Mutations in the serine/threonine protein kinase gene, STK11, in sporadic colorectal cancer

BY

Sonja Teresa Engelbrecht

Dissertation in fulfilment of the requirements for the degree of
Magister Scientiae (M.Sc.) in Human Genetics

In the
Department of Human Genetics and Developmental Biology
Faculty of Health Sciences, University of Pretoria

Supervisor: Professor E.J. van Rensburg

November 2002

ABSTRACT

Colorectal cancer (CRC) is one of the most common forms of cancer in Western nations, it is however uncommon in sub-Saharan Africa. In South Africa there is an approximate ten-fold lower incidence of CRC in black patients compared to Caucasian patients. This could be due to differences in lifestyles and environment that exist between the various population groups. Underlying molecular events could also account for the difference in susceptibility to colorectal cancer.

Mutations in the Peutz-Jeghers syndrome (PJS) gene, *STK11,* predispose to amongst others colorectal cancer. To examine the role of this gene in South African patients with CRC, DNA from 208 tumours (104 black patients, 104 Caucasian patients) was screened for *STK11* mutations via PCR-SSCP analysis. In total 8 novel missense mutations, one of which was germline, were identified in seven tumours (~3.4% 7/208) from 5 black and two Caucasian patients. One tumour from a Caucasian patient was found to be a compound heterozygote. Peutz-Jeghers syndrome was thus diagnosed in 0.96% (1/104) of black patients via a germline mutation.

Thus 4.8% (5/104) of tumours from black patients and 1.9% (2/104) of tumours from Caucasian patients harbour *STK11* missense mutations. In addition, 3 synonymous and 5 intronic mutations were detected in a further 73 tumours from black patients, whereas only 3 synonymous and 5 intronic mutations were detected in 25 tumours from Caucasian patients. The present study is the sixth to suggest that somatic mutation of the *STK11* gene in sporadic colorectal cancer of Caucasians is an infrequent event. However, this is only the second study of a non-Western population to show somatic mutations in sporadic cases of CRC. Furthermore with regard to the anatomic site of tumours with somatic missense mutations, the present study found that for black patients 7.69% (2/26) of the left-sided tumours, 2% (1/50) of rectal tumours and 4.54% (1/22) of right-sided tumours harboured mutations. Thus the frequency of missense mutations of left-sided CRC tumours compared to

right-sided tumours was not significantly elevated ($\chi 2$ –test , 1df, p = 0.881) in the black population.

This study represents the first investigation into the role of the *STK11* gene in putative sporadic cases of CRC from both black and Caucasian South African patients. The observed mutations clearly show that mutations of the *STK11* gene are infrequent in the CRCs of the South African Caucasian population, and more frequent in the South African black population. This may be a reflection of the differences in lifestyle and incidence of CRC in the different populations.

UITTREKSEL

Kolorektale kanker is een van die mees algemeenste vorms van kanker in Westerse nasies. Dit is egter nie so algemeen in sub-Sahara Afrika nie. In Suid Afrika is daar 'n ongeveer tien maal laer insidensie van kolorektale kanker in swart pasiënte in vergelyking met wit pasiënte. Hierdie verskil kan wees as gevolg van verskille in lewenstyl en omgewingstoestande wat heers tussen die twee groepe. Onderliggende molekulêre gebeurtenisse kan ook 'n rol speel by die verskille in vatbaarheid vir kolorektale kanker.

Mutasies in die Peutz-Jeghers syndroom (PJS) geen, *STK11*, lei tot 'n vatbaarheid vir onder andere kolorektale kanker. Om die rol van hierdie geen in Suid Afrikaanse pasiënte te ondersoek was DNA van 208 tumore (104 swart pasiënte, 104 wit pasiënte) deur middel van die polimerase ketting reaksie en enkel draad konformasie polimorfisme analise ondersoek. In totaal was daar 8 missense mutasies, waarvan een kiemlyn, geïdentifiseer in sewe tumors (~3.4% 7/208) van 5 swart en twee wit pasiënte. Een tumor was 'n saamgestelde heterosigoot. Dus was Peutz-Jeghers sindroom gediagnoseer in 0.96% (1/104) van die swart pasiënte deur middel van 'n kiemlyn mutasie. Dus het 4.8% (5/104) van tumore van swart pasiënte en 1.9% (2/104) van tumore van wit pasiënte *STK11* missense mutasies. Addisioneel was daar 3 sinonieme en 5 intronies mutasies geidentifiseer in 'n verdere 73 tumore van swart pasiënte, waar daar slegs 3 sinonieme en 5 introniese mutasies geidentifiseer was in 25 tumore van wit pasiënte. Die huidige studie is die sesde om voor te stel dat somatiese mutasies van die *STK11* geen in sporadies kolorektale kankers in wit pasiënte 'n rare verskynsel is. Dit is egter slegs die tweede studie van 'n non-westerse populasie om somatiese mutasies te kry in sporadiese kolorektale gevalle. Verder in verband met die anatomiese posisie van tumore met somatiese missense mutasies het hierdie studie gevind dat in die swart pasiënte 7.69% (2/26) van die linker tumore, 2% (1/50) van rektale tumore en 4.54% (1/22) van regter tumore mutasies het. Dus was daar geen merkwaardige veskil in die frekwensie van missense mutasies wat gevind was in linkerkantse an regterkantse kolorektale kanker tumore in die swart populasie nie ($\chi 2$ –test , 1df, $p = 0.881$).

Hierdie studie is die eerste ondersoek om te bepaal of die *STK11* geen 'n rol speel in beide wit en swart Suid Afrikaners met vermoede sporadiese kolorekatale kanker. Die gevonde mutasies dui daarop dat mutasies in die *STK11* geen 'n rare veskynsel is in die wit Suid Afrikaanse populasie en meer gereeld voorkom in die swart Suid Afrikaanse populasie. Dit kan 'n refleksie wees van die verskille wat heers tussen die lewensstyl en voorkoms van kolorektale kanker in die verskillende populasies.

Table of contents

ACKNOWLEDGEMENTS

I would like to express my appreciation and thanks to the following persons and institutions:

The department of Human Genetics and Developmental Biology for the opportunity to study at this institution.

Prof. E. J van Rensburg my supervisor, for guidance, assistance and encouragement throughout the project.

Prof. A Christianson for good advice as well as his support for the duration of the project.

The University of Pretoria for granting an UP post-graduate bursary.

M. de la Rey, C. Schlebusch and S. Strümpher for their assistance.

To Darren for making life interesting and enjoyable.

My parents and family for their love and support.

Karen Briedenhann for her artistic contribution.

LIST OF ABBREVIATIONS

a	adenine (in primer sequence)
α	alpha
APS	ammonium persulphate
bp	base pair
BPB	bromophenol blue
BSA	Bovine serum albumin
c	cytosine (in primer sequence)
°C	degrees centigrade
CC	Carney complex
CD	Cowden disease
cm	centimetre
CRC	colorectal cancer
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytidine-5'-triphosphate
ddATP	2'3'-dideoxiadenosine-5'-triphosphate
ddCTP	2'3'-dideoxicytidine-5'-triphosphate
ddGTP	2'3'-dideoxiguanosine-5'-triphosphate
ddNTP	2'3'-dideoxinucleotide
ddTTP	2'3'-dideoxithymidine-5'-triphosphate
del	deletion
df	degrees of freedom
dGTP	2'-deoxyguanosine-5'-triphosphate
DNA	deoxyribonucleic acid
dTTP	2'-deoxythymadine-5'-triphosphate
EDTA	ethylenediamine tetra-acetic acid
EGF	epidermal growth factor
EGFP	Enhanced green fluorescent protein
EGFr	epidermal growth factor receptor
ESE	Exonic splicing enhancers
EtBr	ethidium bromide
g	grams
g	guanine (in primer sequence)
gDNA	genomic DNA
HCl	hydrochloric acid
IGF-1	Insulin-like growth factor-1
ins	insertion
JPS	Juvenile polyposis syndrome
kb	kilo (10^3) base pair
KCl	potassium chloride
kDa	kilodalton
LIP1	LKB1 interacting protein
LOH	Loss of heterozygosity
M	molar
μ	micro
μg	microgram
μl	microlitre
MDE	mutation detection enhancement gel

Mes	4-morpolinethanesulfonic acid
$MgCl_2$	magnesium chloride
ml	millilitre
mLkb1	Mouse Lkb1
mM	millimolar
min	minute
MSK1	Mitogen and stress stimulated protein kinase
NaCl	sodium chloride
$NaClO_4$	sodiumperchloride
Na_2EDTA	sodium ethylenediamine tetra-acetic acid
NLS	Nuclear localisation signal
^{32}P	phosphate isotope
$p90^{RSK}$	p90 ribosomal S6 kinase
PCR	polymerase chain reaction
pH	indicates acidity
PJS	Peutz-Jeghers syndrome
RFLP	Restriction fragment length polymorphism
rpm	revolutions per minute
S6K1	p70 ribosomal S6 kinase
SDS	sodium dodecyl sulphate
sec	seconds
SSCP	single stranded conformational polymorphism
STK11	serine threonine kinase
t	thymidine (in primer sequence)
Taq polymerase	DNA deoxinucleotidyltransferase from *Thermus aquaticus*
TBE	89.15 mM Tris (pH8.0), 88.95 mM boric acid, 2.498 mM Na_2EDTA
TEMED	N,N,N',N'-tetramethylendiamine
U	units
UV	ultraviolet
V	volt
W	watt
XC	xylene cyanol
XEEK	*Xenopus* egg and embryo kinase

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 GENETIC ASPECTS OF CANCER

Cancer arises from the abnormal and uncontrolled division of cells. These cells then invade and destroy the surrounding tissues. (Oxford concise medical dictionary, 1998) At the cellular level such abnormal behaviour of cells can be contributed to genetic alterations. These genetic mutations can be passed through the germline, but are predominantly somatic in origin. Such mutations can be acquired in the somatic cells during normal cell division or via environmental carcinogens such as cigarette smoking, radiation and toxic chemicals. The importance of inherited mutations should, however not be overlooked as the study of mutations in the genes responsible for heritable cancers provide valuable insights into the molecular origins of both inherited as well as sporadic forms of cancer (Hodgson and Maher, 1993).

There appears to be two types of genes involved in the different forms of cancer. Firstly where mutations in the gene lead to a gain in function (oncogenes) and secondly genes in which both alleles have loss of function (tumour suppressor genes).

1.1.1 Oncogenes

The word oncogene is derived from the Greek *onkas* for mass or tumour. In their native state oncogenes do not possess the ability to transform cells and are referred to as proto-oncogenes. Oncogenes are thus mutated proto-oncogenes that induce or maintain cell transformation (Hodgson and Maher, 1993; Latchman, 1997).

The biochemical mechanisms by which proto-oncogenes act can be classified into four groups. Proto-oncogenes act as growth factors, receptors to growth factors, can phosphorylate proteins with serine, threonine and tyrosine as

substrates, they also transmit signals by GTPases, and have also been found to encode transcription factors thus controlling gene expression (Bishop, 1991).

1.1.2 Tumour suppressor genes

Tumour suppressor genes restrain cell growth under normal conditions. However, when both the alleles of this gene are inactivated cells can grow and divide in an abnormal manner. The classic "two-hit" hypothesis was proposed by Alfred Knudson in 1971 to explain the above- mentioned occurrence. The so-called first hit is when a mutational event is present in the germline or somatic cell of one allele while the second hit is a subsequent somatic alteration in the normal allele of the same cell. The loss of both alleles then gives rise to cancer (van Rensburg, 1997).

To date several tumour suppressor genes have been identified and it has become a rather difficult task to clearly classify the role of these genes in the carcinogenetic pathway. Kinzler and Vogelstein (1998) proposed the following three epithets to categorise the different types of tumour suppressor genes; 'gatekeeper', 'caretaker' and 'landscaper' tumour suppressor genes.

'Gatekeeper' tumour suppressor genes

Gatekeepers are the subset of tumour suppressor genes that prevent cancer through direct control of cellular growth. Inactivation of these genes lead to new and abnormal growth, subsequent restoration of the gene function shows marked suppression of neoplastic growth. Hence the name 'gatekeeper', as the gene functions to prevent runaway growth (Kinzler and Vogelstein., 1998; Macleod, 2000).

'Caretaker' tumour suppressor genes

It is a well-known fact that DNA replication is error prone and that a repair system exists to correct such mistakes. The genes that perform this function are referred to as 'caretaker' genes. If these genes were mutated and consequently produced faulty products, incapable of repairing the mismatched nucleotides, mutations

would accumulate at various sites throughout the genome that could potentially lead to cancerous growth. Unlike the 'gatekeeper' genes, restoration of function after the initiation of neoplastic growth does not suppress the growth as these genes have an indirect affect on the cancer causing mutations (Kinzler *et al.*, 1998; Macleod, 2000; van Rensburg, 1997).

