TUBEROUS SCLEROSIS COMPLEX 1

Gene Identification and Characterisation

Marjon van Slegtenhorst

TUBEROUS SCLEROSIS COMPLEX 1

Gene Identification and Characterisation

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr P.W.C. Akkermans M.A.
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
woensdag 16 september 1998 om 13:45

door

Marjon Annette van Slegtenhorst

geboren te Gorinchem

Promotiecommissie

Promotor:	Prof. dr D. Lindhout
Overige leden:	Prof. dr J.H.J. Hoeijmakers
	Dr E.C. Zwarthof
	Prof. dr P.J. Willems
Co-Promotor:	Dr D.J.J. Halley

Omslag: Herlof Schürmann, Amsterdam.

The studies described in this thesis were performed at the MGC-Department of Clinical Genetics at the Erasmus University Rotterdam. This project was financially supported by a grant of the Dutch Organisation for Scientific Research (Human Genome Analysis, Medical Sciences, NWO, grant nr. 960-10-834).

Print: Offsetdrukkerij Ridderprint B.V. te Ridderkerk.

Voor mijn vader

TABLE OF CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

GENERAL INTRODUCTION

1.1 Clinical aspects of TSC

Tuberous sclerosis complex (TSC) was first recognised by Désiré-Magloire Bourneville in 1880. The name of the disease originates from the characteristic sclerotic tubers (hamartomas) present in many patients. Other names describing the disease are Bourneville's disease (in honour of the French neurologist) or epiloia (epilepsy, low intelligence and adenoma sebaceum), which was the official name describing the classical triad of symptoms seen in 30-40% of the patients (McKusick, 1990). TSC is usually classified as one of the phakomatoses (Van der Hoeve, 1933), a group of disorders which also includes neurofibromatosis types 1 (NF1) and 2 (NF2) (Phillips and Rye, 1994), von Hippel-Lindau disease (Bernstein *et al.*,1987) and Sturge-Weber syndrome (Prieto *et al.*,1997). All five diseases show apparently randomly distributed patches of abnormal tissue, but they are distinct from each other with respect to the types of lesions and the affected tissues.

TSC is characterised by the growth of a variety of benign tumours (hamartomas) and malformations (hamartias) in one or more organs (Gomez, 1988). The disease is clinically variable and almost every organ and tissue can be affected. The organs most frequently involved are the heart, skin, brain and kidneys. The variability is reflected by the type and number of symptoms and the severity of the disorder and is seen not only between patients from different families, but also between affected relatives within the same family.

The first symptoms that can be indicative of TSC are cardiac rhabdomyomas, which have been detected by fetal echocardiography as early as in the 26th week of gestation. They are usually multiple, may be associated with fetal cardiac arrhythmia, but often remain clinically silent (Watson,1991). A number of cardiac rhabdomyomas spontaneously regress after birth, suggesting that their prenatal occurrence is partly sex steroid-dependent.

TSC patients can display a wide range of skin signs. Hypomelanotic macules (white spots) are often present at birth and appear in about 90% of the patients. They are usually multiple in TSC patients but are also detected in the normal population, so in themselves they are not sufficient for the diagnosis of TSC. The facial angiofibromas (adenoma sebaceum) are

pathognomonic for TSC and appear during the first years of life in 50-70% of the patients (Ahlsén *et al.*,1994). Ungual fibromas are mostly seen in women from puberty on, whereas shagreen patches and fibrous forehead plaques are predominantly present in older TSC patients.

Lesions in the brain are associated with severe manifestations of TSC. Epileptic seizures occur in about 85% of the patients and they often start in the first year of life with infantile spasms and partial motor seizures (Gomez, 1988) (figure 1.1). With increasing age, the seizures may become of a more generalised type. About 50% of the children with seizures develops mental retardation (Gomez, 1988). There is some correlation between the age of onset and the severity of generalised seizures, and the number, size and location of the brain lesions and degree of mental retardation (Curatolo et al, 1991). Characteristic lesions in the central nervous system are cortical tubers, subependymal nodules and subependymal giant cell astrocytomas. Behavioural problems are quite common among children with TSC. Autism is present in approximately 50% of the patients (Hunt and Shepherd, 1993).

Figure 1.1 Salaam cramp. Form of epilepsy often seen in small children; the name originates from the movements the child is making, while it is bending the head and lifting the arms.

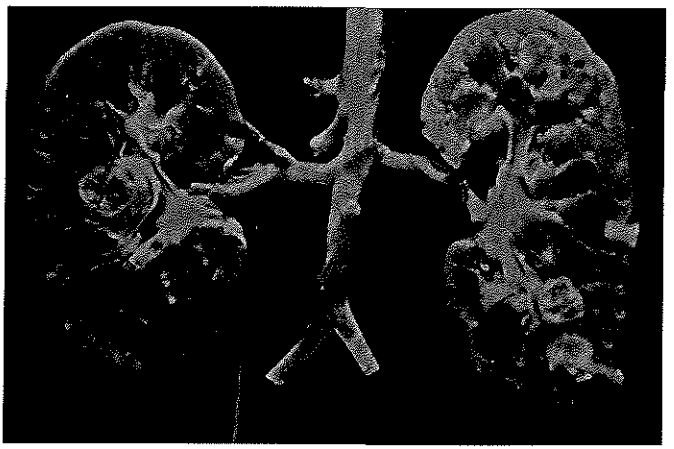


Figure 1.2 Kidneys affected with multiple cysts and angiomyolipomas.

In the second and third decade of life, renal problems are found in 40-80% of TSC patients. The most characteristic renal abnormalities are cysts and angiomyolipomas, generally occurring bilaterally (figure 1.2). Occasionally a renal cell carcinoma develops in patients with TSC (Bjornsson *et al.*, 1996; Cook *et al.*, 1996).

Many other organs may be affected in the pathogenesis of TSC, including the eyes, lungs, skeleton and endocrine glands. Involvement of the lungs in the form of pulmonary lymphangioleiomyomatosis is infrequent. This complication is almost exclusively confined to women with TSC and is treated with anti-estrogens, suggesting a role for steroid hormones and their respective receptors in the development of these tumours (Lie, 1991).

An overview of the criteria for the diagnosis of TSC is listed in table 1.1.

Table 1.1 Current criteria for the diagnosis of TSC

Primary features	
	Facial angiofibromas
	Multiple ungual fibromas
	Cortical tuber[1]
	Subependymal nodule or giant cell astrocytoma[1]
	Multiple calcified subependymal nodules protruding into the ventricle[2]
	Multiple retinal astrocytoma
Secondary features	
	Affected first-degree relative[4]
	Cardiac rhabdomyoma[1,2,3]
	Other retinal hamartoma or achromic patch
	Cerebral tubers[2]
	Noncalcified subependymal nodules[2]
	Shagreen patch
	Forehead plaque
	Pulmonary lymphangiomyomatosis[1]
	Renal angiomyolipoma[1,2,3]
	Renal cysts[1]
Tertiary features	
	Hypomelanotic macules
	"Confetti " skin lesions
	Renal cysts[2,3]
	Randomly distributed enamel pits
	Hamartomatous rectal polyps[1]
	Bone cysts[2]
	Pulmonary lymphangiomyomatosis[2]
	Cerebral white-matter or heterotopias[2]
	Gingival fibroma
	Hamartoma of other organs[1]
	Infantile spasms

The different types of lesions in TSC patients are subdivided into three different categories. A single primary feature is sufficient for the diagnosis of TSC, while a combination of two secondary or one secondary plus two tertiary from the other categories is regarded necessary for a certain diagnosis of TSC. [1]histologically confirmed, [2]radiographic evidence, [3]ultrasound, [4]the affection status of relatives is not taken into acount in our linkage studies. (Roach *et al.*,1992; Neuman and Kandt, 1993)

1.2 Histological and cellular aspects of TSC lesions

The pathogenesis of TSC is poorly understood. The types of lesions most commonly seen in TSC patients are hamartomas and malformations affecting tissues of mesodermal and ectodermal derivation (Gomez, 1988). Histologically, the hamartomas display a disorganised and excessive cell or fiber proliferation without malignant transformation.

In the lesions from TSC patients the normal cellular organisation is often lost and cells are either not correctly differentiated, or are of the wrong type and in the wrong location (Johnson *et al.*,1991). In the brain, cortical tubers contain large cells of unknown origin. The other two brain lesions, subependymal nodules and subependymal giant cell astrocytomas, are histologically identical and they display disordered hypertrophic neurons and enlarged astrocytes.

Most skin lesions consist of a variety of (vascularised or non-vascularised) hamartomatous connective tissue, often characterised by the presence of large neuron-like cells (N-cells). N-cells are large, slowly dividing, dendritic cells, that arise from a primitive precursor of both neurons and glia-cells (Johnson *et al.*, 1991). The hypomelanotic macules are distinct from the hamartomatous skin lesions. The pathology shows a reduction in size, number and pigmentation of the melanosomes (Fitzpatrick, 1991).

The two most frequent kidney lesions are histologically different from each other. The angiomyolipoma consists of vascular, fatty and smooth muscle tissue and also these lesions often contain N-cells. The cyst is a cloved epithelium-lined cavity, filled with fluid. Renal histopathology of cysts from TSC patients resembles autosomal dominant polycystic kidney disease (ADPKD) (Torres *et al.*,1994), however clinical onset is often early (Webb *et al.*,1993) and significant cystic kidney disease in TSC frequently reflects additional mutational involvement of the PKD1 gene (Sampson *et al.*,1997).

In summary, most TSC lesions contain abnormal cells, which are often in wrong locations. It has been suggested, therefore, that TSC is a disease of abnormal cellular growth, migration, differentiation and organisation (Johnson *et al.*,1991).

1.3 Treatment and life expectancy of TSC patients

The life expectancy of TSC patients depends largely on the complications caused by the lesions in the brain and the kidneys (Shepherd *et al.*, 1991). Treatment of TSC patients is dependent on the type of lesion and affected organ system, and is usually symptomatic. Seizures can often be suppressed by medication, but 50% of the children with epilepsy develop cognitive dysfunction. Complications arising from brain lesions cause a higher

mortality rate amongst young TSC patients. Most skin lesions do not need treatment, but laser therapy is often applied to facial angiofibroma for cosmetic reasons. For symptomatic hamartomatous lesions in organ systems, surgery may be the method of choice. However, recent studies indicate that a conservative 'wait and see' policy may be better than early invasive surgery (Jozwiak, 1996). Complications arising from the renal lesions are the most frequent causes of death in TSC patients at adult age. Renal angiomyolipomas can be treated by selective embolisation, which helps to prevent fatal bleeding and postpone progressive renal insufficiency (Fleury, 1989; van Baal *et al.*,1994). In mildly affected patients, many symptoms remain unnoticed until far in adulthood, and these patients have 'a normal life span'.

1.4 Genetics of TSC

Tuberous sclerosis was first recognised as a hereditary disorder in 1913 by Berg (Gomez, 1988). The pattern of inheritance is autosomal dominant with high penetrance but an extremely variable expression. The prevalence of the disease has been subject of study since 1935 and most recent data suggest that it may be as high as 1:6000 (Osborne *et al.*,1991). The prevalence of TSC is probably underestimated because of the existence of very mild clinical phenotypes which are not recognised as TSC.

TSC is a genetically heterogeneous disorder with loci on human chromosomes 9q34 (TSC1) and 16p13.3 (TSC2). About half of the multiplex families are linked to the chromosome 16 locus and the other half to chromosome 9, suggesting an equal proportion of TSC1 versus TSC2 mutations (Kwiatkowski *et al.*,1993). There seems to be no clear correlation between the phenotype and the TSC locus involved. At least 60% of the TSC patients have non-affected parents, representing sporadic cases with a *de novo* mutation (Sampson *et al.*, 1989; Osborne *et al.*,1991). Quite recently, a few cases of somatic mosaicism have been observed (Verhoef *et al.*,1995; van den Ouweland, personal communication). In these families, parents (either apparently unaffected or affected with TSC) of a TSC patient were shown to carry the mutation in part of their leukocytes. However, little is known about the frequency of somatic mosaicism in TSC.

1.5 Tumour suppressor genes

In 1971, Knudson proposed a model for tumour suppressor genes, in which the development of a tumour requires two hits. In familial cases, the first mutation is in the germline and the second hit is a somatic mutation, while in sporadic cases both mutations are somatic. Tumour suppressor gene products constitute key points in many complex cellular pathways that regulate proliferation, differentiation, apoptosis and response to genetic damage (Haber and Harlow, 1997). TP53 (p53) is considered to be the most frequently mutated gene in human cancers. Patients with Li-Fraumeni syndrome show a germline mutation in this gene (Li *et al.*, 1988), but the gene is also mutated in more than 50% of all human cancers (Levine, 1997; Helin and Peters, 1998). Tumour suppressor genes have been implicated in several Mendelian tumour syndromes, which are summarised in table 1.2, but they are also involved in the progression of several common, nonheritable forms of cancer, such as non-familial colorectal cancer (Stanbridge, 1990).

Table 1.2 Tumour Suppressor Genes

Gene (gene product)	Possible function	Familial syndrome
RB1 (p110)	cell cycle regulation	retinoblastoma
WT1	zinc finger protein	Wilm's tumour
TP53 (p53)	cell cycle regulation	Li-Fraumeni syndrome
NF1 (neurofibromin)	GTPase activating protein	neurofibromatosis type I
NF2 (schwannomin)	actin-cytoskeleton organisation	neurofibromatosis type II
DCC	cell surface interactions	colorectal cancer
APC	transcriptional regulator	polyposis colorectal cancer
BRCA1	transcriptional regulator	breast and ovarium cancer
BRCA2	unknown	breast and ovarium cancer
PTEN/MMAC1	novel phosphatase	Cowden disease
VHL (elongin)	mRNA processing	von Hippel-Lindau disease
TSC1 (hamartin)	unknown	tuberous sclerosis complex
TSC2 (tuberin)	GTPase activating protein	tuberous sclerosis complex
MEN1 (menin)	unknown	multiple endocrine neoplasia type 1
STK11	serine threonine kinase	Peutz-Jeghers syndrome

Although the precise cellular defect in TSC is still unknown, the multiple, random, focal distribution of TSC lesions suggests that both TSC genes act as tumour suppressor genes. In the case of TSC, a first hit in the germline results in a mutation in all somatic cells, and a single secondary, postzygotic mutation is supposed to be necessary for tumour formation (figure 1.3). This second somatic hit is often detected in lesions associated with the disease by

loss of heterozygosity (LOH) at polymorphic marker loci in the vicinity of the disease gene (table 1.3).

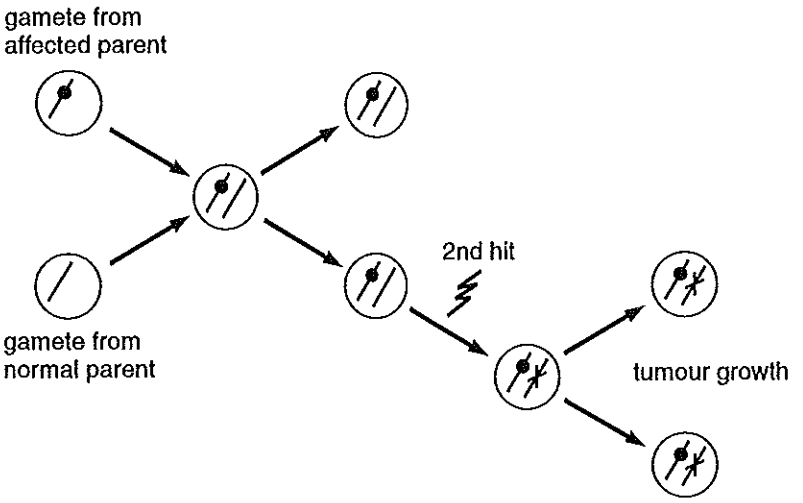


Figure 1.3 Tumour growth in TSC patients. In one of the gametes from a parent, affected with TSC, the gene is mutated. A second somatic mutation (2nd hit) in the homologous wild-type copy of the gene results in complete loss of the gene and will lead to uncontrolled growth.

Table 1.3 LOH frequency in different TSC lesions

	number investigated	LOH at 16p13.3	LOH at 9q34	no LOH detected (%)
Angiomyolipoma	79	37	6	36 (45%)
SEGA	23	5	2	16 (70%)
Cortical tuber	20	3	1	16 (80%)
Facial angiofibroma	10	3	0	7 (67%)
Cardiac rhabdomyoma	9	4	0	5 (56%)
Renal cell carcinoma	7	1	4	2 (28%)
Shagreen patch	2	1	0	1 (50%)

SEGA=subependymal giant cell astrocytoma; data depicted from Green *et al.* (1994a-1994b), Carbonara *et al.* (1994), Henske *et al.* (1995a), Bjornsson *et al.* (1996) and Sepp *et al.* (1996),

The growths in TSC patients, with the exception of renal cell carcinoma, are mostly benign (hamartomas). Other multiple hamartomatous syndromes are Cowden disease (Mallory, 1995; Liaw *et al.*,1997) and Peutz-Jeghers syndrome (Westerman *et al.*,1997; Jenne *et al.*,1998).

LOH is most frequently observed in the angiomyolipomas (AMLs). Since not all lesions have been investigated for LOH, it remains to be elucidated whether all lesions associated with TSC develop by means of two hits. Overall, LOH is more often found at 16p13 than at 9q34. One exception may be renal cell carcinoma (RCC), which shows more often loss at 9q34, although only a limited number of cases have been investigated. LOH in the TSC2 region has also been detected in isolated AML (Henske *et al.*,1995b) and SEGAs (Gutman *et al.*, 1997), not associated with the tuberous sclerosis complex. Whether the TSC1 gene is involved in isolated tumours needs to be investigated.

1.6 Towards the identification of the TSC genes

Mapping of the TSC genes

The first claim of linkage of TSC was to the ABO bloodgroup locus on 9q34 in 1987 (Fryer *et al.*,1987); hence this locus has been designated TSC1. However, subsequent analysis of families by other groups showed no evidence for linkage to 9q34 (Northrup *et al.*,1987; Renwick.,1987; Kandt *et al.*,1989). The most likely explanation for this discrepancy was that gene defects at one or more additional loci may also cause TSC, a phenomenon known as locus heterogeneity. Proof came from a series of linkage studies in a large number of additional families from all over the world which not only confirmed a TSC1 locus on chromosome 9q34, but also indicated the existence of a second locus (Sampson *et al.*,1989; Janssen *et al.*,1990; Haines *et al.*,1991; Povey *et al.*,1991; Northrup *et al.*,1992). Additional candidate loci were indicated by linkage analysis on chromosome 14 (Kandt *et al.*,1991) and by the detection of chromosomal rearrangements in combination with linkage studies including a trisomy of a portion of chromosome 11q (Clark *et al.*,1988; Smith *et al.*,1990) and a translocation event involving chromosomes 3p and 12q (Fahsold *et al.*,1991a/b). However, these loci could not be confirmed in subsequent studies (Sampson *et al.*,1992). A genome search on a subset of families which did not show linkage to chromosome 9, finally yielded a second major TSC locus on chromosome 16p13.3 (TSC2) (Kandt *et al.*, 1992).

Additional studies in a large number of TSC families using methods for linkage analysis under locus heterogeneity defined only these two TSC loci without significant evidence for a third locus (Kwiatkowski *et al.*,1993; Janssen *et al.*,1994; Povey *et al.*, 1994).

Identification of the TSC2 gene

The TSC2 gene was cloned in 1993 (European TSC2 Consortium, 1993), one year after linkage had been found. This was greatly facilitated by the availability of patients with gross rearrangements of the TSC2 region on chromosome 16. This included an unbalanced translocation in a family with TSC and ADPKD, and a few large deletions involving the tip of the short arm of chromosome 16 in non-TSC patients. The 5.5 kb TSC2 transcript contains very short 5' and 3' untranslated regions. The 41 coding exons cover approximately 45 kb of genomic DNA and exons 25 and 31 are alternatively spliced (Maheshwar *et al.*,1996). The TSC2 gene shows a diverse mutational spectrum including large rearrangements, deletions, insertions, and nonsense- and missense- mutations (Brook-Carter *et al.*, 1994; Kumar *et al.*, 1995a-1995b-1997; Verhoef *et al.*, 1995-1998;Vrtel *et al.*, 1996; Wilson *et al.*, 1996; Maheshwar *et al.*, 1997; Au *et al.*, 1997-1998; Wang *et al.*, 1998).

The TSC2 gene encodes a 1807 amino acid protein, designated tuberin, with a predicted molecular mass of 200 kDa. Analysis of the amino acid sequence of tuberin indicated that a region close to the carboxy-terminus (aa 1593-1631) showed homology to the GTPase activating domain of rap1GAP (European TSC2 Consortium, 1993). Other possibly functional domains include a N-terminal leucine zipper (aa 81-102), a C-terminal nuclear localisation signal (aa 1434-1451) (Tsuchiya *et al.*,1996) and three potential transcriptional activation domains (exons 30-32 and 41) (Tsuchiya *et al.*,1996).

1.7 Aims of the study

Tuberous sclerosis shows a complex clinical phenotype and the prospect for finding suitable therapy is hampered by the lack of knowledge about the underlying biochemical defect. After the TSC2 gene had been cloned in 1993, one of the first clinical applications was mutation analysis for diagnostic purposes. In our laboratory, about 1/3 of the TSC2 gene has been screened and mutations have been detected in 15% of the TSC patients. This provides molecular diagnosis, including a prenatal test to TSC2 families. Secondly, the first steps have been undertaken to learn about the function of tuberin, but understanding why a defect in the TSC2 gene causes TSC does not only involve the function of tuberin, but also the gene product of TSC1 and possibly other proteins.

Therefore, the main goal of my project was to identify the TSC1 gene on chromosome 9 using a positional cloning approach, allowing mutation analysis in patients and functional studies including both the TSC1 and TSC2 proteins.

When this project started, the TSC1 gene had been mapped on 9q34 between the markers D9S149 and D9S114, a critical region of approximately 3 cM. A consortium, including groups from Boston, Cardiff, London and Rotterdam, was formed with the aim to construct a cosmid contig covering the TSC1 candidate region and to identify new markers from the region to narrow down the TSC1 region. In addition, different gene isolation techniques were combined to identify as many positional candidate genes as possible from the critical region.

CHAPTER 2

POSITIONAL CLONING

POSITIONAL CLONING

Disease genes can be identified using different approaches. The two most common cloning strategies are functional cloning (Ruddle,1984) and positional cloning (Collins, 1992). When the primary protein defect is known, the corresponding gene can be cloned using antibodies raised against the protein or oligonucleotide probes against the deduced cDNA sequence (functional cloning). However, in most hereditary diseases hardly anything is known about the protein or biochemical defect, and positional cloning is applied to identify the gene of interest. One of the first genes identified by this approach was the Duchenne muscular dystrophy (DMD) gene (Monaco *et al.*, 1986). In the past 10 years, the human genome project has contributed many DNA polymorphic markers, appropriate physical maps, expressed sequence tags (ESTs) and genomic sequence data that facilitates positional cloning. Nowadays the positional cloning strategy is often combined with the positional candidate approach (Ballabio, 1993). It is expected that a shift will take place in the next decade towards studies that investigate the function of all these new genes, how they are regulated and how their products interact.

2.1 Positional cloning in general

Genetic mapping

Positional cloning comprises different steps and starts with genetic mapping. During this process the inherited trait is localised to a chromosome locus and a candidate region is defined. In some cases, cytogenetically visible chromosomal abnormalities, for example translocations, can give a direct indication of the chromosomal region involved. In most cases, the genome needs to be screened with polymorphic markers (linkage analysis) in multiplex families to find a marker close to the disease gene locus. During linkage analysis, individual meioses are analysed to test whether the trait segregates with any of the polymorphic markers.

After the chromosomal position has been defined, refined genetic mapping is initiated, which involves a detailed study of the most useful recombinant events. Recombination events can be very helpful in defining and narrowing down the critical region. In addition, refinement of the region can also result from large rearrangements, which are usually detected by

fluorescence in situ hybridisation (FISH) and longe range mapping by pulsed field gel electrophoresis (PFGE). The resolution of fine mapping is usually limited to about 1 cM (Collins, 1992), depending on the number of available informative meioses.

Physical mapping

The second step of the positional cloning process involves physical mapping, during which genomic clones are isolated and constructed in an overlapping contig. Some of the most commonly screened libraries consist of yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), P1 clones and cosmid clones. Nowadays, many different genomic gridded libraries are available. Clones up to 100kb like cosmids, P1s and BACs can be fingerprinted by shared restriction fragments. YACs in general have inserts too large to construct detailed restriction and transcript maps and are physically mapped by the use of sequence tagged sites (STSs) or YAC fragmentation (Pavan *et al.*, 1991). More recently FISH-derived methods have been developed that allow finer mapping at the level of the extended single DNA fiber, collectively called FiberFISH (Heiskanen *et al.*, 1996; Heng *et al.*, 1997). An important application of FiberFISH is to order individual genomic clones, and to estimate the size of gaps within contigs, because a very high resolution of 1-400 kb can be obtained (Michalet *et al.*, 1997).

Identification of transcripts

During and after the mapping studies, transcripts can be isolated from the critical region using different techniques. Some of the most commonly used methods are cDNA screening, cDNA selection (expression dependent) and exon trapping (not expression dependent).

The most traditional method to isolate genes from a candidate region is to screen cDNA libraries (cDNA screening). A large selection of probes can be applied to this method. The most widely used probes are of genomic origin, for example single copy fragments, whole cosmid clones and CpG island probes. The technique is simple, but is too labour-intensive for the generation of a transcript map from an extensive candidate region.

Nowadays, the most commonly used large scale gene isolation techniques are cDNA selection (Lovett *et al.*,1991; Parimoo *et al.*,1991) and exon trapping (Duyk *et al.*,1990;

Buckler *et al.*,1991). cDNA selection is based on the hybridisation of a selection of immobilised genomic DNA to a pool of cDNAs. Several direct selection strategies have been described, differing predominantly in the type and preparation of genomic DNA, type of cDNA and whether the hybridisation is performed in solution or on membrane (Lovett *et al.*,1991). Figure 2.1 represents the end ligation coincident sequence cloning (EL-CSC) method (Brookes *et al.*,1994).

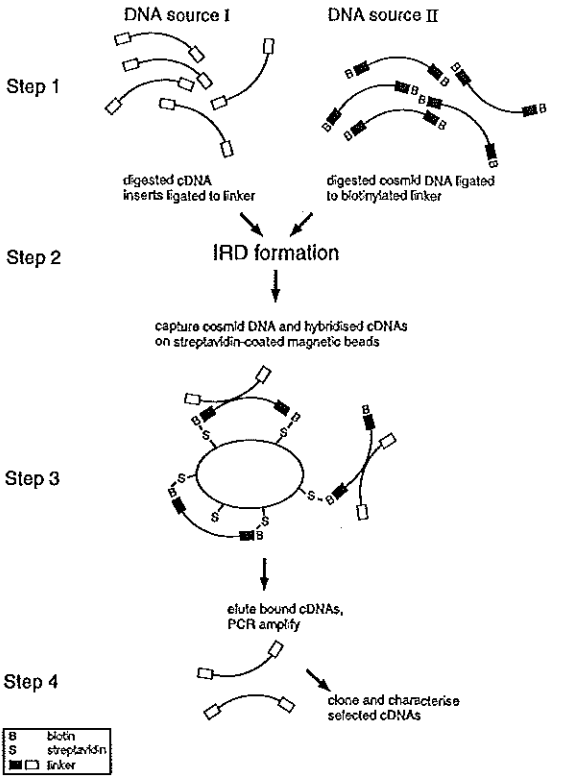


Figure 2.1 cDNA selection protocol (Brookes *et al.*,1994)
Step 1. addition of biotinylated synthetic oligonucleotides (catch-linkers) to the input DNA source II
Step 2. inter-resource duplex (IRD) formation after preblocking of high copy repeat sequences
Step 3. isolation of the complex via the biotin moieties and steptavidin coated magnetic beads
Step 4. end-ligation reaction and PCR amplification

In brief, a selection of cosmids from the candidate region are hybridised in solution to a pool of amplified cDNA from a library or tissue. The cosmids are captured on beads via biotin

moieties and after several washing steps, the cDNAs are eluted, PCR amplified, cloned and characterised.

Exon trapping is based on the detection of coding sequences within genomic DNA, which are selected for by functional splice sites present in the DNA using specific exon-trap vectors. Neither cDNA selection nor exon trapping are sufficient to isolate all the transcripts from a region, and therefore, both techniques are usually combined during a positional cloning effort.

EST mapping and genomic sequencing

The goal of the human genome project (HGP) is to unravel the DNA code from the 24 different human chromosomes. An important aspect is the large-scale sequencing of random cDNAs from various tissues. These expressed sequence tags (ESTs) are partial 5' and 3' sequence of a cDNA clone and are deposited in a public database named dbEST. ESTs that have already been assembled into contigs are present in a separate database called Unigene. Only a limited number of the ESTs have been mapped to chromosomal regions, but ESTs can contribute to the traditional positional cloning strategy by selecting the ones mapping to the region of interest or the ones with an interesting homology. Large scale sequencing of contigs from specific regions of interest is also a development from the last few years and this gives the most detailed information about a candidate region. The raw DNA sequence can be analysed to predict coding sequences (exons) and promotors in the region (GRAIL; Xu *et al.*,1994).

Testing candidate genes

In principle, every gene isolated from a candidate region should be tested for mutations in patients, using different techniques. Southern blot analysis is applied to screen for larger mutations. This method allows for testing a large collection of patients in a short period of time. Smaller mutations are usually screened for by single strand conformation polymorphism analysis (SSCP) (Orita *et al.*,1989) , heteroduplex analysis (HD) (Ganguly *et al.*, 1993) and direct sequencing. Both HD and SSCP rely on changes in electrophoretic mobility due to differences in 3D-structure of the DNA molecules and the majority of small mutations in a gene will be detected by these methods.

Sequence changes leading to a premature stop (deletions, insertions and nonsense mutations) are usually regarded to represent disease-causing mutations. Missense mutations are more difficult to distinguish from polymorphic changes and additional support is required. This can be provided in several steps:

1) In sporadic cases, both parents are tested for the sequence change. A *de novo* change in the patient is usually considered enough evidence.

2) In familial cases, it is required that the mutation segregates with the affected persons in the family and 100-200 unrelated control chromosomes are tested for the same sequence change to evaluate the possibility that the mutation represents a relatively frequent polymorphism.

2.2 Positional cloning applied to the TSC1 gene

2.2.1 Genetic mapping in 9q34

Linkage for TSC was found with the ABO bloodgroup locus and the Abelson oncogene on 9q34 in 1987 (Fryer *et al.*,1987; Connor *et al.*,1987). This locus was denoted TSC1 and subsequently, more markers from this specific region were isolated and tested to find informative markers close to the disease gene (figure 2.2).

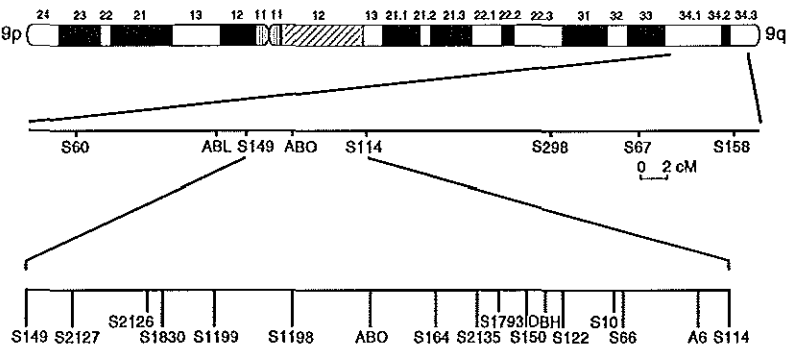


Figure 2.2 The TSC1 region on chromosome 9. G-banded methaphase chromosome 9, with the TSC1 region in 9q34. A detailed map of polymorphic markers in the region is included.

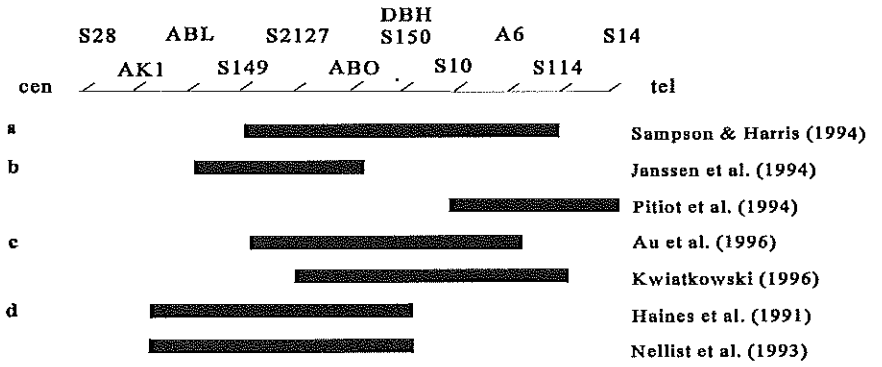


Figure 2.3: Candidate TSC1 intervals, defined by recombination events in families with TSC. The markers defining the different candidate regions are listed at the top (not on scale, cen=centromere; tel=telomere) A) Critical region as defined in 1994 in a review article; B) conflicting critical intervals defined by recombination events in affected individuals. Both recombinants were withdrawn because of an initial false-positive clinical diagnosis (Pitiot) and misclassification of marker data (Janssen); C) independent recombinants in affected TSC patients; D) independent recombinations identified in two families in unaffected family members. The closed boxes represent the critical interval defined by the different recombination reports.

For the TSC1 gene, a region spanning 3-4 cM between markers D9S114 and D9S149 was defined to be the critical region (Sampson and Harris,1994) (figure 2.3a). This interval was supported by a somatic LOH event at 9q34, in which D9S114 and D9S149 were the most proximal and distal markers retained in an astrocytoma from a patient from a TSC1-linked family (Carbonara *et al.*,1996).

Refining the critical region in affected individuals proved to be difficult, because two groups reported conflicting data in chromosome 9-linked families (Janssen *et al.*,1994; Pitiot *et al.*,1994) (figure 2.3b). Novel markers helped to refine the TSC1 region; D9S2127 and A6 were considered to be the centromeric and telomeric boundaries of the interval (Kwiatkowski, personal communication; Au *et al.*,1996) (figure 2.3c).

In addition, two independent recombinations had been described in unaffected individuals in two different TSC families (Nellist *et al.*,1993; Haines *et al.*,1991) (figure 2.3d). In each family, two individuals from successive generations with no evidence of TSC disease carried the same recombinant chromosome. Although the clinical phenotype can be very mild in

some patients, the penetrance of TSC is almost complete. Thus, this observation was sufficient to focus our search on the centromeric part of the region between D9S150 and D9S2127.

2.2.2 Physical mapping in 9q34

Isolation of genomic clones from 9q34

Attempts at YAC cloning revealed a poor representation of this region in various libraries. An added disadvantage was that most of the YACs isolated were chimeric, deleted or rearranged (Nellist, 1994). At that time, a chromosome 9-specific cosmid library became available (Lawrence Livermore Chromosome 9 Cosmid Library; kindly provided by Pieter de Jong) and a cosmid-based strategy was followed as an alternative to build a contiguous contig.

Contig assembly

Five different markers from across the candidate region (D9S149-ABO-DBH-D9S10-D9S114) were chosen as starting points and several cosmid contigs were constructed (chapter 2.6.1). Newly isolated cosmids were *EcoR*I fingerprinted and the original cosmid probe was hybridised to the filter to orientate the cosmids in the contig. End fragments and inter alu PCR products generated from the cosmids were used in a new screening. Another method, *Hinf*I fingerprinting, was employed by our collaborators in London (Nahmias *et al.*,1995). The two different fingerprinting methods resulted in the construction of a 1.7 Mb contiguous cosmid contig, spanning the entire candidate region between D9S149 and D9S114 (Hornigold *et al.*,1997) (chapters 2.5 and 2.6.2).

