaworski J, Jozwiak S, Jansen FE, Kotulska K; EPISTOP Consortium members, Kwiatkowski DJ. TSC2 pathogenic variants are predictive of severe clinical manifestations in TSC infants: results of the EPISTOP study. Genet Med 2020; 22(9): 1489-1497. doi: 10.1038/s41436-020-0823-4

Peron A, Vignoli A, La Briola F, Volpi A, Montanari E, Morenghi E, Ghelma F, Bulfamante G, Cefalo G, Canevini MP. Do Patients with Tuberous Sclerosis Complex have an Increased Risk for Malignancies? Am J Med Genet A 2016; 170(6): 1538-1544. doi: 10.1002/ajmg.a.37644

Peron A, Canevini MP, Ghelma F, Di Marco F, Vignoli A. Healthcare transition from childhood to adulthood in Tuberous Sclerosis Complex. Am J Med Genet C Semin Med Genet 2018a; 178(3): 355-364. doi: 10.1002/ajmg.c.31653

Peron A, Vignoli A, La Briola F, Morenghi E, Tansini L, Alfano RM, Bulfamante G, Terraneo S, Ghelma F, Banderali G, Viskochil DH, Carey JC, Canevini MP. Deep phenotyping of patients with tuberous sclerosis complex and no mutation identified in *TSC1* and *TSC2*. Eur J Med Genet 2018b; 61(7): 403-410. doi: 10.1016/j.ejmg.2018.02.005

Peron A, Au KS, Northrup H. Genetics, genomics, and genotype-phenotype correlations of TSC: Insights for clinical practice. Am J Med Genet C Semin Med Genet 2018c; 178(3): 281-290. doi: 10.1002/ajmg.c.31651

Prentzell MT, Rehbein U, Cadena Sandoval M, De Meulemeester AS, Baumeister R, Brohée L, Berdel B, Bockwoldt M, Carroll B, Chowdhury SR, von Deimling A, Demetriades C, Figlia G; Genomics England Research Consortium, de Araujo MEG, Heberle AM, Heiland I, Holzwarth B, Huber LA, Jaworski J, Kedra M, Kern K, Kopach A, Korolchuk VI, van 't Land-Kuper I, Macias M, Nellist M, Palm W, Pusch S, Ramos Pittol JM, Reil M, Reintjes A, Reuter F, Sampson JR, Scheldeman C, Siekierska A, Stefan E, Teleman AA, Thomas LE, Torres-Quesada O, Trump S, West HD, de Witte P, Woltering S, Yordanov TE, Zmorzynska J, Opitz CA, Thedieck K. G3BPs tether the TSC complex to lysosomes and suppress mTORC1 signaling. Cell 2021; 184(3): 655-674.e27. doi: 10.1016/j.cell.2020.12.024

Qin W, Kozlowski P, Taillon BE, Bouffard P, Holmes AJ, Janne P, Camposano S, Thiele E, Franz D, Kwiatkowski DJ. Ultra deep sequencing detects a low rate of mosaic mutations in tuberous sclerosis complex. Hum Genet 2010; 127: 573-582. doi: 10.1007/s00439-010-0801-z

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine 2015; 17(5); 405-424. doi: 10.1038/gim.2015.30

Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, Raca G, Ritter DI, South ST, Thorland EC, Pineda-Alvarez D, Aradhya S, Martin CL. Technical standards for the interpretation

and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med 2020; 22(2): 245-257. doi:10.1038/s41436-019-0686-8

Roach E, Gomez M, Northrup H. Tuberous sclerosis complex consensus conference: revised clinical diagnostic criteria. J Child Neurol 1998; 13: 624-628. doi: 10.1177/088307389801301206

Roth J, Roach ES, Bartels U, Al. E. Subependymal Giant Cell Astrocytoma: Diagnosis, Screening, and Treatment. Recommendations From the International Tuberous Sclerosis Complex Consensus Conference 2012. Pediatr Neurol 2013; 49: 439-444. doi: 10.1016/j.pediatrneurol.2013.08.017.