'Landscaper' tumour suppressor genes

The last class of tumour suppressor genes also act in an indirect manner. The 'landscaper' gene is postulated to control the microenvironment in which the tumour growth takes place. It is proposed that this is accomplished via the regulation of extra cellular matrix proteins, adhesion proteins, secreted growth factors and cell surface proteins. If the function of the 'landscaper' gene is lost the microenvironment will change to promote abnormal growth of adjacent epithelial cells thus increasing the risk of neoplastic transformation (Kinzler *et al.*, 1998; Macleod, 2000).

1.2 COLORECTAL CANCER

Colorectal cancer (CRC) is one of the most common forms of cancer in Western nations, it is however uncommon in sub-Saharan Africa. American white males have the highest incidence of CRC estimated at 46.5 affected individuals per 100 000 population. Black Americans do not have a much lower rate with 38.6/100 000. White South African males have an incidence of 24.7/100 000, whereas black South African males have only 2.1 affected individuals per 100 000. On average females tend to have a lower incidence of CRC. White South African females have a incidence of 19.3/100 000 and the Black South African female population has a very low incidence of 1.6/100 000 (Sitas *et al.*, 1998).

Of the types of cancer occurring within each population CRC ranks third in white South African males and only tenth in black South African males. In females a very similar situation is noted with CRC being the second most common form of cancer within white females and seventh most common in black South African females (Sitas *et al.*, 1998).

The approximately ten-fold lower incidence of CRC in black South Africans could be due to differences in lifestyles and environment that exist between the various population groups. Underlying molecular events could also account for the difference in susceptibility to colorectal cancer between the white and black South Africans.

1.2.1 Hereditary colon cancer

More than 90 percent of all colon cancers that are reported each year are sporadic. Hereditary nonpolyposis colorectal cancer accounts for 5 to 10 percent of colorectal cancer and polyposis syndromes only one percent. Hodgson and Maher, 1993. Patients with hereditary colorectal cancer should be aware of the family history to ascertain the risk of contracting colorectal cancer. Table 1.1 summarises the empirical risk estimates for individuals with a family history of colorectal cancer.

Table 1.1 Risk of colorectal cancer

Family history	Lifetime risk
One first degree relative	1 in 17
One first degree relative under 45 years	1 in 10
One first and one second degree relative	1 in 12
Both parents	1 in 8.5
Two first degree relatives	1 in 6
Three first degree relatives	1 in 2

Adapted from Hodgson and Maher, 1993

Clinically a number of distinct syndromes, which predispose to colorectal cancer, are recognised. The majority of which are inherited in an autosomal dominant manner thus putting offspring at a 50% risk of inheriting the predisposing mutation. Three groups of syndromes are recognised. Syndromes associated with large amounts of adenomatous polyps, each having malignant potential. The *APC* gene is known to be associated with these syndromes. Mismatch repair genes such as *hMSH2* are associated with the second type of syndrome with a smaller number of adenomatous polyps, but which seem to have greater malignant potential. The third type of syndrome predisposes patients to

hamartomatous polyps, such syndromes appear to be associated with a less severe increase in risk of developing CRC (Bishop and Hall, 1994). Peutz-Jeghers syndrome falls into this last category of syndromes displaying a predisposition to colorectal cancer. (Rustgi, 1994).

CHAPTER 2

PEUTZ-JEGHERS SYNDROME

2.1 INTRODUCTION

Peutz-Jeghers syndrome (PJS), a rare autosomal dominant disorder was probably described for the first time in the late 19th century by a London surgeon, Jonathan Hutchinson. He reported the phenotype of identical twin sisters with spots on the lips and the buccal mucosa. He however did not recognise the connection between polyps and these spots. J.L.A. Peutz made this link in 1921 after which H. Jeghers made a further contribution by giving a more complete description of the syndrome in 1944. Hence the syndrome was dubbed Peutz-Jeghers syndrome (Hemminki, 1999; McGarrity *et al.*, 2000).

2.2 CLINICAL MANIFESTATIONS OF THE PEUTZ-JEGHERS SYNDROME

Peutz-Jeghers syndrome has several clinical manifestations, which is characterised by skin and mucosal lentigines, hamartomatous polyps of the gastrointestinal tract and an increased risk of cancer.

2.2.1 Hyperpigmentation

Melanocytic macules occur primarily on the lips, bucal mucosa, eyes, nostrils and perianal areas of patients affected with this syndrome. A photographic representation of these spots is given in Figure 2.1A. These pigmented blemishes can also be found on the fingers and dorsal and ventral aspects of the hands and feet. The skin pigmentation can vary in colour ranging from bluish black to blue to dark brown (Figure 2.1B). They usually tend to fade with increasing age. Pigmentation in the buccal mucosa occasionally persists to adulthood. To date this hyper-pigmentation has not been found to undergo malignant transformation (Hemminki, 1999; McGarrity *et al.*, 2000).

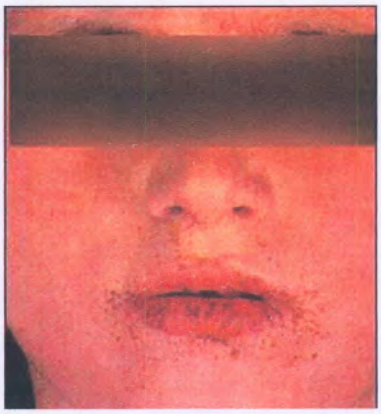


Figure 2.1 Melanin pigmentation of the lips and surrounding areas of a patient with Peutz-Jeghers syndrome.

2.2.2 Gastrointestinal polyps

Patients with PJS present with hamartomatous polyps of the gastrointestinal tract. Such polyps occur mainly in the small intestine with the jejunum being the most common site followed by the ileum and duodenum (Figure 2.2). Polyps also occur in the colon and stomach but at a lower frequency. These polyps can become very large in size often causing intussusception (McGarrity *et al.* 2000). Intussusception is the vagination of one part of the bowel into another resulting in abdominal pain and rectal bleeding (Oxford concise colour medical dictionary, 1998).

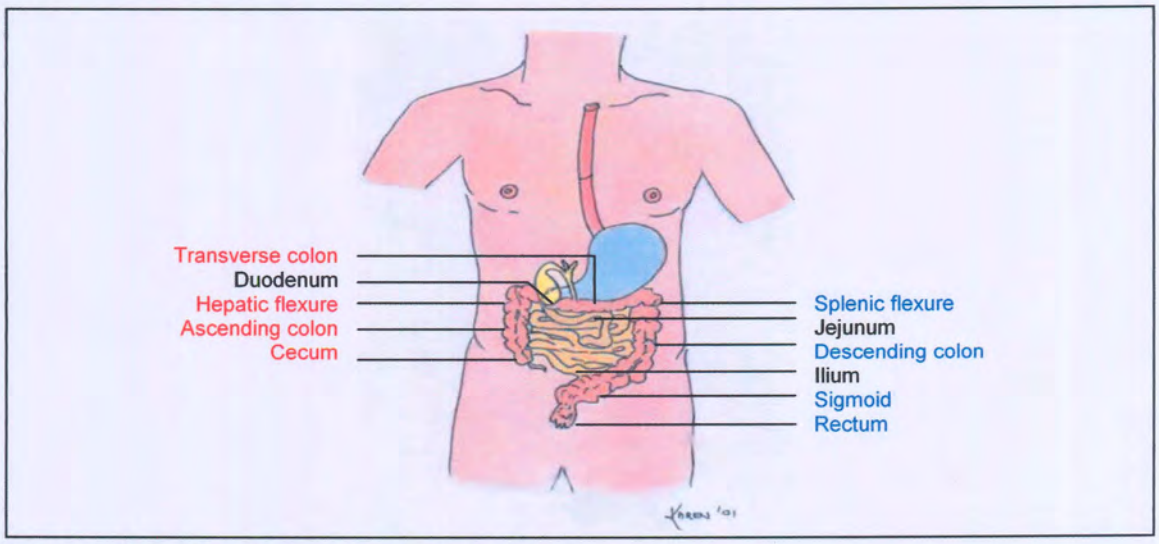


Figure 2.2 The colon and related organs

The colon can be divided into the rectum, sigmoid, descending colon, splenic flexure, transverse colon, hepatic flexure, ascending colon and cecum. The rectum up to the splenic flexure forms the left-sided colon (indicated in blue) and the transverse colon to the cecum the right-sided colon (indicated in red).
Picture courtesy of Karen Briedenhann

2.2.3 Increased risk of cancer in Peutz-Jeghers syndrome patients

An 18-fold relative risk of intestinal as well as extra-intestinal cancer was reported by Gardiello *et al.* (1987). If cancer is contracted there is an estimated 48% chance that the patient will succumb to this disease by the age of 57 (Spigelman *et al.*, 1989). Cancer of the colon, rectum, stomach, small intestine, thyroid, pancreas, breasts, ovaries, sex cord and testicles have been reported in PJS patients.

The clinical features listed above are seemingly clear-cut, this is however not the case as many of these symptoms overlap with those of other syndromes with different genetic loci. The distinction between PJS and Carney complex (CC, chromosome 2p16) for instance can be rather challenging as both have the same pigmentation anomalies, both present with thyroid, ovarian and testicular cancers. In PJS as well as CC Sertoli cell tumours (rare type of testicular tumour) lead to feminisation. Single cases of intestinal polyposis and pancreatic tumours have also been reported in CC patients. Furthermore both syndromes are inherited in an autosomal dominant fashion. There are however certain tumours that occur exclusively in CC patients these are; heart myxomas, skin myxomas, neurofibromas and tumours of the anterior pituitary gland, (leading to an overproduction of growth hormone and thus an enlargement of the hands, feet and the face). These tumours may however not be present in all the patients with CC (Stratakis *et al.*, 1998).

Similar phenotypic overlap can be seen between PJS and Cowden disease (CD, cromosome 10q23). Mucocutaneous pigmentation occur in both syndromes along with thyroid, ovarian, breast and testicular cancers (Luuko *et al.*, 1999). Brain tumours occur in several patients with CD, but has not been reported in PJS. Both these syndromes are associated with hamartomatous polyps, but in CD these hamartomas lack the smooth muscle infiltration of polyp stroma as is characteristic of PJS (Hemminki, 1999). Despite the overlap in phenotypic presentation it is difficult but possible to make a distinction between PJS, CD and CC.

2.3 THE *STK11/LKB1* GENE

Linkage of PJS to two loci, 19p13.3 and 19q13.4, was reported (Mehenni *et al.*, 1997). This pointed to the possibility of heterogeneity. In 1998 the gene on chromosome 19p13.3 was identified independently by two groups. (Hemminki *et al.*, 1998; Jenne *et al*, 1998). Germline mutations in this gene variously known as *STK11* or *LKB1* were identified. Subsequent to cloning various other authors found mutations within the gene and are summarised in Appendix A.

However not all PJS families have mutations within this gene. One group noticed a marked increased risk of primitive biliary adenocarcinoma in patients with PJS, but without *STK11* alterations, providing further evidence for the presence of a second locus (Olschwang *et al.*, 2001). Since this paper was published no further advances have been made regarding the second locus therefore leaving the main focus on *STK11*.

The *STK11* gene extends over 23 kb of genomic DNA and consists of nine exons. The 3'-UTR extends over 1449 bp and the 5'-UTR is approximately 338 bp long. This gene is spliced in an unusual U12 snRNA dependant manner as a result of the following unusual characteristics. Intron 2 of this gene begins with **AT**ATCCTT and ends with CCC**AC** thus deviating from the usual **GT/AG** splice junctions (Jenne *et al.*, 1998). A graphical representation of the *STK11* gene is given in Figure 2.3.

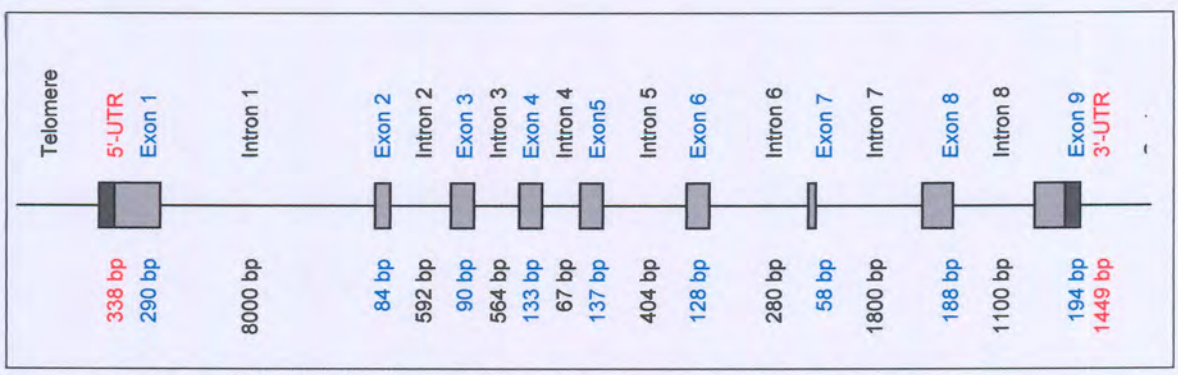


Figure 2.3:Schematic illustration of *STK11* genomic structure
Black boxes indicate the 5'-UTR and the 3'-UTR. The grey boxes represent exons.