The gap between the cosmids 180F1 and 50D9 remained uncloned and the size of the gap was determined using dynamic molecular combing (DMC) in combination with fluorescent hybridisation (Michalet *et al.*,1997). With this technique human genomic DNA is being fixed as parallel DNA fibers, aligned in a single direction, on a cover slip. Cosmid probes on both sides of the gap were hybridised to the DNA and the size of the gap was estimated to span 30 kb of DNA (figure 2.4).

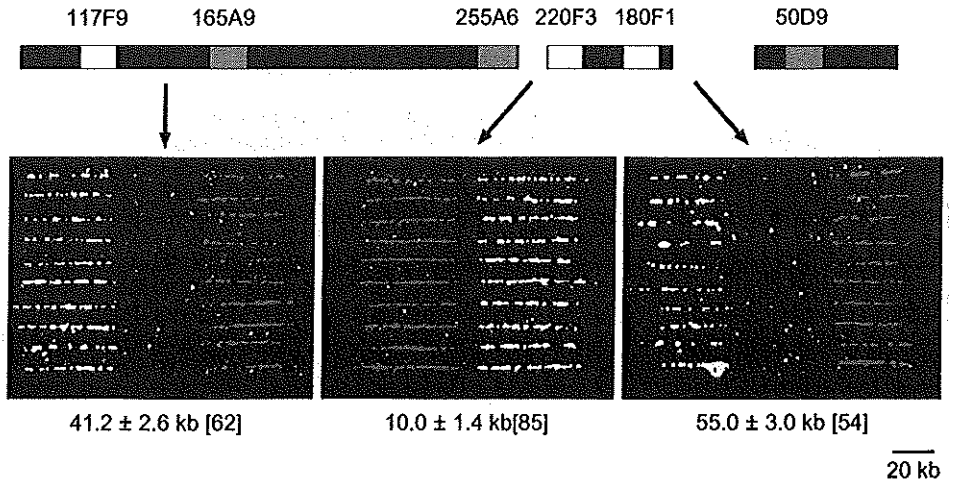


Figure 2.4 Genomic map of part of the 9q34 region. Distances and gaps in the TSC1 cosmid contig were measured on combed total human genomic DNA (Michalet *et al.*,1997). Cosmid probes 117F9 and 165A9 were used as controls, because the distance between those cosmids was known. The gap between cosmids 255A6 and 220F3 and between cosmids 180F1 and 50D9 was estimated to be 10 and 30 kb, respectively.

2.2.3 Identification of transcripts from the cosmid contig

Different complementary gene isolation methods were combined in the different groups in the search for the TSC1 gene. In our group we isolated genes using expression-dependent cDNA identification methods. Genes were isolated by cDNA screening with cosmid-derived hybridisation probes and by cDNA selection (figure 2.1). Other groups in the consortium used expression-independent gene identification techniques, like exon trapping (London, Boston) and genomic sequencing (Whitehead Institute).

cDNA screening

Two different transcripts were isolated from the centromeric part of the critical interval by screening a gridded infant brain cDNA library (gift from Bento Soares). B11 was picked up using a single copy *EcoR*I fragment of 4.5 kb from cosmid 99b11. More sequence data were derived from additional cDNA clones, which had been isolated from a fetal brain library and by computer EST analysis. The B11 gene represented a novel gene, since no sequence homology with any known protein or motif in the database was detected. A second gene, B2, was cloned by hybridising cosmid 115b2 to the infant brain library. The isolation of the full length transcript will be discussed in more detail in chapter 2.2.4, because mutations detected

in this gene identified it as the TSC1 gene.

cDNA selection

Genes isolated by this technique were the carboxyl ester lipase gene (CEL), RalGDS, 9b9b and C10. RalGDS (Humphrey *et al.*,1997) and 9b9b were also isolated by exon trapping. CEL represented a known gene, previously mapped to 9q34 (Taylor *et al.*,1991). We were able to locate this gene more precisely at 9q34 in the cosmid contig, just proximal of ABO. C10 and 9b9b were new transcripts of unknown function, but RalGDS, a member of the ras superfamily of small GTPases, seemed a particulary strong candidate, because the TSC2 gene contains a putative GTPase activating domain. This gene was extensively tested for mutations in TSC patients, but no abnormalities were detected (Humphrey *et al.*, 1997).

Exon trapping

Exon trapping was performed by two of the collaborating groups (Boston and London) on a large part of the cosmid contig. Many exons were isolated from the region and are summarised in the transcript map in chapter 2.5.

Genomic sequencing of the cosmids

The ESTs that mapped to 9q34 and the ones with an interesting homology (good candidates for the TSC1 gene on the basis of homology in the database to other genes or domains) were selected for more precise mapping to the candidate TSC1 interval. The sequencing of our contig resulted in additional ESTs and putative new transcripts in the region. Cosmid sequences were analysed using GRAIL2 (Xu *et al.*, 1994). This program predicts exons, CpG islands, promotor regions and poly-A sites from genomic sequence. About 80% of the 9q34 cosmid contig between markers D9S2127 and A6 has been sequenced and the sequence has been deposited in GenBank.

2.2.4 Testing candidate genes for mutations

All genes isolated by the different groups in the consortium were tested for mutations in TSC patients. In principle, the mutation detection scheme consisted of the search for larger mutations by Southern blotting and the detection of small deletions and point mutations by SSCP, HD and direct sequencing.

Southern blot analysis

Our initial choice of mutation analysis was the search for deletions and insertions using Southern blots. In about 10% of our collection of TSC families and sporadic cases, a mutation has been detected in the TSC2 gene using this method (van den Ouweland, personal communication). DNA from over 200 unrelated patients, not having a TSC2 mutation, was tested for larger mutations in the TSC1 region on 9q34 on Southern blots using four different restriction enzymes *(EcoRI, HindIII, TaqI and PstI)*. All isolated cDNA clones were tested in our set of patient DNAs, but none of the candidate genes showed an aberrant banding pattern.

SSCP and HD analysis

The sequencing project enabled exon predictions in the region to be made and a new strategy was developed to screen every exon from the region for small mutations in a selection of TSC patients from chromosome 9-linked families by SSCP or HD analysis. In one of the largest exons from the region, several shifts were detected in DNA of TSC patients by Kwiatkowski and colleagues. After sequencing, all shifts were shown to reflect truncating mutations and were therefore considered to represent disease-causing mutations. This exon was part of the previously identified B2 gene (see chapter 2.2.3).

2.3 Identification of the TSC1 gene

The 3' end of the B2 gene originated from cosmid 115b2, and serial screenings resulted in the isolation of a 4.5 kb fetal brain cDNA clone. No open reading frame was present, but a major 8.6 kb transcript could be detected on a Northern blot in every tissue tested. The KIAA0243 EST clone of 6.8 kb (Nagase *et al.*,1996) showed 100% homology to the fetal brain clone and contained an open reading frame of 2 kb. Once it was shown that B2 encoded the TSC1 gene by the identification of mutations in TSC patients, the complete cDNA was determined. The 5' end of the TSC1 gene was amplified from fetal brain by 5'RACE (Rapid Amplification of cDNA Ends). All intron-exon boundaries could be identified from the genomic sequence from two cosmids containing the 5' end of the gene (63g10) and the 3' end of the gene (115b2) that had been sequenced. The TSC1 gene is comprised of 23 exons, of which 21 are coding, and spans a genomic distance of 45 kb. The direction of transcription is towards the centromere (figure 2.5).

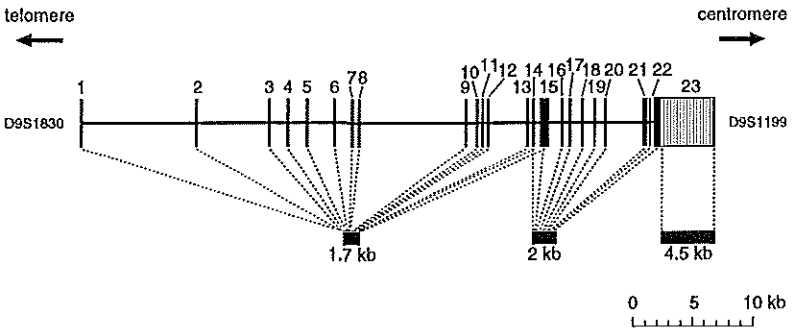


Figure 2.5 Genomic structure of the TSC1 gene. Representation of the TSC1 gene with its 23 exons, of which 21 are coding (not on scale). The startcodon is in exon 3. Three cDNA clones are indicated at the bottom by closed boxes; a 1.7 kb 5' RACE product (exon 1-15); an RT-PCR product of 2 kb (exon14-23); a 4.5 kb fetal brain clone (3' untranslated region, striped box).The TSC1 gene spans about 45 kb of genomic DNA.

(A)

[exon 1a] [exon 1b] [exon 2]

5'-tggagggactgtgag **gtaaacagctgaggggggaggagacggtgg** tgaccatgaaagacaccaggttgacagcactggaaactgaagtaccagttg tcgctagaacag **tttggtagtggccccaatgaagaaccttcagaacctgtagcacac**.....-3'

[exon 3]

Different splice forms;

(B)

[exon 1a] [exon 2]

5'-gtgctgtacgtccaagatggcggcgcctgtaggctggagggactgtgag tgaccatgaaagacaccaggttgacagcactggaaactgaagtaccagttgt cgctagaacag **tttggtagtggccccaatgaagaaccttcagaacctgtagcacac**.....-3'

[exon 3]

(C)

[exon 1b] [exon 3]

5'-**cagctgaggggggaggagacggtgg tttggtagtggccccaatgaagaaccttcagaacctgtagcacac**.....-3'

Figure 2.6 Different splice variants in the 5'UTR. (A) Most extended construct; exon1a and 1b are contiguous on the genomic map. The startcodon of the TSC1 gene is in exon3. **(B)** part of exon 1 (1b) is spliced out. **(C)** exon 2 is spliced out.

Several alternative splice forms that are variable at the 5' end have been isolated (figure 2.6; unpublished results). The functional significance of these alternative forms remains to be investigated.

2.4 Mutational spectrum of the TSC1 gene

SSCP analysis of the 21 coding exons was performed in our collection of 225 unrelated TSC patients and 29 mutations were detected. All types of small mutations leading to a premature stop have been observed: small deletions/insertions, nonsense mutations and splice-site mutations (figure 2.7), as well as a small number of potential missense mutations.The majority of the mutations are clearly inactivating and about half of the mutations are clustered in exons 15 and 17. The disease-causing mutation has been detected in 17% of our familial cases (small families and families linked to 9q34) and in 11% of sporadic TSC patients. After clinical evaluation of the patients with a TSC1 mutation, we find no evidence for a genotype-phenotype correlation in TSC1 disease (chapter 2.6.4).

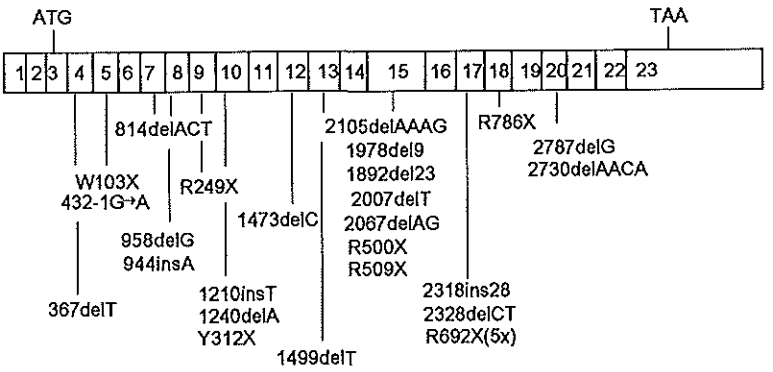
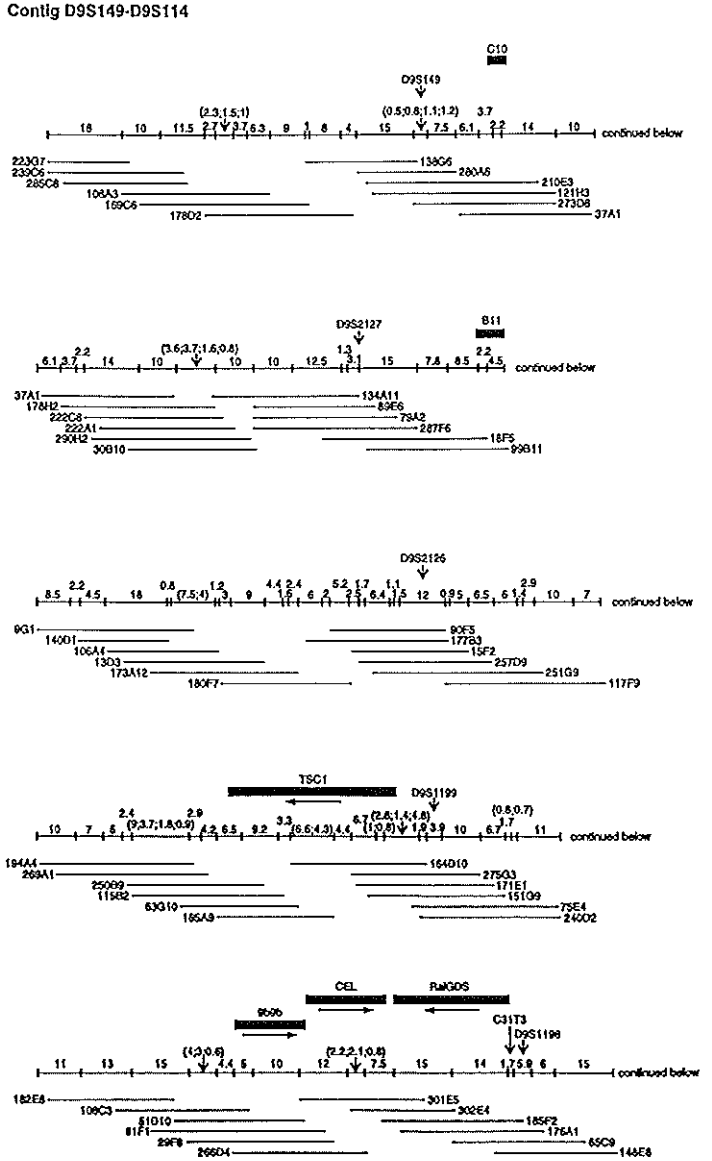
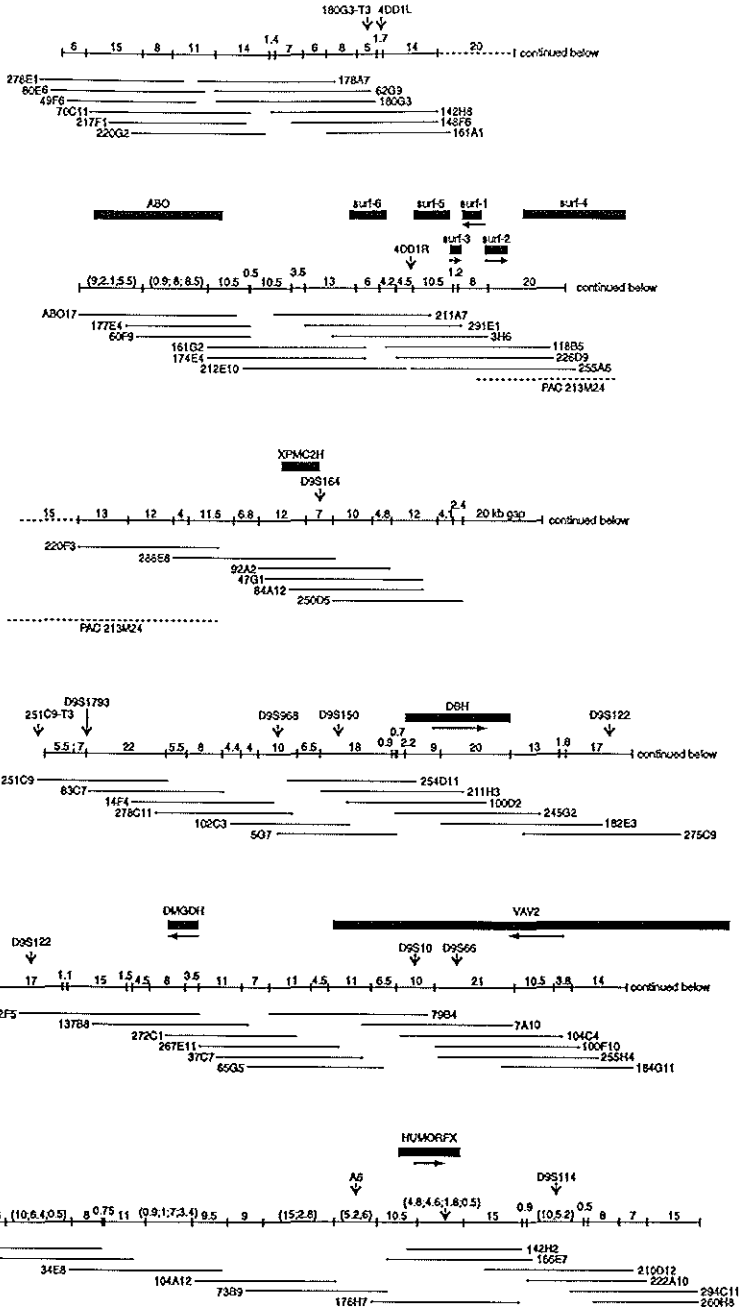


Figure 2.7 Distribution of mutations in the TSC1 gene The nomenclature of the mutations is according to the Ad Hoc Committee on Mutation Nomenclature. (1996). The clinical data of the patients with a mutation in the TSC1 gene are summarised in chapter 2.6.4.

2.5 A transcript map in 9q34



Detailed *EcoRI* restriction map of the 1.7 Mb cosmid contig covering the TSC1 region in 9q34. Cosmids are shown below the *EcoRI* map. Vertical arrows represent RFLP markers, STSs and microsatellites. Genes are shown above the restriction map as thick bars and the direction of transcription, when known, is indicated by an arrow.


180G3-T3
4DD1L
continued below
278E1
178A7
80E6
62G9
49F6
180G3
7CC11
142H8
217F1
148F6
220G2
161A1
ABO
surf-6
surf-5
surf-1
surf-4
surf-3
surf-2
4DD1R
continued below
ABO17
211A7
177E4
291E1
60F9
3H6
161G2
118B5
174E4
226D9
212E10
255A5
PAC 213M24
XPMC2H
D9S164
20 kb gap
continued below
220F3
285E8
92A2
47G1
84A12
250D5
PAC 213M24
D9S1793
251C9-T3
D9S968
D9S150
D9H
D9S122
continued below
251C9
254D11
83C7
211H3
14F4
100D2
278C11
245G2
102C3
182E3
5G7
275C9
D9S122
D9S10
D9S66
continued below
152F5
79B4
137B8
7A10
272C1
104C4
267E11
100F10
37C7
255H4
65G5
184G11
HUMORFX
A6
D9S114
127G4
142H2
122D6
165E7
34E8
210D12
104A12
222A10
73B9
294C11
176H7
260H8

CHAPTER 2.6 Publications

CHAPTER 2.6.1

Cosmid Contigs from the Tuberous Sclerosis Candidate Region on Chromosome 9q34

Marjon van Slegtenhorst, Bart Janssen, Mark Nellist, Sarvan Ramlakhan, Caroline Hermans, Arjenne Hesseling, Ans van den Ouweland, David Kwiatkowski, Bert Eussen, Julian Sampson, Pieter de Jong and Dicky Halley.

Published in: Eur J Hum Genet 1995;3;78-86

Reprinted with permission

Abstract

Tuberous Sclerosis (TSC) is a heterogeneous multisystem disorder, with loci on 9q34 (TSC1) and 16p13.3 (TSC2). The TSC2 gene has recently been isolated, while the TSC1 gene has been mapped to a 5-cM region between the markers D9S149 and D9S114. In our effort to localise and clone TSC1, we have obtained three adjacent cosmid contigs that cover the core of the candidate region. The 3 contigs comprise approximately 600 kb and include 80 cosmids, 2 P1 clones, 1 YAC, 5 anonymous markers and 4 sequence-tagged sites. The ABO blood group locus, the Surfeit gene cluster, the dopamine ß hydroxylase gene (DBH) and VAV2, a homologue of the *vav* oncogene, have all been mapped within the contigs. Exon trapping and mutation screening experiments, aimed at identifying the TSC1 gene, are currently in progess.

Introduction

Tuberous sclerosis (TSC) is an autosomal dominant multisystem disorder. The brain, skin, heart and kidneys are often affected and almost all other tissues and organs may be involved, except muscle syncytia [1]. The disease shows a high penetrance with variable expression and is known for its locus heterogeneity, with one locus mapping to chromosome 9q34 (TSC1) and another to chromosome 16p13.3 (TSC2) [2]. The number of families linked to each locus is approximately equal and there is no significant evidence for a third locus [3]. The TSC2 gene has been isolated [4] and both genes may act as growth- or tumour-suppressors, since loss of heterozygosity (LOH) has been demonstrated on 9q34 [5-7] or 16p13 [7] in various hamartomatous tissues from patients with TSC.

The chromosome 9 locus for tuberous sclerosis, TSC1, is tightly linked to the ABO blood group locus [8] and maps in a gene-rich region on chromosome 9q34. Since the initial linkage report by Fryer et al. [8], the TSC1 region has been refined to a region of 5 cM between D9S149 and D9S114 [3,9-14]. However, there is no consensus on the exact position of TSC1 within this interval, since some groups have found recombinants in favour of a position proximal to ABO and the dopamine ß hydroxylase gene (DBH), while other groups have presented data supporting a location distal of these markers [15]. The conflicting observations have several possible causes, including misclassification of individuals with only minor clinical findings or non-linkage of one or more families.

Several genes have been mapped within the TSC1 candidate region, including ABO,

DBH, the Surfeit gene cluster and VAV2 [16-18], while other disorders genetically linked to ABO include torsion dystonia [19] and nail patella syndrome [20-21].

In this paper we present the results of a contig assembly and gene mapping effort, focused on part of the TSC1 candidate region around ABO and DBH. Our detailed map spans 600 kb, corresponding to more than 2 cM of the TSC1 critical region. Eight genes and several known and novel genetic markers have been precisely positioned on a genomic *EcoRI* map between D9S149 and D9S114.

Materials and Methods

Libraries

The ICI YAC library [22] was accessed through the UK Human Genome Mapping Resource Centre and sets of primary, secondary and tertiary pools for PCR screening were provided by R. Elaswarapu. Primary pools from the St. Louis YAC library [23] were supplied by J. den Dunnen from the Department of Human Genetics in Leiden. The P1 library was made from human foreskin fibroblast DNA [24]. The library was gridded into 125 96-well plates with approximately 12 different P1 clones per well and pools were made for PCR screening. The chromosome 9-specific cosmid library LL09NC01"P" was constructed at the Biomedical Sciences Division, LLNL, Livermore, CA 94550, USA under the auspices of the National Laboratory Gene Library Projects sponsored by the US Department of Energy. The library was replicated on gridded filters as described [25] at the YAC screening core of the Department of Human Genetics in Leiden. Two sets of membranes were used to make pools for PCR screening [26]. The nomenclature of the cosmids in the contigs is the same as the nomenclature of the library source from which they were obtained. Cosmid ABO17 was provided by J. Wolfe.

Cosmid Library Screening

Hybridisation probes were generated by inter Alu PCR [27] using primers CL1 and CL2 [28] or by isolating end fragments from cosmids in low-melting agarose. Probes were randomly labeled, competed with total human DNA, hybridised to nylon filters and washed using standard procedures [29]. Cosmid library screening by PCR was performed by screening two dimensional pools of clones as described by Green and Olson [30].

YACs, P1 and Cosmid Clones

Cosmid and P1 DNA was prepared, isolated and fingerprinted using standard techniques [29]. YACs, P1 and cosmid clones were mapped back to 9q34 by fluorescence in situ hybridisation (FISH) [31].

Sequence-Tagged Sites (STSs)

STSs were developed by YAC end rescue inverse PCR or direct sequencing of cosmids. YAC end rescue was performed as described by Silverman et al. [32] and the products were sequenced directly. The sequence was derived from the cosmid clones by cycle sequencing (Pharmacia) with the appropriate vector primers.

Result and Discussion

Strategy

We aimed to isolate a significant part of the TSC1 critical region between the markers D9S149 and D9S114 on 9q34. Several additional markers are known to map between these two, but have not been convincingly associated with genuine recombination events. The initial strategy was to isolate the region in YACs, P1 and cosmid clones. However, attempts to obtain YACs were hampered by underrepresentation of the region in the available libraries. This prompted us to follow an alternative strategy which consisted of cosmid walking complemented by screening P1 libraries.

YAC Library Screening

Two YACs from the ICI library, 4DD1 (120kb) and 25DG9 (320kb), were identified with primers specific for the ABO locus. STS mapping using primer pairs from both ends of the YACs indicated that the left ends of both inserts overlapped, however inter-Alu PCR in combination with hybridisation experiments suggested that the region of overlap was small. FISH analysis confirmed the localisation of both YACs to chromosome 9q34, however 25DG9 gave an additional signal on chromosome 18 indicating chimerism. This clone was not investigated further. No additional YACs were identified in the ICI library using the end clone STSs from 4DD1 or 25DG9, or with additional markers from the TSC1 candidate region (D9S10, D9S66, DBH). An STS derived from the left arm of YAC 4DD1 was used to screen the St. Louis YAC library and identified two duplicate clones, 51A7 and 61A10

(200kb). FISH analysis mapped 51A7 to 9qter, however STS mapping experiments using primers derived from the right arm of this clone suggested that it contained a large deletion (data not shown) and it was not investigated further. It is interesting to note that the TSC2 locus on chromosome 16 was also found to be underrepresented in YAC libraries (unpubl. results).

Contig Assembly

Starting points for cosmid contig assembly were ABO, DBH and D9S10 (fig. 1). Cosmids were identified with both the left and right end clones of YAC 4DD1 and two contigs were constructed of 110kb and 130kb respectively (fig. 2, contig A and B). The orientation of the cosmid contigs was consistent with results from YAC inter-Alu PCR screening of the cosmid library and with the YAC STS mapping experiments. No cosmids could be identified distal of cosmid 255A6 (contig B). Only a single non-rearranged cosmid and a single P1 clone were detected at the ABO locus, and no clone could be detected linking the two contigs. However, from the size of the 4DD1 YAC and direct visual hybridisation (DIRVISH) experiments of streched DNA [33] (unpubl. results), we estimated that the gap is approximately 30kb.

Cosmids were identified with the DBH cDNA and probe pMCT136 from the D9S10 locus. DBH and D9S10 map 1 and 2 cM distal of ABO, respectively, and were linked by chromosome walking, covering a physical distance of 150 kb (contig C). This indicates that the genetic versus physical distance ratio in this region is large. The contig was extended proximal of DBH by 125 kb, but could not be extended further towards ABO. We did isolate a P1 clone with an STS from the proximal end of 251C9, but could not bridge the gap. The distance between clone 251C9 (contig B) and 255A6 (contig C) could not be resolved by interphase FISH, indicating that the gap between contig B and C is small. DIRVISH DNA mapping experiments are in progress to estimate the size of the gap.

In regions of overlap the contigs presented here were consistent with the cosmid contigs constructed by *HinfI* fingerprinting as described by Nahmias et al. [34]. They need at least 50% overlap between cosmids before the clones are joined in a contig. Our data are more detailed and detect smaller overlaps. Additional cosmids have been isolated from the flanking locus D9S149. Chromosome-walking experiments are currently focussed on closing the gap between D9S149 and the most proximal ABO contig (contig A).

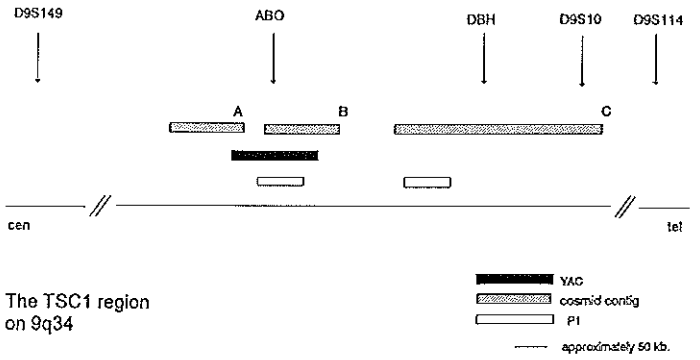


Fig. 1. Schematic representation of the TSC1 region. The starting points for YAC and cosmid walking are indicated, together with the YAC 4DD1, the cosmid contigs and P1 clones.

Mapping of Markers and Genes in the Contigs

RFLPs and unique STSs are listed in tables 1 and 2. The STSs 180G3-T3 and 4DD1L map to adjacent EcoRI fragments in contig A. Two additional STSs, 4DD1R and 251C9-T3 were mapped to contigs B and C respectively. Existing minisatellite repeats (D9S122 and D9S150) [13] were precisely positioned within this contig (fig. 2, contig C) and a *HindIII* polymorphism (D9S968) was detected immediately proximal of DBH.

Table 1. List of RFLPs in the region

Locus	Enzyme	Probe	Fragment sizes, kb	Heterozygosity,%
D9S10	*PstI**	MCT136	2.5 and 2.3	50
	HindIII	MCT136	2.2 and 2.0	50 (200 chromosomes)
	(RFLPs show linkage disequilibrium)			
D9S968	*HindIII*	RD560	4.5 and 2.6	14 (115 chromosomes)
DBH	(several RFLPs* + $(CA)_n$, all listed in GDB)			
VAV2	*PstI*	5' VAV2 (bases 1-865)	5, 4.2 and 2.2	48 (>100 chromosomes)

All RFLPs marked with an asterisk are already listed in GDB. The heterozygosity percentages of the new RFLPs (without asterisk) have been determined in at least 100 chromosomes from Causasians. The map position of each locus is indicated in figures 1 and 2. The VAV2 RFLP maps within the VAV2 gene, distal to the end of the cosmid contig.

Table 2. List of STSes in the region

STS	Primer sequences	Product length	Map position
180G3-T3	5' GGTGT GGTTC TCCCA AGGG 3' GAGAG AGGCT TCCTG CTTGC	128 bp	distal part of contig A
4DD1L	5' CCAAG GGAAG CTGGA GAAGT 3' CCAGA CCCAG CCTAC ATTTC	97 bp	left arm of YAC 4DD1
4DD1R	5' CATGC TGTTG GCACT GTTGTA 3' TTTCT CTTTG GCTTC CCTCTT	135 bp	right arm of YAC 4DD1
251C9-T3	5' GGAAA GAGGA GCGAG GAAG 3' CACAA TCTCA CAGTG AATGCC	152 bp	proximal end of contig C

A number of polymorphic STSs at ABO, DBH, VAV2, D9S149, D9S150, D9S122, D9S66 and D9S114 have been described previously and are therefore not included in the list.

The position and orientation, where known, of genes identified within the contigs are indicated in figure 2. The role and expression pattern of the ABO blood group transferase indicate that it is not a good candidate for TSC1. The Surfeit gene cluster had been previously mapped by in situ hybridization telomeric to the *c-abl* and *can* genes on 9q34 [35]. A oligonucleotide derived from the Surf-3 cDNA sequence detected a 1.2 kb *EcoRI* fragment in several cosmids, slightly distal to ABO in contig B. Cosmid 255A6 was digested with *XbaI* to orientate the cluster in the map. In the mouse this cluster consists of 6 house keeping genes, which are unrelated by sequence homology [35]. To date the Surfeit genes form the tightest gene cluster known in mammals. Since these genes are in the critical region of TSC1 and not much is known about their function, mutation analysis in TSC patients must be considered.

Our *EcoRI* mapping data from the DBH locus is consistent with that of Kobayahi et al. [36]. The direction of transcription is towards the telomere. The role of DBH in the conversion of dopamine to noradrenaline and the neurological manifestations of TSC led to the proposal that DBH could be a candidate for the TSC1 gene [37]. However, more recent results suggest that TSC1 maps either distal or proximal of DBH and consequently DBH is not such an attractive candidate.

Exon trapping [38] efforts using our cosmids from the D9S10 locus identified a gene homologous to the *vav* oncogene [16]. This gene, designated VAV2, was considered a good candidate for the TSC1 gene. However, intensive screening failed to identify any mutations, and VAV2 was consequently excluded as a candidate gene for TSC1 [16-17].

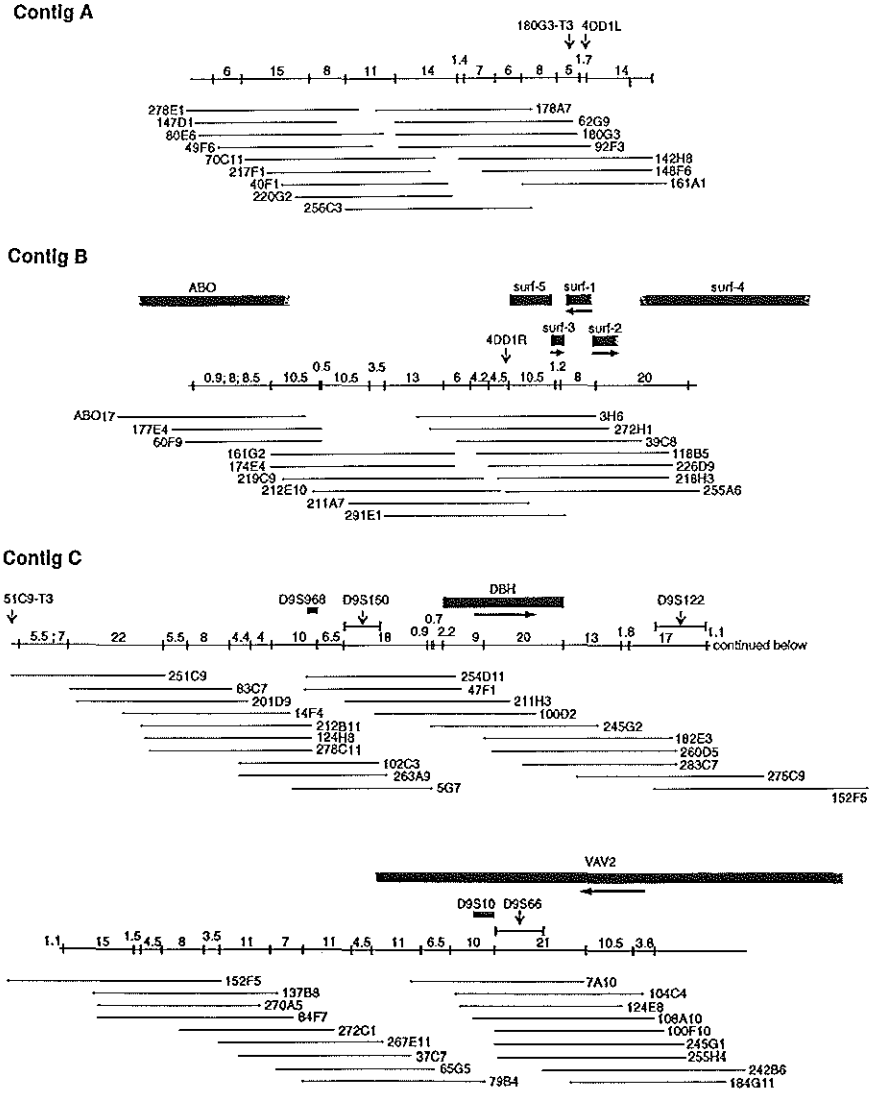


Fig. 2. Detailed *EcoRI* restriction map of the three contigs described in this paper. Cosmids are shown below the the *EcoRI* map. Thin bars represent RFLP markers and vertical arrows indicate STSs and microsatellites. Genes are shown above the restriction map as thick bars. The size of the bars indicates the maximal genomic extent. The direction of transcription is indicated by arrowheads. For DBH, *surf-1*, *surf-2*, *surf-3* and VAV2, the gene structure was studied by Nahmias et al. [34], Yon et al. [35] and Kwiatkowski et al. [16]. The position and orientation of the genes in the cosmid contigs were deduced from our experiments and previously published restriction maps [34,35]

Eight different genes could be placed on the map. The region is gene dense and although some genes map extremely close to each other, we can not exclude the presence of other, as yet unidentified, expressed sequences in the same region. Experiments to identify and characterise additional genes from the TSC1 candidate region are in progress.