Sauter M, Belousova E, Benedik MP, Carter T, Cottin V, Curatolo P, Dahlin M, D'Amato L, d'Augères GB, de Vries PJ, Ferreira JC, Feucht M, Fladrowski C, Hertzberg C, Jozwiak S, Lawson JA, Macaya A, Marques R, Nabbout R, O'Callaghan F, Qin J, Sander V, Shah S, Takahashi Y, Touraine R, Youroukos S, Zonnenberg B, Jansen A, Kingswood JC; TOSCA investigators. Rare manifestations and malignancies in tuberous sclerosis complex: findings from the TuberOus SClerosis registry to increAse disease awareness (TOSCA). Orphanet J Rare Dis 2021; 16(1): 301. doi: 10.1186/s13023-021-01917-y

Treichel AM, Hamieh L, Nathan NR, Tyburczy ME, Wang JA, Oyerinde O, Raiciulecu S, Julien-Williams P, Jones AM, Gopalakrishnan V, Moss J, Kwiatkowski DJ, Darling TN. Phenotypic distinctions between mosaic forms of tuberous sclerosis complex. Genet Med 2019; 21(11): 2594-2604. doi: 10.1038/s41436-019-0520-3

Tyburczy ME, Dies KA, Glass J, Camposano S, Chekaluk Y, Thorner AR, Lin L, Krueger D, Franz DN, Thiele EA, Sahin M, Kwiatkowski DJ. Mosaic and Intronic Mutations in TSC1/TSC2 Explain the Majority of TSC Patients with No Mutation Identified by Conventional Testing. PLoS Genet 2015; 11: e1005637. doi: 10.1371/journal.pgen.1005637

Vannicola C, Tassi L, Barba C, Boniver C, Cossu M, de Curtis M, De Palma L, D'Errico I, Didato G, Guerrini R, La Briola F, Luisi C, Mai R, Mari F, Marras C, Mastrangelo M, Peron A, Specchio N, Toldo I, Turner K, Vignoli A, Canevini MP. Seizure outcome after epilepsy surgery in tuberous sclerosis complex: Results and analysis of predictors from a multicenter study. J Neurol Sci 2021; 427: 117506. doi: 10.1016/j.jns.2021.117506

van Slegtenhorst M, de Hoogt R, Hermans C, Nellist M, Janssen B, Verhoef S, Lindhout D, van den Ouweland A, Halley D, Young J, Burley M, Jeremiah S, Woodward K, Nahmias J, Fox M, Ekong R, Osborne J, Wolfe J, Povey S, Snell RG, Cheadle JP, Jones AC, Tachataki M, Ravine D, Sampson JR, Reeve MP, Richardson P, Wilmer F, Munro C, Hawkins TL, Sepp T, Ali JB, Ward S,

Green AJ, Yates JR, Kwiatkowska J, Henske EP, Short MP, Haines JH, Jozwiak S, Kwiatkowski DJ. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. Science 1997; 277(5327): 805-808. doi: 10.1126/science.277.5327.805

Vignoli A, La Briola F, Peron A, Turner K, Vannicola C, Saccani M, Magnaghi E, Scornavacca GF, Canevini MP. Autism spectrum disorder in tuberous sclerosis complex: searching for risk markers. Orphanet J Rare Dis 2015; 10: 154. doi: 10.1186/s13023-015-0371-1

von Recklinghausen F. Die Lymphelfasse und ihre Beziehung zum Bindegewebe [German]. Berlin: A. Hirschwald; 1862.

Zenner K, Jensen DM, Cook TT, Dmyterko V, Bly RA, Ganti S, Mirzaa GM, Dobyns WB, Perkins JA, Bennett JT. Cell-free DNA as a diagnostic analyte for molecular diagnosis of vascular malformations. Genet Med 2021 Jan; 23(1): 123-130. doi: 10.1038/s41436-020-00943-8

8 ACKNOWLEDGEMENTS

For and foremost, I want to thank my family (mom, dad, and Andrea) for always supporting me, believing in me, and also patiently putting up with me.

A huge thank you goes to all my colleagues at the TSC Clinic of San Paolo Hospital in Milan: together we have done great things! Thank you to my mentor Prof. Maria Paola Canevini, to Prof. Gaetano Bulfamante, and to Dr. Rosa Maria Alfano for the NGS analysis and for always responding to my questions, even to the craziest ones, and seconding all my requests.

Thank you to all my colleagues at the University of Utah, especially to Prof. David Viskochil, who has taught me to be inquisitive - always and no matter what - , and to Drs. Barry Moore and Brent Pedersen: thank you for patiently teaching me how to look at genome data.

Thank you to the Medical Genetics team at the Department of Health Sciences of the University of Milan for assisting with DNA extraction and for always having a smile for me anytime I walked down the hallway in a hurry: Prof. Cristina Gervasini, Dr. Patrizia Colapietro, and Dr. Elisabetta di Fede.

Thank you to all my mentors who were by my side on this long journey and to those who are now my great friends after so many years. It was a privilege to work with you and learn from you!

Finally, I would like to acknowledge the Foundation Blanceflor Boncompagni Ludovisi née Bildt for supporting my research abroad, and the Chan Soon-Shiong Family Foundation for providing funding for genome sequencing.