2.4 THE STK11 PROTEIN

STK11 encodes a 60-kDa serine threonine kinase consisting of 434 amino acid residues. This protein is expressed in all tissues with varied levels, with expression being particularly high in the testis and fetal liver. This widespread expression is consistent with the inflated risk of several cancer types associated with PJS. Expression studies of STK11 in the small intestine have been carried out. Intestinal epithelium consist of stem cells localised near the bottom of the crypt, where cells divide, differentiate and migrate up the villli. Near the tip of the villi cells can either extrude into the lumen or undergo apoptosis. Apoptic cells occur mainly in the crypt. STK11 is detected in both the apoptotic regions, at particularly high levels in dying cells, thus indicating involvement of STK11 in apoptosis of intestinal epithelial cells. (Reviewed by Yoo *et al.*, 2002)

The kinase domain (amino acids 50-319) of the STK11 protein shows homology to a large number of human and other proteins (Hemminki,1999; Tiainen *et al.*, 1999). This is not surprising as the eukaryotic protein kinases make up one of the largest super-families in gene and protein homology (Hanks and Hunter, 1995). There are however two proteins that have significant homology to human STK11 both within the kinase domain as well as outside this region. These are *Xenopus* egg and embryo kinase (XEEK1) and Mouse *Lkb1*. Mouse *Lkb1* (*mLkb1*) shows a higher level of homology particularly in the kinase domain with 96,2% identity (Smith *et al.*, 1999).

The STK11 protein acts as a tumour suppressor gene. Loss of heterozygosity (LOH) was found in 70% of patients with PJS with a *STK11* germline mutation. (Miyaki *et al*, 2000; Gruber *et al*, 1998) Further evidence for the tumour suppressor function was provided by a study conducted on *mLkb1* (Tiainen *et al.*, 1999). In this study tumour cell lines were identified showing severely reduced mRNA levels and impaired Lkb1 kinase activity. When reintroduced into these cells *Lkb1* had a growth suppression activity shown to be mediated through a G1 cell cycle block. Note should be made however that normal proliferating cells also express *Lkb1* and thus expression of this gene per se is not sufficient to suppress cellular growth (Tiainen *et al.*, 1999).

An exception to the observation that mutational events result in inactivation of the protein was found where a mutational event did not result in loss of kinase activity, but could still be considered disease causing as the sub-cellular location of the protein was altered. Wild type STK11 is localised both in the cytoplasm as well as in the nucleus. This particular mutation (9 bp del, codon 303-306 as given in Appendix A) resulted in accumulation of the product in the nucleus with no protein in the cytoplasm thus indicating a loss of cytoplasmic retention ability, which could be vital for the physiological function of STK11. This loss of cytoplasmic retention ability of STK11 is speculated to be due to the inability of the protein to interact with and phosphorylate its substrate protein (Nezu *et al.*, 1999). The nuclear accumulation of this mutant led to the search for possible nuclear localisation domains. By fusing Enhanced Green Fluorescent Protein (EGFP) to four deleted regions of the protein (amino acids 88-433, 1-88, 1-45 and, 43-88) as well as wild type protein the cellular location of the proteins could be shown. Accumulation in the nucleus of EGFP fused to amino acids 43-88 when expressed in COS7 cells indicated that the nuclear localisation signal (NLS) is located in this region (Nezu *et al.*, 1999). In another study, Smith *et al.* (1999) followed a similar approach with *mLkb1*, but first identified a region analogous to a known NLS. The region identified spans from amino acids 37-42 comprising of amino acids PRRKRA and is conserved between the *mLkb1*, human *LKB1* and frog homolog XEEK. This domain is similar to PKKKRKV, which is a known NLS in SV40 large T antigen. When the sequence QPRRRKRAK containing the presumed NLS was fused with the N-terminus EGFP and expressed in COS cells it resulted in nuclear localisation (Smith *et al.*, 1999). There seems to be two potential NLS located at amino acids 38-43 and amino acids 81-84. The latter study however postulates that the concentration of EGFP in the nucleus most likely results from passive entry of NLS-EGFP into the nucleus and does not reflect the normal function of the putative mLkb1 NLS sequence.

STK11 seems to be regulated by phophorylation at Ser^{431}, which is critical for the protein to perform its function. Ser^{431} is located six amino acids away from the C-terminal in the sequence Lys-Xaa-Arg-Arg-Xaa-Ser (Xaa refers to any amino acid). This sequence is highly conserved between all known mammalian LKB1

sequences as well as in XEEK1. This particular Serine was shown to be phosphorylated *in vitro* by cAMP-dependant protein kinase (PKA), p90 ribosomal S6 kinase ($p90^{RSK}$), Mitogen and stress stimulated protein kinase (MSK1), and p70 ribosomal S6 kinase (S6K1), which are all members of the AGC kinase group (Collins *et al.*, 2000; Sapkota *et al.*, 2001). This was demonstrated by producing STK11 mutants via the substitution of Ser^{431} with either Ala or Asp and subsequently testing the ability to suppress cell growth of G361 melanoma cells. These cells are known not to contain STK11 and reintroduction of wild type protein arrests cellular growth. The mutated versions (S431D and S431A) failed to have the same effect. This finding could imply that phosphorylation of Ser^{431} is necessary for STK11 to perform its function as a growth suppressor. *In vivo* $p90^{RSK}$ and PKA rather than MSK1, and S6K1 mediate the phosphorylation of Ser^{431}. This phosphorylation by $p90^{RSK}$ and PKA occurs in response to agonists identified as Forskolin (an adenylated cyclase activator), 12-O-tetradecanoylphorbl-13 acetate (TPA), epidermal growth factor (EGF), and insulin-like growth factor-1 (IGF-1). These signal transduction pathways are demonstrated in Figure 2.4.

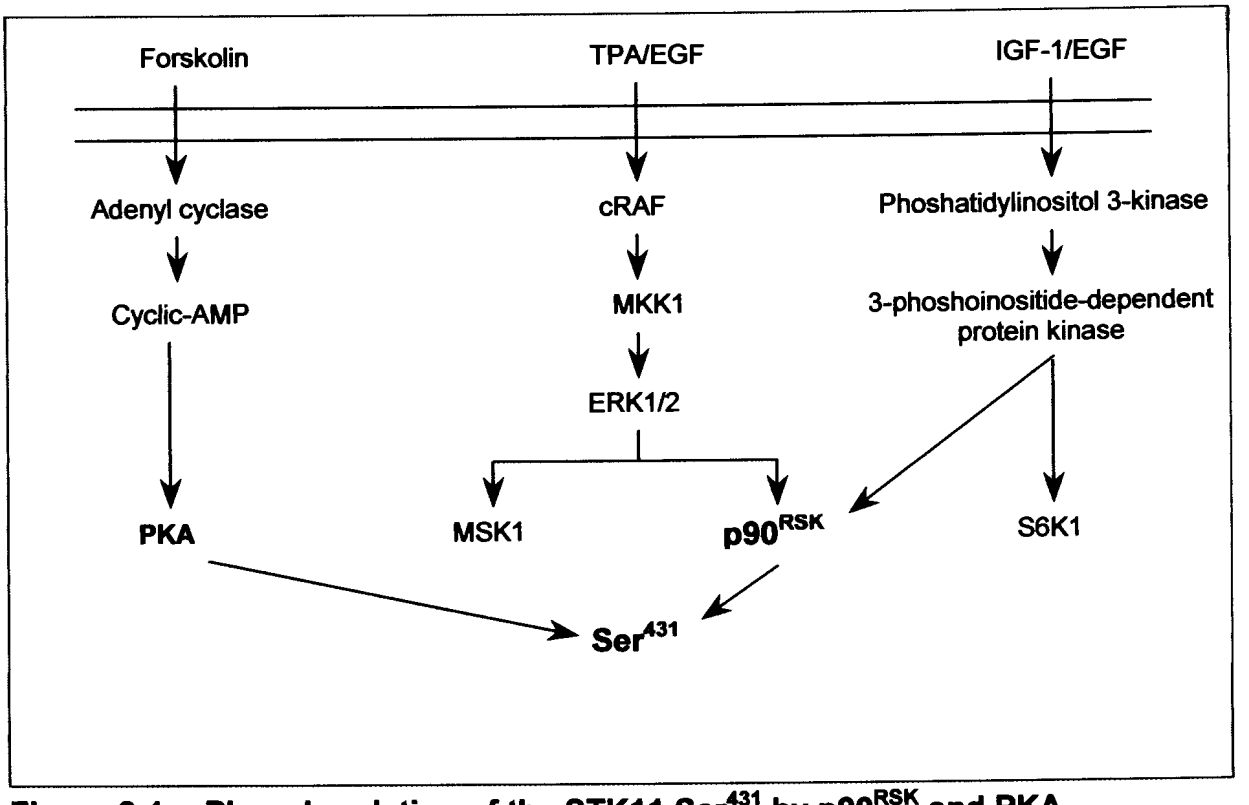


Figure 2.4 Phosphorylation of the STK11 Ser^{431} by $p90^{RSK}$ and PKA

In the human homologue Ser^{431} is followed by Ala-Cys-Lys-Gln-Gln. The Cys located two residues closer to the C-terminal than Ser is located in a consensus sequence known as the *CAAX* motif where C is the amino acid cysteine (cys 433), A is any aliphatic amino acid except alanine and X represents the C-terminal amino acid. The *CAAX* motif mediates the prenylation of many proteins (Sapkota *et al.*, 2001). Where prenylation is the addition of the 15 carbon farnesyl group or the 20 carbon geranylgeranyl group to acceptor proteins. Both farnesyl as well as geranylgeranyl are isopropenoid groups derived from the cholesterol biosynthetic pathway (Campbell, 1995). The cysteine in STK11 is no exception to this rule and is prenylated by the addition of a farnesyl group. This prenyalation however was not found to be essential for STK11's ability to suppress cell growth, but it was suggested that it could play a role in cellular location, it's interaction with a regulatory substrate or perhaps the stability of the protein (Sapkota *et al.*, 2001).

2.4.1 Suppression of cell proliferation

Thus far four proteins have been identified which could act as substrates for STK11. The tumour suppressor protein p53 was the only protein out of 50 routinely tested proteins that was phophorylated in vitro by wild-type *STK11*, but not a catalytically inactive form of this protein. (Sapkota *et al.*, 2001) In a further study the Peutz-Jeghers gene product LKB1 appears to mediate p53-dependent apoptosis. (Karuman *et al.*, 2001) Functional p53 is also required in STK11 mediated induction of p21. Cytoplasmic STK11 induces G1 arrest by initiating a signal cascade, which encroaches upon proper cell cycle function. This has been shown to be due to initiation of the promoter of p21, which was found to be specifically upregulated in STK11 transfected cells (Tiainen *et al.*, 2002).

Another more promising substrate is the very recently identified LKB1 interacting protein (LIP1). The *LIP1* gene located on chromosome 2q36 is expressed in a wide range of tissues including the colon. This gene consists of 25 exons and yields a protein of 1099 amino acids with a molecular weight of 121.4 kDa. LIP1 is a cytoplasmic protein, whereas STK11 is found mainly in the nucleus. However when these two proteins are co-expressed the presence of *STK11* in the

cytoplasm increases dramatically, indicating that LIP1 may be involved in controlling the sub-cellular location of STK11. Further evidence of a functional link between the proteins is that they both induce anteriorly truncated secondary body axis when ectopically expressed in *Xenopus* embryos. They could therefore also regulate the same signalling pathway. As the TGFβ superfamily are responsible for the induction of secondary body axis both STK11 and LIP1 may be involved in the regulation of this particular pathway. This notion is further supported by the interaction of LIP1 and the TGFβ-regulated transcription factor SMAD4. Through this interaction LIP1 forms a bridge between STK11 and SMAD4. Interestingly SMAD4 mutations account for approximately one third of cases of juvenile polyposis syndrome (JPS), which like PJS is characterised by gastrointestinal harmatomatous polyps and an increased risk of gastrointestinal cancer. This observation may suggest a mechanistic link between PJS and JPS. (Smith *et al.*, 2001)

STK11 also interacts and regulates brahma-related gene 1 (*Brg1*) an essential component of chromatin remodelling complexes. The association occurs between the amino terminus of STK11 and the helicase domain of Brg1 and is required for Brg1-induced growth arrest in the Brg1/retinoblastoma signalling pathway. (Marignani *et al.*, 2001)

2.5 STK 11 AND SPORADIC CANCER

To date six studies of relatively small sample size (23-75) have been conducted to determine the involvement of the *STK11* gene in sporadic colorectal cancer. Four of these studies concluded that somatic mutations are rare in the *STK11* gene (Avizienyte *et al.*, 1998; Forster *et al* 2000; Resta *et al.*, 1998; Launonen *et al.*, 2000; Wang *et al.*, 1998) The remaining study, by Dong *et al.* (1998), screened 23 CRCs and 26 adenomas of Korean patients and concluded that somatic mutations are frequent in the *STK11* gene. In this study seven mutations were detected in colon tumours and two mutations in the adenomas which all occurred in the left side of the colon. The authors thus came to the conclusion that the anatomic site of the tumour, i.e. left-sided CRC, appears to coincide with the occurrence of mutations.