Further efforts are directed towards extending the contigs and screening TSC patients for mutations by pulsed-field gel electrophoresis using novel probes derived from our cloned material. The identification of large deletions at the TSC2 locus made a significant contribution to the rapid isolation of the TSC2 gene [4].

In conclusion we have identified 80 cosmids, 2 P1 clones and a single non-rearranged YAC from the TSC1 candidate region on 9q34. We have constructed a detailed restriction map of three adjacent cosmid contigs and oriented the maps with respect to known and previously unidentified genes and DNA markers. We have shown that DBH and D9S10, previously estimated to be 1 cM apart, are separated by less than 300 kb, and estimate that the physical distance between ABO and DBH is less than 300 kb.

In conjunction with the accompanying article [34] we have shown that cosmid walking, using a large chromosome specific cosmid library can provide almost complete coverage of a large genomic region. This minimises the need to search for non-chimeric non-rearranged YAC clones, which have been difficult to obtain from the TSC1 region. Moreover, our contigs and the associated maps provide a good tool for generating novel markers and cloning additional genes from this region. It would be of great help to get more excluding data on the recombinants within the region, so that the search for TSC1 can be restricted to a smaller area. LOH studies in tumours of patients and the development of new polymorphic CA repeats in the area, especially between ABO and D9S149, could help reduce the critical region. Ultimately it is hoped that this work will lead to the identification of the TSC1 gene.

Acknowledgements

We are grateful to Professor H. Galjaard and Professor D. Lindhout for their continuous support. We like to thank Dr. S. Povey, Dr. J. Wolfe, Joseph Nahmias and co-workers for sharing unpublished data and their cooperation. This work was funded by the Dutch Organisation for Scientific Research (NWO).

References

1 Gomez MR: Tuberous sclerosis 2nd edition. New York, Raven Press 1988.

2 Kandt RS, Haines JL, Smith M, Northrup H, Gardner RJM, Short MP, Dumars K, Roach ES, Steingold S, Wall S, Blanton SH, Flodman P, Kwiatkowski DJ, Jewell A, Weber JL, Roses AD, and Pericak-Vance MA: Linkage of an important gene locus for tuberous sclerosis to a chromosome 16 marker for polycystic kidney disease. Nature Genet 1992, 2:37-41.

3 Janssen LAJ, Sampson J, van der Est M, Deelen W, Verhoef S, Daniels I, Hesseling A, Brook-Carter P, Nellist M, Lindhout D, Sandkuijl L, Halley D: Refined localization of TSC1 by combined analysis of 9q34 and 16p13 data in 14 tuberous sclerosis families. Hum Genet 1994;94:437-440.

4 The European Chromosome 16 Tuberous Sclerosis Consortium: Identification and characterisation of the tuberous sclerosis gene on chromosome 16. Cell 1993;75:1305-1315.

5 Carbonara C, Longa L, Grosso L, Borrone C, Garrè M, Brisigotti M, Migone N: 9q34 loss of heterozygosity in a tuberous sclerosis astrocytoma suggests a growth suppressor-like activity also for the TSC1 gene. Hum Mol Genet 1994;3:1829-1832.

6 Green AJ, Johnson PH, Yates JRW: The tuberous sclerosis gene on chromosome 9q34 acts as a growth suppressor. Hum Mol Genet 1994;3:1833-1834.

7 Green AJ, Smith M, and Yates JRW: Loss of heterozygosity on chromosome 16p13.3 in hamartomas from tuberous sclerosis patients. Nature Genet 1994;6:193-196.

8 Fryer AE, Chalmers A, Connor JM, Fraser I, Povey S, Yates AD, Yates JRW, Osborne JP: Evidence that the gene for tuberous sclerosis is on chromosome 9. Lancet i 1987:659-661.

9 Connor JM, Pirrit LA, Yates JRW, Fryer AE, Ferguson-Smith MA: Linkage of the tuberous sclerosis locus to a DNA polymorphism detected by *c-abl*. J Med Genet 1987;24:544-546.

10 Haines JL, Short MP, Kwiatkowski DJ, Jewell A, Andermann E, Bejjani B, Yang C-H, Gusella JF, Amos JA: Localization of one gene for tuberous sclerosis within 9q32-9q34, and further evidence for heterogeneity. Am J Hum Genet 1991;49:764-772.

11 Northrup H, Kwiatkowski DJ, Roach ES, Dobyns WB, Lewis RA, Herman GE, Rodriguez E, Daiger SP, Blanton SH: Evidence for genetic heterogeneity in tuberous sclerosis: one locus on chromosome 9 and at least one locus elsewhere. Am J Hum Genet 1992;51:709-720.

12 Sampson JR, Janssen LAJ, Sandkuijl LA and the Tuberous Sclerosis Collaborative Group: Linkage investigation of three putative tuberous sclerosis determining loci on chromosome 9q, 11q and 12q. J Med Genet 1992;29:861-866.

13 Povey S, Armour J, Farndon P, Haines JL, Knowles M, Olopade F, Pilz A, White JA, members of the Utah Genome Center Genetic Marker and Mapping group, Kwiatkowski DJ: Report on the Third International Workshop on Chromosome 9. Ann Hum Genet 1994;58:177-250.

14 Janssen LAJ, Sandkuyl LA, Merkens EC, Maat-Kievit JA, Sampson JR, Fleury P, Hennekam RCM, Grosveld GC, Lindhout D, Halley DJJ: Genetic heterogeneity in tuberous sclerosis. Genomics 1990;8:237-242.

15 Sampson JR, Harris PC: The molecular genetics of tuberous sclerosis. Hum Mol Genet 1994;3:1477-1480.

16 Kwiatkowski DJ, Short MP, Jozwiak S, Bovey CM, Ramlakhan S, Haines JL, Henske EP. Identification of VAV2 on 9q34 and its exclusion as the tuberous sclerosis gene TSC1. Ann Hum Genet 1995;59:25-37.

17 Smith M, Brickey W, Handa K, Gargus JJ: Sequence analysis of MCT 136 (locus D9S10) in the TSC gene region reveals amino acid domains with sequence homology to RAS activating signal molecules VAV and HSOS. Ann Hum Genet 1994;58:235-236.

18 Woodward KJ, Nahmias J, Hornigold N, West L, Pilz AJ, Kwiatkowski DJ, Benham F, Wolfe J, Povey S: Mapping chromosome 9q34 by FISH using metaphase chromosomes with specific translocation breakpoints. Ann Hum Genet 1994;58:241-242.
19 Ozelius L, Kramer PL, Moskowitz CB, Kwiatkowski DJ, Brin MF, Bressman SB, Schuback DE, Falk CT, Risch N, de Leon D, Burke RE, Haines J, Gusella JF, Fahn S, Breakfield: Human gene for torsion dystonia located on chromosome 9q32-q34. Neuron 1989;2:1427-1434.
20 Renwick JH, Lawler SD: Genetical linkage between the ABO and nail-patella loci: Ann Hum Genet 1955;19:312-331.
21 Renwick JH, Schulze J: Male and female recombination fractions for the nail patella: ABO linkage in man. Ann Hum Genet 1965;28:379-392.
22 Anand R, Riley JH, Butler R, Smith JC, Markham AF: A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. Nucleic Acids Res 1990;18:1951-1956.
23 Burke DT, Carle GF, Olson MV: Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 1987;236:806-812.
24 Pierce JC, Sauer B, Sternberg N: A positive selection vector for cloning high molecular weight DNA by the bacteriophage P1 system: Improved cloning efficacy. Proc Natl Acad Sci USA 1992;89:2056-2060.
25 Bentley DR, Todd C, Collins J, Holland J, Dunham I, Hassock S, Bankier A, Giannelli F: The development and application of automated gridding for efficient screening of yeast and bacterial ordered libraries. Genomics 1992;12:534-541.
26 Kwiatkowski DJ, Zoghbi HY, Ledbetter SA, Ellison KA, Chinault AC: Rapid identification of yeast artificial chromosome clones by matrix pooling and crude lysate PCR. Nucleic Acids Res 1993;18:7197-7203.
27 Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT: Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. Proc Natl Acad Sci USA 1989;86:6686-6690.
28 Lengauer C, Riethman HC, Speicher MR, Taniwaki M, Konecki D, Green ED, Becher R, Olson MV, Cremer T: Metaphase and interphase cytogenetics with Alu-PCR-amplified yeast artificial chromosome clones containing the BCR gene and the protoonocogene *c-raf-1*, *c-fms* and *c-erbB-2*. Cancer Res 1992;52:2590-2596.
29 Sambrook J, Fritsch EF, Maniatis T: Molecular cloning: A Laboratory Manuel 2nd edn. Cold Spring Harbour Press 1989, New York.
30 Green ED, Olson MV: Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. Proc Natl Acad Sci USA 1990;87:1213-1217.
31 Breen M, Arveiler B, Murray I, Gosden JR, Porteous DJ: YAC mapping by FISH using Alu-PCR-generated probes. Genomics 1992;13:726-730.
32 Silverman GA, Jockel JI, Domer PH, Mohr RM, Taillon-Miller P, Korsmeyer SJ: Yeast artificial chromosome cloning of a two-megabase-size contig within chromosomal band 18q21 establishes physical linkage between BCL2 and plasminogen activator inhibitor type 2. Genomics 1991;9:219-228.
33 Wiegant J, Kalle W, Mullenders L, Brookes S, Hoovers JMN, Dauwerse JG, van Ommen GJB, Raap AK: High-resolution *in situ* hybridization using DNA halo preparations. Hum Mol Genet 1992;1:587-591.
34 Nahmias J, Hornigold N, Fitzgibbon J, Woodward K, Pilz A, Griffin D, Henske EP, Nakamura Y, Graw S, Florian F, Benham F, Povey S, Wolfe J: Cosmid contigs spanning 9q34 including the TSC1 candidate region. Eur J Hum Genet 1995;3:65-77.
35 Yon J, Jones T, Garson K, Sheer D, Fried M: The organization and conservation of the human Surfeit gene cluster and its localization telomeric to the *c-abl* and *can* proto-oncogenes at chromosome band 9q34.1. Hum Mol Genet 1993;2:237-240.

36 Kobayashi K, Kurosawa Y, Fujita K, Nagatsu T: Human dopamine beta hydroxylase gene: two mRNA types having different 3' terminal regions are produced through alternative polyadenylation. Nucleic Acids Res 1989;17:1089-1102.
37 Janssen LAJ, Nellist M, Eussen BE, Ramlakhan S, Sampson JR, Hesseling-Janssen ALW, Verhoef S, Lindhout D, Halley DJJ: The map position of three candidate genes for tuberous sclerosis 1: XPAC, DBH and TAN1. Cyt Cell Genet 1993;64:115.
38 Buckler AJ, Chang DD, Graw SL, Brook JD, Haber DA, Sharp PA, Housman DE: Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc Natl Acad Sci USA 1991;88:4005-4009.

CHAPTER 2.6.2

A 1.7 Megabase Sequence-Ready Cosmid Contig Covering the TSC1 Candidate Region in 9q34

N. Hornigold, M. van Slegtenhorst, J. Nahmias, R. Ekong, S. Rousseaux, C. Hermans, D. Halley, S. Povey, J. Wolfe.

Published in: Genomics 1997;41;385-389

Reprinted with permission

GENOMICS 41, 385–389 (1997)
ARTICLE NO. GE974681

A 1.7-Megabase Sequence-Ready Cosmid Contig Covering the TSC1 Candidate Region in 9q34

N. Hornigold,* M. van Slegtenhorst,† J. Nahmias,‡ R. Ekong,‡ S. Rousseaux,‡ C. Hermans,† D. Halley,† S. Povey,‡ and J. Wolfe*,[1]

*The Galton Laboratory, Department of Biology and ‡MRC Human Biochemical Genetics Unit, University College London, London, United Kingdom; and †Department of Clinical Genetics, Erasmus University, Dr. Molewaterplein 50, NL-3015, GE Rotterdam, The Netherlands

Received September 25, 1996; accepted January 30, 1997

The disease gene TSC1 has been genetically mapped to human chromosome region 9q34, in a 4-cM interval between the markers D9S149 and D9S114. Within this interval there is conflicting genetic evidence as to the finer localization of the gene. We have used fingerprinting methods and hybridization to produce a 1.7-Mb overlapping clone map covering the TSC1 candidate region, with a single gap of 20 kb. We have localized 12 previously cloned genes and 17 genetic markers on this map and have confirmed the order of the genetic map. This deep set of overlapping clones is now ready to be used for candidate gene isolation, for transcription studies, or for sequencing. © 1997 Academic Press

INTRODUCTION

Tuberous sclerosis (TSC) is an autosomal dominant disorder characterized by a variety of skin signs, by mental handicap, and by seizures, but the severity of the symptoms is very variable. The incidence of TSC may be as high as 1 in 6000 live births (Osborne *et al.*, 1991). TSC shows locus heterogeneity, with approximately 50% of cases being caused by each of two genes (Povey *et al.*, 1994a). TSC1 shows close genetic linkage to the ABO blood group locus in 9q34 (Fryer *et al.*, 1987). TSC2 is on chromosome 16 and has been isolated (The European Chromosome 16 Tuberous Sclerosis Consortium, 1993).

Analysis of meiotic breakpoints in families where TSC shows clear linkage to chromosome 9 suggests that the disease gene must lie in a 4-cM interval between the loci D9S149 and D9S114 (Povey *et al.*, 1994b). Recombinations have been published that place TSC1 distal to DBH (Kwiatkowski *et al.*, 1993; Gilbert *et al.*, 1993) and distal to D9S122 (Pitiot *et al.*, 1994). Other reports have suggested that TSC1 lies proximal to DBH (Nellist *et al.*, 1993, and J. R. Sampson, Cardiff, pers. comm., 1996). Because of this uncertainty we have decided to construct a physical map across the whole of the region from D9S149 to D9S114 to confirm the marker order given by the genetic data and to provide a resource for further work to isolate the gene.

A number of contiguous arrays of cosmid clones (contigs) in this region have been described in Nahmias *et al.* (1995) and in van Slegtenhorst *et al.* (1995). These contigs have now been extended by probing the LL09NC01"P" library with inter *Alu* PCR or YAC end rescue products from selected YAC and PAC clones and with further cosmid end probes. This has allowed us to construct a deep overlapping clone map, consisting mainly of cosmid clones but including three YAC clones, one P1 clone, one BAC clone, and three PAC clones, covering the TSC1 candidate interval between D9S149 and D9S114, with a single gap of 20 kb, between D9S164 and D9S150.

MATERIALS AND METHODS

Sources of clones. Clones previously in contigs came from a number of sources (Nahmias *et al.*, 1995, van Slegtenhorst *et al.*, 1995). YAC clones 35HG8 and 15GF1 and cosmids derived from the YAC C11-C (Zhou *et al.*, 1995) were provided to us by R. Furlong. PAC clones 63F12, 145N8, and 213M24 were obtained by PCR screening of the PAC human genomic library constructed by Pieter de Jong, with primers for D9S164 and for P6 (M. Smith, unpublished primer sequences). PCR pools and clones were accessed via the HGMP Resource Centre. P1 clone DMPC-HFF1-0529F4, which was positive for ABO, was provided by H. Clausen (Bennet *et al.*, 1995). BAC clone 9E21 was provided by J. Korenburg. All other new cosmid clones were obtained by screening the LL09NC01"P" chromosome 9-specific library.

Cosmid library screening and contig construction. Hybridization probes from YAC and PAC clones were generated by inter *Alu* PCR using primers AluIV (Cotter *et al.*, 1991) and 5R′ (Nahmias *et al.*, 1995). End clones from cosmids were generated either by isolating end fragments from agarose gels or by vectorette PCR using primers designed from vector sequence. End probes from YACs were made according to the method of Silverman *et al.* (1991). Contig assembly by *Hin*fI fingerprinting and *Eco*RI restriction fragment analysis has been described elsewhere (Nahmias *et al.*, 1995, van Slegtenhorst *et al.*, 1995).

[1] To whom correspondence should be addressed at Wolfson House, 4 Stephenson Way, London NW1 2HE UK. Telephone: 0171 380 7417. Fax: 0171 383 2048. E-mail: jwolfe@galton.ucl.ac.uk.

0888-7543/97 $25.00

FISH onto metaphase chromosomes. Metaphase spreads were prepared from the cell lines SD1 and 9T12, with karyotypes 46XX, t(9;22)(q34.1;q11)(Dhut *et al.*, 1991) and 46XY, t(9;20)(q34.3;q11.2)(Zhou *et al.*, 1992, Woodward *et al.*, 1995), respectively. Cosmid DNA was biotin labeled by nick-translation using BioNick kits (Gibco BRL) and ethanol precipitated with a 40-fold excess of human Cot1 DNA. Two hundred nanograms of biotinylated probe was used for each slide,and detection was by FITC-avidin (Vector Laboratories) after the method of Pinkel *et al.* (1986).

Strand FISH. Target genomic DNA for the three-color strand FISH of clones around ABO was prepared from normal human peripheral blood lymphocytes in agarose blocks and extended onto poly-L-lysine slides (BDH) as described by Heiskanen *et al.* (1995). Biotinylated probes were made by the same method as was used for the FISH onto metaphase spreads. Digoxigenin-11-UTP-labeled probes were made using a Boehringer Mannheim nick-translation kit. Hybridization was performed with 200 ng of each probe in 50% formamide, 2× SSC, 10% dextran Sulfate. The yellow probe was achieved using equal quantities of biotinylated and digoxigenin-labeled probe. Images were captured with a confocal laser scanning microscope system (Bio-Rad MRC 600).

RESULTS AND DISCUSSION

The TSC1 Candidate Interval

Genetic mapping currently places the TSC1 gene in the 4-cM interval between D9S149 and D9S114. Many such areas of the genome are now covered by overlapping YAC clones; however, 9q34 generally is very underrepresented in the currently existing YAC libraries. The YAC contig map produced by the Whitehead Institute has no YACs containing markers distal to D9S164. Thus the map has no coverage at all in the distal part of the TSC1 candidate interval, and there are problems correlating the YAC contig data with the genetic data. The YAC contig map produced by Généthon includes contigs containing markers from the proximal and distal parts of the TSC1 interval, but there are clearly problems with the YACs in 9q34 since some of the mapping positions contradict other available physical mapping evidence in 9q34 and the order of the genetic map. Our own experience with CEPH YACs that were picked with D9S164 showed that these all mapped onto chromosome 6. This observation may stem from the paralogy seen between 9q34 and 6p21.3 (Katsanis *et al.*, 1996) and illustrates one of the problems with using whole-genome YAC libraries to map a particular area. Our more focused efforts allowed us to produce a better detailed map. A physical map of the region distal to DBH in overlapping YAC clones has been produced by Murrell *et al.* (1995). Most of the YACs that they used were derived from a chromosome 9-specific YAC library made in their laboratory especially for the experiment. They used the YACs to isolate cosmids both from the LL09NC01"P" library and from their own cosmid library, and these cosmids were binned according to hybridization data from the YAC clones and from STSs that they developed in the region. The cosmids from their own library were not available for our mapping, but those from the LL09NC01 library were fingerprinted and added to our database. Most of these map to the expected region, but some are placed elsewhere, showing that the YACs used to isolate these clones contain rearrangements. The existence of a fast and efficient fingerprinting method for the mapping of cosmid clones, coupled with an already high coverage of this area in cosmid contigs, persuaded us to map the whole of the TSC1 candidate interval in overlapping cosmids. A deep cosmid contig map of the TSC1 candidate interval is more detailed than a map based on YACs and prevents problems with rearranged clones.

Isolation of Clones

PAC clones 145N8 and 213M24 were isolated using primers for the genetic marker D9S164. They proved also to be positive with primers from the gene SURF4. PAC 63F12 was isolated with primers for the marker P6 in the ORFX gene (M. Smith, unpublished primer sequences). Inter *Alu* PCR probes have been produced from two of the YACs on the map, 15GF1 and 35HG8, and also from the PACs 63F12, 145N8, and 213M24. End probes were produced from 4DD1, 15GF1, and 35HG8 and from many cosmids. ^{32}P-labeled probes were hybridized individually to filters of the gridded LL09NC01"P" library. A total of 189 cosmids were picked.

Contig Assembly

The cosmid clones were analyzed in two ways: by *Hin*fI fingerprinting and by *Eco*RI restriction mapping, and positions were verified by hybridizations with cosmid end fragments. The positional data produced by these two techniques were broadly in agreement even though the lengths and positions of *Hin*fI fingerprinted clones are only estimated from the number of bands in the fingerprints. There were slight differences (about 2% of clone length) caused by variations in the density of *Hin*fI sites. Three continuous cosmid contigs have been produced covering the TSC1 candidate region, and they are separated by a small gap between SURF and D9S164 and by a second small gap between D9S164 and D9S150. The first of these gaps is covered by two PAC clones that both contain SURF4 as well as D9S164, but no clones have been detected covering the second gap. The map now contains 410 cosmids, 3 YACs, 1 P1, 1 BAC, and 3 PACs, and the average depth of the contigs is 10 clones. It spans approximately 1.7 Mb from a point 125 kb proximal to D9S149 to 110 kb distal to D9S114. PAC, P1, BAC, and YAC clones have been positioned on this map according to their patterns of hybridization to the cosmid clones. A representative section of the contig is shown in Fig. 1. The data from the *Hin*fI fingerprinting and probe hybridizations have been put into an ACeDB database. Complete details of the contig in AceDB format are available from the chromosome 9 home page, http://www.gene.ucl.ac.uk/chr9/, or they can be obtained from the authors.

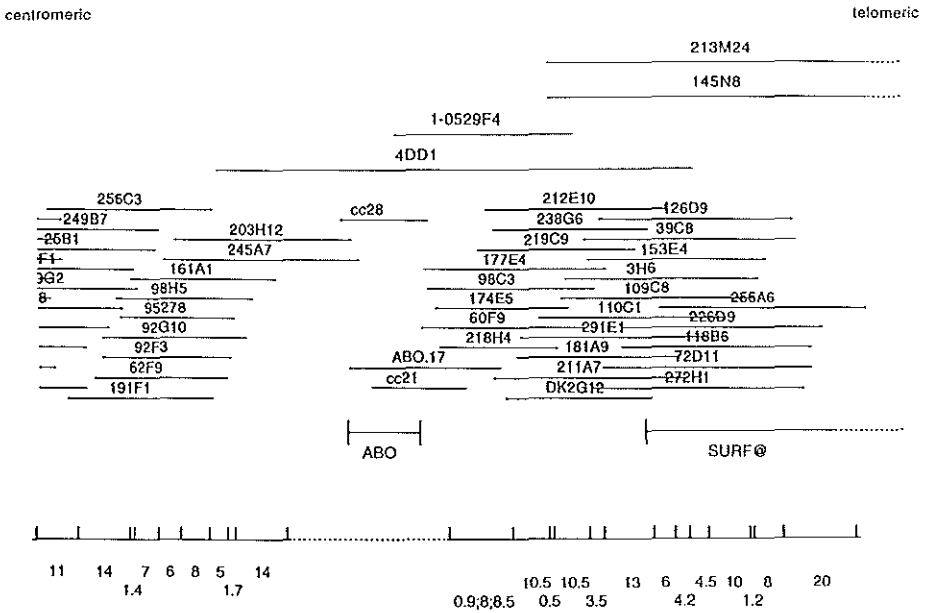


FIG. 1. A sample section of the contig, including the region of reduced depth of coverage proximal to ABO. Clones are reresented by horizontal lines, with the clone names written above the line. YAC, P1, and PAC clones are shown above the cosmid clones. The approximate positions of ABO and of the SURF cluster of genes are shown below the contig, and the *Eco*RI restriction map of this region is shown on the lowest line of the figure.

Testing Results by FISH to Metaphase Chromosomes

A large number of those contigs that had been constructed with the clones picked using inter *Alu* PCR from irradiation hybrids were positioned relative to translocation breakpoints, by FISH onto metaphase chromosomes (Nahmias *et al.*, 1995. Woodward *et al.*, 1995). Two cell lines in particular carry translocations that define an interval within which the TSC1 region lies. These are SD1, which carries the Philadelphia breakpoint in the first intron of ABL, and 9T12, which carries a breakpoint that has been mapped to between D9S114 and D9S298. The TSC1 candidate region covers the distal half of this interval. As the anchored contigs in the TSC1 interval were expanded, they incorporated a number of other contigs that had been assigned by FISH to between these two breakpoints. Further cosmids were tested against the breakpoints as the work progressed. The map now contains 24 clones that have been mapped to between SD1 and 9T12 in this way. The mapping of these clones by FISH was helpful during contig construction, in suggesting which potential overlaps would be most useful to investigate further, and in confirming the integrity of the contig. One clone, in the region just proximal to ABO, gave signals on both derivative chromosomes on the cell line 9T12. Further investigation by hybridization to an STS developed from the YAC 4DD1 and by careful examination of the fingerprint data showed that this result was not an artifact, but represents a part of the chromosome that is duplicated proximal to ABO and in distal 9q34 near the gene PAEP.

Strand FISH

In the region around and proximal to ABO the depth of coverage of our contig is reduced from about 10 cosmids deep in the majority of the contig to only 3 clones deep (see Fig. 1). This region appears to be absent from the LL09NC01"P" library but is covered by 2 cosmids subcloned from the ICI YAC C11-C and with YAC 4DD1 (Zhou *et al.*, 1995) and 1 cosmid derived from a human whole genome library (ABO.17; Cachon-Gonzalez, 1992). To confirm the integrity of the contig in this region, we have performed strand FISH, hybridizing cosmid probes to a single strand of genomic DNA on a slide, using 3 clones simultaneously (Fig. 2). These clones were shown to map in the same order and at approximately the same distances from each other that we had predicted from our contig. No statistical analysis of the distances between the cosmids was possible because there were too few signals on the slide.

The size of the gaps in the cosmid contig between SURF and D9S164 and of those between D9S2135 and D9S1793 was estimated by two-color strand FISH. Cosmids for this work were carefully chosen to avoid those

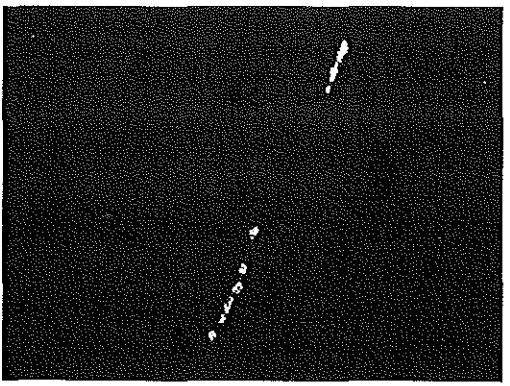


FIG. 2. Strand FISH experiment confirming the positions of cosmids 203H12 (green), ABO.17 (red), and 255A6 (yellow).

containing large amounts of repetitive DNA. The size of the gap between D9S2135 and D9S1793 was also measured using PFGE by Dr. J. R. Sampson (unpublished data). Two-color FISH of cosmids from either side of the gaps onto combed genomic DNA (Ekong *et al.*, in preparation) has also provided estimates of their sizes. The more proximal gap is about 10 kb, and the distal gap is about 20 kb. Probes made from the immediate proximal edge of this distal gap hybridize to a great many clones in the LL09NC01"P" library. FISH on metaphase chromosomes using these cosmids shows a large signal at the centromere of chromosome 9 and further scattered signals on 9q (data not shown). This indicates that there is a copy of a chromosome 9-specific repeat at this point in 9q34 and that most of the copies of this repeat are found near the centromere of chromosome 9. Both of the other places where our overlapping clone map is reduced in depth contain DNA that is duplicated elsewhere on chromosome 9. The existence of this repetitive DNA may explain our difficulty in finding clones covering these areas.

Genes and Polymorphic Markers Placed on the Map

Twelve previously known genes have been positioned on this physical map according to the cosmids or *Eco*RI fragments that contain them. These are, in order, from centromere to telomere: CEL–RalGDS–ATSV–ABO–SURF5–SURF3–SURF1–SURF2–SURF4–DBH–VAV2–ORFX. The gene ATSV has not been mapped on the LL09NC01"P" library but is positioned on our map by its reported position in cosmids cC31 and cC11 (Zhou *et al.*, 1995; Furlong *et al.*, 1996). Seventeen polymorphic loci have also been positioned on the map, from centromere to telomere: D9S149–D9S2127–D9S2126–D9S1199–D9S1198–ABO–SURF–D9S164–D9S2135–D9S1793–D9S150–DBH–D9S122–D9S10–D9S66–P6–D9S114 (P6; Smith and Handa, 1996) (see Fig. 3). This confirms the previously deduced order of the genetic map in this region. Detailed knowledge of the physical positions of these markers should facilitate investigation into relationships between allelic association and physical distance, and these studies are in progress. The 4-cM interval (sex averaged) between D9S149 and D9S114 represents 1.5 Mb of DNA. It has been known for some time that the recombination rate is elevated in 9q34, similar to that in other telomeric regions, and this result conforms to the estimate reported by Ozelius *et al.* (1992) that 1 cM is equivalent to approximately 400 kb of DNA in this part of the chromosome.

Apart from the small gap already mentioned this 1.7-Mb contig is now sequence ready and can be used for gene-hunting strategies. Work on cDNA selection and on trapping and analyzing exons from the cosmids of this contig has been started.

ACKNOWLEDGMENTS

We thank the Tuberous Sclerosis Association, the U.K. Human Genome Mapping Project, the Dutch Organisation for Scientific Research (NWO), and the Wellcome Trust for grants that supported parts of this work. The chromosome specific gene library LL09NC01"P" used in this work was constructed at the Biomedical Scien ces Division, Lawrence Livermore National Laboratory, Liv-

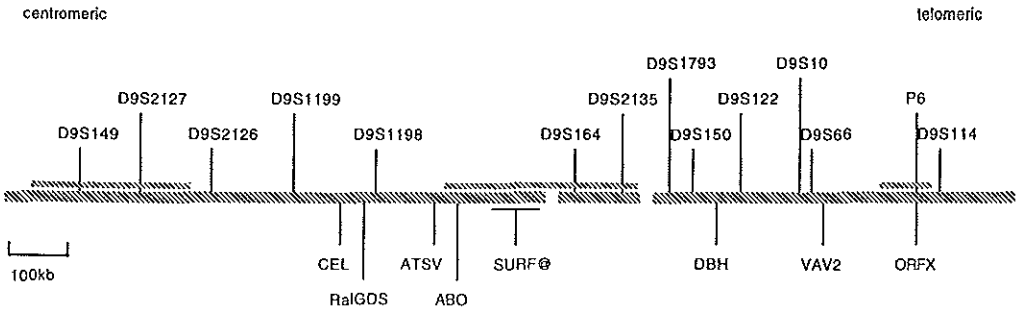


FIG. 3. The area covered by the cosmid contig (wide hatched bar) and by YAC, PAC, or BAC clones or contigs (narrow hatched bar). Genetic markers are shown above the bars, and genes are shown below.

ermore, California under the auspices of the National Laboratory Gene Library project sponsored by the U.S. Dept. of Energy. We are very grateful to Rob Furlong who provided YACs from this region and cosmids derived from one of the YACs and to Julia Korenburg for BAC clones that had been mapped to 9q34. We are indebted to David Kwiatkowski for unpublished information about genetic markers in this region, to Julian Sampson for sharing PFGE data, and to Moyra Smith for providing unpublished primer sequences. We also thank other members of the MRC HBGU for their technical assistance, computer support, and many helpful discussions.

REFERENCES

Bennet, E. P, Seffensen, R., Claussen, H., Weghuis, D. O., and van Kessel, A. G. (1995). Genomic cloning of the human histo-blood group ABO locus. *Biochem. Biophys. Res. Commun.* 206: 318–325.

Cachon-Gonzalez, M. B. (1992). "Linkage Analysis in Familial Adenomatous Polyposis Families in the United Kingdom, and a Search for Highly Polymorphic Markers." Ph.D. thesis, University of London.

Cotter, F. E., Das, S., Douek, E., Carter, N. P., and Young, B. D. (1991). The generation of DNA probes to chromosome 11q23 by *Alu* PCR on small numbers of flow-sorted 22q- derivative chromosomes. *Genomics* 9: 473–480.

Dhut, S., Gibbons, B., Chaplin, T., and Young, B. D. (1991). Establishment of a lymphoblastoid cell line, SD-1, expressing the p190 *bcl-abl* chimeric protein. *Leukemia* 5: 49–55.

European Chromosome 16 Tuberous Sclerosis Consortium (1993). Identification and characterisation of the tuberous sclerosis gene on chromosome 16. *Cell* 75: 1305–1315.

Fryer, A. E., Chalmers, A., Connor, J. M., Fraser, I., Povey, S., Yates, A. D., Yates, J. R. W., and Osborne, J. P. (1987). Evidence that the gene for tuberous sclerosis is on chromosome 9. *Lancet* 1: 659–661.

Furlong, F. A., Zhou, C. H., Ferguson-Smith, M. A., and Affara, N. A. (1996). Characterization of a kinesin-related gene ATSV, within the tuberous sclerosis locus (TSC1) candidate region on chromosome 9q34. *Genomics* 33: 421–429.

Gilbert, J. R., Kandt, R. S., Lennon, F., Gardner, R. J. M., Roses, A. D., and Pericak-Vance, M. A. (1993). Linkage and cross-over analysis in tuberous sclerosis (TSC). *Am. J. Hum. Genet.* 53: A1004.

Heiskanen, M., Hellsten, E., Kallioniemi, O.-P., Makela, T. P., Alitalo, K., Peltonen, L., and Palotie, A. (1995). Visual mapping by fibre-FISH. *Genomics* 30: 31–36.

Janssen, B., Sampson, J., van der Est, M., Deelen, W., Verhoef, S., Daniels, I., Hesseling, A., Brook-Carter, P., Nellist, M., Lindhout, D., Sandkuijl, L., and Halley, D. (1994). Refined localisation of TSC1 by combined analysis of 9q34 and 16p13 data in 14 tuberous sclerosis families. *Hum. Genet.* 94: 437–440.

Kataanis, N., Fitzgibbon, J., and Fisher, E. M. C. (1996). Paralogy mapping: Identification of a region in the human MHC triplicated onto human chromosomes 1 and 9 allows the prediction and isolation of novel PBX and NOTCH loci. *Genomics* 35: 101–108.

Kwiatkowski, D., Armour, J., Bale, A., Fountain, J., Goudie, D., Haines, J., Knowles, M., Pilz, A., Slaugenhaupt, S., and Povey, S. (1993). Report and abstracts of the Second International Workshop on Human Chromosome 9 Mapping. *Cytogenet. Cell Genet.* 64: 93–121.

Murrell, J., Trofatter, J., Rutter, M., Cutone, S., Stotler, C., Rutter, J., Long, K., Turner, A., Deaven, L., Buckler, A., and McCormick, M. K. (1995). A 500-kilobase region containing the tuberous sclerosis locus (TSC1) in a 1.7-megabase YAC and cosmid contig. *Genomics* 25: 59–65.

Nahmias, J., Hornigold, N., Fitzgibbon, J., Woodward, K., Pilz, A., Griffen, D., Henske, E. P., Nakamura, Y., Graw, S., Florian, F., Benham, F., Povey, S., and Wolfe, J. (1995). Cosmid contigs spanning 9q34 including the candidate region for TSC1. *Eur. J. Hum. Genet.* 3: 65–77.

Nellist, M., Brook-Carter, P. T., Connor, J. M., Kwiatkowski, D. J., Johnson, P., and Sampson, J. R. (1993). Identification of markers flanking the tuberous sclerosis locus on chromosome 9 (TSC1). *J. Med. Genet.* 30: 224–227.

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

Ozelius, L. J., Kwiatkowski, D. J., Schuback, D. E., Breakfield, X. O., Wexler, N. S., Gusella, J. F., and Haines, J. L. (1992). A genetic linkage map of human chromosome 9q. *Genomics* 14: 715–720.