2.6 AIM OF THE STUDY

This study forms part of an investigation into the molecular aetiology of colorectal cancer in black and white South African patients. Since the *STK11* gene predisposes PJS patients to an increased risk of, amongst others, colorectal cancer this study focuses on the involvement of this gene in sporadic colorectal cancer patients. The objectives of this study are thus to determine the extent to which this gene is involved in the CRCs (of both black and white patients) as well as the nature of the mutations. In addition an assessment will be made of whether the mutational events occur in left- or right-sided colon cancers.

CHAPTER 3
MATERIALS AND METHODS

Unless otherwise stated all chemicals used in this study were obtained from E Merck Darmstadt, BDH Laboratory supplies or United States Biochemicals.

3.1 PATIENTS

This is a retrospective study of 208 tumour samples from patients with CRC attending; Kalafong and Pretoria Academic Hospitals (174) during the period 1985-1995, and Chris Hani-Baragwanath Hospital (34) from 1990-1995. Both races were equally represented with 104 patients each. The study contained more females (122) than males (83). The gender of the remaining 3 patients was unknown. The Department of Histopathology (SAIMR) and the Department of Anatomical Pathology (University of Pretoria) collected these samples for routine diagnostic purposes. Prior to analysis the anonymity of the samples were ensured.

3.1.1 Black patients

The 104 patients, 50 male and 53 females ranged in age from 16 years to 83 and have a mean age of 52.99 (± 16.43).

3.1.2 Caucasian patients

The 104 caucasian patients comprised, 35 males and 69 females with an age range of 19 to 88 (mean age 65.22 ± 13.47).

3.2 DNA ISOLATION FROM PARAFFIN EMBEDDED TISSUE

Tumour and normal tissues were dissected from formalin-fixed paraffin embedded sections with surgical blades. To extract the DNA, the flakes of tumour and normal tissue were placed in separate eppendorf tubes to which 200µl of

extraction buffer (20mM Tris, pH8,4; 50mM KCl, 0,45% NP40, 0,45% Tween 20 was added. Proteinase K was added to a final concentration of 200 µg/ml. The samples were then incubated overnight at 54°C in a heating block. In order to inactivate the proteinase K the samples were heated to 94°C for 10 min quenched on ice and spun down. The supernatant was transferred to clean eppendorf tubes and stored at 4°C. Throughout the procedures care was taken to avoid sample to sample contamination.

3.3 SINGLE STRANDED CONFORMATIONAL POLYMORPHISM

Using the primers listed in Table 3.1 the complete coding sequences and splice junctions of the *STK11* gene was amplified (13 fragments) for single stranded conformational polymorphism (SSCP) analysis. Originally exon 1 was amplified in three overlapping fragments and later only two fragments. By combining the forward primer of the first fragment with the reverse primer of the second fragment the combined first fragment was obtained. For the combined second fragment the forward primer of the second fragment was used with the reverse primer of the third fragment Table 3.1. SSCP analysis was conducted by performing radiolabeled PCR, followed by electrophoresis on a mutation detection enhancement gel (MDE).

3.3.1 PCR procedure

5'-End labelling of primers

The primer pairs described in Table 3.1 were 5'-end labelled with ^{32}P according to the method of Ausubel *et al* (1995). The reaction mixture with a total volume of 10µl contained 20pmole primer, 10x Kinase buffer, 12 units T_4-polynucleotide kinase (Amersham) and 42µCi [γ-^{32}P] ATP (7000Ci/mmol, ICN). The mixture was incubated at 37°C for 1 hour. Thereafter the T_4 enzyme was inactivated at 65°C for 10 min, quenched on ice, spun down and stored at -20°C until required.

PCR amplification of *STK11* exons

Annealing temperatures (T_{ann}) were calculated for each primer (obtained from Integrated DNA Technologies), using the formula:

T_{ann} = 0.41(%GC) + 34.9°C (Eeles and Stamps, 1993)

Amplification conditions were then further optimised for each primer pair (Table 3.1)

Table 3.1: STK11 primers and PCR conditions

Primer[a]	Primer sequence	Product size	T_{ann}	[$MgCl_2$]
STK11-1AF[b]	5'-ctcagggctggcggcgggact-3'	163 bp	62°C	1,0mM
STK11-1AR[b]	5'-cttgcggcgcggctggtagatga-3'			
STK11-1BF[b]	5'-acgttcatccaccgcatcgac-3'	150 bp	60°C	1.5mM
STK11-1BR[b]	5'-gcacagcgtctccgagtccag-3'			
STK11-1CF[b]	5'-tcttacggcaaggtgaaggaggtg-3'	140 bp	58°C	1.5mM
STK11-1CR[b]	5'-ccgaccccagcaagccatactta-3'			
STK11-1AF[b]	5'-ctcagggctggcggcgggact-3'	259 bp	62°C	1.5mM
STK11-1BR[b]	5'-gcacagcgtctccgagtccag-3'			
STK11-1BF[b]	5'-acgttcatccaccgcatcgac-3'	245 bp	58°C	1.5mM
STK11-1CR[b]	5'-ccgaccccagcaagccatactta-3'			
STK11-2F[c]	5'-gaggtacgccacttccacag-3'	288 bp	58°C	1.5mM
STK11-2R[c]	5'-cttcaaggagacgggaagag-3'			
STK11-3F[d]	5'-tgagctgtgtgtccttagcg-3'	196 bp	58°C	1.0mM
STK11-3R[d]	5'-agtgtggcctcacggaaa-3'			
STK11-5F[b]	5'-cctgagggctgcacggcacc-3'	213 bp	66°C	1.0mM
STK11-5R[b]	5'-cccctcggagtgtgcgtgtggt-3'			
STK11-7F[b]	5'-tcacccagggcctgacaacagag-3'	193 bp	66°C	1.5mM
STK11-7R[b]	5'-gcagcctcggccccactg-3'			
STK11-8F[e]	5'-gacaggcgccactgcttctg-3'	271 bp	60°C	1.0mM
STK11-8R[e]	5'-ggacatcctggccgagtcag-3'			
STK11-9F[f]	5'-gtaagtgcgtccccgtggtg-3'	357 bp	59°C	1.0mM
STK11-9R[f]	5'-gtggcatccaggcgttgtcc-3'			

a) F and R indicate forward and reverse primers respectively. b) Dong *et al.*, 1998. c) Avizienyte *et al.*, 1998. d) Avizienyte *et al.*, 1999. e) Wang *et al.*, 1998. f) Jenne *et al.*, 1998.

Exons four and six were optimised using PCR enhance system (In vitrogen Gibco BRL) which contained Enhance buffer with and without enhance solution. (Table 3.2)

Table 3.2: STK11 primers and PCR conditions of exons 4 and 6

Primer[a]	Primer sequence	Product size	T_{ann}	[Enhance solution]	[$MgSO_4$]
STK11-4F[b]	5'-cggccccaggacgggtgt-3'	218 bp	60°C	0x	2.0mM
STK11-4R[b]	5'-ctcagggagtgcccgggagg-3'				
STK11-6F[b]	5'-gaccacgcctttcttccctccc-3'	213 bp	60°C	0,5x	1.0mM
STK11-6R[b]	5'-cacaaaagccccgcctccct-3'				

a) F and R indicate forward and reverse primers respectively. b) Dong *et al.*, 1998.

All exons except for exons 3, 5 and 9 were directly amplified in a 10µl reaction containing radiolabelled primers. Due to the low yield of PCR product exons 3, 5 and 9 were subjected to two rounds of PCR. The second round was performed with the addition of 10% DMSO and radiolabelled primers.

Briefly, the PCR reaction mixtures contained10 X PCR buffer [20mM Tris.HCl (pH 8.4), 50mM KCl], appropriate concentration $MgCl_2$, 25µM of each dNTP, 2pmole of each primer, 0.2pmole ^{32}P-end labelled forward and reverse primers, 1µg bovine serum albumin (BSA), 0.5U *Taq* DNA polymerase, dH_2O and 2µl DNA in a final reaction volume of 10µl.

The samples were subjected to thermal cycling (MJ-Research PTC100) which consisted of an initial denaturation for 4 min at 94°C, followed by 30 cycles of denaturation at 94°C for 1 min, annealing for 1 min, and extension at 72°C for 1 min. A final extension step was performed at 72°C for 7 min.

3.3.2 Mutation detection enhancement gel electrophoresis

To perform the SSCP analysis a 1:10 dilution of PCR product was made with formamide loading buffer (95% formamide, 12,5mM EDTA, 0.25% bromophenol blue). The diluted PCR product was then denatured at 95°C for five minutes, cooled on ice prior to loading on an MDE gel.

The PCR products were separated on 0,6 x mutation detection enhancement gels (MDE) (40cm x 30cm). This method relies on migrational differences of the

variable single stranded DNA conformations created via SSCP within the matrix of the gel. The gels were electrophoresed in 0.6 x TBE (pH 8.3 10 x TBE; 89mM Tris, 89mM boric acid, 20mM EDTA) at various Watts in a cold room (4°C). Exons two and nine were electrophoresed at room temperature on MDE gels where the TBE buffer was substituted with 1 x TME (pH 6.8, 30mM Tris, 35mM Mes, 1mM EDTA) buffer for enhanced separation. Furthermore a fan was placed in front of the gel in order to keep the gel at room temperature. The different electrophoretic conditions of the exons of the STK11 gene are represented in Table 3.3.

Table 3.3 MDE gel electrophoresis conditions for SSCP analysis of the *STK11* gene

Exon	Power	Run time
STK11-1A	8 W	14 hrs
STK11-1B	8 W	9 hrs
STK11-2	10 W	BPB 5 cm from bottom of plate [a]
STK11-3	8 W	9 hrs
STK11-4	12 W	9 hrs
STK11-5	8 W	14 hrs
STK11-6	12 W	9 hrs
STK11-7	8 W	9 hrs
STK11-8	8 W	15 hrs
STK11-9	10 W	BPB 5cm from bottom of plate [a]

a) Gels were electrophoresed until the bromophenol blue was situated approximately 5 cm from the bottom of the plate.

After electrophoresis was completed the gels were transferred to 3 MM Whatman®[1] filter paper, covered with cling wrap and vacuum dried for 2 hours at 80°C. The gels were exposed to X-ray film [Fuji RX-U] at -70°C using an intensifying screen and the images developed after adequate exposure was obtained.

[1] Whatman® is a registered trademark of Whatman Scientific Ltd., Kent, England

3.4 SEQUENCING

All DNA samples which displayed aberrant electrophoretic banding patterns via SSCP analysis were sequenced to determine the mutations. Prior to sequencing PCR product was treated with 10 units Exonuclease I and 2 units Shrimp alkaline Phosphatase at 37°C for 15 min. Thereafter the samples were inactivated at 80°C for 15 min.

Sequencing was performed according to the chain termination principle described by Sanger *et al.* (1977), utilising the Sequenase™[2] PCR product sequencing kit from Amersham using either forward or reverse primer and ^{35}S dATP. The sequencing reaction was carried out according to the manufacturer's specifications with some modification. The PCR products were obtained using the reaction described in section 3.3

Samples were heated at 75°C for 3 min prior to loading 3µl on a 6% polyacrylamide gel and electrophoresed at 60 Watts (W) in 1X TBE buffer until sufficient separation of the fragments was achieved. The gels were transferred to 3MM Whatman filter paper after completion of electrophoresis, covered with cling wrap and vacuum dried for 2 hours at 80°C. The cling wrap was removed from the dry gels prior to X-ray film [Fuji RX-U] exposure. The images on the film were developed when adequate exposure was obtained.

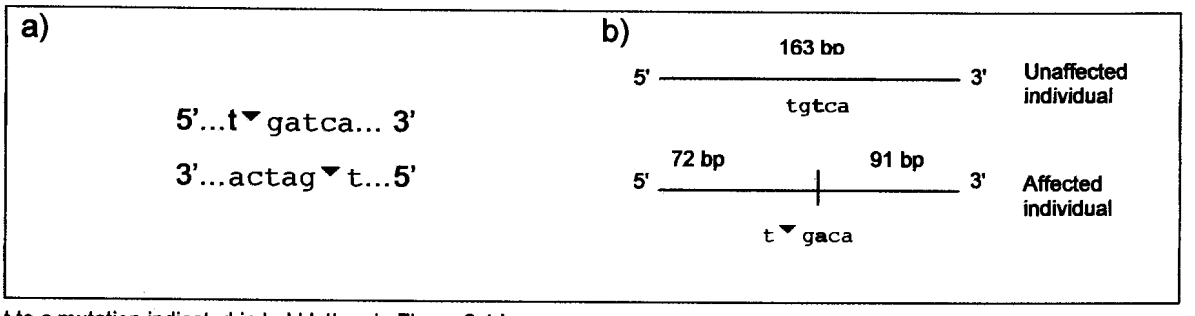
3.5 PCR RESTRICTION ENZYME DIGESTION

Restriction enzyme digestion was employed to determine whether two of the missense mutation F12I is present in a control group.

[2] Sequenase™ is a trademark of United States Biochemical Corporation, Cleveland, Ohio, U.S.A.

Detection of F12I with *Bcl* I

For the T→A mutation (F12I) the restriction enzyme *Bcl* I can be used to differentiate between individuals harbouring the mutation and individuals who don't. This mutation introduces a *Bcl* I restriction site. The recognition site of this restriction enzyme is indicated in Figure 3.1a. Digestion with *Bcl* I results in two fragments one of 72 bp and the other 91 bp (Figure 3.1b).



t to c mutation indicated in bold letters in Figure 3.1 b.

Figure 3.1 Recognition sequence for *Bcl* I and fragments created in affected and unaffected individuals

The reaction mixture for the digestion was set up as follows: 18µl PCR product, 2,5µl Buffer (10mM Tris-HCl, 50mM NaCl, 10mM $MgCl_2$, 10mM DTT), 10U *Bcl* I, dH_2O to a final reaction volume of 25µl. Digestion took place at 50°C for 1 hour. Inactivation of the enzyme was accomplished by heating the samples to 95°C for a period of 5 min. Loading buffer was added to each tube prior to loading 25µl of the digested fragments onto a 3% agarose gel (Hispanagar) and run at 70V in 1 x TBE.

CHAPTER 4
RESULTS AND DISCUSSION

4.1 PATIENTS

4.1.1 Age and gender

Two hundred and eight patients were included in the study with an equal representation of white and black patients. The age of diagnosis of two black patients and one Caucasian patient is unknown. There is a significant difference in the mean age between the black and Caucasian group (Student's t-distribution df 203, $p<0.0001$). The mean age of the black patient population was 52.99 (± 16.43 years) and the Caucasian population was 65.22 (± 13.47 years). The age ranges are 16 to 83 in the black population and 19 to 88 in the Caucasian patients. In the black population 50 males and 53 females were included. The 104 Caucasian patients comprised, 35 males and 69 females.

4.1.2 Cancer site

Tumours located distal to the splenic flexure (excluding the rectum) were classified as left-sided and those proximal to the splenic flexure as right-sided.

The tumours of the black patients consisted of 25.3% (26/104) left-sided, 21% right-sided (22/104) and 48% (50/104) rectal cancers. The remaining 5.7% (6/104) of tumours were of unknown colorectal origin. Twenty percent (21/104) of the tumours from the Caucasian patients occurred in the right colon, 34.3% (35/104) in the left colon and 40% (42/104) in the rectum. (Figure 4.1)

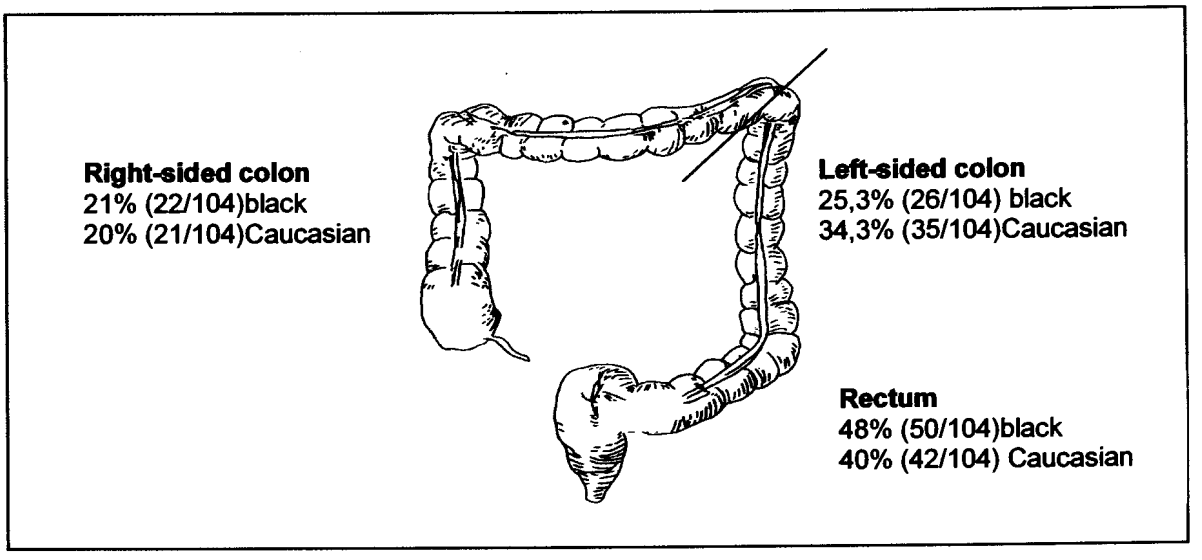
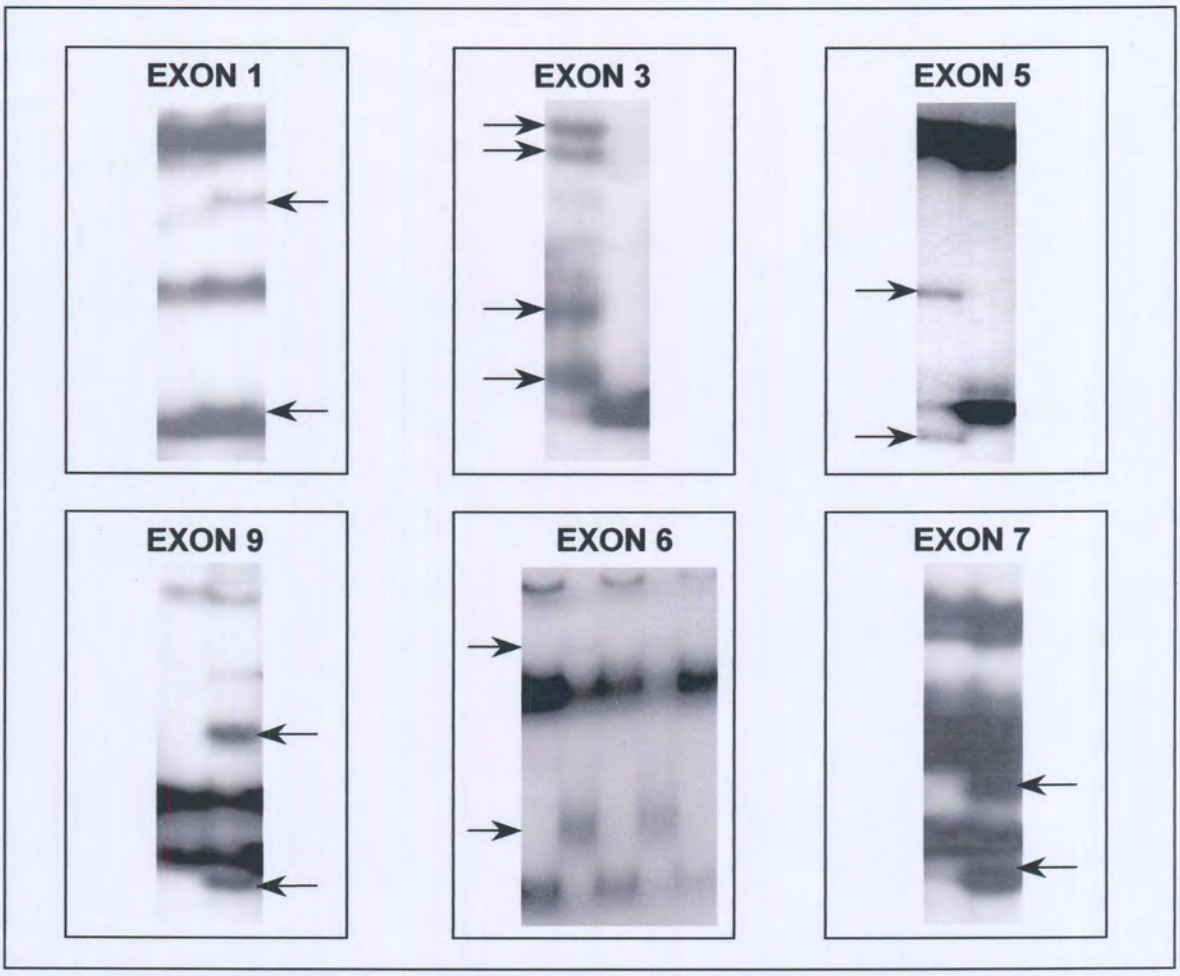


Figure 4.1 Anatomic site of cancer in black and Caucasian patients

In the group as a whole (208 patients), 26.9% (56/208) had tumours in the left colon, 46.63% had tumours in the rectum (97/208) and 20.67% (43/208) were located in the right colon. The location of the remaining 5,8% of tumours is unknown. Thus 47.6% (99/208) of the patients had tumours of the colon and 46.63% (97/208) of the patients had tumours in the rectum.

4.2 SSCP ANALYSIS

STK11 mutation analysis was carried out using exon-by-exon PCR SSCP analysis. A number of samples showed aberrant migrating bands in various exons. Representative autoradiographs of some of the exons are presented in figure 4.2. Samples showing bands with altered mobility were sequenced.



Arrows indicate the position of the aberrant bands.

Figure 4.2 Representative examples of SSCP shifts detected in some of the exons of the *STK 11* gene

4.3 MUTATIONS DETECTED

Of the mutations detected the vast majority (21) were single nucleotide changes of which 8 were missense mutations, 5 synonymous mutations and 8 intronic variants. The remaining mutation was a deletion in the intronic region. Of the single base pair substitutions 6 were transversions and the remaining 15 were transition mutations (Table 4.1), thus transitions were approximately 1,7 times more frequent than transversions, which is in keeping with the expectation of transition mutations to outweigh transversions. (Strachan and Read, 1997).

Table 4.1 Transitions and transversions in the *STK11* gene of colorectal cancer patients

Type of mutation	Nt change	Transition/transversion
Missense (F12I)	c.34T->A	Transversion
Missense (E16K)	c.40G->A	Transition
Missense (K83N)	c.249G->T	Transversion
Missense (I88S/R)	c.263T->G	Transversion
Missense (V150M)	c.448G->A	Transition
Missense (D207N)	c.619G->A	Transition
Missense (A398V)	c.1193C->T	Transition
Missense (P411S)	c.1231C->T	Transition
Silent (I88I)	c.264C->A	Transversion
Silent (Q123Q)	c.369G->A	Transition
Silent (T272T)	c.816C->T	Transition
Silent (N393N)	c.1179C->T	Transition
Silent (A414A)	c.1251C->T	Transition
Intronic	IVS2+24G->T	Transversion
Intronic	IVS3+20G->A	Transition
Intronic	IVS3+29G->A	Transition
Intronic	IVS3+35G->A	Transition
Intronic	IVS3+37G->A	Transition
Intronic	IVS5+16C->T	Transition
Intronic	IVS6+20G->A	Transition
Intronic	IVS6+22C->A	Transversion

Nucleotide positions given refer to STK11 cDNA sequence (Genbank U63333)

Six of the transition mutations were C to T changes and 9 were G to A. Transitions are thought to occur more frequently than transversions partly due to the de-amination of the 5-methylcytosines occurring in CpG islands to give thymidine. This same mechanism also leads to G→A transitions when the de-amination of C→T occurs on the antisense DNA strand, followed by miscorrection of G→A on the sense strand. (Antonarakis *et al.*, 2002) Transitions can also be favoured over transversions due to differential repair of mispaired bases by the sequence dependent proofreading activities of the relevant DNA

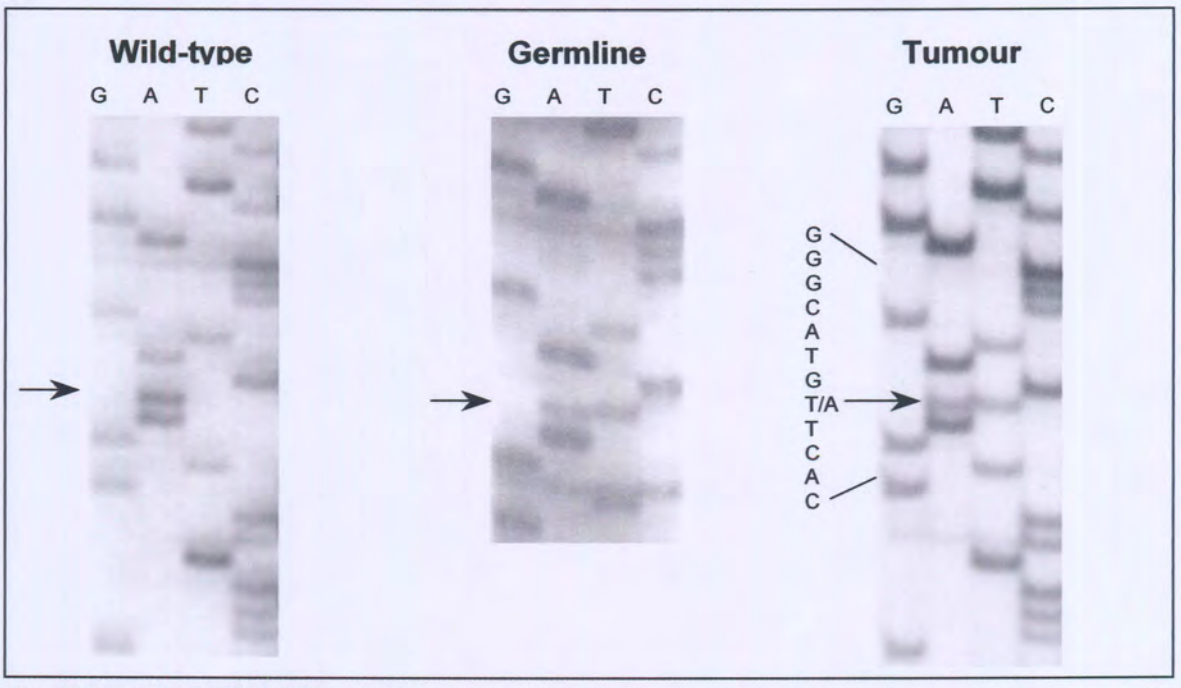
polymerases. Furthermore transitions result in a more conserved polypeptide sequence. (Strachan and Read, 1997)

4.3.1 Missense mutations

All 8 of the missense mutations detected in 7 tumours are novel. Of these only one is a germline change. These mutations occurred in tumours from 5 black and two Caucasian patients. Four of the missense mutations occurred in exon 1, one each occurred in exon 3 and 5 and two missense mutations occurred in exon 9.