Pinkel, D., Straume, T., and Gray, J. M. (1986). Cytogenetic analysis using quantitative high sensitivity fluorescent hybridisation. *Proc. Natl. Acad. Sci. USA* 83: 2934–2938.

Pitiot, G., Waksman, G., Bragado-Nilsson, E., Jobert, S., Cornelis, F., and Mallet, J. (1994). Linkage analysis places the TSC1 gene distal to D9S10. *Ann. Hum. Genet.* 58: 232–233.

Povey, S., Burley, M. W., Attwood, J., Benham, F., Hunt, D., Jeremiah, S. J., Franklin, D., Gillet, G., Malas, S., Robson, E. B., Tippet, P., Edwards, J. H., Kwiatkowski, D. J., Super, M., Mueller, R., Fryer, A., Clarke, A., Webb, D., and Osborne, J. (1994a). Two loci for tuberous sclerosis: One on 9q34 and one on 16p13. *Ann. Hum. Genet.* 58: 107–127.

Povey, S., Armour, J., Farndon, P., Haines, J. L., Knowles, M., Olopade, F., Pilz, A., White, J. A., Members of the Utah Genome Centre Genetic Marker and Mapping Group, and Kwiatkowski, D. J. (1994b). Report on the Third International Workshop on Chromosome 9. *Ann. Hum. Genet.* 58: 177–260.

Silverman, G. A., Jockel, J. I., Domer, P. H., Mohr, R. M., Taillon-Miller, P., and Korsmeyer, S. J. (1991). Yeast artificial chromosome cloning of a two-megabase-size contig within chromosome band 18q21 establishes physical linkage between BCL2 and plasminogen activator inhibitor type 2. *Genomics* 9: 219–228.

Smith, M., and Handa, K. (1996) Analysis of the VAV2 and bromodomain homeotic gene region on chromosome 9q34. Abstract presented at the Fifth International Workshop on Chromosome 9.

van Slegtenhorst, M., Janssen, B., Nellist, M., Ramlakhan, S., Hermans, C., Hesseling, A., van den Ouweland, A., Kwiatkowski, D., Eussen, B., Sampson, J., de Jong, P., and Halley, D. (1995). Cosmid contigs from the tuberous sclerosis candidate region on chromosome 9q34. *Eur. J. Hum. Genet.* 3: 78–86.

Woodward, K., Nahmias, J., Hornigold, N., West, L., Pilz, A., Benham, F., Kwiatkowski, D., Fitzgibbon, J., Wolfe, J., and Povey, S. (1995). Regional localisation of 64 cosmid contigs, including 18 genes and 14 markers, to intervals on human chromosome 9q3 4. *Genomics* 29: 257–260.

Zhou, C., Goudie, D., Carter, N., Furlong, R., Lyall, J., Yuille, M., Affara, N., and Ferguson-Smith, M. A. (1992). Assignment of CA repeat polymorphisms and other markers to a thirteen-interval map of chromosome 9. *Ann. Hum. Genet.* 56: 215–216.

Zhou, C. Y., Wu, K. Y., Leversha, M. A., Furlong, R. A., Ferguson-Smith, M. A., and Affara, N. A. (1995). Physical analysis of the tuberous sclerosis region in 9q34. *Genomics* 25: 304–308.

CHAPTER 2.6.3

Identification of the Tuberous Sclerosis Gene TSC1 on Chromosome 9q34

Marjon van Slegtenhorst, Ronald de Hoogt, Caroline Hermans, Mark Nellist, Bart Janssen, Senno Verhoef, Dick Lindhout, Ans van den Ouweland, Dicky Halley*
Janet Young, Mariwyn Burley, Steve Jeremiah, Karen Woodward, Joseph Nahmias, Margaret Fox, Rosemary Ekong, John Osborne, Jonathan Wolfe, Sue Povey*
Russell G. Snell, Jeremy P. Cheadle, Alistair C. Jones, Maria Tachataki, David Ravine, Julian R. Sampson*
Mary Pat Reeve, Paul Richardson, Friederike Wilmer, Cheryl Munro, Trevor L. Hawkins*
Tiina Sepp, Johari B. M. Ali, Susannah Ward, Andrew J. Green, John R. W. Yates*
Jolanta Kwiatkowska, Elizabeth P. Henske, M. Priscilla Short, Jonathan H. Haines, Sergiusz Jozwiak, David J. Kwiatkowski.

Published in: Science 1997;277;805-808

Reprinted with permission

Tuberous sclerosis complex (TSC) is an autosomal dominant disorder characterized by the widespread development of distinctive tumors termed hamartomas. TSC-determining loci have been mapped to chromosomes 9q34 (*TSC1*) and 16p13 (*TSC2*). The *TSC1* gene was identified from a 900-kilobase region containing at least 30 genes. The 8.6-kilobase *TSC1* transcript is widely expressed and encodes a protein of 130 kilodaltons (hamartin) that has homology to a putaive yeast protein of unknown function. Thirty-two distinct mutations were identified in *TSC1*, 30 of which were truncating, and a single mutation (2105delAAAG) was seen in six apparently unrelated patients. In one of these six, a somatic mutation in the wild-type allele was found in a TSC-associated renal carcinoma, which suggests that hamartin acts as a tumor suppressor.

TSC is a systemic disorder in which hamartomas occur in multiple organ systems, particularly the brain, skin, heart, lungs, and kidneys (1,2). In addition to its distinct clinical presentation, two features serve to distinguish TSC from other familial tumor syndromes. First, the tumors that occur in TSC are very rare in the general population, such that several TSC lesions are, by themselves, diagnostic of TSC. Second, TSC hamartomas rarely progress to malignancy. Only renal cell carcinoma occurs at increased frequency in TSC (~2.5%) and with earlier age of onset; it appears to arise in TSC renal hamartomas, termed angiomyolipomas (3). Nonetheless, TSC can be a devastating condition, as the cortical tubers (brain hamartomas) frequently cause epilepsy, mental retardation, autism, or attention deficit-hyperactive disorder, or a combination of these conditions (1,4).

TSC affects about 1 in 6000 individuals, and ~65% of cases are sporadic (5). Linkage of TSC to chromosome 9q34 was first reported in 1987, and this locus was denoted *TSC1* (6). Later studies provided strong evidence for locus heterogeneity (7) and led to the identification of chromosome 16p13 as the site of a second TSC locus (denoted *TSC2*) (8). The *TSC2* gene was identified by positional cloning and the encoded protein, denoted tuberin, contains a domain near the COOH-terminus with homology to a guanosine triphosphatase (GTPase) activating protein (GAP) for rap1, a Ras-related GTPase (9).

The focal nature of TSC-associated hamartomas has suggested that *TSC1* and *TSC2* may function as tumor suppressor genes. The occurence of inactivating germline mutations of

TSC2 in patients with tuberous sclerosis (9-11) and of loss of heterozygosity (LOH) at the *TSC2* locus in about 50% of TSC-associated hamartomas (12-14) supports a tumor suppressor function for *TSC2*. In contrast, LOH at the *TSC1* locus has been detected in <10% of TSC-associated hamartomas (13,14), suggesting the possibility of an alternative mechanism for lesion development in patients with *TSC1* disease.

As part of a comprehensive strategy to identify *TSC1*, we identified 11 microsatellite markers from the 1.4-Mb *TSC1* region and developed an overlapping contig (with only a single gap of 20 kb) of cosmid, P1 artificial chromosome (PAC), and bacterial artificial chromosome (BAC) clones (15). Figure 1 shows the *TSC1* region (16,17), including limiting centromeric and telomeric markers, as derived from analysis of affected individuals (solid arrows) from families with lod scores of >2 (18). These limits are also consistent with the information available from LOH studies (13). Two additional recombination events were identified in unaffected individuals (open arrows), also from families with lod scores of >2 (19). In each of these families, two individuals from different generations carried the same recombinant chromosome, and all four had no evidence of TSC in nearly 100% (2), we concentrated our search within the 900-kb region between markers D9S2127 and DBH.

In a search for further positional information, we looked for large deletions and rearrangements by means of pulsed-field gel electrophoresis (Fig. 1) (9) and through analysis of patient-derived hybrid cell lines retaining a single chromosome 9 bearing a *TSC1* mutation (20). No abnormalities were detected, and we therefore began a systematic gene-by-gene analysis.

Several techniques were used to identify genes in the *TSC1* region, which proved to be relatively gene-rich. Using a combination of exontrapping (21), cDNA selection, expressed sequence tag (EST) mapping, and whole-cosmid hybridization (22), we identified 142 exons and 13 genes between D9S1199 and D9S114. In all, 30 genes were identified or mapped to the 900-kb critical region.

In parallel, we began sequencing the entire contig (23). We used the polymerase chain reaction (PCR) to amplify putative (24) and confirmed exons found in 208 kb of sequence on a screening set of 60 DNA samples from 20 unrelated familial TSC cases with linkage to 9q34, and 40 sporadic TSC cases (18). Amplification products were analyzed for hetero-duplex formation using weakly denaturing polyacrylamide gels (25). The 62nd exon screened demonstrated mobility shifts in 10 of the 60 patient samples (Fig. 2A).

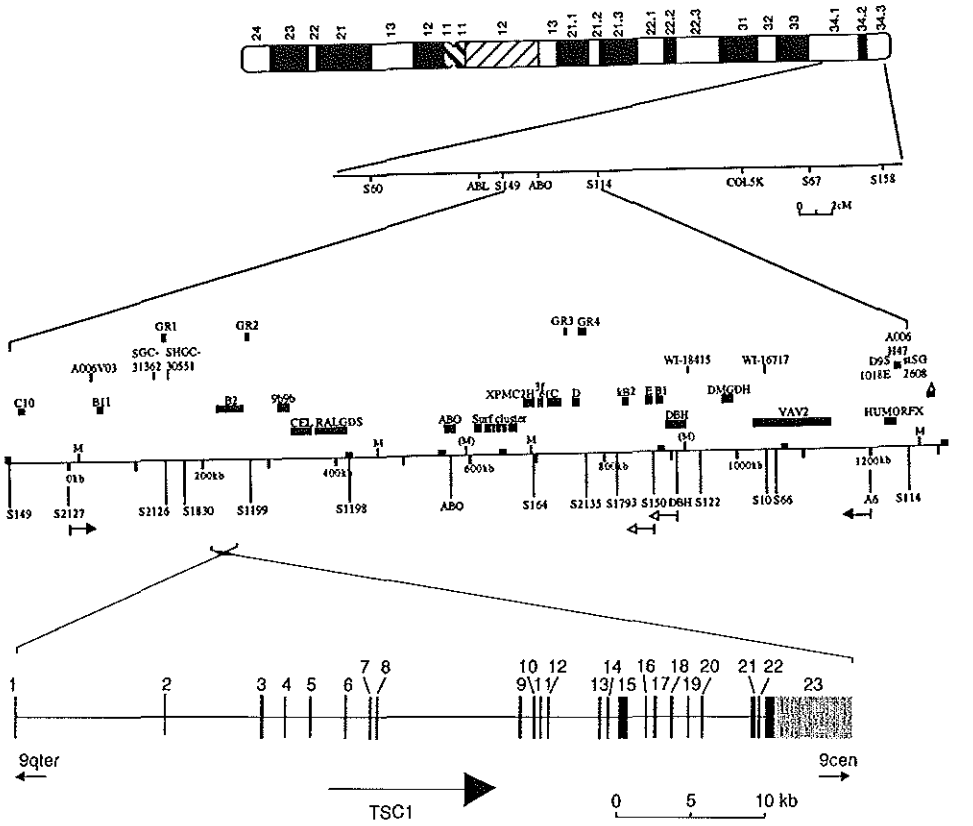
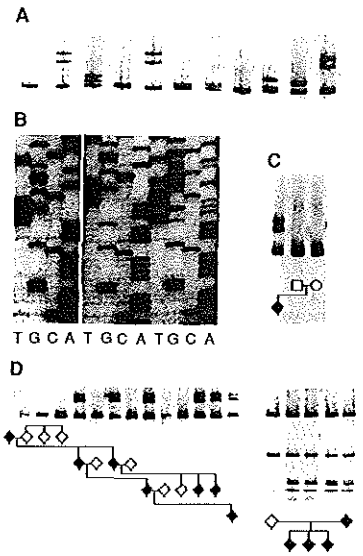


Fig. 1. The *TSC1* region on chromosome 9. The ideogram (top) represents a normal G-banded metaphase chromosome 9, with the *TSC1* region located at 9q34. The male genetic map (next line) shows selected anchor polymorphic loci mapped to 9q34. The detailed physical map of the candidate region (next level) shows the positions of polymorphic markers and key recombination events in affected members (filled arrows) and unaffected members (open arrows) of families showing linkage to TSC to 9q34; the approximate positions of *MluI* (M) sites (with sites that partially cut in genomic DNA shown in parentheses) and of probes used to screen the region for rearrangements in patients with TSC by means of pulsed-field gel electrophoresis (orange boxes); genes previously mapped to the *TSC1* candidate region (blue boxes); novel cDNAs isolated from the region (red boxes); ESTs mapped to the region (green); and additional putative genes predicted by GRAIL analysis of genomic sequence (light blue boxes). There was a single 20-kb gap in the contig near D9S1793. The map of the *TSC1* gene (bottom) shows the 23 exons, of which exons 3 to 23 are coding. B2 is the *TSC1* gene.

Sequence analysis revealed seven small frameshifting deletions (three identical), one nonsense mutation, one missense change, and one polymorphism that did not change the encoded amino acid (Fig. 2B). Eight of the nine mutations were from the familial cases tested, and only one mutation was seen among the 40 sporadic cases (Fig. 2C). Analysis of samples from other family members confirmed that each of the familial mutations segregated with TSC and that a frameshift mutation had occured de novo in the sporadic case (Fig. 2D). The recurrent mutation, 2105delAAAG, was identified in two apparently unrelated familial cases and a sporadic case. Haplotype analysis of the families, using markers flanking the mutation (D9S2126, D9S1830, and D9S1199, Fig. 1), confirmed that the three mutations were of independent origin.



Flg. 2. Identification of mutations in *TSC1* exon 15 **(A)** Heteroduplex analysis. Control sample (left) is followed by 10 samples with a shift. **(B)** Sequence analysis demonstrating 2105delAAAG mutation. The sequence reactions were done in antisense orientation, so that reading from the top down (b2083 to 2124 of the normal allele sequenced is shown), the allele sequenced on the left has the deletion, the middle sequence is a normal allele, and the sequence on the right is the heteroduplex product with both alleles. **(C)** In a sporadic case, the heteroduplex mobility shift is not present in either parent. **(D)** Segregation of heteroduplex mobility shifts in a large family with TSC (left) and digestion of amplification products with *Mwo I* in another family (right) demonstrates segregation of the 2105delAAAG mutation with the disease.

The exon with mutations was part of a transcriptional unit identified by earlier gene discovery efforts (26). The full sequence of the *TSC1* gene was determined by comparison of genomic sequence and cDNA clone sequence, including clones obtained by 5' rapid amplification of cDNA ends (RACE). The *TSC1* gene consists of 23 exons, of which the last 21 contain coding sequence and the second is alternatively spliced (Fig. 1, bottom). The open reading frame (ORF) of the longest transcript begins at nucleotide 162, and the likely initiator

ATG codon is at nucleotide 222. The first stop codon is at nucleotide 3738, leaving a 4.5-kb 3' untranslated region. Northern (RNA) blot analysis with a coding region probe (nucleotides 1100 to 2100) revealed a major 8.6-kb transcript that was widely expressed and was particularly abundant in skeletal muscle (Fig. 3).

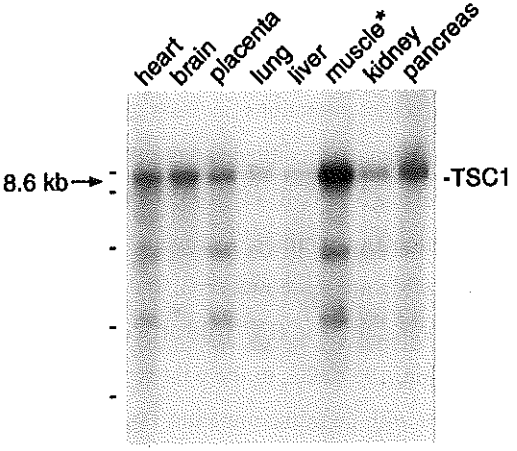


Fig. 3. Norhern blot analysis of *TSC1* expression. Each lane contained 2μg of polya-denylated RNA from adult human organs, and the probe consisted of base pairs 1100-2100 of the *TSC1* gene. Minor hybridization signals of size 4 and 2.5 kb are also seen. *= skeletal.

The predicted TSC1 protein, which we call hamartin, consists of 1164 amino acids with a calculated mass of 130 kD (Fig. 4). The protein is generally hydrophilic and has a single potential transmembrane domain at amino acids 127-144 (27) as well as a probable 266-amino acid coiled-coil region beginning at position 730 (28). Database searches identified a possible homolog of TSC1 in the yeast *Schizosaccharomyces pombe* (GenBank accession numberr Q09778), a hypothetical 103-kD protein, but there were no strong matches with vertebrate proteins (29).

```
MAQQANVGKL LAMLDSFKLG VREDVTAVFK ENLNSDRGFM LVNTLVDYYL ETSSQPALHI   60
LYTLQEPHDK HLLDRINEYV GKAATRLSIL SLUGNVIRLQ FSWKHKLSQA PLLESLLKCL  120
FMDTDVKVLT TGVLVLLTML EMIRQSGFQH LLDFFDIFGR LSSWCLKKPG HVAEVYLVHL  180
RASVIALFHR LYGMYFCNFV SFLESHYSMK ENLETFEEVV FFMMEHVRIH PELVTGSKDH  240
ELDFRENERL ETHDVVIECA KISLDPTEAS YEDQYSVSHQ ISAFFHRSA DVTTSFTADT  300
QNSTGCATST FYSTSRLMLL KMFQQLPQTL SSPSTRLITE FPQATLWSPS KVCGMTTPPT  360
SPGNVFFDLS HFFSKVFGTT AGKKDTFLGT PATSFFPAPL CHSDDYVHIS LPQATVIPPR  420
EEEEMDSARP CLHRQHHLLN DRGSEEPPGS FGSVTLSDLP GFLGDLASEE DSIEKDEEEA  480
AISEELSEIT TAEAEFVVFR GSFDSFFYED SLPGSQRKTH SAASSSQQAS VNFEPLHSSL  540
DKLGPDTFFQ AFTPIDLPCG SADESPAGGR ECQTSLSTST FTFSPCKIPP PTRVGFGSGQ  600
EPPYDHLFEV ALPKIAHHFV IRKTEELLKK AKGNTEEDGV FSISFMEVLD RLTQQGADAH  660
SKELNKLPLP SKSVDWTHFG GSPPSDEIRT LSDQLLLLHN QLLYERFKEQ QRALRNRLLL  720
REVIKAAALS EIQAAMEDQL KLQEFDICMV KVSLGKECAR YKQLQEQRDT KYTELHSQIR  780
QLGHDREEFY NQSQELQTKL EDCRNMIAEL RIELKKANNK VCHTELLLSQ VSQKLSSSES  840
YQQCHEFLNR QLLVLGEVNE LYLEQLQNYH SDTTKEVEMM FAAYREELEK KRSHVLQQTQ  900
RLDTSQKRIL ELESHLAKED HLLLEQKKYL EDVELQAAGQ LQAAESRYEA QKRITQYTEL  960
EILDLYGRLE YDGLLKKLEE KYAEAAEAAE ERLDCCNDGC SDSMVGHNEE ASGHNGETKT 1020
PRPSSAEGSS GSRGGGGSSS SSSELSTPEK PPHQPAGPFS SRWETTMGEA SASIPTTVGS 1080
LPSSFSFLGM KARELFENFS ESQCDEDGNT SSLSESLKTE LGKDLGVEAK IPLNLDGPHP 1140
SPPTPDSVGQ LHIMDYNETH HEHS                                       1164
```

Fig. 4. Predicted amino acid sequence of the *TSC1* protein, hamartin. A potential trans-membrane domain (amino acids 127 to 144) and a coiled-coil domain (amino acids 730 to 965) are underlined. The *TSC1* genomic sequence and the cDNA sequence have been deposited in GenBank (accession numbers AC002096 and AF013168, respectively).

Because the initial screen identified a high frequency of mutations in exon 15, we studied this exon in a large sample of patients. Mutations in exon 15 [559 base pairs (bp), 16% coding region] were identified in 8 of 55 (15%) familial DNA samples with linkage to the *TSC1* region, and in 15 of 607 (2.5%) DNA samples from sporadic patients or families uninformative for linkage (Table 1). A screen for mutations in all coding exons in 20 familial cases and 152 sporadic patients yielded eight mutations in each group (40% and 5%, respectively). In total, 19 mutations were found in coding exons other than exon 15. No mutations have been detected thus far in exons 3 to 6, 8, 11 to 14, 16 or 21 to 23. Of the 32 distinct mutations seen in 42 different patients or families, five were recurrent. Thirty were predicted to be truncating, one was a possible missense mutation, and one was a splice site mutation. Analysis of a renal cell carcinoma from a TSC patient with germline mutation 2105delAAAG, revealed a somatic mutation, 1957delG, in the wild-type *TSC1* allele (30). A giant cell astrocytoma from another patient with germline mutation 1942delGGinsTTGA had retained the mutant allele but lost the wild-type allele.

Table 1. All mutations found in *TSC1*. Both heteroduplex and single-strand conformation polymorphism (*33*) gels were used to search for mutations after the initial screening. F, familial; S, sporadic.

Exon	Number of patients screened*		Mutations	Patients
	F	S		
7	20	152	865delTT	1S
9	39	230	966delA	1S
			970T → G, L250X	1F
			993G → T, E258X	1F
			1112T → G, Y297X	1F
10	20	152	1207delCT	1S
15	55	607	1746C → T, R509X	1F, 2S
			1750delCA	1S
			1801delAG	1F
			1892del23	1S
			1929delAG	2S
			1942delGGinsTTGA	1S
			1981A → G, K585R	1F
			2009delT	1S
			2041delTT	1F
			2060delA	1S
			2105delAAAG	4F, 2S
			2122delAC	2S
			2126delAG	1F
			2176delTG	1F
17	45	296	2295C → T, R692X	1F
			2324dupIGTTACTC	1F
			2332delAT	1S
			2395insA	1S
18	45	296	2448C → T, Q743X	1S
			2519del23bp	1S
			2540delC	1F
			2577C → T, R786X	1F, 1S
			2583G → T, E788X	1F
19	39	230	2691delAC	1S
20	39	230	2724-1 G → T	1F
			2730insA	1S

*Families are defined as those with linkage to the *TSC1* region and negative linkage to the *TSC2* region. Sporadics include both sporadic cases and cases from families without linkage information. Exon structure and primer information are provided at http://expmed.bwh.harvard.edu/projects/tsc/.

Our results support the hypothesis that *TSC1* functions as a tumor suppressor gene. First, the majority of mutations are likely to inactivate protein function. Second, in two TSC-associated tumors we have shown that loss of the wild-type *TSC1* allele occurred through LOH or intragenic somatic mutation. The paucity of LOH for the *TSC1* region found in patient lesions (13, 14) may reflect the same mutational spectrum seen in the germline of TSC patients with a high frequency of small mutations causing inactivation of the second allele. It is also possible that there is a greater frequency of *TSC2*- versus *TSC1*- associated disease among the sporadic cases providing the lesions analyzed. This is suggested by the low frequency of mutations we have detected in *TSC1* in sporadic cases. However, in families suitable for linkage analysis, about half show linkage to *TSC1* and half to *TSC2* (16, 31).

The mutations observed in *TSC1* consist of small deletions, small insertions, and point mutations. No genomic deletions or rearrangements in *TSC1* were detected by Southern (DNA) blot analysis of 250 TSC patients. This restricted mutational spectrum may reflect an intrinsic tendency for this type of mutation in this region of the genome. Alternatively, it may reflect selection against more disruptive mutations such as large deletions, which would involve neighboring genes.

The mechanism by which loss of hamartin expression produces TSC lesions is unknown. It is likely that hamartin and tuberin participate in the same pathway of cellular growth control, because the clinical features of *TSC1* and *TSC2* disease are so similar (31). Tuberin has modest GAP activity for both rap1 and rab5, members of the Ras superfamily of small GTPases. The physiological function of the rap1 GTPase is not understood, whereas rab5 is thought to be involved in aerly endosomal transport. Tuberin-deficient rat embryo fibroblasts display increased endocytosis, which suggests that the rab5 interaction of tuberin has physiological relevance (32). It is unclear how a deficiency of GAP activity for rap1 or rab5, if that is the critical function of tuberin, leads to tumor development. The sequence homology of hamartin to a putative S. pombe protein suggests that it may participate in an evolutionarily conserved pathway of eukaryotic cell growth regulation. The identification of *TSC1* will enable analysis of the functions of both hamartin and tuberin, and may permit further insight into the molecular pathogenesis of TSC.

REFERENCES AND NOTES

1. M. Gomez, *Tuberous Sclerosis* (Raven, New York, 1988); M. R. Gomez, *Ann. N. Y. Acad. Sci.* **615**, 1 (1991).
2. D.J. Kwiatkowski and M.P. Short, *Arch. Dermatol.* **130**, 348 (1994); E. Roach *et al., J. Child Neurol.* **7**, 221 (1992); D. W. Webb and J. P. Osborne, *J. Med. Genet.* **28**, 417 (1991).
3. J. Bjornsson, M. P. Short, D. J. Kwiatkowski, E. P. Henske, *Am. J. Pathol.* **149**, 1201 (1996); J. A. Cook, K. Oliver, R. F. Mueller, J. Sampson, *J. Med. Genet.* **33**, 480 (1996).
4. A. Hunt and C. Shepherd, *J. Autism Dev. Disord.* **23**, 323 (1993); S. L. Smalley, P. E. Tanguay, M. Smith, G. Gutierrez, *ibid.* **22**, 339 (1992).
5. J. P. Osborne, A. Fryer, D. Webbs, *Ann. N.Y. Acad. Sci.* **615**, 125 (1991); J. R. Sampson, S. J. Scahill, J. B. Stephenson, L. Mann, J. M. Connor, *J. Med. Genet.* **26**, 28 (1989).
6. A. E. Fryer *et al., Lancet* i, 659 (1987).
7. L.A.J. Janssen *et al. Genomics* **8**, 237 (1990); H. Northrup *et al., Lancet* ii, 804 (1987); J. R. Sampson *et al., J. Med. Genet.* **26**, 511 (1989); R. Kandt *et al., Exp. Neurol.* **104**, 223 (1989).
8. R. Kandt *et al., Nature Genet.* **2**, 37 (1992).
9. European Tuberous Sclerosis Consortium, *Cell*, **75**, 1305 (1993).
10. P. T. Brook-Carter *et al., Nature Genet.* **8**, 328 (1994).
11. P. J. Wilson *et al., Hum. Mol. Genet.* **5**, 249 (1996); A. Kumar *et al., ibid.* **4**, 1471 (1995); A. Kumar *et al., ibid.*, p. 2295; S. Verhoef *et al., Hum Mutat.*, in press; R. Vrtel *et al., J. Med Genet.* **33**, 47 (1996); Q. Wang *et al., Hum. Mutat.*, in press; A. Kumar *et al., ibid.* **9**, 64 (1997).
12. A. J. Green, M. Smith, J. R. Yates, *Nature Genet.* **6**, 193 (1994).
13. T. Sepp, J. R. W. Yates, *J. Med. Genet.* **33**, 962 (1996); E. P. Henske *et al., Am. J. Hum. Genet.* **59**, 400 (1996).
14. C. Carbonara *et al., Genes Chromosomes Cancer* **15**, 18 (1996).
15. N. Hornigold *et al., Genomics* **41**, 385 (1997)
16. D. J. Kwiatkowski *et al., Cytogenet. Cell Genet.* **64**, 93 (1993); S. Povey *et al., Ann. Hum. Genet.* **58**, 177 (1994).
17. K. S. Au, J. Murrell, S. H. Buckler, S. H. Blanton, H. Northrup, *J. Med. Genet.*, **33**, 559 (1996).
18. The diagnosis of tuberous sclerosis was made according to standard diagnostic criteria (2). Blood samples (obtained after informed consent) were used for DNA preparation, either directly or after creation of immortalized Epstein-Bar virus-transformed lymphoblastoid cell lines. Linkage to the *TSC1* region was inferred if a family demonstrated obligate recombination with markers within 2 centimorgans (cM) of *TSC2* and had positive lod scores (logarithm of the odds ratio for linkage) in analyses with 9q34, and haplotype analysis was performed manually to identify the site of recombination.
19. J. L. Haines *et al., Am. J. Hum. Genet.* **49**, 764 (1991); M. Nelist *et al., J. Med. Genet.* **30**, 224 (1991).
20. C. Jones, F. T. Kao, R. T. Taylor, *Cytogenet. Cell Genet.* **28**, 181 (1980).
21. Cosmid, BAC, or PAC DNA was digested with PstI or BAMHI, and a library of subclones was prepared in pSPL3 []. Exons identified in this manner were used to identify cDNA clones (i) by screening cDNA libraries by conventional methods, (ii) by screening GenBank and dbEST databases, and (iii) in reverse transcription PCR experiments. IMAGE clones were obtained from Research Genetics or the UK Human Genome Mapping Project (HGMP) Resource Centre.
22. Complementary DNA selection was performed using the end-ligation coincident sequence cloning method []. A normalized infant brain cDNA library [] was screened using whole cosmids as probes. A human fetal brain cDNA library (Clontech) was screened by standard phage plating and filter lift methods.
23. Cosmid (15) DNA was sheared and subcloned into M13mp18. Single clear plaques were picked using an automated picking device (PBA Technologies, Cambridge, UK) and expanded with JM101, and phage supernatant was collected. M13 DNA isolation was performed with the Sequatron robotic system [T.L. Hawkins et al., Science 276, 1887 (1997)] following the solid-phase reversible immobilization protocol [T.L. Hawkins et al., Nucleic Acids Res. 22, 4543

(1994)]. Dye primer DNA sequencing used energy transfer primers and thermosequenase (Amersham), and electrophoresis was performed on Applied Biosystems 377 DNA sequencers. Gel files were extracted, signal-processed, and bases called with the program Trout (available from genome.wi.mit.edu/distribution/ software/trout) and were assembled with Alewife, a sequence assembly package. Typically, 1200 reads from a single cosmid assembled into one to three contigs, which were then finished by directed primer walking and directed selection of reverse reads from existing M13 templates to span sequence gaps. All sequence data and protocols were available during the sequence process from our Web site, http://www-seq.wi.mit.edu.

24. Genomic sequence was analyzed with the program GRAIL2 to identify possible exons and gene models []. Putative transcriptional units were also identified by BLAST searches of public databases and comparison with our own collection of cDNA clones adn exontrapping products.
25. Oligonucleotide primers were designed to be external to exons by 40 to 60 bp were possible. DNA products with mobility shifts on heteroduplex analysis [] were subjected to sequence analysis of both strands.
26. A transcriptional unit denoted B2 was identified from a brain cDNA library by hybridization with a cosmid near D9S1830 (22). A 4.5-kb cDNA clone was sequenced and contained no ORF. Database searches that this clone was the 3' portion of a 6.8-kb cDNA clone (KIAA0243) [], which contained a 2.0-kb ORF. After the discovery of mutations in exon 15 of B2, 5'RACE was performed with the Marathon cDNA kit (Clontech), using oligonucleotides derived from *TSC1* cDNA clones or inferred by analysis of genomic sequence information (24). RACE and other cDNA clones were sequenced fully on both strands by means of Taq cycle sequencing methodology.
27. R. F. Smith, *et al. Genome Res.* **6**, 454 (1996).
28. A. Lupas, M. Van Dyke, J. Stock, *Science* **252**, 1162 (1991).
29. A BLAST search of the NCBI combined protein database using the *TSC1* amino acid sequence gave the best match at a probability of 2.2×10^{-18} with the Q09778 sequence, with several regions of homology throughout the sequence but no areas of particularly strong match. The next lowest probability was 5.2×10^{-8}.
30. Allele-specific amplification was performed on the renal cell carcinoma DNA using a primer specific for the normal allele, not bearing the 2105delAAAG mutation. Sequence analysis of this product indicated that it contained the 1957delG mutation.
31. S. Povey *et al., Ann. Hum. Genet.* **58**, 107 (1994); B. Janssen *et al., Hum. Genet.* **94**, 437 (1994).
32. R. Wienecke, A. Konig, J. E. DeClue, *J.Biol. Chem.*, **207**, 16409 (1995). G.H. Xiao, F. Shoarinejad, F. Jin, E.A. Golemis, R. S. Yeung, *ibid.* **272**, 6097 (1997); R. Wienecke *et al. Oncogene*, **13**, 913 (1996).
33. M. Orita, Y. Suzuki, T. Sekiya, K. Hayashi, *Genomics* **5**, 874 (1989).
34. We thank the hundreds of TSC patients and their families for blood samples and financial support (particularly V. Whittemore and W. Watts), and numerous clinicians (particularly A. Oranje, H. Stroink, M. Wildervanck de Blecourt, H. Kros, B. Prahl-Andersen, and H.L.J. Tanghe) for assistance; H. Galjaard and P. Harper for support; R. Bakker, S. Ramlakhan, D. Humphrey, Q. Wang, M. Maheshwar, A.L.W. Hesseling, N. Thoams, and R. Slomski for technical support; and the UK HGMP Resource Centre and the European Collection of Animal Cell Cultures. Supported by the NIH, NTSA, NWO, STSN, Dutch Kidney Foundation, Westminster Medical School Research Trust (Royds Fund), Bath Unit for Research into Paediatrics, and New Zealand Health Resources Council.

CHAPTER 2.6.4

Mutational spectrum of the *TSC1* gene in a cohort of 225 tuberous sclerosis complex patients; no evidence for a genotype-phenotype correlation

Marjon van Slegtenhorst, Senno Verhoef, Anita Tempelaars, Lida Bakker, Qi Wang, Marja Wessels, Remco Bakker, Mark Nellist, Dick Lindhout, Dicky Halley and Ans van den Ouweland

Submitted for publication

Abstract

Tuberous sclerosis complex is an inherited tumour suppressor syndrome, caused by a mutation either in the *TSC1* or *TSC2* gene. The disease is characterised by a broad phenotypic spectrum that can include seizures, mental retardation, renal dysfunction and dermatological abnormalities. The *TSC1* gene was recently identified and has 23 exons, spanning 45 kb of genomic DNA, and encoding an 8.6 kb mRNA. After screening all 21 coding exons in our collection of 225 unrelated patients, only 29 small mutations were detected, suggesting that *TSC1* mutations are under-represented among TSC patients. Almost all *TSC1* mutations were small changes leading to a truncated protein, except for a splice-site mutation and two in-frame deletions in exon 7 and exon 15. No clear difference was observed in the clinical phenotype of patients with an in-frame deletion or a frameshift or nonsense mutation. We found the disease causing mutation in 12% of our unrelated set of TSC patients, with more than half of the mutations clustered in exons 15 and 17, and no obvious under-representation of mutations among sporadic cases. About 50% of the patients with a *TSC1* mutation were mentally retarded with no detectable position effect of the mutations. In conclusion, we find no support for a genotype-phenotype correlation for the group of *TSC1* patients compared to the overall population of TSC patients.

Tuberous sclerosis complex (TSC) is an autosomal dominant neurocutaneous disorder characterised by the growth of hamartomas in many tissues and organs, including brain, skin, heart and kidney [1]. Common neurological manifestations including seizures and mental retardation have their onset during early childhood, while cysts and angiomyolipomas in the kidney mostly become apparent during adult life. Considerable clinical variation is observed between as well as within families [2]. At least 60% of the TSC patients represent sporadic cases, as they have non-affected parents [3].

Linkage analysis has demonstrated locus heterogeneity for TSC, with one locus on chromosome 9 [4] and a second locus on chromosome 16 [5]. About half of the large families can be linked to the *TSC1* locus on chromosome 9q34 and the other half to the *TSC2* locus on chromosome 16p13 [6,7]. The *TSC1* gene as well as the *TSC2* gene were identified by positional cloning [8,9] and there is abundant evidence that both genes act as tumour suppressor genes [10-13].