Mutation F12I

The germline F12I mutation in exon 1 was identified in the Duke B stage tumour of a 72-year-old black male patient (CRC11) with colon cancer. The T→A transversion results in the replacement of phenylalanine with isoleucine. This is a conserved amino acid change. This amino acid is conserved between human and mouse STK11.



Arrow indicates the position of the mutation

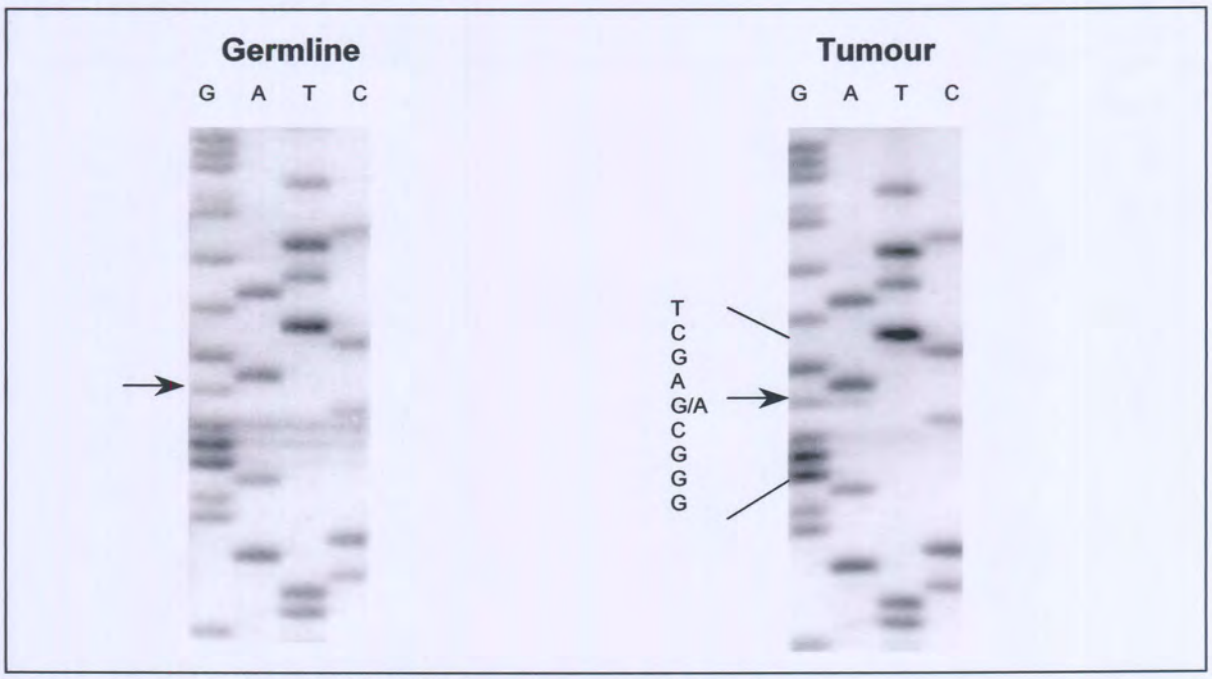
Figure 4.3 Sequence analysis of the anti-sense strand of exon 1 in CRC11

To determine whether this is a polymorphic change, 100 chromosomes from 50 black control individuals were screened via restriction enzyme digestion with *Bcl* I. None of the individuals were positive for the mutation, suggesting that this is not a polymorphism (Cotton and Scriver, 1998).

Mutation E16K

Patient CRW83, a 68 year old white male, had a grade II tumour located in the rectum. Sequence analysis of DNA from tumour and normal tissue revealed a G→A somatic mutation resulting in the missense mutation E16K (Figure 4.4).

This is a non-conservative change from glutamic acid (negatively charged) to lysine (positively charged) and occurs in an evolutionary conserved region. Codon 16 falls outside the kinase domain of the protein.



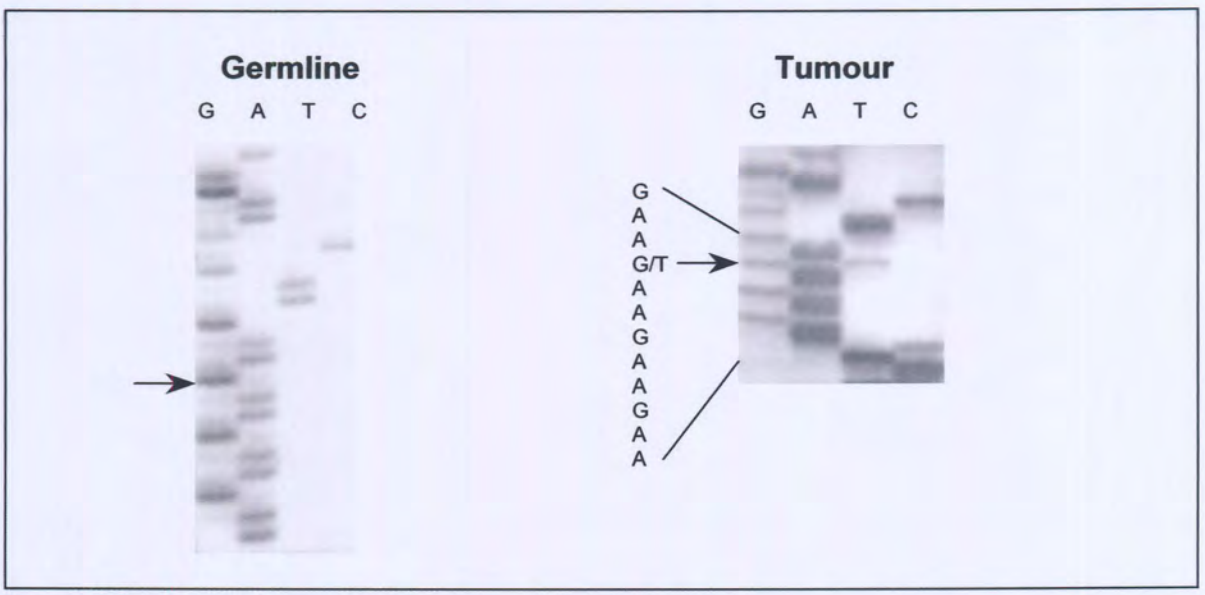
Arrow indicates the position of the of the mutation.

Figure 4.4 Sequence analysis of the sense strand of exon 1 in CRW83

Mutation K83N

A 51-year-old black male patient, CRC12, revealed a somatic missense mutation in his tumour located in the caecum with Duke B tumour stage development (Figure 4.5).

Codon 83 located in the kinase domain as well as the second nuclear localisation signal region codes for lysine, which is evolutionary conserved between human, mouse and *Xenopus* STK11. The nucleotide substitution from G→T results in the replacement of this amino acid with asparagine, which is a non-conservative change.



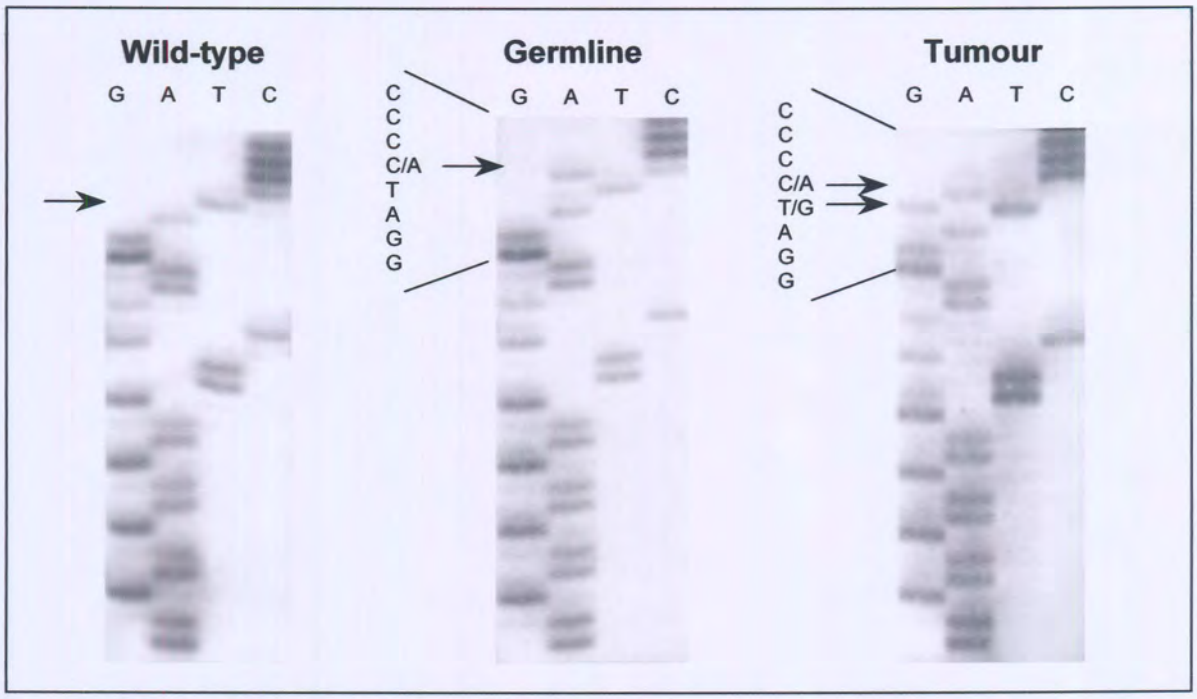
Arrows indicate the position of the mutation.

Figure 4.5 Sequence analysis of the sense strand of exon 1 in CRC12

Mutation I88I and I88S/I88R

A missense mutation in exon 1 was detected in a 42-year-old black female, CRC123, who had a grade II tumour of the rectum. Upon sequencing of the normal and tumour tissue it was found that this patient had a germline C→A transversion in codon 88, which is a synonymous mutation (I88I) that is usually not expected to be pathogenic (Figure 4.6). The tumour however displayed both

this silent mutation as well as a G→T transversion in the same codon. The latter change, if it occurred on the wild type allele (i.e. in trans) will lead to a missense mutation with a serine (uncharged polar amino acid) replacing the isoleucine (non-polar amino acid) (I88S). However if the G→T transversion occurred in cis with the allele harbouring the silent change I88I then this G→T would lead to a missense mutation with arginine replacing the isoleucine in codon 88 (I88R). Of the two possibilities the I88R mutation can be expected to have a greater effect on the STK11 protein as arginine contains a strong basic guanindinium function.



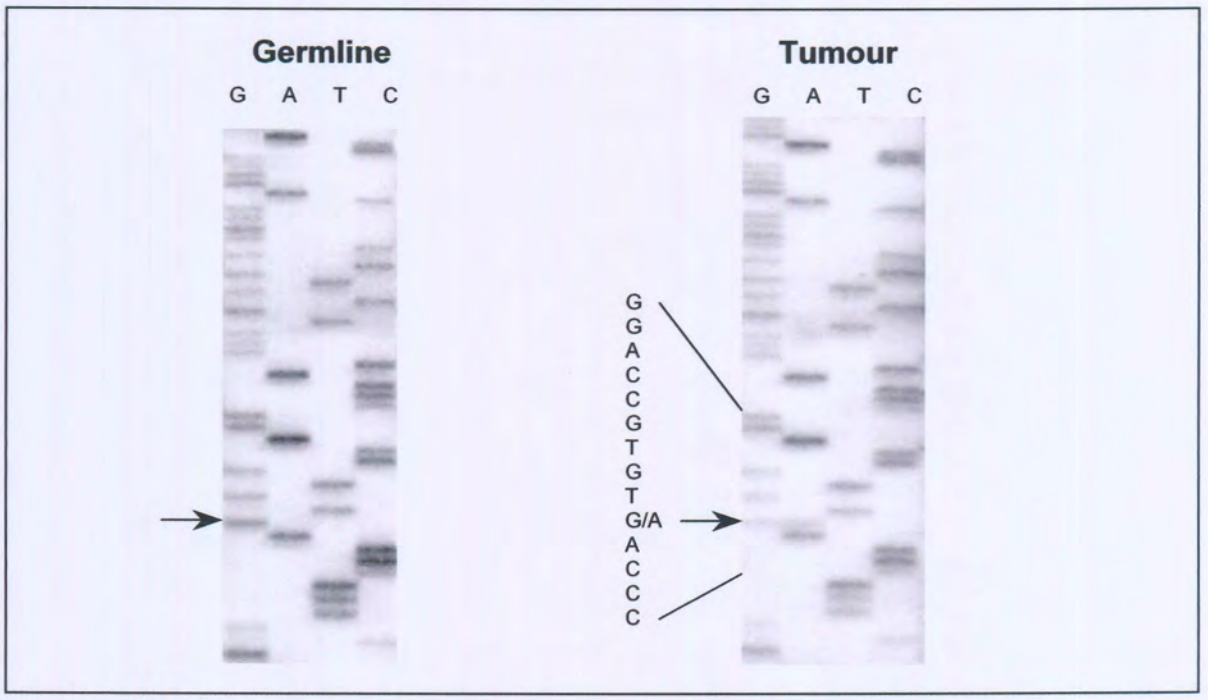
Arrows indicate the position of the mutations

Figure 4.6 Sequence analysis of the sense strand of exon 1 in CRC123

Mutation V150M

An aberrant banding pattern was obtained in exon 3 for patient CRC75. The 62-year-old black patient had a grade II, Duke B2 tumour located in the splenic flexure. Sequence analysis revealed a somatic mutation of a guanine to an adenine in codon 150 (Figure 4.7).