The *TSC2* gene consists of 41 exons, spanning 43 kb of genomic DNA [14]. It encodes a 200 kDa protein, tuberin, which has a putative GAP activity for rab5 [15] and rap1 [16], two members of the ras superfamily of small GTPases. The mutational spectrum of *TSC2* includes a number of large deletions often disrupting the PKD1 gene as well [17,18], but also point mutations [19-27] and a number of missense changes. [28].

The *TSC1* gene contains 23 exons and encodes an 8.6 kb mRNA. It spans 45 kb of genomic DNA and codes for hamartin, a 1164 amino acid protein of 130 kDa. Analysis of the amino acid sequence revealed a potential coiled-coil domain at the C-terminus but no homology to tuberin or any other known vertebrate protein was detected [8].

The first report describing the molecular genetic and phenotypic analysis of the *TSC1* gene suggested that all mutations are small changes, that *TSC1* mutations are less common in the sporadic population and that there is a reduced risk of mental retardation in *TSC1* related disease [29]. The goal of this study was to construct the mutational spectrum of the TSC1 gene in our collection of TSC patient samples by Southern blot- and SSCP- analysis. This would enable testing whether there is a significant difference in the detection rate between familial and sporadic cases and whether there is a genotype-phenotype correlation for the *TSC1* group compared to the overall population of TSC patients.

Patients and methods

PATIENT SELECTION AND DNA ISOLATIONS

In this study, 225 unrelated patients with tuberous sclerosis complex, diagnosed according to Gomez' criteria [1] were included. Eigthy-two patients represented familial cases (36%) with at least one affected parent and/or other affected relatives (first degree). Three large families showed linkage to 9q34 [7]. One hundred and fourty-three patients were designated sporadic cases (64%), in the absence of an apparent family history of TSC. DNA was isolated from peripheral blood cells according to standard procedures [30].

SOUTHERN BLOTTING ANALYSIS

Genomic DNA (6µg) of 200 unrelated patients was digested with four different restriction enzymes (*EcoRI, HindIII, PstI and Taq*I) and run on a 0.7% agarose gel. Southern blotting and hybridisation were performed using standard methods [31]. Three cDNA clones were tested on the blots; a 5'RACE clone which had been amplified from a fetal brain cDNA pool (bp 24-1696)(Clontech), an RT-PCR product (bp 1616-3684) generated from fibroblast RNA

and a fetal brain cDNA clone (bp 4100-8600). The three probes cover the coding sequence as well as the 5'- and 3'-untranslated region of the *TSC1* gene.

SSCP ANALYSIS AND DNA SEQUENCING

Sequences of primers used for amplification of the 21 coding exons of the *TSC1* gene are provided at http://expmed.bwh.harvard.edu/projects/tsc/. For exon 22, a new intronic forward primer was designed (5'-atactaccagcttactttccata-3'). SSCP analysis was performed following Orita *et al.* [32]. 2 µl of the PCR product were applied to the Pharmacia GenePhor Electrophoresis system. Gels were run for 2.5 hours at both 5 and 18 degrees Celsius. Running conditions for two gels were 600 V, 50 mA and 10 W. Subsequently, bands were visualised using a DNA silver staining kit (Pharmacia) in a Hoefer automated gel stainer. Variant patterns were further characterised by direct sequence analysis of the PCR products on an automated DNA sequencer (ABI 377) using the cyclesequencing Dyeprimed kit (Perkin Elmer).

ASO HYBRIDISATION

Oligonucleotides for ASO hybridisation were designed for the mutated and normal sequence. ASO hybridisations were performed at 37° C for 30 minutes. Filters were washed to 0.3XSSC for 10 minutes at 37° C.

Results

SCREENING FOR LARGE ABNORMALITIES

No large insertions or deletions were identified in the *TSC1* gene by Southern analysis in 200 patients. Only in one case (T2965) an aberrant restriction pattern in a *TaqI* digest was detected (figure 1a), but no consistent change was seen with other enzymes. Comparing the genomic sequence of the *TSC1* gene with the size of the extra fragment, we could locate the lost *TaqI* site in exon 15. Sequencing exon 15 of the *TSC1* gene of this patient revealed a C to T substitution at bp position 1719 (figure 1b), resulting in the nonsense mutation R500X. The presence of this mutation was confirmed by allele specific oligonucleotide (ASO) hybridisation (figure 1c).

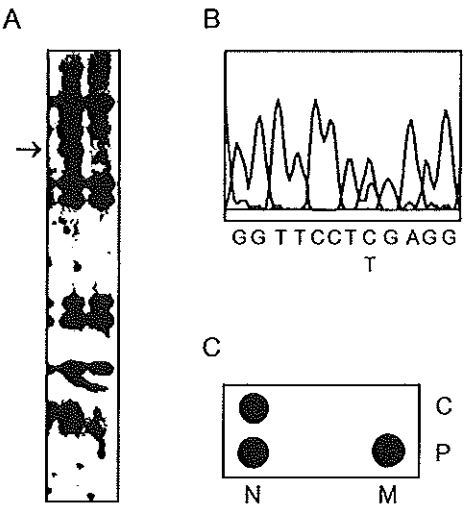


Figure 1; (A) Southern blot of *TaqI* digested DNA from 2 unrelated TSC patients, hybridised with a 5' TSC1 cDNA probe (nt 24-1696). A novel 3 kb fragment is detected in lane 2, indicated by the arrow. (B) Direct sequence analysis reveals the *de novo* nonsense mutation C→T (R500X) in a *TaqI* site in patient T2965. (C) ASO hybridisation of the R500X mutation in patient T2965 (P) and a control (C). N= normal allele; M= mutant allele.

SCREENING FOR SMALLER MUTATIONS

Systematic SSCP analysis was undertaken to screen the 21 coding exons in the *TSC1* gene for small mutations. In total 29 mutations were detected in a set of 225 unrelated patients (table 1). All types of mutations resulting in a truncated protein have been observed: small deletions/insertions, nonsense mutations and splice-site mutations. In addition two in-frame deletions and 9 missense changes (table 2) were detected.

Fourteen of the 29 mutations were small deletions, ranging from 1 to 23 bp. Three of these mutations have been reported previously [8]. In two patients we detected in-frame deletions of 3 and 9 bp respectively. In family T1298, a three basepair deletion in exon 7 segregated with the disease phenotype and resulted in a small amino acid change (Asp-Phe to Ile at position 198) in the protein. The grandparents, having no signs of TSC, tested negative for the mutation. In a sporadic patient (T5913), 9 basepairs were deleted in exon 15, also leading to a different protein product (Cys-Lys-Ile-Pro to Ser at aa position 586). Both parents tested

negative for the mutation. All the other deletions lead to a premature stopcodon. Nonsense mutations were detected in 11 cases; R692X was present in four sporadic cases and in one family. Three insertions were identified: a single basepair substitution in exon 7 and exon 10 in familial cases, and a duplication of 28 bp in exon 17 in a sporadic patient. A substitution at a splice-site (bp postition 432-1) was detected in a sporadic patient, of which the parients tested negative for the change. The most downstream mutation detected is a one basepair deletion in exon 20 in the middle of the coiled-coil domain of hamartin. No mutations were found 3' of the coiled-coil domain (aa 719-998).

Table 1. Mutations identified in the *TSC1* gene.

Exon	Patient code	mutation (nt substitution)	Familial / Sporadic
4	T2545	367delT	Familial
5	T1214	W103X (529G→A)	Familial
	T1817	432-1G→A	Sporadic[2]
7	T1298	814delACT	Familial
8	T4715	944insA	Familial
	T8129	958delG	Sporadic[2]
9	T7806	R249X (966A→T)	Sporadic
10	T1207	1210 insT	Familial
	T3945	1240del	Familial
	T9809	Y312X (1157C→A)	Sporadic
12	T10301	1473delC	Familial
13	T1515[3]	1499delT	Familial
15	T2965	R500X (1719C→T)	Sporadic
	T9886	R509X (1746C→T)*	Sporadic
	T3922	1892del23	Sporadic[2]
	T2067	1929delAG[1]*	Familial
	T5913	1978del9	Sporadic[2]
	T2636	2007delT[1]	Sporadic[2]
	T4124	2105delAAAG[1]*	Sporadic
17	T1197	R692X (2295C→T)*	Familial
	T3838		Sporadic[2]
	T3908		Sporadic
	T5210		Sporadic[2]
	T10816		Sporadic[2]
	T7659	2318ins28	Sporadic
	T4068	2328delCT	Sporadic
18	T2077[3]	R786X (2577C→T)*	Familial
20	T5406	2729delAACA	Familial
	T1295	2787delG	Familial

nt=nucleotide; [1]mutation reported before (van Slegtenhorst *et al.*, 1997); [2]parents tested negative for the mutation; [3]family linked to chromosome 9 (Janssen et al., 1994); *recurrent mutation, also identified by other groups (van Slegtenhorst *et al.*, 1997; Jones *et al.*, 1998).

Table 2. Missense and silent changes in the TSC1 gene.

Exon	Code	nt change	aa change
4	T5768	374A→C	E51D
7	T1486	789C→T	R190S*
	T4712	793T→A	L191H
8	T1524	892T→G	M224R
12	T10383	1429C→T	S403
14	T2083	1556A→G	E445*
15	T5100	2415C→T	H732Y*
	T8775		
21	T1219	3050C→T	A943*

*confirmed polymorphism; E445, H732Y and A943 were reported before (Jones). nt= nucleotide, aa=amino acid

Table 3. Overall summary of clinical features of all patients with mutations in the *TSC1* gene.
A distinction has been made between patients with a mutation 5' (exons 3-14), 3' (exons 15-23) and inframe.

	mutation exon 3-14	mutation exon15-23	inframe deletion	total
Facial angiofibroma	6/10	11/19	3/3	20/32 (63%)
Ungual fibroma	4/8	6/18	2/3	12/29 (41%)
Hypomelanotic macule	9/10	14/16	4/4	27/29 (93%)
Subependymal nodule	8/8	16/20	3/4	27/32 (84%)
Cortical tuber	2/8	8/20	2/4	12/32 (38%)
Epilepsy	8/10	15/20	2/4	25/34 (74%)
Mental retardation	5/8	8/19	1/3	14/30 (47%)
Renal cyst	2/8	2/18	0/4	4/30 (13%)
Renal angiomyolipoma	0/8	1/18	0/4	1/30 (3%)
Cardiac rhabdomyoma	2/8	3/16	1/4	6/28 (21%)

Table 4. Clinical features of patients with mutation R692X.

code age	T1197a 56	T1197b 15*	T3838 14	T3908 54	T5210 6	T10816 3	Total
Facial angiofibroma	+	+	-	+	-	-	3/6 (50%)
Ungual fibroma	+	-	-	+	-	-	2/6 (33%)
Hypomelanotic macule	-	+	+	-	+	+	4/6 (67%)
Subependymal nodules	+	+	-	+	+	+	5/6 (83%)
Cortical tuber	-	-	-	-	+	-	1/6 (17%)
Epilepsy	+	-	+	+	+	+	5/6 (83%)
Mental retardation	-	-	?	-	+	-	1/5 (20%)
Renal cyst	-	-	?	-	-	-	0/5 (0%)
Renal angiomyolipoma	-	-	?	-	-	-	0/5 (0%)
Cardiac rhabdomyoma	-	-	+	-	+	-	2/6 (33%)

*died age 15 of giant cell astrocytoma

MISSENSE AND SILENT CHANGES

Nine different abnormal SSCP patterns were observed representing missense or silent changes (table 2). In four cases the change was also found in the normal population. These polymorphisms were R190S (present in unaffected parent and absent in the affected parent), E445 (allele frequency of 16% in normal population), H732Y (allele frequency of 0.5%) and A943. E445, H732Y and A943 were reported before [29]. None of the additional missense abnormalities have been confirmed to represent the disease causing mutation yet.

TSC1 MUTATIONS IN FAMILIAL AND SPORADIC CASES

Eighty two of our 225 unrelated patients (36%) had other affected family members. Of the 29 mutations, 13 were identified in the 82 familial TSC patients (16%), and 16 in the 143 sporadic cases (11%). Hence, we do not find a significant difference in detection rate between familial and sporadic cases. In about half of the sporadic cases, both parents were available for analysis and tested negative for the mutation. In the other sporadic patients, DNA of both parents was not available, but there was no clinical indication of TSC disease in the family.

CLINICAL SYMPTOMS VERSUS TYPE AND POSITION OF THE MUTATION

Patients with a mutation in the *TSC1* gene were scored for the most frequent skin, brain, kidney and heart lesions detected in TSC. A distinction has been made between type 1 mutations detected in *TSC1* (deletion, insertion, splice site and nonsense mutations) in exons 3-14 versus type 1 mutations in the exons 15-23 (table 3). No positional effect was noted for mental handicap or other TSC symptoms. In addition the clinical phenotype was evaluated of patients with a type 1 mutation versus the in frame deletions (type 2 mutations) in exons 7 and 15. Comparing both types of mutations, no obvious correlation could be detected between the genotype and phenotype in the TSC patients. The missense changes were left out of the analysis, because so far none of them has been confirmed to represent a disease causing mutation.

CLINICAL MANIFESTATIONS IN PATIENTS WITH THE RECURRENT MUTATION R692X

The mutation R692X was present in 4 sporadic TSC patients and in 2 patients from the same family. The clinical data of these 6 patients are summarized in table 4. Almost all patients have a history of epilepsy. No renal lesions were detected, but only 2 patients are older than 15, so these results could be biased by the later onset of the cysts and angiomyolipomas. All

other symptoms were scored once at least. The patients with the R692X mutation do not share an obvious similar phenotype.

Discussion

After screening the 21 coding exons of the *TSC1* gene, 29 mutations were detected, all of them small changes. The only mutation detected by Southern blot analysis was due to the substitution of a C to T in a *TaqI* (TCGA) restriction site, resulting in a stopcodon. Since four different restriction enzymes have been used to test a selection of our TSC patients on Southern blots, it is unlikely that large abnormalities disrupting the *TSC1* gene have remained undetected. Previous mutation studies in the *TSC2* gene have shown a diverse mutational spectrum including large rearrangements, deletions, insertions, and nonsense- and missense mutations. In the *TSC2* gene, approximately 10% of the mutations detected so far are large deletions, often resulting in disruption of the neighbouring PKD1 gene as well. A possible explanation for the lack of large mutations in *TSC1* can be the presence of unknown neighboring- or intragenic- genes that are essential for embryonic development and survival. Although all the mutations detected in the *TSC1* gene were small, we could not confirm any missense change as the disease causing mutation. Only one missense mutation in *TSC1* has been reported before by Jones *et al.*, but this *de novo* change in the gene appeared not to be the disease causing mutation (personal communication). Conversely, a number of missense mutations have been reported in the *TSC2* gene [28]. It remains to be explained why the mutational spectrum of the *TSC1* and *TSC2* gene is different. Despite the differences, most of the mutations in either *TSC1* or *TSC2* lead to a truncated protein, which is in concordance with a loss of function mechanism. Results obtained by interaction studies indicate that hamartin and tuberin function as a complex [33], which supports the phenotypic overlap observed between TSC patients with either a *TSC1* or *TSC2* mutation.

So far we have detected a mutation in the *TSC1* gene in 13% of our unrelated TSC collection screening all of the coding region of the gene by SSCP analysis. It is likely that this technique fails to detect all of the mutations and the promoter region has not been tested yet. We only detected the disease causing mutation in two out of three of our clearly chromosome 9 linked families. This number is too small to give an indication of the ratio of the undetected mutations in the *TSC1* gene. We expect that for the whole group of TSC patients, the majority of the mutations will be found in the *TSC2* gene. The coding region of the *TSC2* gene is twice

as large and mutations have already been detected in 15% of our patients, after screening only 30% of the coding region (unpublished data). We do not observe a significantly larger number of *TSC1* mutations in our familial cases versus the sporadic population, as was found in a recent study [29], although it is possible that some of our patients are misclassified as sporadic, because in only half of these cases material from both parents was available.

We find a clustering of mutations in exon 15 and 17. In these exons, 14 out of 29 identified mutations were present. The high mutation rate in exon 15 was already observed when the TSC1 gene was identified [8] and can be partially explained by the size of the exon (17% of the coding region). The high proportion of recurrent mutations detected in exon 15 and 17 suggests that part of these exons are particularly prone to nucleotide changes.

We did not detect a clear correlation between the location or nature of a *TSC1* mutation and the clinical phenotype. In a previous study, mental handicap was seen only in patients with truncated *TSC1* mutations towards the 3' end of the gene (exons 15-23) [29]. In our study, mental handicap was detected equally in patients with a mutation in exons 3-14 (table 3). In addition the type 1 versus the type 2 mutations in the *TSC1* gene did not show an obvious difference in clinical phenotype in the patients. The six patients with the recurrent mutation R692X also displayed a wide range of clinical symptoms. This is comparable with the clinical differences detected within families with TSC. The phenotypic differences in TSC patients are more likely caused by mechanisms such as a second hit [10-12], somatic mosaicism [18] and modifying genes. The latter has also been proposed to contribute to the complex phenotype in the comparable 'monogenic' disease neurofibromatosis 1 (NF1) [34]. In conclusion we find no support for a different phenotypic spectrum in patients with a mutation in the *TSC1* gene.

We are grateful to the patients and their families for their participation, to the clinicians and the TSC working group for referral of the TSC patients and to Prof. H. Galjaard for continuous support. This work was funded by the Dutch Organization for Scientific Research (NWO), the Dutch Praevention Fund (grant number 28-1723) and the Dutch Kidney Foundation (grant number C93.1313).

1 Gomez MR. Tuberous Sclerosis, 2nd edition. Raven Press, New York 1988.
2 Sampson JR, Scahill SJ, Stephenson JBP, *et al.* Genetic aspects of tuberous sclerosis in the west of Scotland. *J Med Genet* 1989;**26**:28-31.
3 Osborne JP, Fryer A and Webb D. Epidemiology of tuberous sclerosis. *Ann N Y Acad Sci* 1991; **615**:125-127.
4 Fryer AE, Chalmers A, Connor JM, *et al.* Evidence that the gene for tuberous sclerosis is on chromosome 9. *Lancet* 1987;**i**:659-661.
5 Kandt RS, Haines JL, Smith M, *et al.* Linkage of an important gene locus for tuberous sclerosis to a chromosome 16 marker for polycystic kidney disease. *Nature Genet* 1992;**2**: 37-41.
6 Povey S, Burley MW, Attwood J, *et al.* Two loci for tuberous sclerosis: one on 9q34 and one on 16p13. *Ann Hum Genet* 1994;**58**:107-127.
7 Janssen B, Sampson J, van der Est M, *et al.* Refined localization of TSC1 by combined analysis of 9q34 and 16p13 data in 14 tuberous sclerosis families. *Hum Genet*1994;**94**: 437- 440.
8 van Slegtenhorst M, de Hoogt R, Hermans C, *et al.* Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science* 1997;**277**:805-807.
9 The European Chromosome 16 Tuberous Sclerosis Consortium. Identification and characterisation of the tuberous sclerosis gene on chromosome 16. *Cell* 1993;**75**:1305-1315.
10 Green AJ, Smith M and Yates JRW. Loss of heterozygosity on chromosome 16p13.3 in hamartomas from tuberous sclerosis patients. *Nature Genet* 1993;**6**:193-196.
11 Carbonara C, Longa L, Grosso E, *et al.* 9q34 loss of heterozygosity in a tuberous sclerosis astrocytoma suggests a growth suppressor-like activity also for the TSC1 gene. *Hum Mol Genet* 1994;**3**:829-832.
12 Henske LP, Scheithauer BW, Short M, *et al.* Allelic loss is frequent in tuberous sclerosis kidney lesions but rare in brain lesions. *Am J Hum Genet* 1997;**59**:400-406.
13 Kobayashi T, Mitani H, Takahashi RI, *et al.* Transgenic rescue from embryonic lethality and renal carcinogenesis in the Eker rat model by introduction of a wild type TSC2 transgene. *Proc Natl Acad Sci USA* 1997;**94**:3990-3993.
14 Maheshwar M, Sandford R, Nellist M, *et al.* Comparative analysis and genomic structure of the tuberous sclerosis 2 (TSC2) gene human and pufferfish. *Hum Mol Genet* 1996;**5**: 131-137.
15 Xiao G-H, Shoarinejad F, Jin F, *et al.* The TSC2 gene product, tuberin, functions as a Rab5 GTPase activating protein (GAP) in modulating endocytosis. *J biol Chem* 1997; **272**:6097-6100.
16 Wienecke R, Konig A and DeClue JE. Identification of tuberin, the tuberous sclerosis 2 product. *J Biol Chem* 1995;**207**:16409-16414.
17 Brook-Carter PT, Peral B, Ward CJ, *et al.* Deletion of the TSC2 and PKD1 genes associated with severe infantile polycystic kidney disease-a contiguous gene syndrome. *Nature Genet* 1994;**8**:328-332.
18 Verhoef S, Vrtel R, van Essen T, *et al.* Somatic mosaicism and clinical variation in tuberous sclerosis complex. *Lancet* 1995;**345**:202.
19 Vrtel R, Verhoef S, Bouman K, *et al.* Identification of a nonsense mutation at the 5'end of the TSC2 gene in a family with presumptive diagnosis of tuberous sclerosis complex. *J Med Genet* 1996;**33**:47-51.
20 Wang Q, Verhoef S, Tempelaars, AMP, *et al.* Identification of a large insertion and two

novel point mutations (3671del8 and S1221X) in tuberous sclerosis comlex (TSC) patients. *Hum Mut* 1998;**11**:331-332.
21 Kumar A, Wolpert C, Kandt RS, *et al.* A de novo frame-shift mutation in the tuberin gene *Hum Mol Genet* 1995a;**4**:1471-1472.
22 Kumar A, Kandt RS, Wolpert C, *et al.* Mutation analysis of the TSC2 gene in an African-American family. *Hum Mol Genet* 1995b;4:2295-2298.
23 Kumar A, Kandt RS, Wolpert C, *et al.* A novel splice site mutation in the TSC2 gene. *Hum Mut* 1997;**9**:64-65.
24 Verhoef S, Vrtel R, Bakker L, *et al.* Recurrent mutation 4882delTT in the GAP-related domain of the tuberous sclerosis TSC2 gene. *Hum Mut* 1998;**Suppl.** 1:s85-s87.
25 Au KS, Rodriguez JA, Rodriguez E, *et al.* Mutations and polymorphisms in the tuberous sclerosis complex gene on chromosome 16. *Hum Mut* 1997;**9**:23-29.
26 Wilson PJ, Ramesh V, Kristiansen A, *et al.* Novel mutations detected in the TSC2 gene from both sporadic and familial TSC patients. *Hum Mol Genet* 1996;**5**:249-256.
27 Au KS, Rodriguez JA, Finch JL, *et al.* Germ-line mutational analysis of the TSC2 gene in 90 tuberous-sclerosis patients. *Am J Hum Genet* 1998;**62**:286-294.
28 Maheshwar MM, Cheadle JP, Jones A, *et al.* The GAP-related domain of tuberin the product of the TSC2 gene, is a target for missense mutations in tuberous sclerosis. *Hum Mol Genet* 1997;**6**:1991-1996.
29 Jones AC, Daniells CED, Snell RG, *et al.* Molecular genetic and phenotypic analysis reveals differences between TSC1 and TSC2 associated familial and sporadic tuberous sclerosis. *Hum Mol Genet* 1997;6:2155-2161.
30 Miller SA, Dykes, DD and Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;**16**:1215.
31 Sambrook J, Fritsch EF and Maniatis T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY 1989.
32 Orita M, Iwahana H, Kanazawa H, *et al.* Detection of polymorphism of human DNA by gel electrophoresis as single-strand conformation polymorphism. *Proc Natl Acad Sci USA* 1989;**86**:2766-2770.
33 van Slegtenhorst M, Nellist M, Nagelkerken B, *et al.* Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products. *Hum Mol Genet* 1998;7:1053-1057.
34 Easton DF, Ponder MA, Huson SM and Ponder BA. An analysis of variation in expression of Neurofibromatosis (NF) type 1 (NF1): evidence for modifying genes. *Am J Hum Genet* 1993;**53**:305-313.

CHAPTER 3

FUNCTIONAL ANALYSIS OF THE TSC GENE PRODUCTS

FUNCTIONAL ANALYSIS OF THE TSC GENE PRODUCTS

3.1 The TSC1 gene product, hamartin

The TSC1 gene encodes a 1164 amino acid protein of 130 kDa. The protein was called hamartin, after the hamartomatous lesions seen in patients. Analysis of the protein sequence (Smith *et al.*,1996) showed a putative transmembrane domain at the N-terminus (aa 127-144) and an extensive predicted coiled-coil region at the C-terminus (aa 719-998). Coiled-coil domains are alpha-helical structures, which often mediate homomeric interactions and interactions with other proteins (Lupas *et al.*,1991). The TSC1 nucleotide and amino acid sequence was analysed using different Basic Local Alignment Search Tool (BLAST) programs, but no significant homology with the TSC2 protein or any other known vertebrate protein was detected. Only the coiled-coil domain in hamartin matches other coiled-coil structures, but the significance of these homologies needs further characterisation.

3.2 TSC genes in other species

Since hamartin did not show significant homology to any known vertebrate proteins, the next step was to compare the human TSC1 sequence with TSC1 sequences from different species. This is one of the approaches to gain insight into gene evolution and function (Rastan and Beeley, 1997). For example, functional domains in proteins are often represented by regions that are highly conserved through species. The TSC2 gene has been identified in the rat (Yeung *et al.*, 1994; Kobayashi *et al.*, 1995a), the mouse (Kim *et al.*,1995), *Fugu rubripes* (Maheshwar *et al.*,1996) and *Drosophila melanogaster* (1847aa, 30% identity) (Gerald Rubin, personal communication). Thus far, TSC1 sequences have been detected in the genome databases of the mouse, the yeast strain *Schizosaccharomyces pombe* and *D. melanogaster*. Over a distance of 145 nucleotides, the homology with a mouse EST clone (GenBank accession number 604696) is 95% (aa identity is 98%). The homology with a putative *S. pombe* protein (GenBank accession number Q09778) is relatively low and restricted to two regions in hamartin. The TSC1 gene in *D. melanogaster* was presented at the ASHG meeting 1997 (1100aa, 24% identity) (Gerald Rubin, personal communication). The completion of the genome sequencing projects for these different organisms may help to understand the function of the TSC1 and TSC2 gene products in normal development and in disease pathology.

3.3 Expression studies

The TSC2 gene is widely expressed in many different cell lines and tissues when studied by Northern blot analysis (European TSC2 Consortium, 1993). Transfection studies indicate that tuberin is present exclusively in the cytoplasm (chapter 3.6). *In situ* hybridisation (ISH) showed that the expression of TSC2 mRNA is most prominent in cells with a rapid mitotic rate and turnover (Menchine *et al.*,1996). No major differences were detected in TSC2 mRNA expression between TSC2 patients and normal individuals in any organ or tissue tested (Menchine *et al.*,1996). At least six splice variants of tuberin have been reported to be present in humans and rodents, showing both an age-associated and tissue-related expression pattern.

Northern blot analysis for the TSC1 gene showed expression in all tissues tested, including brain and kidney, which was expected from the known multiorgan involvement in patients. All tissues show a major 8.6 kb transcript, with potentially smaller mRNA bands, alhough no tissue-specific expression was detected for any of the bands. The first studies with a polyclonal antiserum, which has been raised against hamartin, show a 130 kDa protein on a Western blot in fibroblasts, COS- and HeLa- cells. When hamartin is overexpressed in COS cells, a punctated cytoplasmic labelling pattern is detected.

3.4 A natural animal model for the TSC2 gene, the Eker rat

After the TSC2 gene had been cloned, two reports linked a model of hereditary renal carcinoma in the Eker rat to a germline mutation in the rat homologue of the TSC2 gene (Yeung *et al.*,1994; Kobayashi *et al.*,1995a). The mutation in the rat is an intragenic insertion of about 4.5 kb and has been identified as a rat intracisternal A-particle (IAP) element (Xiao *et al.*,1995), resulting in a frameshift and premature stopcodon before the GAP related domain. Virtually all heterozygous rats develop renal cell carcinoma by the age of one year and tumours in the spleen, uterus and pituitary also occur (figure 3.1).

Only the mutated transcript is expressed in the tumour, supporting the two-hit model for the tumours seen in the Eker rat. Most of these tumours have not been described in TSC patients, although renal cell carcinoma has been reported in about 30 cases with TSC (Bjornsson *et*

al.,1996). Eker tumour cell lines provide a nice model to test tumour suppressor function of tuberin, since they lack normal TSC2 expression.

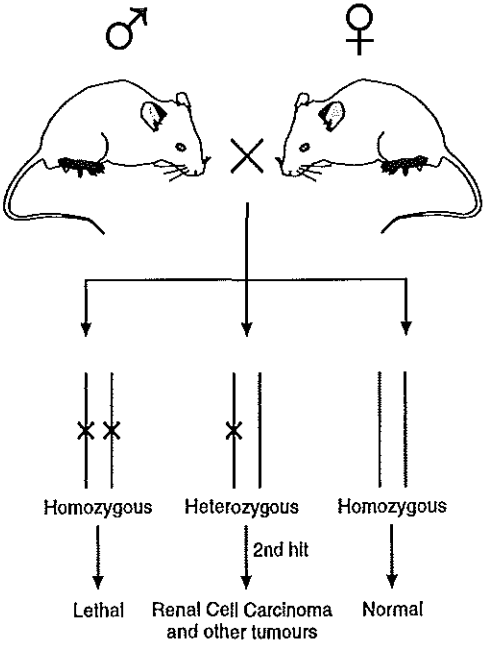


Figure 3.1 Eker rat model (Kobayashi *et al.*, 1995a). Inheritance of two mutant TSC2 alleles is lethal in developing embryos. Heterozygous animals have a strong predisposition to tumour formation and all animals develop RCC in the first year of life, after a second somatic hit.

First experiments showed that re-introduction of TSC2 suppresses the tumourigenicity of these cell-lines, strongly suggesting a role for tuberin in growth control (Jin *et al.*,1996).

3.5 Interaction between the TSC1 and TSC2 gene products

The clinical similarities between patients with TSC1 and TSC2 disease suggests that the proteins function in the same pathway. A widely used system to detect interactions between two proteins is the yeast two-hybrid system (Field and Song, 1989). Interactions between gene products from other genetically heterogeneous disorders have recently been described for ADPKD (Tsiokas *et al.*,1997; Qian *et al.*,1997) and Fanconi's anaemia (Kupfer *et al.*,1997).

Yeast two-hybrid system

The 'original' yeast two-hybrid system is based on the properties of the yeast GAL4 protein, which consists of separate domains responsible for DNA-binding and transcriptional activation. The first application of this system is to screen for unknown interacting proteins with a protein of interest. Identifying interacting proteins is a powerful tool to learn more about the function of a protein. One of the most prominent examples is Huntington's disease. Several interacting proteins have been identified in the last few years (Wankler *et al.*,1997; Kalchman *et al.*,1997; Boutell *et al.*,1998) and the type of proteins suggest an indirect role for huntingtin in vesicle trafficking through the cytoskeleton (Engelender *et al.*,1997) and as a regulator of a ras-related signalling pathway (Colomer *et al.*,1997).

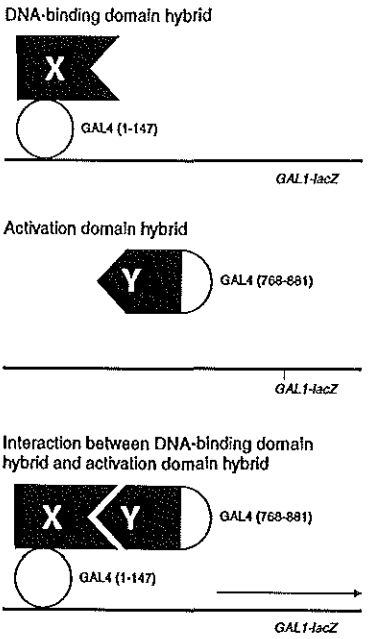


Figure 3.2 Yeast two-hybrid system. The DNA-binding domain of yeast protein GAL4 is used to generate a fusion with target protein X and the GAL4 activation domain is fused to target protein Y. Both domains are required for normal activation of transcription. Interaction of protein X and Y will lead to expression of the reporter gene lacZ

A second, more specific application is to test for an interaction between two proteins of interest, and to define domains of proteins involved in interactions. In this application, plasmids encoding two hybrid proteins are constructed and transformed into yeast. Interaction between the two proteins will lead to transcriptional activation of a reporter gene containing a binding site for GAL4 (figure 3.2). When an interaction has been detected, the interacting site within the proteins can be defined using truncated or otherwise mutated constructs.

The yeast two-hybrid system was applied to test whether hamartin and tuberin can interact. Several TSC1 and TSC2 constructs were tested and it was found that the two TSC proteins can form a complex and that the interaction between hamartin and tuberin is mediated by their predicted coiled-coil domains (chapter 3.6). The specificity of the interaction was further examined by two independent techniques, coimmunoprecipitation of endogenous proteins and immunofluorescence after transfection.

Coimmunoprecipitation

Coimmunoprecipitation is one of the most commonly used techniques to study protein-protein interactions *in vivo*. A major advantage of this method is that the endogenous proteins are studied in their natural cellular environment. This technique was applied to study the interaction between the TSC proteins, and both hamartin and tuberin could be recovered from immunoprecipitates of antisera specific for tuberin or hamartin (chapter 3.6).

Immunofluorescense

A third method to study the interaction between hamartin and tuberin was by immunofluorescense. The endogenous proteins were difficult to detect in cell systems, so both proteins were transfected into mammalian cells. In cotransfection assays, a specific punctated labelling pattern was observed to which hamartin and tuberin colocalised (chapter 3.6.1). A major disadvantage of the transfections is that the overexpression might change the normal behaviour of the proteins. Therefore it is important to interpret the results with caution. However, the results from the coimmunoprecipitations and immunofluorescense were in concordance with those obtained using the two-hybrid system, giving strong evidence that the twoTSC proteins function in a complex.

CHAPTER 3.6

Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products

Marjon van Slegtenhorst, Mark Nellist, Bas Nagelkerken, Jeremy Cheadle, Russell Snell, Ans van den Ouweland, Arnold Reuser, Julian Sampson, Dicky Halley and Peter van der Sluijs

Published in: Hum Mol Genet 1998;7;1053-1057.

Reprinted with permission

Tuberous sclerosis (TSC) is an autosomal dominant disorder caused by a mutation either in the *TSC1* or *TSC2* tumour suppressor gene. The disease is characterised by a broad phenotypic spectrum that can include seizures, mental retardation, renal dysfunction and dermatological abnormalities. *TSC2* encodes tuberin, a putative GTPase activating protein for rap1 and rab5. The *TSC1* gene was recently identified and codes for hamartin, a novel protein with no significant homology to tuberin or any other known vertebrate protein. Here we show that hamartin and tuberin associate physically *in vivo* and that the interaction is mediated by predicted coiled-coil domains. Our data suggest that hamartin and tuberin function in the same complex rather than in separate pathways.

INTRODUCTION

Tuberous sclerosis (TSC) is characterised by the widespread development of hamartomatous growths in many tissues and organs. The brain, eyes, kidneys, heart and skin are frequently affected but the lungs, skeleton and endocrine glands may also be involved (1). The lack of clues as to cellular functional abnormalities has meant that efforts to identify the primary underlying defect in TSC patients have focussed on positional cloning.

TSC is genetically heterogeneous, with loci on chromosome 9q34 *(TSC1)* and 16p13.3 *(TSC2)* (2). The *TSC2* gene was isolated in 1993 (3) and codes for tuberin, a 200 kDa (1807 amino acid) protein. Tuberin contains a relatively hydrophobic N-terminal domain (4) and a conserved 163 amino acid region close to the C-terminus that is homologous to the GTPase activating proteins (GAP) rap1GAP and mSpa1 (5).