This heterozygous mutation brings about a conservative amino acid change from valine to methionine which is in the kinase domain of STK11. The amino acid at this position is conserved between human, mouse and *Xenopus* STK11.

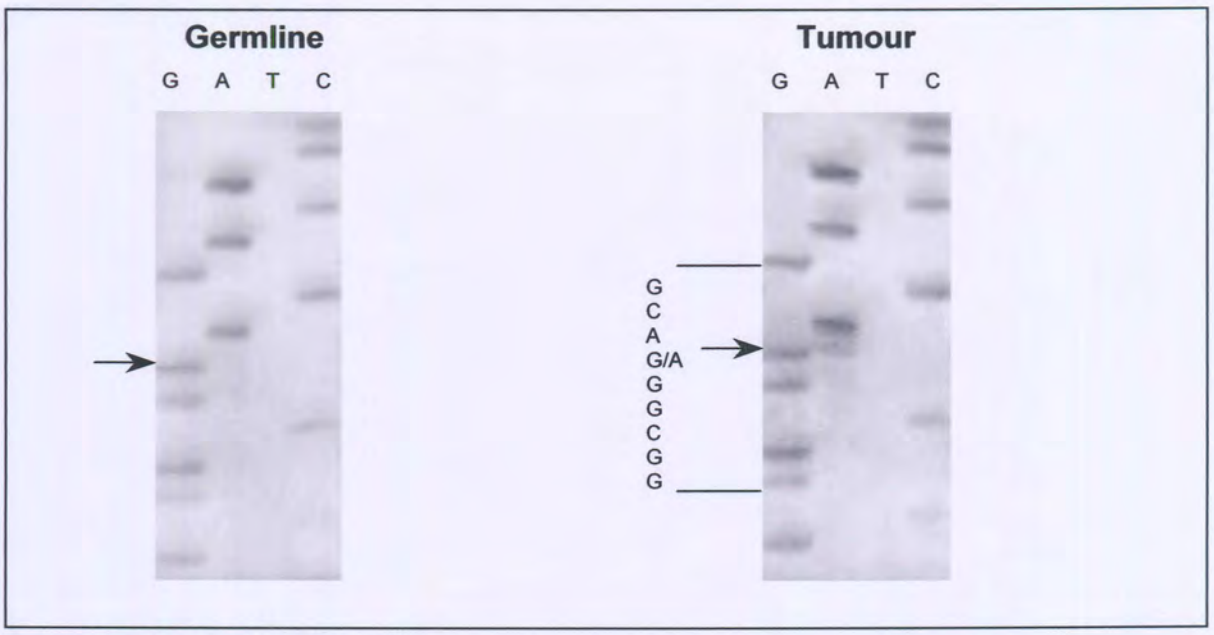


Arrow indicates the position of the mutation

Figure 4.7 Sequence analysis of the sense strand of exon 3 in CRC75

Mutation D207N

The somatic D207N mutation in exon 5 occurred in a black female, aged 46 years (CRC144), who had a GrII, Duke C tumour located in the sigmoid colon. This non-conservative amino acid change falls in the kinase domain of the protein, and is conserved between the human and mouse protein. (Figure 4.8).

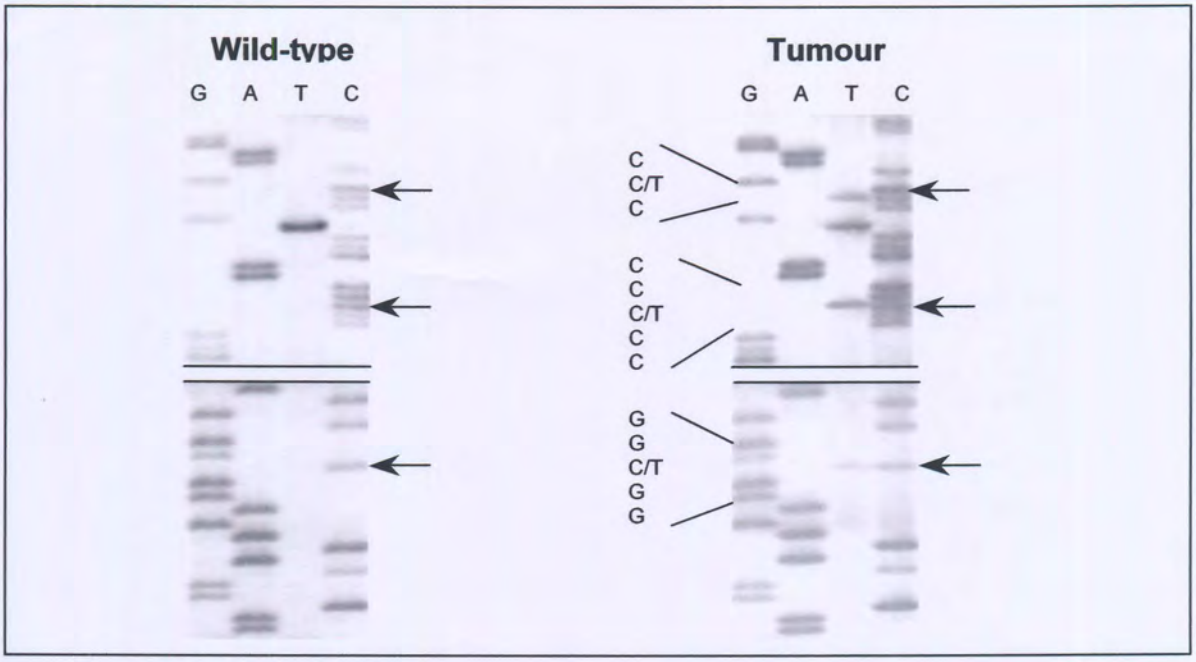


Arrows indicate the positions of the mutation

Figure 4.8 Sequence analysis of the sense strand of exon 5 in CRC144

Mutations A398V, P411S and A 414A

Sequencing DNA of tumour tissue from patient CRW115, an 80-year-old white male, disclosed three nucleotide substitutions in exon 9, two of which resulted in amino acid alterations and the third is a synonymous change (A414A). This tumour, located in the sigmoid region of the colon, is heterozygous for both the missense mutations as well as the synonymous change. All three mutations resulted from the substitution of cytosine with thymidine. This caused the conservative replacement of alanine at codon 398 with valine, and a non-conservative replacement of proline with serine at codon 411 in addition to the synonymous mutation A414A (Figure 4.9). Neither of the affected amino acids are in the kinase domain of the protein and furthermore this is not an evolutionary conserved region of the protein. Sufficient amplification of the normal tissue for sequence analysis couldn't be obtained, thus it cannot be determined if the mutations are germline or somatic. This tumour may possibly be a compound heterozygote, however as no germline sequence data could be generated it is not possible to determine whether these mutations are in cis or trans.



Arrows indicate the positions of the mutation

Figure 4.9 Sequence analysis of the sense strand of exon 9 in CRW115

The pathogenic significance of the missense mutations detected

Cotton and Scriver (1998) suggested criteria to determine the pathologic significance of missense mutations. The criteria suggested for designating a missense mutation as either phenotype- modifying or neutral are;

(i) To determine whether the mutation segregates with the disease.

(ii) To determine the amino acid affected, conserved amino acids are likely to have greater functional importance. Thus mutations that alter conserved residues are likely to affect the structure/function of the protein.

(iii) To determine the prevalence of the mutation. A mutation that occurs on less than 1% of the alleles in a control population, is rare and may be disease causing, whereas a mutation found at higher frequency may be a neutral polymophism.

(iv) Functional analysis of the mutant gene.

Applying these criteria to the 8 missense mutations detected in this study show the following:

(i) Only one germline mutation (F12I) was detected in the present study, but as this is a population based study it is not possible to determine whether this mutation segregates with the disease.

(ii) Codons 12, 16, 83, 88, 150 and 207 are located in evolutionary conserved regions of the protein and thus substitution of these amino acids with amino acids of different physical character (non-conservative change) could potentially affect the function of the protein. Except for the missense mutations in codons 12 and 150 all of the other missense mutations were non-conservative amino acid substitutions. Furthermore codons 83, 88, 150 and 207 reside within the kinase domain (amino acids 50-319) of the STK11 protein. In addition to the effect that the K83N mutation may have on the kinase activity of the protein, this residue is also located in the second NLS region of the protein, which may influence the cellular localisation of the protein. The four non-conservative amino acid substitutions in codons 16, 83, 88 and 207 may therefore be pathogenic.

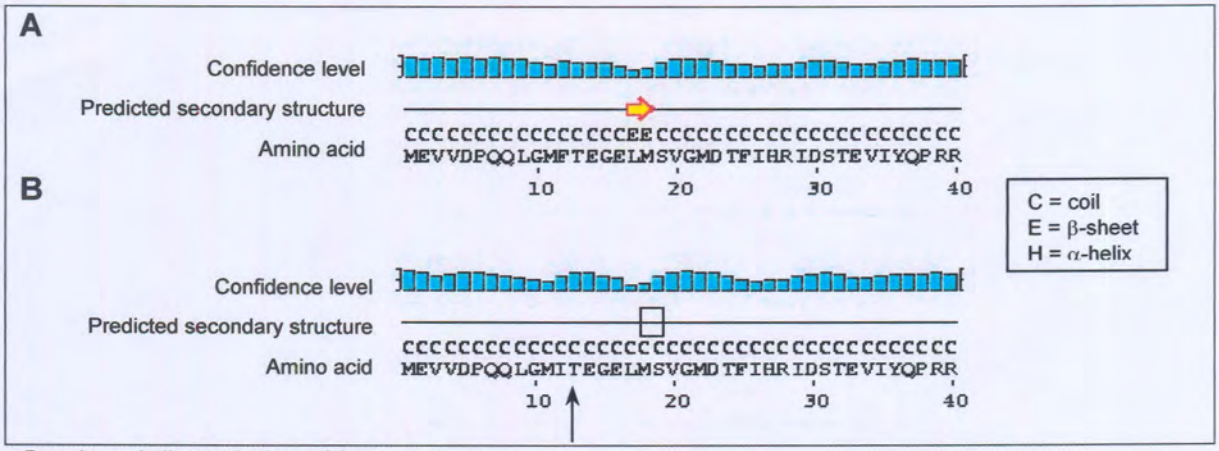
(iii) The germline F12I mutation was not present in 50 unaffected control individuals suggesting that this is not a polymorphism.

(iv) At this time no functional studies could be carried out to determine what the effects of the missense mutations are on the biological activity of STK11.

In addition to the above criteria the effect of the missense mutations on the secondary structure of the STK11 protein was investigated using the PSIPRED program of the University College London (McGuffin *et al.*). The predicted secondary structure of the wild-type STK11 protein was compared to the predicted secondary structures containing the various missense mutations. Mutations F12I, A398V and P411S all displayed changes.

F12I

In figure 4.10A the predicted secondary structure of the wild type protein shows a β-sheet corresponding to amino acids 16 and 17. A prediction of the secondary structure of the same domain of the protein containing the substituted amino acid shows this β-strand to be replaced by a coil indicated by the box in figure 4.10B.



Boxed area indicates the loss of the β-sheet
Figure adapted from; McGuffin *et al.* http://bioinf.cs.ucl.ac.uk/psiform

Figure 4.10 Secondary structure prediction for the first forty amino acids of wild type STK11(A) and the F12I mutation (B)

A398V and P411V

The predicted secondary structure of the protein encompassing the codons 398 and 411 is given in figure 4.11. At the locations of the two mutations no changes in the predicted secondary structure is noted. However four differences in the last fragment of the protein can be seen. The α-helix at amino acids 315-318 is replaced by a coil in the mutant form of the protein. Furthermore the two α-helix structures spanning from 329-333 and 337-340 becomes one large α-helix. Perhaps the most obvious change occurs in the very last fragments of the protein. At amino acid residue 424 a β-sheet spanning 8 amino acids is formed in the protein with the two missense mutations. In this region an α-helix is predicted to occur in the wild type.