The *TSC1* gene was recently identified (6) and codes for hamartin, a 130 kDa (1164 amino acid) hydrophilic protein with no significant homology to tuberin or other known vertebrate proteins.

We tested whether hamartin and tuberin could interact using the yeast two-hybrid system (7) and transfection assays. Further, in human cells in culture, we investigated the association between endogenous hamartin and tuberin by coimmunoprecipitation. Our data demonstrate that hamartin and tuberin associate physically *in vivo* suggesting that both proteins play a closely related role in an as yet undetermined physiological process.

RESULTS

The predicted coiled-coil domain in hamartin interacts with a putative coiled-coil domain in tuberin

The predicted amino acid sequences of hamartin and tuberin were analysed for potential interaction domains. COILS version 2.1 (8) identified a more extensive coiled-coil structure in hamartin than reported previously (6) (amino acids 719-998, window size 28) (figure 1a) while a less stringent analysis (window size 14) of the original tuberin sequence (3) predicted two coiled-coil domains at amino acid positions 346-371 and 1008-1021 (figure 1a). As coiled-coil domains have the capacity to form homophilic and heterophilic protein complexes (8), these domains were made the focus of subsequent yeast two-hybrid experiments.

A construct coding for tuberin (amino acids 1-1784, as described in reference 3) fused to the GAL4 DNA-binding domain was tested against constructs coding for the GAL4 transactivating domain fused to the N-terminal (XB1, amino acids 23-357) and C-terminal (EE1a, amino acids 334-1153) domains of hamartin (figure 1b). A strong, specific interaction was detected between tuberin and EE1a, containing the C-terminal, putative coiled-coil domain of hamartin. No interaction was detected between tuberin and the N-terminal domain of hamartin. Self-activation of GAL4 activity for any of the constructs used in this study was not observed.

To define the binding domain within hamartin more precisely a series of truncated constructs were analysed (figure 1b). Only construct ESA (amino acids 334-673), lacking the entire coiled-coil domain, did not interact with tuberin. Construct EE2 (amino acids 334-788) tested positive, suggesting that the first seven heptad structures in the coiled-coil domain were sufficient for hamartin to interact with tuberin.

The interaction domain in tuberin was also defined using partial constructs (figure 1c). A potential leucine zipper (amino acids 81-121), the two putative coiled-coil domains (amino acids 346-371 and 1008-1021) and a deletion construct (GRD; amino acids 1-41 and 861-1784) containing the GAP related domain were tested against the EE1 (amino acids 334-1153) hamartin construct. Only the most N-terminal coiled-coil construct (amino acids 346-371) tested positive. The specificity of the coiled-coil interaction was investigated in the two-hybrid assay using another coiled-coil containing protein, giantin (9), against hamartin and tuberin. No GAL4 activation occurred (data not shown), indicating that the interaction detected between the coiled-coil domains in hamartin and tuberin was specific.

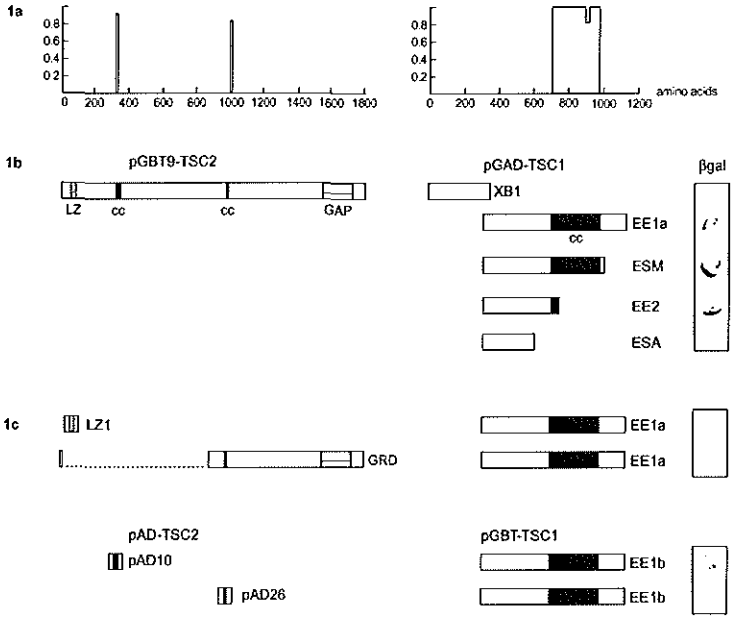


Figure 1. Coiled-coil predictions and mapping of the interacting domains of hamartin and tuberin in the yeast two-hybrid system. **(A)** Plot showing the position and the probability of the coiled-coil motifs in tuberin and hamartin. COILS version 2.1 window size 28 for hamartin and window size 14 for tuberin. **(B)** One N-terminal and four C-terminal *TSC1* constructs, fused to the DNA transactivation domain of GAL4 (XB1, EE1a, ESM, EE2 and ESA), were assayed for interaction with a full-length *TSC2* construct, fused to the DNA binding domain of GAL4. The deletion constructs of EE1a were created with the internal restriction sites SmaI (ESM), EcoRI (EE2) and SalI (ESA). The putative coiled-coil domains (cc; shaded boxes) in hamartin and tuberin and the N-terminal leucine zipper and C-terminal GAP-related domain in tuberin (hatched boxes) are indicated. The bait and prey constructs were cotransformed in yeast strain YGH1 bearing a lacZ reporter. Interaction was detected with the ß-galactosidase assay, positives resulting in a blue colour (last column). **(C)** The N-terminal leucine zipper (LZ1) and C-terminal construct (GRD), fused to the GAL4 DNA-binding domain, and the two putative coiled-coil structures fused to the GAL4 transactivation domain (pAD10 and pAD26) were tested against the EE1 hamartin construct fused to either the GAL4 DNA-binding- (EE1b) or transactivation -domain (EE1a).

Hamartin and tuberin colocalise in transfected mammalian cells

In order to confirm the two-hybrid results, the localisation of hamartin and tuberin in transfected cells was studied using immunofluorescent microscopy. A full-length *TSC1* cDNA in the pcDNA3.1 expression vector was transfected into COS cells. A distinct labelling pattern was observed, consisting of discrete structures in the cytoplasm (figure 2A). In contrast, expression of an epitope tagged full-length TSC2 construct produced a general cytoplasmic labelling pattern (figure 2B). When COS cells were cotransfected with the TSC1 and TSC2 construct, both hamartin and tuberin localised to the same structures as well as to the cell cytoplasm (figure 2C-E). Untransfected COS cells did not stain with the hamartin and tuberin specific antisera. Similar results were obtained in transfected HeLa and CHO cells (data not shown). To investigate whether the colocalisation was due to overexpression of hamartin and tuberin, several control proteins including the fragile X mental retardation protein, acid α-glucosidase and the C-terminal domain of polycystin, containing a predicted coiled-coil structure (10), were co-expressed with hamartin in COS cells. None of the controls colocalised to the hamartin positive structures, confirming that the colocalisation of hamartin and tuberin was specific.

Hamartin and tuberin coimmunoprecipitate *in vivo*

In order to investigate whether the observed association between hamartin and tuberin also occurred *in vivo*, the endogenous proteins were immunoprecipitated from HeLa cells and cultured human fibroblasts. Hamartin could be recovered from the immunoprecipitates of antisera specific for tuberin, while tuberin coimmunoprecipitated with hamartin when an antiserum specific for hamartin was used (figure 3). Identical results were obtained with different anti-hamartin and anti-tuberin antisera (data not shown). Preimmune sera and a control antiserum (against human acid α-glucosidase) were negative, demonstrating that the observed coimmunoprecipitation of tuberin and hamartin was specific and confirming that the association detected by the two-hybrid assay occurs in mammalian cells.

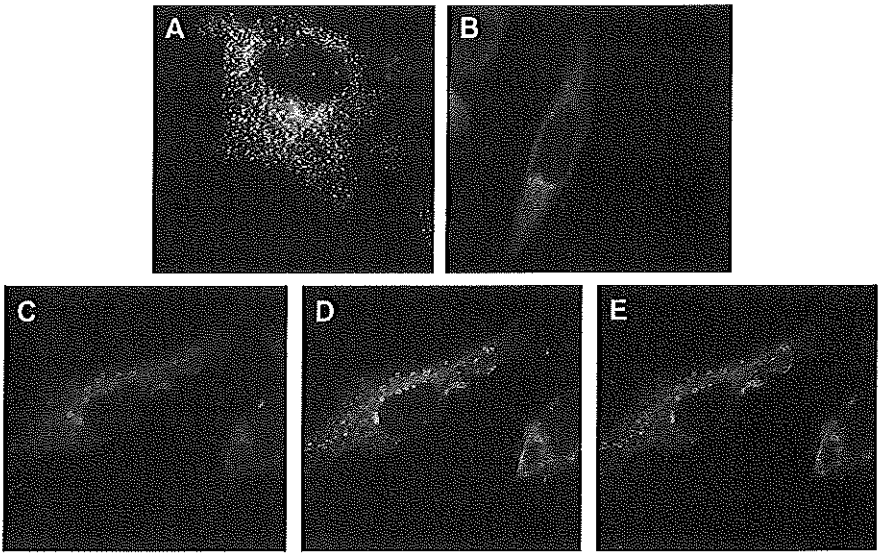


Figure 2. Colocalisation of hamartin and tuberin in COS cells. **(A)** COS cell, transfected with the full-length *TSC1* construct. **(B)** COS cell, transfected with the full-length *TSC2* construct. Both proteins were detected with specific rabbit polyclonal primary antisera, followed by anti-rabbit IgG secondary antisera conjugated to fluorescein (FITC). **(C, D and E)** Cotransfection of full-length *TSC1* and *TSC2* constructs. Cells were double labelled: **(C)** hamartin was with a specific rabbit polyclonal antiserum (as in **(A)**), followed by an anti-rabbit IgG Texas Red coupled secondary antibody and **(D)** tuberin was visualised with a mouse monoclonal against an N-terminal epitope tag sequence, followed by an anti-mouse IgG-FITC conjugate (Xpress; Invitrogen). **(E)** Colocalisation of hamartin and tuberin in the cytoplasm of COS cells (yellow).

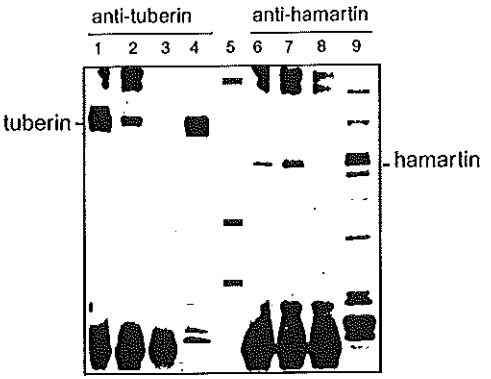


Figure 3. *In vivo* coimmunoprecipitation of tuberin and hamartin. Tuberin and hamartin were immunoprecipitated from fibroblast lysates with tuberin specific antiserum (lanes 1 and 6), hamartin specificantiserum (lanes 2 and 7) and an antiserum against human acid α-gluco-sidase (lanes 3 + 8). Lanes 4 + 9 contain the lysate prior to immuno-precipitation. The molecular weight marker is in lane 5 (from top to bottom: 230 kDa, 100 kDa, 80 kDa). Lanes 1 to 4 were incubated with anti-tuberin antibody and lanes 6 to 9 with anti-hamartin antibody. IgG heavy chains (50kDa) are visible in lanes 1, 2, 3, 6, 7 and 8. The additional lower molecular weight bands in lane 9 are most likely degradation products of hamartin, not con- sistently observed (data not shown).

DISCUSSION

To investigate the molecular mechanism underlying TSC, we tested for an interaction between tuberin and hamartin using three independent methods. In each case we showed that the proteins are partners. In view of the phenotypic overlap observed between TSC patients with either a *TSC1* or *TSC2* mutation, this suggests that inactivation of hamartin or tuberin may prevent the formation of a functional protein complex. A comparable scenario has been reported recently for the PKD1 and PKD2 proteins (10, 11). However, unlike polycystin 1 and 2, no regions of homology between tuberin and hamartin have been detected (6).

The interaction between tuberin and hamartin is mediated by potential coiled-coil domains. A predicted N-terminal coiled-coil domain in tuberin interacts with only a small part of an extensive coiled-coil region in hamartin. Preliminary results indicate that the coiled-coil domain in hamartin can form a homophilic complex (data not shown). We are currently investigating whether the additional coiled-coils in hamartin mediate interactions with additional proteins important in the pathogenesis of TSC.

Overexpression of hamartin in COS cells showed a distinct labelling pattern in the cytoplasm, while tuberin produced a general cytoplasmic labelling. When hamartin and tuberin were cotransfected in mammalian cells, tuberin was recruited in a specific manner to the hamartin positive structures and the proteins colocalised more generally in the cytoplasm. These data, together with the *in vivo* association of hamartin and tuberin detected by coimmunoprecipitations, support the results of the two-hybrid system and provide strong evidence that hamartin and tuberin exist as a complex.

Recently it has been demonstrated by the two-hybrid system that the C-terminal part of tuberin interacts with rabaptin-5, (12). Rabaptin-5 is an 115 kDa cytosolic protein, that is an effector for the endosomal small GTPase rab5 and therefore involved in endocytic fusion events (13). Consistent with the tuberin-rabaptin-5 interaction, tuberin has been shown to act as a GTPase activating protein for rab5 and to reduce the rate of fluid-phase endocytosis (12). It will be important to establish whether binding between hamartin and tuberin regulates the rab5 GAP activity of tuberin and to investigate the effect of hamartin expression on fluid-phase endocytosis and early endosome fusion. We are currently investigating the nature of the hamartin containing structures detected in the transfection experiments. The identification of additional endosomal proteins that interact with either hamartin, tuberin or both may help to clarify whether dysregulation of endocytosis is important in the aetiology of TSC.

MATERIAL AND METHODS

TSC1 and *TSC2* constructs

A full-length *TSC2* cDNA (nucleotides 1 - 5474) was derived from previously identified partial cDNAs (3) and cloned into the pGBT9 (Clontech) and pcDNA3.1HisA (Invitrogen) vectors. The *TSC2* C-terminal GAP domain construct (GRD) was made by digesting with SacII, leading to an in-frame deletion of amino acids 42-860. The full-length *TSC1* cDNA was amplified by RT-PCR with oligonucleotides 5'-TGAGGTAAACAGCTGAGGGG-3' and 5'-AAGGTCAAGAGGCATTTCAA-3' and cloned into pGEM-T Easy (Promega) and subsequently into pcDNA3.1. The remaining *TSC1* and *TSC2* constructs were derived by PCR using primers with linkers for direct restriction site cloning. pAD26, pAD10 and LeuZip were amplified from a *TSC2* cDNA clone using the primer pairs 5'-CTCGAATTC CACGCAGTGGAAGCACTCTG-3' and 5'-CTCGGATCCGGAAGGGTAATCCTTGATG ACC-3' for LeuZip, 5'-GGAATTCCAGACGTCCCTCACCAGTGC-3' and 5'-GCTCTAG AAGCCGTGAAGTTGGAGAAGA-3' for pAD26, 5'-GGAATTCGAGATCGTCCTGTCC ATCAC-3' and 5'-GCTCTAGACGCACATCTCTCCACCAGTT-3' for pAD10. The *TSC1* deletion constructs were amplified by RT-PCR with the primer pairs 5'-CCCGGG GGAC GACGTGACAGCTGTCTTT-3' and 5'-CCCGGGGAGTGGTCATACCACAAACCAT-3' for XB1, 5'-GGATCCCATGATGAGTCTCATTGTAGTC-3' and 5'-GGATCCGACACG GCTGATAACTGAACCA-3' for EE1a, 5'-GGATCCTCATGATGAGTCTCATTGTAGT-3' and 5'-GGATCCCGACACGGCTGATAACTGAACCA-3' for EE1b. pAD26 and pAD10 were cloned into pAD-2.1 (Stratagene); LeuZip and EE1b into pGBT9 (Clontech); and XB1 and EE1a into pGADGH (Clontech). Three deletion constructs (ESM, EE2 and ESA) were generated using internal restriction sites (SmaI, EcoRI and SalI respectively). All constructs were checked by sequencing and where appropriate by in vitro coupled transcription-translation assay (TnT system, Promega).

Yeast two-hybrid assay

Yeast host strain YGH1 was co-transformed with 2.5µg of each plasmid according to the SBEG method (14). Transformants were plated on minimal media lacking Tryp and Leu. After 3 days, 3 colonies per interaction were plated on media lacking His, Tryp and Leu, and growing colonies were tested for ß-galactosidase activity with the filter assay.

Generation of tuberin and hamartin specific antisera

Two fusion proteins containing N-terminal histidine tag sequences and amino acids 1535-1784 from tuberin and 543-1087 from hamartin were overexpressed in bacteria and affinity purified under denaturing conditions according to the manufacturer's protocol (Qiagen Gmbh). The final eluates were concentrated through an Amicon PM-10 filter and dialysed against phosphate buffered saline. New Zealand white rabbits were immunized with 150 µg of purified fusion protein suspended in Freund's complete adjuvant, and boosted at 4 week intervals with 150µg fusion protein in Freund's incomplete. Serum was collected 10 days after injection of the immunogen. The resulting polyclonal sera were checked for specificity by Western blot and transfection experiments.

Immunofluorescence

Expression constructs were transfected into COS cells with lipofectamine, as recommended by the manufacturer (Gibco BRL). For immunocytochemistry, cells were fixed in 3% paraformaldehyde (10 minutes, room temperature), quenched with 50 mM NH_4Cl (10 minutes), and permeabilized in 0.1% Triton X-100 (5 minutes). Cells were incubated with primary antibodies, followed by fluorescein (FITC) or Texas Red (TRITC) coupled secondary antibodies. Images were captured using the Power Gene FISH system on a Leica DM RXA microscope. Images were processed using a filter wheel (Chroma Technology) and the Adobe photoshop software package. In addition to the polyclonal sera against hamartin and tuberin generated as part of this study, antibodies against FMRP, polycystin and α-glucosidase were used for control experiments.

Coimmunoprecipitations

Washed cells (one 10cm plate) were lysed in 700µl TNE buffer (40mM Tris-HCl, pH7.4, 150mM NaCl, 1mM EDTA, 0.5% Nonidet P-40 containing 0.2mM PMSF), according to standard procedures (15) and cleared by centrifugation (10 000g, 10 minutes, 4°C). The supernatant was recovered and incubated with 2µl antisera for 60 minutes on ice before the addition of 30µl 50% protein A-Sepharose suspension. After gentle rotation for 60 minutes at 4°C, the immunoprecipitates were washed extensively with TNE buffer. Immunoblotting was performed according to standard procedures (15) and coimmunoprecipitating proteins were detected using the appropriate antibodies and enhanced chemiluminescence (Amersham).

REFERENCES

1. Gomez, M.R. (1980) *Tuberous Sclerosis*, 2nd edition. Raven Press, New York.
2. Janssen, L.A.J., Sampson, J.R., van der Est, M., Deelen, W., Verhoef, S., Daniels, I., Hesseling, A., Brook-Carter, P., Nellist, M., Lindhout, D., Sandkuijl, L.A. and Halley, D.J.J. (1994) Refined localisation of *TSC1* by combined analysis of 9q34 and 16p13 data in 14 tuberous sclerosis families. *Hum Genet*, **94**, 437-440.
3. The European Chromosome 16 Tuberous Sclerosis Consortium. (1993) Identification and characterisation of the tuberous sclerosis gene on chromosome 16. *Cell*, **75**, 1305-1315.
4. Maheshwar, M., Sandford, R., Nellist, M., Cheadle, J.P., Sgotto, B., Vaudin, M. and Sampson, J. (1996) Comparative analysis and genomic structure of the tuberous sclerosis 2 *(TSC2)* gene human and pufferfish. *Hum. Mol. Genet.*, **5**, 131-137.
5. Maheshwar, M., Cheadle, J.P., Jones, A.C., Myring, J., Fryer, A.E., Harris, P.C. and Sampson, J.R. (1997) The GAP-related domain of tuberin, the product of the TSC2 gene, is a target missense mutations in tuberous sclerosis. *Hum. Mol. Genet.* **6**, 1991-1996.
6. 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, R.G., Cheadle, J.P., Jones, A.C., Tachataki, M., Ravine, D., Sampson, J.R., Reeve, M.P., Richardson, Wilmer, F., Munro, C., Hawkins, T.L., Sepp, T., Ali, J.B.M., Ward, S., Green, A.J., Yates, J.R.W., Kwiatkowska, J., Henske, E.P., Short, M.P., Haines, J.H., Jozwiak, S. and Kwiatkowski, D.J. (1997) Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science*, **277**, 805-807.
7. Fields, S. and Song, O.(1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246.
8. Lupas, A., Van Dyke, M. and Stock, M. (1991) Predicted coiled coils from protein sequences. *Science* **252**, 1162-1164.
9. Lindsted, A.D., Foguet, M., Renz, M., Seelig, H.P., Glick, B.S. and Hauri, H.P. (1995) A C-terminally-anchored Golgi protein is inserted into the endoplasmic reticulum and then transported to the Golgi apparatus. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5102-5105.
10. Qian, F., Germino, F.J., Cai, Y., Zhang, X., Somlo, S. and Germino, G.G. (1997) PKD1 interacts with PKD2 through a probable coiled-coil domain. *Nature Genet.* **16**, 179-183.
11. Tsiokas, L., Kim, E., Arnould, T., Sukhatme, V.P. and Walz, G.(1997) Homo- and heterodimeric interactions between the gene products of PKD1 and PKD2. *Proc. Natl. Acad. Sci. U.S.A.* 94, 6965-6970.
12. Xiao, G-H., Shoarinejad, F., Jin, F., Golemis, E.S. and Yeung, R.S. (1997) The TSC2 gene product, tuberin, functions as a Rab5 GTPase activating protein (GAP) in modulating endocytosis. *J. biol. Chem.* **272**, 6097-6100.
13. Stenmark, H., Vitale, G., Ullrich, O. and Zerial, M. (1995) Rabaptin-5 is a direct effector of the small GTPase Rab5 in endocytic membrane fusion. *Cell* **83**, 423-432.
14. Klebe, R.J., Harriss, J.V., Sharp, Z.D. and Douglas, M.G. (1983) A general method for polyethylene-glycol-induced genetic transformation of bacteria and yeast. *Gene* **25**, 333-341.
15. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbour Laboratory Press, Cold Spring Harbour, NY.

ACKNOWLEDGEMENTS

We are grateful to Bert Eussen and Tom de Vries-Lensch for help with the figures, to Arjenne Hesseling-Janssen and Shelley Idziaszcyk for technical assistance, and Professors H. Galjaard and D. Lindhout for support. We thank Filippo Tamanini and Miriam Goedbloed for providing us with the *FMRP* and *PKD1* constructs and antibodies. This work was funded by the Dutch Organization for Scientific Research (NWO), Dutch Kidney Foundation (Nierstichting), Dutch Cancer Society (NKB), Tuberous Sclerosis Association of Great Britain, National Tuberous Sclerosis Association (USA), Medical Research Council and Cancer Research Wales. Peter van der Sluijs is an investigator of the Royal Netherlands Academy of Arts and Sciences

CHAPTER 4

GENERAL DISCUSSION

GENERAL DISCUSSION

4.1 Positional cloning

New developments and collaborations

Positional cloning of disease genes has accelerated enormously during the past years. On one side, many new techniques have been developed to screen rapidly large regions of the genome for genes, on the other side the Human Genome Mapping Project has led to a large collection of expressed sequencing tags (ESTs) in databases and to the genomic sequence of parts of several chromosomes. The identification of the TSC1 gene benifited from both these developments. It was the systematic application of a whole collection of methods rather than any one particular technique that finally allowed the TSC1 gene to be isolated. In addition, a collaboration with groups in the USA and UK made it possible for each group to concentrate on specific tasks, thus avoiding duplication of effort. A crucial contribution was made by the TSC patients, their families and their Associations by participating in this study.

The time span between localisation and identification of the TSC1 gene

Despite these new developments and collaborations, it took about 10 years from the original discovery of linkage to the isolation of the gene itself. The major rate limiting factors were the locus heterogeneity, which complicated the linkage analysis, conflicting recombinant data, and the absence of large rearrangements in the region.

Several candidate TSC loci were identified but subsequently excluded. Additional difficulties resulted from the high proportion of sporadic cases and small families, which is probably due to reduced reproductive fitness of the patients.

Conflicting recombinant data were mainly caused by false positive clinical diagnosis and mis-interpretation of marker data. Even though TSC is considered to be almost completely penetrant, the variability of clinical expression can hamper a correct phenotypic classification in both affected and unaffected individuals. Ultimately, recombination events in clinically rigorously investigated unaffected individuals in TSC families proved to have been of more value to narrow down the region than the ones in presumably affected persons.

Only a few other genes have been identified by positional cloning without any indication from cytogenetic abnormalities. Comparable examples are the searches for the genes for cystic fibrosis (Rommens *et al.*,1989; Riordan *et al.*,1989; Kerem *et al.*,1989), Huntington's

disease (Huntington's Disease Collaborative Research Group, 1993) and myotonic dystrophy (Harley *et al.*, 1992). The time span between the first linkage reports and the identification of the genes for Huntington's disease and myotonic dystrophy also covered up to ten years.

4.2 First applications after the cloning of the TSC1 gene

Mutation analysis

The first application after the cloning of the gene was mutation analysis in our collection of TSC patients. The detection of the mutation in the TSC1 gene (13% of our cases) or the TSC2 gene (thus far 15% of the cases) enables reliable genetic counseling for relatives at risk and prenatal testing have become options for these patients and their families. Previously, prenatal testing was only possible for a few large families which had been linked to either the chromosome 9 or the chromosome 16 locus, or in families shown to have a mutation in the TSC2 gene. Mutation analysis has the obvious advantage of being a direct method, the application of which is not restricted by the chance of recombination and the lack of informative markers.

It is likely that some of the mutations in the TSC1 gene have remained undetected. Small mutations in regions outside of the coding region, for example in the promotor region and other untranslated regions, which may have a role in controlling expression levels, have not been screened for yet. Furthermore, each mutation detection method, including SSCP, is known to be less than 100% sensitive.

Since there are no data supporting the existence of a TSC3 gene in familial cases, the most likely explanation for the paucity of TSC1 mutations is that the majority of the disease-causing mutations will be located in the TSC2 gene. Preliminary data from our group support this hypothesis (A van den Ouweland, personal communication). Explanations for the excess of TSC2 mutations could include the larger size of the TSC2 gene and the greater variety of mutations detected at this locus. More mutations in the TSC2 gene is consistent with the observation of more frequent LOH at 16p13.3 in TSC-associated lesions.

On the other hand, linkage data in a subset of TSC families suggest an equal distribution of TSC1 and TSC2 mutations, at least in the large TSC families. An explanation could be that the TSC1 mutations are underrepresented among sporadic cases as was concluded in a recently published study (Jones *et al.*, 1997). However, in our own studies, a statistically significant difference in detection rate of TSC1 mutations among familial and sporadic cases

could not be found (chapter 6). Moreover, no obvious differences have been detected so far in the severity of the phenotype between TSC1 and TSC2 disease in sporadic or familial cases.

An alternative hypothesis is that a proportion of the sporadic patients have a mutation in another, as yet unidentified, gene. If this is the case, patients with a mutation in 'TSC3' would be expected to have a more severe phenotype with more severely reduced reproductive fitness, since all large TSC families are either linked to TSC1 or TSC2. After both genes have been completely screened for mutations, the percentage of missed mutations in chromosome 9 and 16 linked families can be extrapolated to the small families and sporadic cases, and this may give an indication whether a putative additional gene in TSC might exist.

Genotype versus phenotype in TSC1 disease

The mutational spectrum in the TSC1 gene is different from the TSC2 gene. In TSC1 mostly small mutations leading to a truncated protein have been detected, while the TSC2 gene also harbours large deletions and missense mutations. The lack of large deletions at a disease locus can often be explained by the presence of adjacent genes that may play an essential role in embryonic development and survival. Potential intragenic transcripts and a putative growth repressor gene have been identified 5' of the TSC1 gene by computer analysis, but their existence and possible functional meaning have not yet been confirmed. Although the spectrum of mutations is different in TSC1 and TSC2, the mechanism by which the mutations operate is the same: 'loss-of-function'.

As described in chapter 2.6.4, no correlation was found between the type and location of the mutation in the TSC1 gene and the clinical phenotype in the Rotterdam patient set. A recent report (Jones *et al.*, 1997) suggested that there was a reduced risk for mental retardation in TSC1 disease, but our data shows that 50% of the patients with a mutation in the TSC1 gene are mentally retarded, which is in agreement with the overall reported prevalence of mental retardation in TSC (Gomez *et al.*, 1988). Since the clinical variation observed within TSC families is wide, the lack of any obvious phenotype-genotype correlation is, perhaps, not surprising. Furthermore, the majority of mutations are loss-of-function mutations leading to the same null effect at the protein level.

The absence of a correlation between genotype and phenotype in TSC1 shows the complexity of this 'monogenic' disease. Studies of other autosomal dominant phakomatoses

neurofibromatosis type 1 (NF1) (Cnossen, 1997) and von Hippel-Lindau disease (Crossey *et al*, 1994) are similar in this respect. The timing and tissue distribution of second hits may be responsible for a large part of the clinical variability. Additional factors that may contribute to the influences of other genetic factors (modifying genes), and somatic mosaicism (Verhoef *et al*, 1995).

Since the clinical phenotype is unlikely to help solve the question whether a patient has a defect in the TSC1 or TSC2 gene, both genes need to be screened when mutation analysis is requested. For the TSC1 gene, all exons can be screened for mutations in one SSCP run. Based on the number of mutations in TSC1 observed with this method so far, it is expected that a TSC1 mutation will be detected in approximately 13% of the patients. Screening for TSC2 mutations has proven to be more laborious and complicated, because different methods have to be employed to detect the different types of mutations in this gene.

To overcome this problem, it would be of great help to develop a test to look at protein expression levels in tumour material from patients, but at present this is not feasible. Restrictions are the limited availability of tumours, and the current lack of data on the normal expression pattern of the TSC gene products. Further it is not known whether loss of hamartin affects the level of tuberin or vice versa.

Automated mutation analysis with the newly developed DNA chip technology (O'Donnell-Maloney and Little, 1996) will enable quick screening of many genes in the near future, but it is not likely that the TSC genes will be included in the first commercial applications of this technique. The developmental costs are relatively high and, because TSC is a relatively rare disorder, the commercial value would be low.

4.3 How do the different lesions develop in TSC patients?

TSC-specific lesions mainly occur in tissues that are normally populated by cells derived from the neural crest and neural tube (Lalier, 1991; Johnson *et al.*, 1991). In humans, both structures develop from a thickened area of embryonic ectoderm called the neural plate. The neural crest is a transient embryonic structure from which cells migrate extensively before they develop into the peripheral nervous system and a variety of differentiated cell types. The neural tube differentiates into the central nervous system thereby giving rise to all neurons and macroglial cells.

Loss of tuberin has been detected in both renal and brain lesions from TSC patients (Henske *et al.*, 1996-1997), suggesting that they only develop after a second, somatic, hit. Interestingly, loss of tuberin does not affect all cell types in this lesion. To illustrate the diversity of the TSC lesions, a schematic model for the development of some of the most common features has been drawn in figure 4.1.

Hamartomas

Cardiac rhabdomyoma is a characteristic benign tumour in TSC patients, and LOH has been detected in 50% of the cases investigated (table 1.3).

Benign TSC lesions in the kidneys (cysts and angiomyolipomas) are commonly multiple and bilateral and the histopathologic findings are practically diagnostic of TSC (Bernstein *et al.*, 1986). LOH has been detected in both lesions, but much more frequently in angiomyolipomas.

Cortical tubers are histologically heterogeneous lesions, composed of abnormally shaped neurons and giant cells. The cells within tubers stain positive for nestin, which is a marker of an immature cellular phenotype (Crino *et al.*, 1996), suggesting that the cells in tubers have retained the molecular phenotype of embryonic or immature neurons due to early disruption of neuronal maturation caused by loss of either tuberin or hamartin. LOH has been demonstrated within tubers (Green *et al.*, 1994), suggesting that a second hit is also required for tuber development. The subependymal nodule is another distinctive neuropathologic feature of TSC. Histologically, they are composed of cells that are similar to those found in the cortical tubers. However, they frequently undergo calcification, while the cortical tubers generally do not (Richardson, 1991).

In summary, most TSC-associated hamartomas show abnormal cells, which are often enlarged and in the wrong location, suggesting that the lesions share a common developmental mechanism. The number and location of hamartomas in a TSC patient are most likely determined by the second hits at the cellular level.

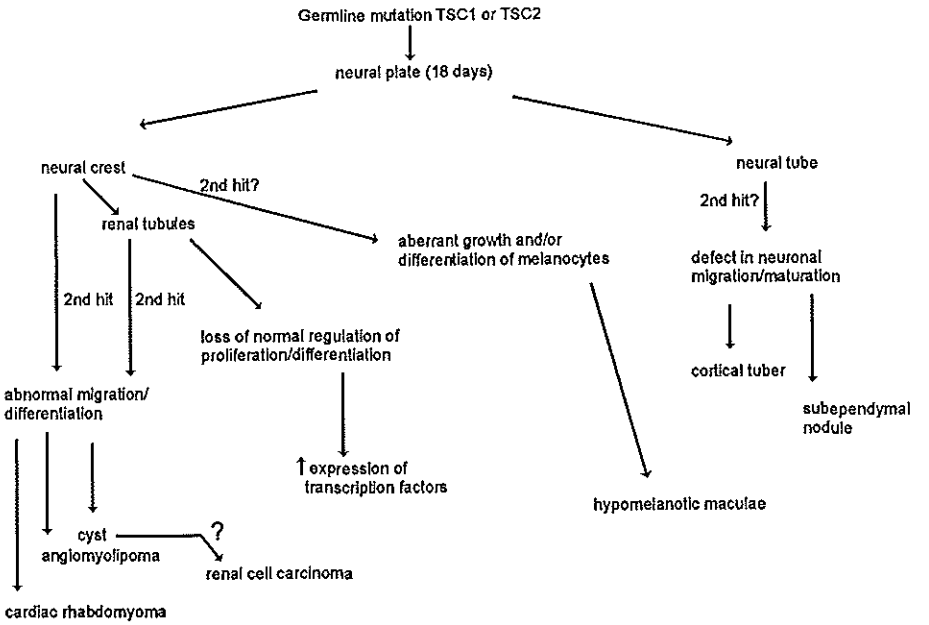


Figure 4.1. Schematic representation of the development of some distinct characteristic TSC lesions. Some common TSC-associated lesions have been selected (cardiac rhabdomyoma, cyst, angiomyolipoma, renal cell carcinoma, hypomelanotic macule, cortical tuber and subependymal nodule), which are likely to develop via distinct pathways. Each phase in the development of the lesion is indicated by an arrow. It is not clear whether all lesions require a second somatic hit.

Renal cell carcinoma

In some TSC patients malignancies occur in the kidneys in the form of renal cell carcinoma (RCC). It appears unlikely that RCCs develop from the atypical cellular changes in angiomyolipomas. Instead renal cysts, with their epithelial cell proliferation, are supposed to be a more probable origin of RCC (Bernstein and Robbins, 1991).

TSC-associated RCC occurs at an earlier age than sporadic RCC. Another difference is that half of the TSC-associated tumours occur bilaterally (Cook *et al.*, 1996). Some of the tumours in TSC patients have been reported to have a different immuno-phenotype than sporadic RCC (Bjornsson *et al.*,1996), and it remains to be investigated whether mutations at 9q34 and 16p13 occur in sporadic RCC.

Studies in the Eker rat show that the RCC develops through multiple stages from phenotypically altered renal tubules to adenomas and carcinomas, and that LOH is already detected in very early preneoplastic lesions/stages (Kubo *et al.*, 1995). Therefore the second

hit in one of the TSC genes is likely to result in loss of normal regulation of proliferation and/or differentiation of the renal tubules.

Several genes with an increased expression in Eker renal carcinoma cells have been identified (Hino *et al.*, 1995; Urakami *et al.*, 1997). Almost all of the genes identified so far with an increased expression in TSC-associated RCC are members of the AP-1 transcription factor family (Lau and Nathans, 1986; Szabo *et al.*, 1996). The increased expression suggests that these genes may have a role in the development of RCC, but the mechanisms by which they are regulated are largely unknown.

Hypomelanotic macule

Hypomelanotic macules are common lesions in TSC patients, but are also frequently detected in the normal population. They are composed of melanocytes deficient in the production of melanin (Fitzpatrick, 1991). Melanocytes are also neural crest derived and apparently migrate without problem to their destination, but fail to differentiate into mature cells. Whether a second hit is involved in this process remains to be investigated.

In conclusion, the mechanism by which mutations in either the TSC1 or TSC2 gene cause the disease is predominantly by homozygous inactivation by a first hit in the germline and a second hit in the abnormal tissue. LOH has been detected in only a minority of the lesions studied, but it could very well be that most of the second hits are small changes. Very recently, the first somatic point mutation in the wild-type TSC1 allele of an RCC from a TSC patient with a known TSC1 germline mutation has indeed been described (TSC1 consortium, 1997).

Hypomelanotic maculae in TSC patients may be an exception to the second hit mechanism. There is no LOH data available for this skin lesion. Furthermore the nature of the lesion is different from the hamartomas whereas it is the most frequent feature in TSC patients. It is possible that reduced levels of the TSC1 or TSC2 gene product, due to the presence of only one normal copy of the gene (haploinsufficiency), leads to the development of this particular lesion.

4.4 The function of the TSC gene products

4.4.1 Function of hamartin

The predicted sequence of the TSC1 protein, hamartin, does not resemble that of any known protein. Two potential functional domains are an N-terminal transmembrane domain and an extensive coiled-coil region at the C-terminus. We have recently demonstrated that hamartin colocalises in the cytoplasm with vesicle-like structures of unknown origin and that tuberin is recruited to these compartments. Since a mutation in either of the two genes causes a similar clinical phenotype, interaction between hamartin and tuberin was one of the first hypotheses to test for. The interaction between hamartin and tuberin was shown to occur *in vivo* and *in vitro* and is mediated by coiled-coil domains (chapter 3.6). Preliminary results indicate that the coiled-coil domain in hamartin can also form a homophilic complex, but whether this is of functional importance needs to be elucidated. Since hamartin interacts with tuberin, hamartin is most likely associated with the function of tuberin in the pathology of TSC.

So far, abundant expression of the TSC1 gene has only been detected on Northern blots with RNA from adult tissues. It will be interesting to study the expression pattern in early stages of development, because some of the TSC lesions develop very early in life. Preliminary data from studying a giant cell astrocytoma from a TSC patient suggest that expression of tuberin is lost, but there is no evidence for abnormal or elevated hamartin expression. Additional expression studies in a large number of TSC-associated tumours should be performed in the near future to test whether there is an effect of the absence of one of the TSC proteins on the expression pattern of the other. Expression studies in both normal and hamartin/tuberin negative cell lines will contribute to knowledge about the function of the TSC gene products.The effect of tuberin or hamartin deficiency on expression levels of other genes is another important approach in this endeavour. Additional valuable data can be collected from studies in cell lines from foetal tissues, which can be induced to differentiate.

Immunocytochemistry and RNA *in situ* hybridisation can be applied to study the differences in the levels of expression of the TSC mRNAs and proteins, thereby giving an indication about the level at which expression is controlled. Further, studying the differences between expression in normal and TSC-associated tissues can be a step forward in unravelling the tumour-suppression function of both hamartin and tuberin.

4.4.2 Function of tuberin

The Eker rat

Apart from LOH studies and the nature of the tumours in TSC (bilateral, multiple and focal), functional evidence for a tumour-suppressor role of the TSC2 gene has now been demonstrated in a naturally occurring animal model, the Eker rat. Re-introduction of the wild-type TSC2 gene in Eker cell lines lacking functional tuberin, suppresses growth and tumourigenicity (Jin *et al.*,1996; Kobayashi *et al.*, 1997). In the first studies in the rat, tumours were detected only in the kidney (renal cell carcinoma), uterus (leiomyomata), spleen (hemangiosarcoma) and pituitary. In contrast, TSC patients display predominantly benign brain-, heart-, and kidney-tumours and a variety of skin lesions. Quite recently, closer examinations of the brain of Eker rats demonstrated the presence of subependymal and subcortical hamartomas (Yeung *et al.*,1997), which provides evidence for remarkable pathological similarities between brain lesions in the Eker rat TSC patients. These findings open possibilities to study cell lines derived from these specific lesions to learn more about the mechanisms by which they develop. It is not clear, however, whether the Eker rat can serve as a model for epilepsy. It would be interesting to check whether cardic rhabdomyomas are present in animals which died in utero or very early in life. Despite the differences between the rat and human, the Eker rat has already proven its importance in studying tumourigenesis and will contribute in future studies to more knowledge about some functional aspects of tuberin.

Apart from the Eker rat, future studies will focus on additional animal models. Transgenic and knock-out mice are being developed at the moment for both TSC1 and TSC2. Targeted transgene expression with inducible promotors (cell-type, tissue and time specific) can contribute to studying the effect of second hits on the development of the different lesions.

Comparison of the conservation of gene sequence and structure through evolution can provide additional insight into gene function. So far, TSC2 homologues have been isolated from mouse, rat, *F. rubripes,* and *D. Melanogaster*. Since many basic biochemical processes are conserved through evolution, TSC1 homologues would also be expected to be present in these and other organisms.

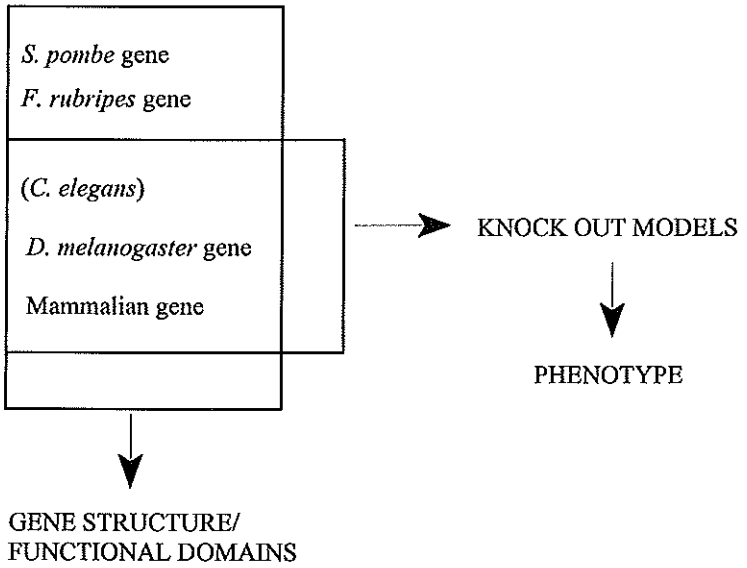


Figure 4.2 Applications of TSC genes in other species
Sequence comparison against sequences of animals at relatively great evolutionary distance (yeast, worm and fish) can indicate the presence of functional domains and important aspects of gene structure. Rodent, fly and worm are most frequently used as knock out models to compare phenotype. These animal models can make contributions to knowledge about the function of a gene product. The *TSC1* and *TSC2* genes have not been detected in *C. elegans* yet.

Genetic pathway studies are most powerful in the worm and fly. Given a particular phenotype, selection for suppressors and enhancers of the phenotype is relatively easy. For example, the rap-dependent signalling pathway, like the ras-signalling pathway, is conserved in the worm (Oliver, 1996), giving opportunities to identify upstream and downstream (target) molecules in this model system. Drosophila offers a large catalogue of mutations with well-defined phenotypes (Gelbart *et al.*, 1997). In Drosophila, both TSC1 and TSC2 homologues have been cloned (Gerald Rubin, personal communication). The TSC2 homologue was identified through a homozygous lethal mutant in the fly and was called gigas, due to the presence of larger wing bristles and eye cells. The TSC1 homologue was cloned by searching GenBank for fly ESTs. These different animal models can provide insight into the structure and functions of the TSC1 and TSC2 genes and into the pathogenesis of TSC itself (figure 4.2).

Putative GAP activity of tuberin

Tuberin displays a variety of putative functional domains. In the past four years the GAP related domain at the C-terminus has been the target of functional studies. Tuberin has been shown to act as a GTPase activating protein (GAP) for rap1 and rab5, two members of the ras superfamily of the small GTPases. The different members of this family play critical roles in the control of cell growth, differentiation and proliferation. They are either in the active GTP-bound form or in the inactive GDP-bound form. Regulation occurs through a variety of distinct factors, including GAPs. The biological role for rap1 has not been identified yet, while rab5 regulates early endosome fusion events.

How does tuberin regulate rap1?

Tuberin has been shown to function as a negative regulator of rap1 (Wienecke *et al.*, 1995). Although not much is known about the precise function of rap1, there is a striking similarity between the effector domains of ras and rap1, suggesting that ras and rap1 might bind to the same down-stream effectors (Noda, 1993). Unlike ras, however, rap1 does not have oncogenic potential, but rather seems to show growth-promoting activity (McCormick, 1995). Some reports have localised rap1 and tuberin to the Golgi-compartment (Wienecke *et al.*, 1996), whereas our studies detected a more general cytoplasmic labelling pattern for tuberin (chapter 7). In addition, the rap1 expression pattern has been shown to depend mainly on the cell type or tissue studied. For example in fibroblasts, rap1 has also been localised to the early and late endosomes (Pizon *et al.*, 1994; Beranger *et al.*, 1991). A role for rap1 has been proposed in a pathway which is antagonistic to ras signalling (Cook *et al.*, 1993; Urano *et al.*, 1996), but little solid evidence exists to support this hypothesis. Putative downstream effector molecules of rap1 include raf1 (Burgering *et al.*, 1993) and RalGDS (Bos *et al.*, 1997). However, it is not clear whether and/or how they mediate rap1 function and only additional studies can clarify this.

Alterations in the rap1 signalling pathway are important in the development of certain sporadic gliomas (Gutman *et al.*, 1997), suggesting that alterations in the rap1 signalling pathway may be associated with the development of human astrocytomas. Mutations in the TSC2 gene may disrupt the normal interaction of tuberin with the rap1 gene product and

thereby cause abnormal growth. Although an interaction between these two proteins has not yet been demonstrated, it is tempting to speculate that rap1 plays a physiologically relevant role in the tumour suppression function of tuberin.

Neurofibromin, the protein product of the neurofibromatosis I (NFI) gene, has been shown to stimulate the intrinsic GTPase activity of ras (Johnson *et al.*,1994), suggesting that neurofibromin may also function as a tumour suppressor through a GAP like activity. The link between ras and rap on one side, and the tumour suppressor genes NF1 and TSC2 on the other side, raise the possibility that the growth inhibitory function of neurofibromin and tuberin have much in common.

One report (Soucek *et al.*, 1997) provides evidence that tuberin is a potent regulator of G0/G1 transition during the cell cycle. It will be interesting to investigate the effect of hamartin on this vital process, e.g. by the use of specific TSC1-antisense oligonucleotides.

Is tuberin an effector molecule for rab5?

Members of the rab family of small GTPases are required in the control of vesicle-mediated transport. More than 30 different rab proteins have been identified and each rab protein is found at a particular stage of a membrane transport pathway (Nuoffer and Balch, 1994).

A putative role for tuberin in the membrane transport pathway was first postulated when rabaptin-5 was isolated by a two-hybrid screen with tuberin as bait (Xiao *et al*, 1997). Rabaptin-5 is a 100kDa cytosolic protein involved in endocytic fusion events and behaves as an effector for the endosomal GTPases rab5 and rab4 (Stenmark *et al.*, 1995). Also, tuberin has been shown to stimulate the GTP hydrolysis of rab5 and to behave as a negative regulator of endocytosis (Xiao *et al.*, 1997).

A simplified model for a role the TSC genes might play in rap1 and rab5 signalling is proposed in figure 4.3.

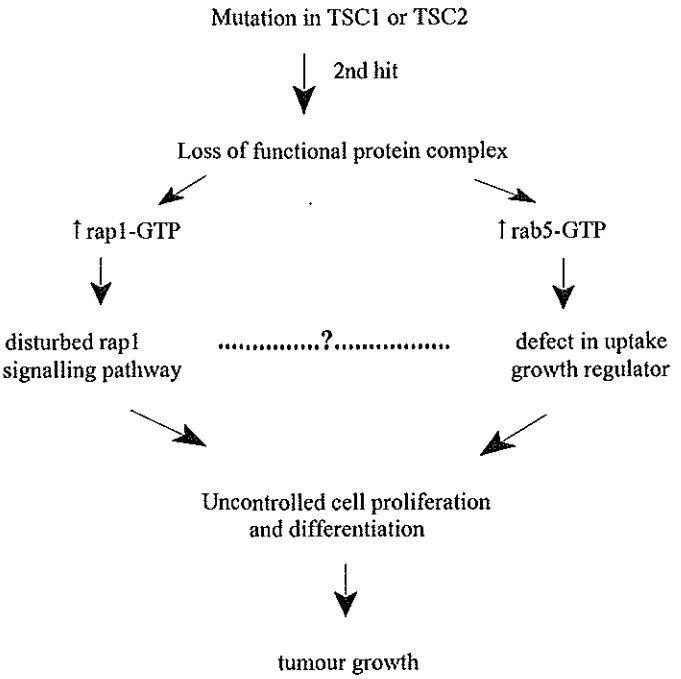


Figure 4.3 Simplified model for rap1/rab5 signalling pathway in TSC. Loss of a functional tuberin/hamartin complex by a mutation in either of the genes, accompanied by a second somatic mutation leads to more rap1 and rab5 in the active (GTP-bound) form.Tumour growth is caused by uncontrolled cell proliferation and differentiation.

Unfortunately, we have not been able to confirm the interaction between tuberin and rabaptin-5, either by yeast two-hybrid screening or coimmunoprecipitation studies. The identification of additional interacting proteins for hamartin and tuberin may help to clarify whether dysregulation of endocytosis and/or the rap1 signalling pathway are important to the aetiology of TSC. Preliminary results indicate that hamartin and tuberin are part of a larger complex (Mark Nellist, unpublished results), so it is expected that additional interacting proteins will be identified in the near future.

Cell fractionation studies indicated that hamartin as well as tuberin are associated with membranes (unpublished results), however neither hamartin nor tuberin colocalised with any early- and late-endosomal, lysosomal or Golgi markers. At present, the nature of the hamartin-positive structures, detected in the transfection assays, are the subject of further investigation. Electron microscopic examination of these structures will be facilitated by the construction of cell lines stably transfected with TSC1 and TSC2. Whether hamartin functions as an effector molecule of tuberin can be studied using in vitro GAP assays and by measuring the effect of hamartin and tuberin expression on fluid-phase endocytosis.

Some major topics of interest in the near future will be the study of the role of hamartin and tuberin in cell cycle control and endocytosis. A combinational approach will be necessary to study the different putative pathways and the links between them.

In summary, a few examples of experimental approaches have been discussed, of which some have already been initiated. They are all likely to contribute to an insight into the TSC pathology, but it is difficult to predict how and when the results might be used to improve the treatment of TSC patients.

References

Ad Hoc Committee on Mutation Nomenclature. (1996) Update on nomenclature for human gene mutations. Hum. Mut., 8, 197-202.

Ahlsén, G., Gillberg, I.C., Lindblom, R. and Gillberg, C. (1994) Tuberous sclerosis in Western Sweden. Arch. Neurol., 51, 76-81.

Au, K.S., Murrell, J., Buckler, S.H. et al. (1996) Report of a critical recombination further narrowing the TSC1 region. J. Med. Genet., 33, 559.

Au, K.S., Rodriguez, J.A., Rodriguez, E. et al (1997) Mutations and polymorphisms in the tuberous sclerosis complex gene on chromosome 16. Hum. Mut., 9, 23-29.

Au, K.S., Rodriguez, J.A., Finch, et al. (1998) Germ-line mutational analysis of the TSC2 gene in 90 tuberous-sclerosis patients. Am. J. Hum. Genet., 62, 286-294.

Ballabio, A. (1993) The rise and fall of positional cloning? Nature Genet., 3, 277-279.

Beranger, F., Goud, B., Tavitian, A. and De Grunzburg, J. (1991) Association of the Ras-antagonistic Rap1/ Krev-1 proteins with the Golgi complex. Proc. Natl. Acad. Sci. U.S.A., 88, 1606-1610.

Bernstein, J., Robbins, T.O. and Kissane, J.M. (1986) The renal lesions of tuberous sclerosis. Semin. Diagn. Pathol., 3, 97-105.

Bernstein, J., Evans, A.P. and Gardner, K.D. (1987) Human cystic kidney diseases: epithelial hyperplasia in the pathogenesis of cysts and tumors. Pediatr. Nephrol., 1, 393-396.

Bernstein, J. and Robbins, T.O. (1991) Tuberous sclerosis and allied disorders: Renal involvement in tuberous sclerosis. Ann. N.Y. Acad. Sci., 615, 36-49.

Bjornsson, J., Short, M.P., Kwiatkowski, D.J. et al. (1996) Tuberous Sclerosis-Associated Renal Cell Carcinoma. Am. J. Pathol., 149, 1201-1208.

Bos, J.L., Franke, B., M'Rabet, L. et al. (1997) In search for a function for the Ras-like GTPase Rap. FEBS Letters, 410, 59-62.

Boutell, J.M., Wood, J.D., Harper, P.S. and Jones, A.L. (1998) Huntingtin interacts with cystathionine beta-synthase. Hum. Mol. Genet., 7, 371-378.

Brook-Carter, P.T., Peral, B., Ward, C.J. et al. (1994) Deletion of the TSC2 and PKD1 genes associated with severe infantile polycystic kidney disease-a contiguous gene syndrome. Nature Genet., 8, 328-332.

Brookes, A.J., Slorach, E.M., Morrison, K.E. et al. (1994) Cloning the shared components of complex DNA resources. Hum. Mol. Genet., 3, 2011-2017.

Buckler, A.J., Chang, D.D., Graw, et al. (1991) Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc. Natl. Acad. Sci. U.S.A., 88, 4005-4009.

Burgering, B.M.T., Pronk, G.J., van Weeren, P. et al. (1993) cAMP antagonizes p21-ras-directed activation of extracellular signal-regulated kinase 2 and phosphorylation of mSos nucleotide exvhange factor. EMBO J., 12, 4211-4220.

Carbonara, C., Longa, L., Grosso, et al. (1994) 9q34 loss of heterozygosity in a tuberous sclerosis astrocytoma suggests a growth suppressor-like activity also for the TSC1 gene. Hum. Mol.Genet., 3, 829-832.

Carbonara, C., Loga, L., Grosso, E. et al. (1996) Apparent preferential loss of heterozygosity at TSC2 over TSC1 chromosomal region in tuberous sclerosis hamartomas. Genes, chromosomes and cancer, 15, 18-25.

Clark, R.D., Smith, M., Pandolfo et al. (1988) Tuberous sclerosis in a liveborn infant with trisomy due to t(11q23.3;22q11.2) translocation: Is neural cell adhesion molecule a candidate gene for tuberous sclerosis? Am. J. Hum. Genet., 43, A44.

Cnossen, M.H. (1997) Neurofibromatosis type 1: A clinical and molecular study. Thesis

Erasmus University Rotterdam.
Collins, F.S. (1992) Positional cloning: let's not call it reverse anymore. Nature Genet., 1, 3-6.
Colomer, V., Engelender, S., Sharp, A.H. et al. (1997) Huntingtin-associated protein (HAP1) binds to a Trio-like polypeptide, with a rac1 guanine nucleotide exchange factor domain. Hum. Mol. Genet., 6, 1519-1525.
Connor, J.M., Pirrit, L.A., Yates, J.R.W., Fryer, A.E. and Ferguson-Smith, M.A. (1987) Linkage of the tuberous sclerosis locus to a DNA polymorphism detected by c-abl. J. Med. Genet.,24, 544-546.
Cook, S.J., Rubinfeld, B., Albert, I. and McCormick, F. (1993) RapV12 antagonizes Ras-dependent activation of ERK1 and ERK2 by LPA and EGF in rat-1 fibrobalsts. EMBO J., 12, 3475-3485.
Cook, J.A., Oliver, K., Mueller, R.F. and Sampson, J. (1996) A cross sectional study of renal involvement in tuberous sclerosis. J. Med. Genet., 33, 480-484.
Crino, P.B., Trojanowski, J.Q., Dichter, M.A. and Eberwine, J. (1996) Embryonic neuronal markers in tuberous sclerosis: Single-cell molecular pathology. Proc. Natl. Acad. Sci. U.S.A., 93, 14152-14157.
Crossey, P.A., Richards, F.M., Foster, K. et al. (1994) Identification of intragenic mutations in the von Hippel-Lindau disease tumour suppressor gene and correlation with disease phenotype. Hum. Mol. Genet., 3, 1303-1308.
Curatolo, P., Chiron, C. et al. (1991) Tuberous sclerosis and allied disorders: Neuropsychiatric aspects of tuberous sclerosis. Ann. N.Y. Acad. Sci., 615, 8-16
Duyk, G.M., Kim, S., Myers, R.M. and Cox, D.R. (1990) Exon trapping: A genetic screen to identify randomly transcribed sequences in cloned mammalian genomic DNA. Proc. Natl. Acad. Sci. U.S.A., 87, 8995-8999.
Engelender, S., Sharp, A.H., Colomer, V. et al. (1997) Huntingtin-associated protein (HAP1) interacts with the p150[Glued] subunit of dynactin. Hum. Mol. Genet. 6, 2205-2212.
European Chromosome 16 Tuberous Sclerosis Consortium. (1993) Identification and characterisation of the tuberous sclerosis gene on chromosome 16. Cell, 75, 1305-1315.
Fahsold, R., Rott, H-D. and Lorenz, P. (1991a) A third locus for tuberous sclerosis is closely linked to the phenylalanine hydroxylase locus. Hum. Genet., 88, 85-90.
Fahsold, R., Rott, H-D., Claussen, U. and Schmalenberger, B. (1991b) Tuberous sclerosis in a child with de-novo translocation t(3;12) (p26.3;q23.3). Clin. Genet., 40, 326-328.
Fields, S. and Song, O. (1989) A novel genetic system to detect protein-protein interactions. Nature 340, 245-246.
Fitzpatrick, T.B. (1991) Tuberous sclerosis and allied disorders: History and significance of white macules, earliest visible sign of tuberous sclerosis. Ann. N.Y. Acad. Sci., 615, 26-35.
Fleury, P. (1989) Tubereuze sclerose, veranderingen in de diagnostische en de therapeutische benadering. Tijdschr. Kindergeneeskd., 57, 158-167.
Fryer, A.E., Chalmers, A., Connor, J.M. et al. (1987) Evidence that the gene for tuberous sclerosis is on chromosome 9. Lancet, 1(8534), 659-61.
Ganguly, A., Rock, M.J. and Prockop, D.J. (1993) Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in doubl-stranded PCR products and DNA fragments. Proc. Natl. Acad. Sci. U.S.A., 90, 10325-10329.
Gelbart, W.M., Crosby, M., Matthews, B. et al. (1997) Flybase: a Drosophila database. The flybase consortium. Nucleic Acids Res., 25, 63-66.
Gomez, M. (1988) Tuberous Sclerosis. Raven Press, New York,

Green, A.J., Smith, M. and Yates, J.R.W. (1994a) Loss of heterozygosity on chromosome 16p13.3 in hamartomas from tuberous sclerosis patients. Nature Genet., 6, 193-196.
Green, A.J., Johnson, P.H., Yates, J.R.W. (1994b) The tuberous sclerosis gene on chromosome 9q34 acts as a growth suppressor. Hum. Mol. Genet., 3, 1833-1834.
Gutman, D. H., Saporito-Irwin, S., DeClue, J.E, et al (1997) Alterations in the rap1 signaling pathway are common in human gliomas. Oncogene, 15, 1611-1616.
Haber, D. and Harlow, E. (1997) Tumour-suppressor genes; evolving definitions in the genomic age. Nature Genet., 16, 320-322.
Haines, J.L., Short, M.P., Kwiatkowski, D.J., et al. (1991) Localization of one gene for tuberous sclerosis within 9q32-9q34, and further evidence for heterogeneity. Am. J. Hum. Genet., 49, 764-772.
Harley, H.G., Brook, J.D., Rundle, S.A. et al (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature, 355, 545-546.
Heiskanen, M., Kallioniemi, O., Palotie, A. et al. (1994) Fiber-FISH: experiences and a refined protocol. Genet. Anal., 12, 179-184.
Helin, K. and Peters, G. (1998) Tumor suppressors; from genes to function and possible therapies. TIG, 14, 8-9.
Heng, H.H., Spyropoulos, B., Moens, P.B. et al (1997) FISH technology in chromosome and genome research. Bioassays, 19, 75-84.
Henske, E.P., Wessner, L.L., Golden, J. et al. (1995a) Loss of tuberin in both subependymal giant cell astrocytoma and angiomyolipoma supports a two-hit model for the pathogenesis of tuberous sclerosis. Am. J. Pathol., 151, 1639-1647.
Henske, E.P., Neuman, H.P., Scheithauer, B.W. et al. (1995b) Loss of heterozygosity in the tuberous sclerosis (TSC2) region of chromosome band 16p13 occurs in sporadic as well as TSC-associated renal angiomyolipomas. Genes Chromosomes Cancer, 13, 295-298.
Henske, E.P., Short, M.P., Jozwiak, et al. (1995c) Identification of VAV2 on 9q34 and its exclusion as the TSC1 gene. Ann. Hum. Genet., 59, 25-37.
Henske, E.P., Scheithauer, B.W., Short, M.P., et al. (1996) Allelic loss is frequent in tuberous sclerosis kidney lesions but rare in brain lesions. Am. J. Hum. Genet., 59, 400-406.
Hino, O., Kobayashi, T. et al. (1995) Renal carcinogenesis in the Eker rat. J. Cancer Res. Clin. Oncol., 121, 602-605.
Hornigold, N., van Slegtenhorst, M., Nahmias, J., et al. (1997) A 1.7-Megabase Sequence -Ready Cosmid Contig Covering the TSC1 Candidate Region in 9q34. Genomics , 41, 385-389.
Humphrey, D., Kwiatkowska, J., Henske, E.P. et al (1997) Cloning and evaluation of RalGDS as a candidate for the tuberous sclerosis gene TSC1. Ann Hum Genet, 61, 299-305.
Hunt and Shepherd (1993) A prevalence study of autism in tuberous sclerosis. J. Autism. Dev. Disord., 23, 323-339.
Huntington's Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell, 72, 971-983.
Janssen, L.A.J., Sandkuijl, L.A., Merkens, E.C. et al. (1990) Genetic heterogeneity in tuberous sclerosis. Genomics, 8, 237-242.
Janssen, L.A.J., Sampson, J.R., van der Est, M., et al. (1994) Refined localisation of TSC1 by combined analysis of 9q34 and 16p13 data in 14 tuberous sclerosis families. Hum Genet, 94, 437-440.
Jenne, D.E., Reiman, H., Nezu, J. (1998) Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. Nature Genet., 18, 38-43.

Jin, F., Wienecke, R. and Xiao, G-H (1996) Suppression of tumorigenicity by the wild-type tuberous sclerosis 2 (Tsc2) gene and its C-terminal region. Proc. Natl. Acad. Sci. U.S.A., 93, 9154-9159.

Johnson, W.G. and Gomez, M.R. (1991) Tuberous sclerosis and allied disorders: Clinical, cellular and molecular studies. Published by the New York Academy of Sciences.

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

Johnson, M.R., DeClue, J.E., Felzman, S. et al. (1994) Neurofibromin can inhibit Ras-dependent growth by a mechanism independent of its GTPase-accelerating function. Mol. Cell. Biol., 14, 614-645.

Jones, A.C., Daniells, C.E.D., Snell, R.G. et al. (1997) Molecular genetic and phenotypic analysis reveals differences between TSC1 and TSC2 associated familial and sporadic tuberous sclerosis. Hum. Mol. Genet., 6, 2155-2161.

Jozwiak, S. (1996) Long-term evaluation of inner organ lesions in children with tuberous sclerosis. TSC international symposium abstract, Bath.

Kalchman, M.A., Brook Koide, H., McCutcheon, K et al. (1997) HIP, a human homologue of S. Cerevisiae Sla2p, interacts with membrane associated huntingtin in the brain. Nature Genet., 16, 44-53.

Kandt, R.S., Pericak-Vance, M.A., Hung, W-Y et al. (1989) Absence of linkage of ABO blood group locus to familial tuberous sclerosis. Exp. Neurol. , 104, 223-228.

Kandt, R.S., Pericak-Vance, M.A., Hung, W-Y. et al. (1991) Linkage studies in tuberous sclerosis: chromosome 9?, 11? Or maybe 14! Ann. N.Y. Acad. Sci., 615, 284-297.

Kandt, R.S., Haines, J.L., Smith, M., et al. (1992) Linkage of an important gene locus for tuberous sclerosis to a chromosome 16 marker for polycystic kidney disease. Nature Genet., 2, 37-41.

Kerem, B., Rommens, J.M. et al. (1989) Identification of the cystic fibrosis gene: genetic analysis. Science, 245, 1073-1080.

Kim, K.K. Pajak, L., Wang, H. et al. (1995) Cloning, developmental expression and evidence for alternative splicing of the murine tuberous sclerosis (TSC2) gene product. Cell. Mol. Biol. Res., 41, 515-526.

Knudson A.G.J. (1971) Mutation and cancer; statistical study of retinoblastoma. Proc. Natl. Acad. Sci. U.S.A., 68, 820-823.

Kobayashi, T., Hirayama., Y., Kobayashi, E. et al. (1995a) A germline mutation in the tuberous sclerosis (Tsc2) gene gives rise to the Eker rat model of dominantly inherited cancer. Nature Genet., 9, 70-74.

Kobayashi, T., Nishizawa, M., Hirayama., Y. et al. (1995b) cDNA structure, alternative splicing and exon-intron organization of the predisposing tuberous sclerosis (Tsc2) gene of the Eker rat model. Nucleic Acid Res., 23, 2608-2613.

Kobayashi, T., Mitani, H., Takahashi, R.I., et al. (1997) Transgenic rescue from embryonic lethality and renal carcinogenesis in the Eker rat model by introduction of a wild type TSC2 transgene. Proc. Natl. Acad. Sci. USA., 94, 3990-3993.

Kubo, Y., Klimek, F., Kikuchi, Y. et al. (1995) Early detection of Knudson's two-hits in preneoplastic renal cells of the Eker rat model by laser microdissection procedure. Cancer Res., 55, 989-990.

Kumar, A., Wolpert, C., Kandt, R.S. et al. (1995a) A de novo frame-shift mutation in the tuberin gene. Hum. Mol. Genet., 4, 1471-1472.

Kumar, A., Kandt, R.S., Wolpert, C. et al (1995b) Mutation analysis of the TSC2 gene in

an African-American family. Hum. Mol. Genet., 4, 2295-2298.
Kumar, A., Kandt, R.S., Wolpert, C. et al. (1997) A novel splice site mutation in the TSC2 gene. Hum. Mut., 9, 64-65.
Kupfer, G.M., Naf, D., Suliman, A. et al. (1997) The Fanconi anaemia proteins, FAA and FAC, interact to form a nuclear complex. Nature Genet., 17, 487-490.
Kwiatkowski, D.J., Armour, J., Bale, A.E. et al. (1993) Report and abstracts of the Second International Workshop on Human Chromosome 9 Mapping 1993. Cytogenet. Cell. Genet., 64, 93.
Lau, L.F and Nathans, D. (1986) Expression of a set of growth-related immediate early genes in BALB/c 3T3 cells: coordinate regulation with c-fos or c-myc. Proc. Natl. Acad. Sci. USA., 84, 1182-1186.
Lallier, T.E. (1991) Tuberous sclerosis and allied disorders: Cell lineage and cell migration in the neural crest. Ann. N.Y. Acad. Sci., 615, 158-171.
Levine, A.J. (1997) p53, the cellular gatekeeper for growth and division. Cell, 88, 323-331.
Li, F.P. (1988) Cancer families: Human models of susceptibility to neoplasia. Cancer Res., 48, 5381-5386.
Liaw, D., Marsh, D.J., Li, J. et al. (1997) Germline mutation of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nature Genet., 16, 64-67.
Lie, J.T. (1991) Tuberous sclerosis and allied disorders: Cardiac, pulmonary, and vascular involvement in tuberous sclerosis. Ann. N.Y. Acad. Sci., 615, 58-70.
Lovett, M., Kere, J. and Hinton, L.M. (1991) Direct selection; a method for the isolation of cDNAs encoded by large genomic regions. Proc. Natl. Acad. Sci. U.S.A., 88, 9628-9632.
Lupas, A., Van Dyke, M. and Stock, M. (1991) Predicted coiled coils from protein sequences. Science 252, 1162-1164.
Maheshwar, M., Sandford, R., Nellist, M., et al. (1996) Comparative analysis and genomic structure of the tuberous sclerosis 2 (TSC2) gene human and pufferfish. Hum. Mol. Genet., 5, 131-137.
Maheshwar, M., Cheadle, J.P., Jones, J., et al. (1997) The GAP-related domain of tuberin, the product of the TSC2 gene, is a target for missense mutations in tuberous sclerosis. Hum. Mol. Genet. 6, 1991-1996.
Mallory, S.B. (1995) Cowden syndrome (multiple hamartoma syndrome). Dermatol. Clin., 13, 27-31.
McCormick, F. (1995) Ras-related proteins in signal transduction and growth control. Mol. Reprod. Dev., 42, 500-506.
McKusick, V.A. (1990) Mendelian Inheritance in man. 9th edition (Baltimore and London), The Johns Hopkins University Press.
Menchine, M., Emelin, K., Mischel, P.S. et al. (1996) Tissue and cell-type specific expression of the tuberous sclerosis gene, TSC2, in human tissues. Mod. Pathol., 11, 1071-1080.
Michalet, X., Ekong, R., Fougerousse, F., et al. (1997) Dynamic Molecular Combing: Stretching the Whole Human Genome for High Resolution Studies. Science, 277, 1518-1523.
Monaco., A.P., Neve, R.L., Colletti-Feener, C., et al. (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature, 323, 646-650.
Nagase, T., Seki, N. et al. (1996) Prediction of the coding sequences of unidentified human genes. DNA Res., 3, 321.
Nahmias, J., Hornigold, N., Fitzgibbon, J., et al. (1995) Cosmid contigs spanning 9q34 including the TSC1 candidate region. Eur. J. Hum. Genet., 3, 65-77.
Nellist, M., Brook-Carter, P.T., Connor, J.M. et al. (1993) Identification of markers

flanking the tuberous sclerosis locus on chromosome 9. J. Med. Genet., 30, 224-227.
Nellist, M. (1994) Molecular genetic investigation of tuberous sclerosis. Thesis Institute of Medical Genetics, Cardiff.
Neuman, H.P.H. and Kandt, R.S. (1993) Klinik and Genetic der tuberosen Sklerose. Dtsch. Med. Wschr., 118, 1577-1583.
Noda, M. (1993) Structures and functions of the K rev-1 transformation suppressor gene and its relatives. Biochim. Biophys. Acta, 1155, 97-109.
Northrup, H., Beaudet, A.L., O'Brien, W.E. et al. (1987) Linkage to tuberous sclerosis to ABO blood group. Lancet, 2, 804-805.
Northrup, H., Kwiatkowski, D.J., Roach, E.S., et al. (1992) Evidence for genetic heterogeneity in tuberous sclerosis: one locus on chromosome 9 and at least one locus elsewhere. Am. J. Hum. Genet., 51, 709-720.
Nuoffer, C., and Balch, W.E. (1994) GTPases: Multifunctional molecular switches regulating vesicular traffic. Annu. Rev. Bioch., 63, 949-990.
O'Donnell-Maloney, M.J. and Little, D.P. (1996) Microfabrication and array technologies for DNA sequencing and diagnosis. Genet. Anal., 13, 151-157.
Oliver, S.G. (1996) From DNA sequence to biological function. Nature, 379, 597-600
Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. (1989) Detection of polymorphism of human DNA by gel electrophoresis as single-strand conformation polymorphism. Proc. Natl. Acad. Sci. USA., 86, 2766-2770.
Osborne, J.P., Fryer, A, and Webb, D. (1991) Epidemiology of tuberous sclerosis Ann. N.Y. Acad. Sci., 615, 125-127.
Parimoo, S., Patanjali, S.R., Shukla, H. et al. (1991) cDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. Proc. Natl. Acad. Sci. USA., 87, 1300-1304.
Pavan, W.J., Hieter, P. and Reeves, R.H. (1991) Generation of deletion derivatives by targeted transformation of human-derived yeast artificial chromosomes. Proc. Natl. Acad. Sci. USA., 87, 1300-1304.
Phillips, C.M. and Rye, B. (1994) Neurofibromatosis and tuberous sclerosis: a case of double phakomatosis. J. Am. Acad. Dermatol., 31, 799-800.
Pitiot, G., Waksman, G., Bragado-Nillson, E. et al. (1994) Linkage analysis places the TSC1 gene distal to D9S10. Ann. Hum.Genet., 58, 232-233.
Pizon, V., Desjardins, M., Bucci, C. et al. (1994) Association of Rap1a and Rap1b proteins with late endocytic/phagocytic compartments and Rap2a with the Golgi complex. J. Cell. Science, 107, 1661-1670.
Povey, S., Attwood, J., Janssen, L.A.J. et al. (1991) An attempt to map two genes for tuberous sclerosis using novel two-point methods. Ann. NY. Acad. Sci., 615, 298-305.
Povey, S., Burley, M.W., Attwood, J., et al. (1994) Two loci for tuberous sclerosis: one on 9q34 and one on 16p13. Ann. Hum.Genet., 58, 107-127.
Prieto, M.L., de Juan, J., Anton, M. et al. (1997) Sturge-Weber syndrome with atypical calcifications. Rev. Neurol., 25, 1411-1413.
Qian, F., Germino, F.J., Cai, Y., et al. (1997) PKD1 interacts with PKD2 through a probable coiled-coil domain. Nature Genet. 16, 179-183.
Rastan, S. and Beeley, L.J. (1997) Functional genomics; going forwards from the databases. Curr. Opin. Genet. Dev., 7, 777-783.
Renwick, J.H. (1987) Tuberous sclerosis and ABO. Lancet, 2, 1096-1097.
Richardson, E.P. (1991) Tuberous sclerosis and allied disorders: Pathology of tuberous sclerosis, neuropathologic aspects. Ann. N.Y. Acad. Sci., 615, 128-139.

Riordan, J.R., Rommens, J.M. et al. (1989) Identification of the cystic fibrosis gene: cloning and characterisation of complementary DNA. Science, 245, 1066-1073.
Roach, E.S., Smith, M., Huttenlocher, P. et al. (1992) Report of the diagnostic criteria committee of the National Tuberous Sclerosis Association. J. Child Neurol., 7, 221-224.
Rommens, J.M., Iannuzzi, M.C. et al. (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. Science, 245, 1059-1065.
Ruddle, F.H. (1984) The William Allan memorial award address: reverse genetics and beyond. Am. J. Hum. Genet., 36, 944-953.
Sampson, J.R., Scahill, S.J., Stephenson, J.B.P., et al. (1989) Genetic aspects of tuberous sclerosis in the west of Scotland. J. Med. Genet., 26, 28-31.
Sampson, J.R., Yates, J.R., Pirrit, L.A.et al. (1989) Evidence for genetic heterogeneity in tuberous sclerosis. J. Med. Genet., 26, 511-516.
Sampson, J.R., Janssen, L.A.J., Sandkuijl, L.A. and the Tuberous Sclerosis Collaborative Group. (1992) Linkage investigation of three putative tuberous sclerosis determining loci on chromosome 9q, 11q and 12q. J. Med. Genet., 29, 861-866.
Sampson, J.R. and Harris, P.C. (1994) The molecular genetics of tuberous sclerosis. Hum. Mol. Genet., 3, 1477-1480.
Sampson, J.R., Maheshwar, M.M., Aspinwall, R. et al. (1997) Renal cystic disease in tuberous sclerosis; role of polycystic kidney disease 1 gene. Am. J. Hum. Genet., 61, 843-851.
Sepp, T., Yates, J.R.W. and Green, A.J. (1996) Loss of heterozygosity in tuberous sclerosis hamatomas. J. Med. Genet., 33, 962-964.
Shepherd, C.W., Gomez, M.R. and Crowson, C.S. (1991) Causes of death in patients with tuberous sclerosis. Mayo Clin. Proc., 66, 792-796.
Smith, M., Smalley, S., Cantor, R. et al. (1990) Mapping of a gene determining tuberous sclerosis to human chromosome 11q14-11q23. Genomics, 6, 105-114.
Smith, R.F., Wiese, B.A., Wojzynski, M.K. et al. (1996) BCM Search Launcher-an integrated interface to molecular biology data base search and analysis services available on the World Wide Web. Genome Res., 6, 454-462.
Soucek, T., Pusch, O., Wienecke, R. et al. (1997) Role of the tuberous sclerosis gene-2 product in cell cycle control. J. Biol. Chem., 272, 29301-29308.
Stanbridge, E.J. (1990) Human tumor suppressor genes. Ann. Rev. Genet., 24, 615-657.
Stenmark, H., Vitale, G., Ullrich, O. and Zerial, M. (1995) Rabaptin-5 is a direct effector of the small GTPase Rab5 in endocytic membrane fusion. Cell, 83, 423-432.
Szabo, E., Riffe, M.E., Steinberg, S.M. et al. (1996) Altered cJun expression: an early event in human lung cancinogenesis. Cancer Res., 56, 305-315.
Taylor, A.K., Zambaux, J.L., Kliasak, I. et al. (1991) Carboxyl ester lipase: a highly polymorphic locus on human chromosome 9qter. Genomics, 10, 425-431.
Torres, V.E., Keith, D.S., Offord, K.P. et al. (1994) Renal ammonia in autosomal dominant kidney disease. Kidney Int., 45, 1745-1753.
TSC1 consortium (1997) Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. Science, 277, 805-807
Tsiokas, L., Kim, E., Arnould, T., et al. (1997) Homo- and heterodimeric interactions between the gene products of PKD1 and PKD2. Proc. Natl. Acad. Sci. U.S.A. 94, 6965-6970.
Tsuchiya, H., Orimoto, K., Kobayashi, T. and Hino, O. (1996) Presence of potent transcriptional activation domains in the predisposing tuberous sclerosis (Tsc2) gene product of the Eker rat model. Cancer Res., 36, 429-433.
Urakami, S., Tsuchiya, H., Orimoto, K. et al. (1997) Overexpression of members of the

AP-1 transcriptional factor family from early stage of renal carcinogenesis and inhibition of cell growth by AP-1 gene antisense oligonucleotides in the Tsc2 gene mutant (Eker) rat model. Biochem. Biophys. Res. Commun., 219, 70-75.
Urano, T., Emkey, R. and Feig, L.A. (1996) Ral-GTPases mediate a distinct pathway fom Ras that facilitates cellular transformation. EMBO J, 15, 810-816.
Van Baal, J.G., Smits, N.J., Keeman, D., et al (1994) The evolution of renal angiomyolipomas in patients with tuberous sclerosis. J. Urol., 152, 35-38.
Van der Hoeve, J. (1933) Les phakomatoses de Bourneville, de Recklinghausen et de von Hippel-Lindau. J. Belge Neurol. Psychiat., 33, 752-762.
Verhoef, S., Vrtel, R., van Essen, T., et al. (1995) Somatic mosaicism and clinical variation in tuberous slcerosis complex. Lancet 345, 202.
Verhoef, S., Vrtel, R., Bakker, L. et al. (1998) Recurrent mutation 4882delTT in the GAP-related domain of the tuberous sclerosis TSC2 gene. Hum. Mut., Suppl. 1, s85-s87.
Vrtel, R., Verhoef, S., Bouman, et al. (1996) Identification of a nonsense mutation at the 5'end of the TSC2 gene in a family with presumptive diagnosis of tuberous sclerosis complex. J. Med. Genet., 33, 47-51.
Wang, Q., Verhoef, S., Tempelaars, A.M.P. et al.(1998) Identification of a large insertion and two novel point mutations (3671del8 and S1221X) in tuberous sclerosis comlex (TSC) patients. Hum. Mut., 11, 331-332.
Wankler, E.E., Rovira, C., Scherzinger, E. et al. (1997) HIP-1: A huntingtin interacting protein isolated by the two-hybrid system. Hum. Mol. Genet., 6, 487-495.
Watson, G.H. (1991) Tuberous sclerosis and allied disorders: Cardiac rhabdomyomas in tuberous sclerosis. Ann. N.Y. Acad. Sci., 615, 50-57.
Webb, D., Super, M., Normand, C. and Osborne, J. (1993) Tuberous sclerosis and polycystic kidney disease. B.J.M., 306, 1258-1259.
Westerman, A.M., van Velthuysen, M.L., Bac, D.J. et al. (1997) Malignancy in Peutz-Jeghers syndrome. J. Clin. Gastro. Enterol., 25, 387-390.
Wienecke, R., Konig, A. and DeClue, J.E. (1995) Identification of tuberin, the tuberous sclerosis 2 product. J.Biol. Chem., 207, 16409-16414.
Wienecke, R., Maize Jr, J.C., Shoarinejad, F. et al. (1996) Co-localization of the TSC2 product tuberin with its target Rap1 in the Golgi apparatus. Oncogene, 13, 913-923.
Wilson, P.J., Ramesh, V., Kristiansen, A. et al. (1996) Novel mutations detected in the TSC2 gene from both sporadic and familial TSC patients. Hum. Mol. Genet., 5, 249-256.
Xiao, G-H., Jin, F. and Yeung, R.S. (1995) Germ-line Tsc2 mutation in a dominantly inherited cancer model defines a novel family of rat intracisternal-A particle elements. Oncogene, 11, 81-87.
Xiao, G-H., Shoarinejad, F., Jin, F., et al. (1997) The TSC2 gene product, tuberin, functions as a Rab5 GTPase activating protein (GAP) in modulating endocytosis. J. biol. Chem., 272, 6097-6100.
Xu, Y., Einstein, J.R., Mural, R.J. et al. (1994) An improved system for exon recognition and gene modeling in human DNA sequences. ISBM, 2, 376-384.
Yeung, R.S., Xiao, G-H., Jin, F. et al (1994) Predisposition to renal cell carcinoma in the Eker rat is determined by germ-line mutation of the tuberous sclerosis 2 (Tsc2) gene. Proc. Natl. Acad. Sci. U.S.A., 91, 11413-11416.
Yeung, R.S., Katsetos, C.D. and Klein-Szanto, A. (1997) Subependymal astrocytic hamartomas in the Eker rat model of tuberous sclerosis. Am. J. Pathol., 151, 1477-1486.

This thesis describes the identification and the characterisation of the gene involved in tuberous sclerosis complex (TSC) on chromosome 9 (TSC1). For this purpose we used the positional cloning approach, a strategy by which many other disease genes, including the TSC2 gene on chromosome 16, have been isolated in the past years. The identification of the TSC1 gene has taken 10 years of search and involved the cloning of the entire 1.5 Mb candidate region.

Tuberous sclerosis is characterised by the widespread development of distinctive tumours (hamartomas) in many different tissues, and a broad phenotypic spectrum which may often include disturbed mental function, renal problems and dermatological abnormalities. TSC has an estimated prevalence of 1/6000 and occurs when either one of the TSC1 or TSC2 tumour suppressor genes is inactivated. Mutations in the TSC1 and TSC2 genes cause a very similar clinical phenotype, suggesting that both genes play a closely related role in a still undetermined biological process.

Evidence for linkage between TSC and markers on chromosome 9q34 had already been found in 1987. A consensus TSC1 interval was defined in 1994, spanning approximately 1.5 Mb of genomic DNA. Refining the critical interval in affected individuals proved to be difficult, because of conflicting recombinant data in TSC families.

Therefore, the entire 1.5 Mb candidate region was cloned in a contiguous cosmid contig, which served many applications (chapter 2.6.2). New polymorphic markers and single copy probes were isolated from the contig and they were used to refine the genetic and physical map position of the TSC1 locus. Furthermore, several genes were isolated from the cosmids and a detailed transcript map was constructed. A start was made with sequencing the entire cosmid contig and, ultimately, one of the candidate genes was proven to be the TSC1 gene (chapter 2.6.3).

The coding region of the gene has already been screened for mutations in our collection of 225 unrelated TSC patients, and the mutaional spectrum consists of predominantly small changes, detected in 13% of the patients (chapter 2.6.4). The majority of mutations are clearly inactivating and we find no support for a genotype-phenotype correlation in the group of TSC1 patients. We have detected a small difference in detection rate between sporadic and familial cases, but it remains to be proven whether TSC1 mutations are truely under-represented among sporadic cases. Although only part of the TSC2 has been analysed, TSC2

mutations are found more frequently and therefore we expect that the majority of the mutations will be in the TSC2 gene.

The TSC1 gene is expressed in a wide variety of tissues, which was expected from the multi-organ involvement in patients. The TSC1 gene encodes a 1164 aa protein, hamartin, with a potential coiled-coil domain at the C-terminus. The protein sequence of the TSC1 gene product did resemble the sequence of tuberin or any other vertebrate protein, thereby giving no direct indication for a possible function for hamartin. So far, putative TSC1 sequences have been detected in the genome databases of the mouse, the yeast strain *S. Pombe* and the fly *D. Melanogaster* (chapter 3.2).

In order to start functional analysis of hamartin, polyclonal antibodies have been raised against different parts of hamartin. Immunohistochemical studies indicate that hamartin is predominantly detected in the cytoplasm in mammalian cells. We have investigated putative interactions between the TSC1 and TSC2 gene products using several approaches including the yeast two-hybrid system, immunoprecipitation and transfection assays in mammalian cells. Initial results indicate that hamartin and tuberin can form a complex in the cytoplasm and that the interaction is mediated by coiled-coil domains (chapter 3.6). The labelling pattern is punctated, in which the hamartin-tuberin complex seems to colocalise with vesicle-like structures. The nature of these vesicles is under investigation and future research will hopefully resolve the pathways through which hamartin and tuberin function.

The identification of the TSC1 gene is an important step forward towards a better understanding of the disease pathology of TSC. Abundant evidence that both TSC genes act as tumour suppressors has been accumulated. However, it is not clear whether all TSC-associated lesions require a second somatic hit to develop. It could very well be that the hypomelanotic macule develop by a different mechanism, for example haplo-insufficiency. On the other hand, most data suggests that a second hit, probably during early development, results in defects in migration and differentiation processes, thereby giving rise to tumour growth. Interestingly, malignant transformation in TSC patients is only very rarely observed and occurs predominantly in the kidneys, resembling the tumourigenesis in a natural animal model for the TSC2 gene, the Eker rat.

Tuberin has been shown to act as a putative GTPase activating protein (GAP) for rap1 and rab5, suggesting a role for the TSC2 gene product in the rap1 signalling pathway and/or early endosome fusion. A link between the rap1 signalling pathway and uncontrolled growth in

TSC patients is plausible, although not much is known about the precise functions of rap1. The second putative GAP activity of tuberin, for rab5, suggests a role for the TSC proteins in endocytosis. A model could include a defect in the uptake of molecules that play a role in growth regulation and/or signal transduction. Whether there is a link between vesicle trafficking and controlling specific signalling pathways in TSC will be a challenge to unravel in the future.

The hypothesis arising from the interaction studies is that tuberin can only function correctly when it is associated with hamartin. Therefore, a plausible explanation for the similar phenotype caused by a mutation in either the TSC1 or TSC2 gene is that hamartin regulates the GAP activity of tuberin. Whether there are additional interacting proteins and what role they may play in the etiology of TSC is one of the pressing questions at the moment.

The ultimate aim of TSC research is a complete understanding of the disease process at the cellular and molecular levels, as a pre-requisite for future therapeutic intervention. The isolation of both TSC genes has been a critical step forward in this endeavour.

Dit proefschrift beschrijft het kloneren en het karakteriseren van het gen voor de ziekte tubereuze sclerosis complex (TSC) op chromosoom 9 (TSC1). Omdat er niets bekend is over het biochemische defect bij patiënten met deze aandoening, hebben we de 'positional cloning' techniek gekozen. Al in 1987 was bekend dat het TSC1 gen op chromosoom 9 moest liggen, vlakbij het locus voor de ABO bloedgroep, maar het heeft in totaal 10 onderzoeksjaren gekost voordat het gen gevonden is.

TSC is een autosomaal dominant overervende aandoening, met een geschatte prevalentie van 1:6000. De ziekte wordt gekenmerkt door goedaardige tumoren, hamartomas, die zich in vrijwel alle organen en weefsels kunnen ontwikkelen en vaak leiden tot epilepsie, mentale achterstand en nierproblemen in patiënten. De ziekte wordt veroorzaakt door een mutatie in het TSC1 of het TSC2 gen en beide genen opereren als 'tumor-suppressor genen'. Dit betekent dat beide kopieën van het gen uitgeschakeld moeten zijn (2 hits), voordat een tumor kan gaan groeien. Er is klinisch geen onderscheid te maken tussen patiënten met een mutatie in het TSC1 of TSC2 gen, en daarom vervullen beide genen waarschijnlijk een verwante rol in een nog onbekend biologisch proces in de cel.

In 1987 werd het TSC1 gen gelinkt met de ABO bloedgroeplocus op chromosoom 9. De volgende stap omvatte het definiëren van een kandidaat gebied aan de hand van recombinaties die opgetreden waren binnen TSC families. Er werd overeenstemming bereikt over een gebied tussen de markers D9S149 en D9S114 in 1994. Dit gebied was vrij groot, 1500000 baseparen, en het is door de jaren heen moeilijk gebleken om het gebied te verkleinen, omdat er door verschillende groepen data gepresenteerd werden die elkaar tegenspraken. Een belangrijke stap was dan ook de formatie van een consortium met groepen uit Londen, Cardiff en Boston, met als doel het TSC1 gen te kloneren.

Het in kaart brengen van het kandidaat gebied is gedaan met overlappende cosmide klonen (hoofdstuk 2.6.2.). De cosmiden zijn gebruikt voor verschillende doeleinden. Nieuwe polymorfe markers zijn geïdentificeerd, verschillende technieken zijn toegepast om kandidaat genen te isoleren en er is een start gemaakt met het sequencen van de contig. Om het TSC1 gen te identificeren is bekeken welk gen in TSC patiënten gemuteerd is. Een groot aantal kandidaat genen uit het gebied zijn getest voor mutaties, en uiteindelijk zijn er mutaties aangetoond in één van de geïsoleerde genen (hoofdstuk 2.6.3).

Nadat de genomische structuur (exon-intron grenzen) van het TSC1 gen opgehelderd was, zijn de 21 coderende exonen getest voor mutaties in de 225 onafhankelijke in Rotterdam bestudeerde TSC patiënten. In totaal hebben we 29 mutaties in het TSC1 gen gevonden (13%), waarvan de helft terug te vinden is in de exonen 15 en 17. Bijna alle mutaties leiden tot een vroege stop in het eiwit, behalve 2 in-frame deleties (hoofdstuk 2.6.4.). In tegenstelling tot het TSC2 gen hebben we geen grote deleties gevonden in het TSC1 gen en geen van de missense veranderen kon worden bevestigd als de ziekte veroorzakende mutatie. In onze groep patiënten met een TSC1 mutatie hebben we geen relatie aan kunnen tonen tussen het type mutatie (genotype) en de klinische symptomen (fenotype). Na 30% van het TSC2 gen gescreend te hebben, zijn er mutaties gedetecteerd in 15% van onze TSC patienten. We verwachten daarom dat de meerderheid van de mutaties in het TSC2 gen gevonden gaat worden.

Het TSC1 gen komt tot expressie in vrijwel alle weefsels en codeert voor een eiwit van 1164 animozuren (hamartine). In de voorspelde TSC1 aminozuur volgorde is een coiled-coil domein geïdentificeerd. Dit is een structuur die interacties aan kan gaan met andere eiwitten. Hamartine vertoond geen overeenkomsten met bekende vertebrate eiwitten, maar er zijn wel homologe sequenties gevonden in de muis, in de vlieg *(Drosophila)* en mogelijk in gist (*S. Pombe*).

Om een start te kunnen maken met functionele studies zijn polyklonale antisera geproduceerd. We hebben kunnen aantonen dat hamartine voornamelijk tot expressie komt in het cytoplasma van cellen en vervolgens is gestart met interactieproeven om te testen of hamartine kan binden aan het genprodukt van het TSC2 gen, tuberine. Met behulp van het 2-hybrid systeem in gist hebben we aangetoond dat coiled-coil domeinen in beide eiwitten een interactie aangaan. De interactie tussen tuberine en hamartine is bevestigd *in vivo* door coïmmunoprecipitatie proeven. Tenslotte hebben we beide eiwitten tot overexpressie gebracht in fibroblasten en COS cellen, en daaruit blijkt dat er een colokalisatie optreedt in het cytoplasma met nog ongedefiniëerbare blaasachtige structuren (hoofdstuk 3.6).

De identificatie van het TSC1 gen is een belangrijke stap voorwaarts om een beter begrip te krijgen hoe de ziekte zich in patiënten ontwikkelt. Het is nu bekend dat beide genen als 'tumor suppressor genen' werken (2-hit mechanisme), maar het is nog niet duidelijk of alle verschillende lesies zich via dit mechanisme ontwikkelen. Tevens worden er zelden kwaadaardige tumoren in TSC patiënten gevonden. Een uitzondering zijn 'renal cell

carcinomas', die identiek zijn aan tumoren die in een natuurlijk diermodel voor TSC 'de Eker rat' gevonden worden.

Het TSC2 gen is geïdentificeerd in 1993 en codeert voor tuberine, een potentiëel GTPase activerend eiwit (GAP) voor rap1 en rab5. Dit kan een aanwijzing zijn voor een rol voor tuberine in de celcyclus en in endocytose. Een rol voor tuberine in de celcyclus is aannemelijk, omdat veel tumoren onstaan doordat dit proces ontregeld is. Een model voor tuberine in endocytose zou gelinkt kunnen zijn aan een defect in de opname van molekulen die betrokken zijn bij de regulatie van groei of het geven van signalen in de cel.

De hypothese die voortkomt uit de interactie studies is dat tuberine alleen correct kan functioneren als het in een complex is met hamartine. Een aannemelijke verklaring voor het identieke fenotype in TSC patienten, onafhankelijk van welk gen gemuteerd is, is dat hamartine de GAP activiteit van tuberine reguleert. Of er naast hamartine en tuberine nog meer eiwitten betrokken zijn bij het ontstaan van tubereuze sclerosis is een belangrijke vraag die hopelijk in de nabije toekomst opgelost gaat worden.

Het einddoel van TSC onderzoek is om het ziekteproces op het nivo van cellen en molekulen te begrijpen, wat hopelijk zal leiden tot betere therapie mogelijkheden voor de patiënten. Met het kloneren van beide TSC genen is een belangrijke stap voorwaarts gezet.

Curriculum vitae van Marjon van Slegtenhorst

Date of birth
1 December 1966

Education

1985-1986	First year Exam Social Sciences, Utrecht University.
1986-1992	Doctoral Exam in Biomedical Sciences (Medical Faculty of Leiden University). Electives: Genetics.

Work experience

1991-1992	MGC Department of Human Genetics, Leiden University, Prof. Dr. G.J.B. van Ommen. Homologous recombination of YACs in yeast to reconstruct the 2.4 Mb Duchenne Muscular Dystrophy gene.
1992-1993	Institute for Molecular Genetics, Baylor College of Medicine, Houston, Texas, Prof. Dr. A. Ballabio. Identification of genes at Xp22.3.
1994-1998	PhD at the MGC Department of Clinical Genetics, Erasmus University Rotterdam, Dr. D.J.J Halley. Identification of the tuberous sclerosis gene on 9q34.

Courses and Workshops:

MGC courses;

*Strategy for purification, sequencing and structural analysis of proteins
*Oxford examination in English as a foreign language EFL higher level
*Working with laboratory animals; art. 9

International courses

*HGMP computing course, Cardiff
*EMBO workshop 'Trafficking of proteins', Crete

1998-	Postdoc at the Mayo Clinic Jacksonville, Florida, Dr. M. Hutton. Fronto-Temporal Dementia.

List of publications

M.A. van Slegtenhorst, M.T. Bassi, G. Borsani, M.C. Wapenaar, G.B. Ferrero, L. de Conillis, E.L. Rugarli, A. Grillo, B. Franco, H.Y. Zoghni, A. Ballabio.
A gene from the Xp22.3 region shares homology with voltage-gated chloride channels.
Hum Mol Genet 1994;3:547-552.

M. van Slegtenhorst, B. Janssen, M. Nellist, S. Ramlakhan, C. Hermans, A. Hesseling, A. van den Ouweland, D. Kwiatkowski, B. Eussen, J. Sampson, P. de Jong, D. Halley.
Cosmid Contigs from the Tuberous Sclerosis Candidate Region on Chromosome 9q34.
Eur J Hum Genet 1995;3:78-86.

E. Henske, B. Scheithauer, P. Short, R. Wollmann, J. Nahmias, N. Hornigold, M. van Slegtenhorst, C. Welsh, D. Kwiatkowski.
Allelic Loss is Frequent in Tuberous Sclerosis Kidney Lesions but Rare in Brain Lesions.
Am J Hum Genet. 1996;59:400-406.

N. Hornigold, M. van Slegtenhorst, J. Nahmias, R. Ekong, S. Rousseaux, C. Hermans, D. Halley, S. Povey, J. Wolfe.
A 1.7-Megabase Sequence -Ready Cosmid Contig Covering the *TSC1* Candidate Region in 9q34.
Genomics 1997;41:385-389.

D. Humphrey, J. Kwiatkowska, E.P. Henske, J. L. Haines, D. Halley, M. van Slegtenhorst and D. J. Kwiatkowski.
Cloning and evaluation of RalGDS as a candidate for the tuberous sclerosis gene TSC1.
Ann Hum Genet 1997;61:299-305.

X. Michalet, R. Ekong, F. Fougerousse, S. Rousseaux, C. Schurra, N. Hornigold, M. van Slegtenhorst, J. Wolfe, S. Povey, J. Beckmann, A. Bensimon.
Dynamic Molecular Combing: Stretching the Whole Human Genome for High Resolution Studies.
Science 1997;277:1518-1523.

Marjon van Slegtenhorst, Ronald de Hoogt, Caroline Hermans, Mark Nellist, Bart Janssen, Senno Verhoef, Dick Lindhout, Ans van den Ouweland, Dicky Halley* Janet Young, Mariwyn Burley, Steve Jeremiah, Karen Woodward, Joseph Nahmias, Margaret Fox, Rosemary Ekong, Jonathan Wolfe, Sue Povey*John Osborne* Russell G. Snell, Jeremy P. Cheadle, Alistair C. Jones, Maria Tachataki, David Ravine, Julian R. Sampson* Mary Pat Reeve, Paul Richardson, Friederike Wilmer, Cheryl Munro, Trevor L. Hawkins* Tiina Sepp, Johari B. M. Ali, Susannah Ward, Andrew J. Green, John R. W. Yates* M. Priscilla Short, Jonathan H. Haines, Sergiusz Jozwiak, Jolanta Kwiatkowska, Elizabeth P. Henske, David J. Kwiatkowski.
Identification of the Tuberous Sclerosis gene *(TSC1)* on chromosome 9q34.
Science 1997;277:805-808.

Marjon van Slegtenhorst, Mark Nellist, Bas Nagelkerken, Jeremy Cheadle, Russell Snell, Ans van den Ouweland, Arnold Reuser, Julian Sampson, Dicky Halley and Peter van der Sluijs.
Interaction between hamartin and tuberin, the *TSC1* and *TSC2* gene products.
Hum Mol Genet 1998;7;1053-1057.

Marjon van Slegtenhorst, Senno Verhoef, Anita Tempelaars, Lida Bakker, Qi Wang, Marja Wessels, Remco Bakker, Mark Nellist, Dick Lindhout, Dicky Halley and Ans van den Ouweland.
Mutational spectrum of the *TSC1* gene in a cohort of 225 tuberous sclerosis complex patients; no evidence for a genotype-phenotype correlation.
Submitted for publication.

Dankwoord

Kloneren doe je niet alleen (never clone alone). Niet in de laatste plaats omdat er bergen werk verzet moeten worden, maar ook omdat het soms moeilijk is om uit de vele dalen weer omhoog te klimmen. Dit proefschrift is dan ook het resultaat van vele experimenten, samenwerkingen en discussies, waarvoor ik een aantal mensen wil bedanken.

In de eerste plaats mijn co-promotor en dagelijkse begeleidster Dicky. Beste Dicky, ik waardeer het bijzonder dat jij mij zoveel vrijheid hebt gegeven om onderzoek te doen en het vertrouwen dat ik op de goede weg zat. Ik heb niet alleen in het lab veel kennis opgestoken, maar daarnaast ook van de rol die ik mocht vervullen als beginnende OIO tussen alle labhoofden in het consortium.

Het is natuurlijk een voorrecht om te mogen werken op een vrij nieuwe, goed draaiende afdeling als klinische genetica. Hiervoor ben ik Professor Galjaard dank verschuldigd. Het TSC onderzoek is waar het tot uitgegroeid is mede dankzij mijn promotor Professor Lindhout. Beste Dick, ik heb je de afgelopen maanden meerdere malen belaagd met enorme stukken tekst, die ik meestal binnen een paar dagen gecorrigeerd al weer terug had. Veel dank! De professoren-doctoren Patrick Willems, Jan Hoeijmakers en Ellen Zwarthof wil ik bedanken voor het kritisch lezen van het manuscript.

Toen ik vier en een half jaar geleden begon, was het TSC2 gen net boven tafel en er was al veel voorbereidend werk verricht voor TSC1, waardoor ik relatief snel de draad op kon pikken. Dit was natuurlijk mede dankzij het TSC team dat toen in het lab aan ploeteren was (geweest): Arjenne, Bart, Mark en Sarvan .

De eerste twee jaren heb ik lief en leed gedeeld met Caroline. Het was niet altijd makkelijk om, toen ik zelf nog bezig was op te starten en de weg in het lab te verkennen, meteen al iemand te moet begeleiden. Daar was dan ook geen sprake van, maar gelukkig was je gewend zelfstandig te werken. Ik ben blij dat we het cosmide banken screenen konden afwisselen en dat er iemand orde op zaken hield.

Jawel, Ronald, tevens paranimf, kon zich een jaar storten op het screenen van cDNA banken. Ik ken niemand die zo droog uit de hoek kan komen. Je bent ook iemand die de diepte in wil en heel precies en hard te werk gaat en waar iedereen naar toe kwam om te leren hoe je nu het beste banken uit kon platen en poso's opwerken. Je wordt nog steeds gemist! Goed dat we zoveel werk gestoken hebben in B2 (ook al zaten we in het begin met 4.5 kb niet coderende sequentie) en dat je op de valreep het feesten nog hebt meegemaakt.

Dan niet te vergeten Mark (in het Nederlands). Ik denk dat ik van jouw de afgelopen jaren het meest heb opgestoken. Vooral je scherpe opmerkingen en inzicht in de materie zullen me bij blijven. Vaak heb ik toch mijn zin doorgedreven, maar kwam er in de meeste gevallen al snel achter dat jij het bij het juiste eind had gehad.

De overige vaste leden van het TSC team wil ik bedanken voor de toch wel grote bijdrage aan het mutatiewerk: Anita, Ans, Lida en Senno (in alfabethische volgorde). Ik hoop dat ik niet te veel DNAtjes en blotjes onvindbaar heb gemaakt. Ik ruim nog op voordat ik vertrek! Remco was mijn meest succesvolle (en enige) student. Jij hebt in korte tijd massa's blotjes geproduceerd. Jammer dat we hier de mutatie niet in gevonden hebben. En niet te vergeten de 'tijdelijke' en 'ondersteunende' TSC leden; Arnold, Bert E., Radek, Qi, Javuz en Marja.

Nu ben ik er natuurlijk nog niet, want het lab is nog veel groter. Ben, 'als nieuw hoofd'. Mijn eerste generatie kamergenoten E* en Peter en tweede generatie kamergenoten Carola en Henk (duidelijk minder volwassen, maar niet minder leuk). E*, ik heb nog steeds dagelijks plezier en minder buikkrampen van jouw koffiezetapparaat. Peter, je stond altijd klaar om de vaak juiste 'prikkelende' adviezen te geven, zowel tijdens als na het werk. Carola en Henk zal ik vooral missen (is hier nog iets te eten?) om de levendige competities Pacmannen.

Verder wil ik mijn favoriete Bobo-sitters Robert en Patrizia bedanken, zodat ik zonder schuldgevoel het land uit kon. Marijke, die mij heeft leren vechten met de ABI en Jeltje voor het leren rolschaatsen. Laangenootjes Elke-Marjon-Jorine-Arjenne en Caroline, voor het tolereren van mijn 'vrolijke' bende en alle andere leden van diagnostiek en DNA, die ik maar niet bij naam ga noemen, anders vergeet ik er zeker een paar.

Alleen dan nog Filippo, my favourite room mate op congressen. I will never forget all the important talks we missed, because you wanted to play tennis.

En natuurlijk niet te vergeten alle mensen die niet in het lab staan, maar wel erg belangrijk zijn voor het functioneren van de 24ste: van de fotografie Tom (zie blz 37-38!) en Ruud. De secretariaten, vooral Jeannette. Het team in de spoelkeuken (kleine wasjes, grote wasjes) en Rein-Melle-Peter v V. voor materiële support.

Uit Utrecht wil ik Bas en Peter bedanken, die me snel op weg hebben geholpen om de interactie proeven lopend te krijgen. Ik kon altijd terecht, vooral in de weekeinden. En niet te vergeten uit Leiden, Birgitta, mijn andere paranimf (ja leuk, wat is dat?). Jij wist me ervan te overtuigen dat ik echt een maand in Afrika moest uitrusten voordat ik begon te schrijven.

the collaborators from the consortium, especially David, Sue, Jonathan, Nick, Joseph, Janet, Russell, Jeremy and Julian. It has been a very open and successful teamwork. I really enjoyed spending some time in London, when the airconditioning was being fixed in our department. Those were important weeks, because we managed to finish the contig! Janet, thanks for the 'exciting' cycling tour over the London Bridge during traffic jam. I hope we keep in touch. This work would not have been possible without the help of all the TSC patients, their Associations and many clinicians. I would like to mention the STSN (Els den Hollander), the NTSA (Vicky Whittemore) and the British TSA.

Abbreviations

aa	amino acid
ADPKD	autosomal dominant polcystic kidney disease
AML	angiomyolipoma
BAC	bacterial artificial chromosome
cM	centiMorgan
DBH	dopamine β hydroxylase gene
DCC	deleted in colorectal cancer
DNA	deoxyribonucleic acid
DMC	dynamic molecular combing
DMD	duchenne muscular dystrophy
EST	expressed sequence tag
FISH	fluorescence *in situ* hybridisation
GAP	GTPase activating protein
HD	hetroduplex, not to be confused with Huntington's Disease
HGP	human genome project
IRD	inter-resource duplex
YAC	yeast artificial chromosome
kb	kilobase
LOH	loss of heterozygosity
Mb	megabase
N-cells	neuron-like cells
NF	neurofibromatosis
nt	nucleotide
PFGE	pulsed field gel electrophoresis
RACE	rapid amplification of cDNA ends
RCC	renal cell carcinoma
SEGA	subependymal giant cell astrocytoma
SSCP	single strand conformation polymorphism
STS	sequence-tagged sites
TSC	tuberous sclerosis complex