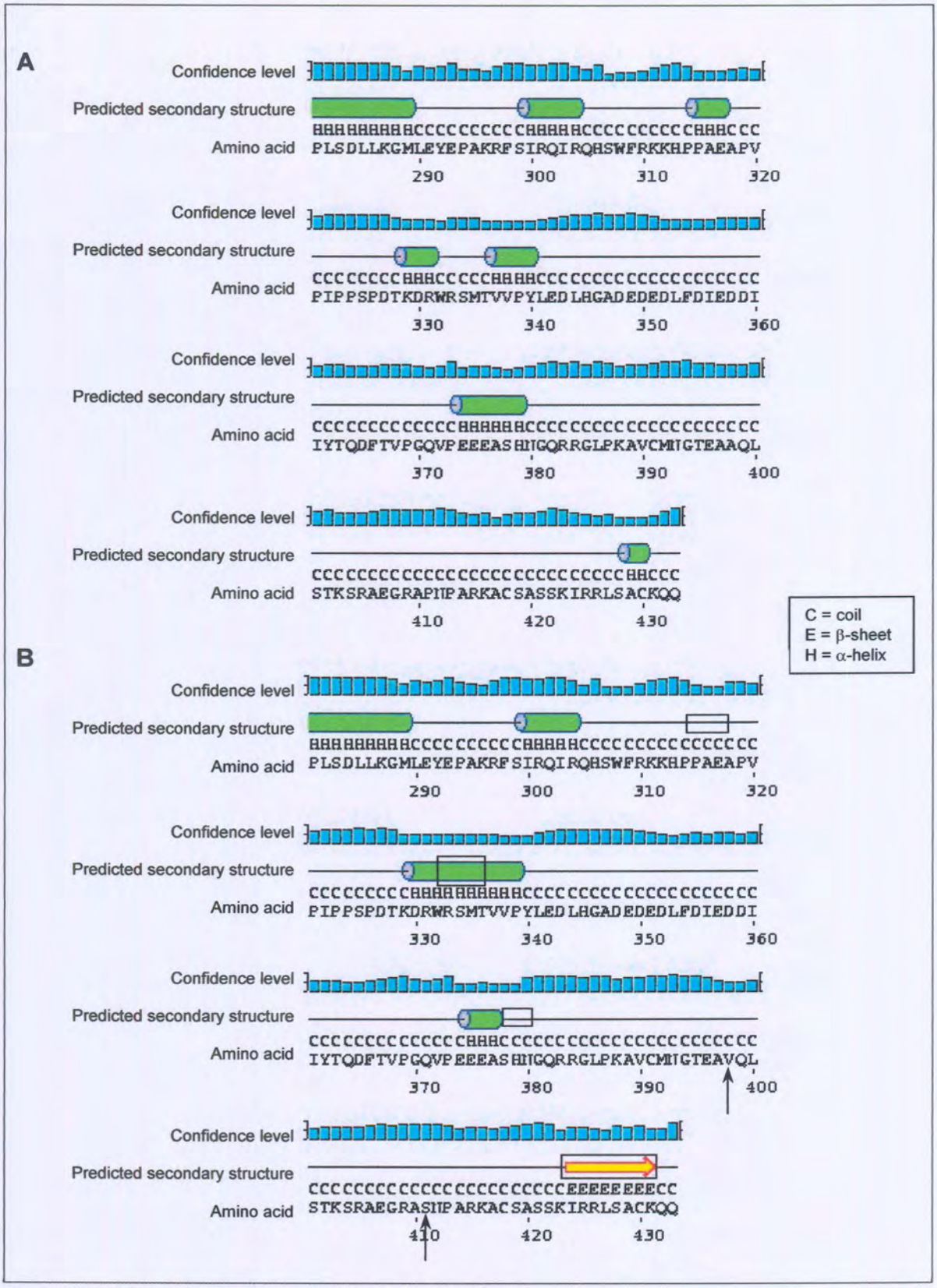


Figure adapted from; McGuffin *et al.* http://bioinf.cs.ucl.ac.uk/psiform

Figure 4.11 Secondary structure prediction for amino acids 280-433 of wild type STK11 (A) and the A398V and P411S mutations (B)

Taking all of the above into account it would appear that 4 of the missense mutations could possibly have pathogenic significance (Table 4.2)

Table 4.2 Pathogenic significance of the missense mutations

Exon	Patient	Nt change (cDNA sequence)	Designation	Pathogenic significance
1	CRC11	34T >A	F12I *	?
1	CRW83	40G >A	E16K	Possible?
1	CRC12	249G >T	K83N	Possible?
1	CRC123	263T >G	I88S/R	Possible?
3	CRC75	448G >A	V150M	?
5	CRC144	619G >A	D207N	Possible?
9	CRW115	1193C >T	A398V	?
9	CRW115	1231C >T	P411S	?

Nucleotide positions given refer to STK11 cDNA sequence (Genbank U63333)
* germline mutation

If one was to assume that all of the missense mutations identified in this study are of pathogenic importance it would imply the involvement of *STK11* in 3.36% (7/208) of sporadic CRC. Compared to the study of Dong *et al.*, (1998) who identified 7 mutations in 23 CRC (30.43%) the present study shows an approximate ten-fold lower mutation rate. Interestingly when the present results are separated into the two population groups, 4.8% (5/104) of the tumours from black patients and 1.92% of tumours (2/104) from the Caucasian patients had *STK11* missense mutations. Although this is a 2.5-fold difference it is considered not significant (χ^2-test p=0.442).

Anatomic site of tumours with missense mutations

Of the 8 missense mutations, which occurred in the seven tumours, two occurred in tumours in the rectum (2/7, 28.57%), three in the left colon (3/7, 42.86%), and one in the right colon (1/7, 14.28%) (Figure 4.12). Unfortunately the anatomic site of the F12I mutation is unknown.

Of the total 208 patients, 97 had tumours of the rectum, 56 had tumours in the left colon and 43 in the right colon. Thus 2.06% (2/97) of rectal tumours, 5.36% (3/56) of left-sided tumours and 2.33% (1/43) of right-sided tumours had missense mutations. There is no significant difference between the ratio of missense mutations in left and right-sided tumours (χ^2-test, 1 df, p=0.807).

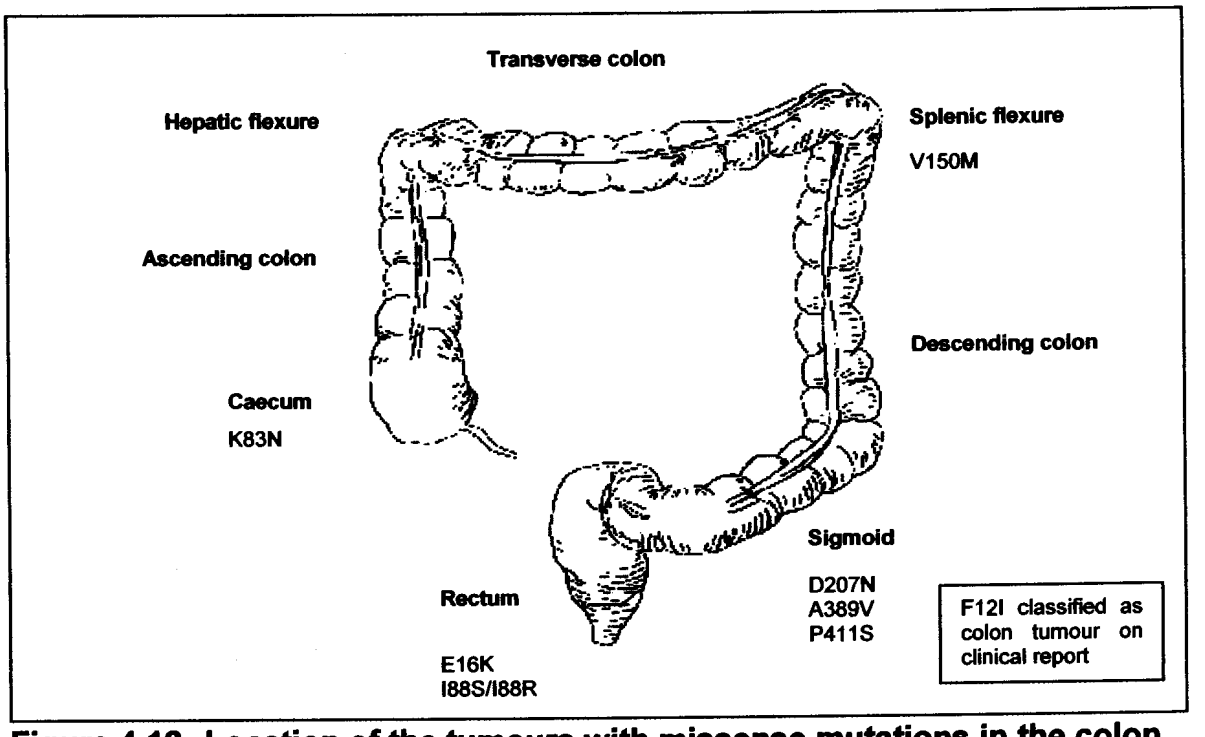


Figure 4.12 Location of the tumours with missense mutations in the colon

When separating the results into the two population groups 11.54% (3/26) of the left-sided tumours, 2% (1/50) of rectal tumours and 4.54% (1/22) of right-sided tumours harbored mutations. These differences are not significant (χ^2-test, 1 df, p=0.727). In the Caucasian group 2.86% (1/35) of left-sided tumours, 2.38% (1/42) of rectal tumours and none of the right-sided tumours had *STK11* mutations.

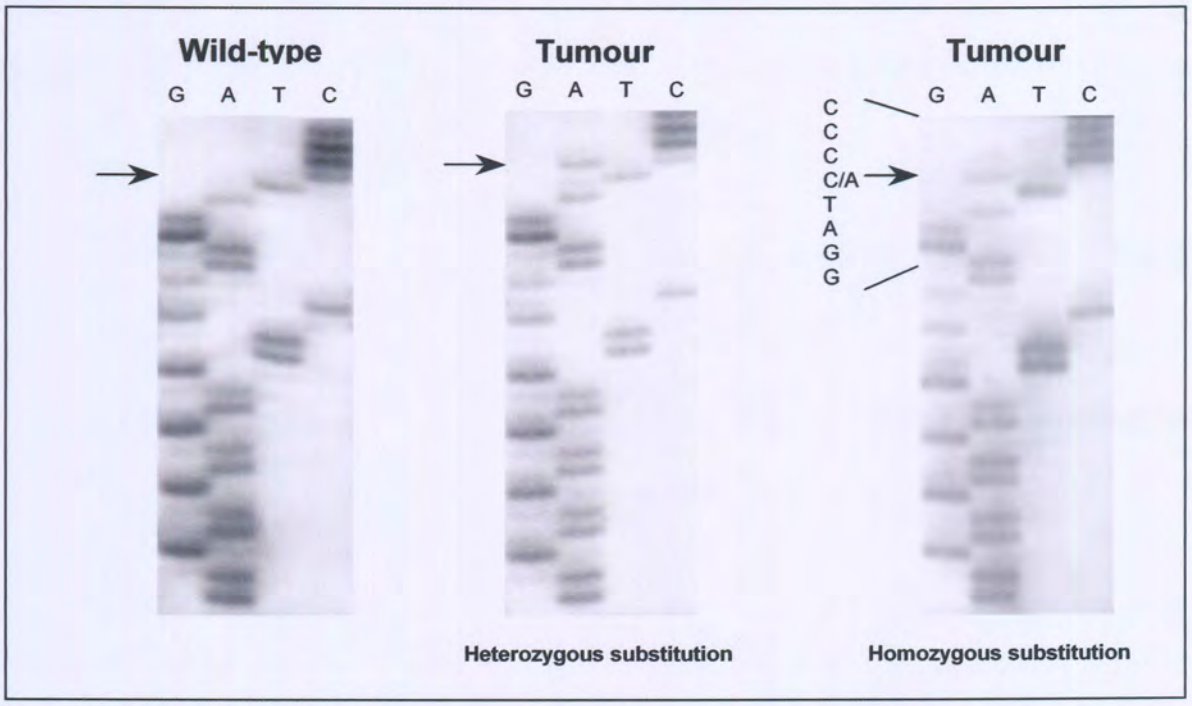
4.3.2 Synonymous mutations

In this study five novel synonymous mutations were identified, I88I, Q123Q, T272T, N393N and A414A. There is increasing evidence that many human

disease genes harbour translationally silent mutations that affect pre-mRNA splicing. These mutations are believed to alter cis-elements that are important for correct splicing. (Cartegni *et al.*, 2002)

Mutation I88I

Twenty tumours were found to harbour a single nucleotide change C→A in codon 88. This is a synonymous change thus not affecting the amino acid (isoleucine) (Figure 4.13).



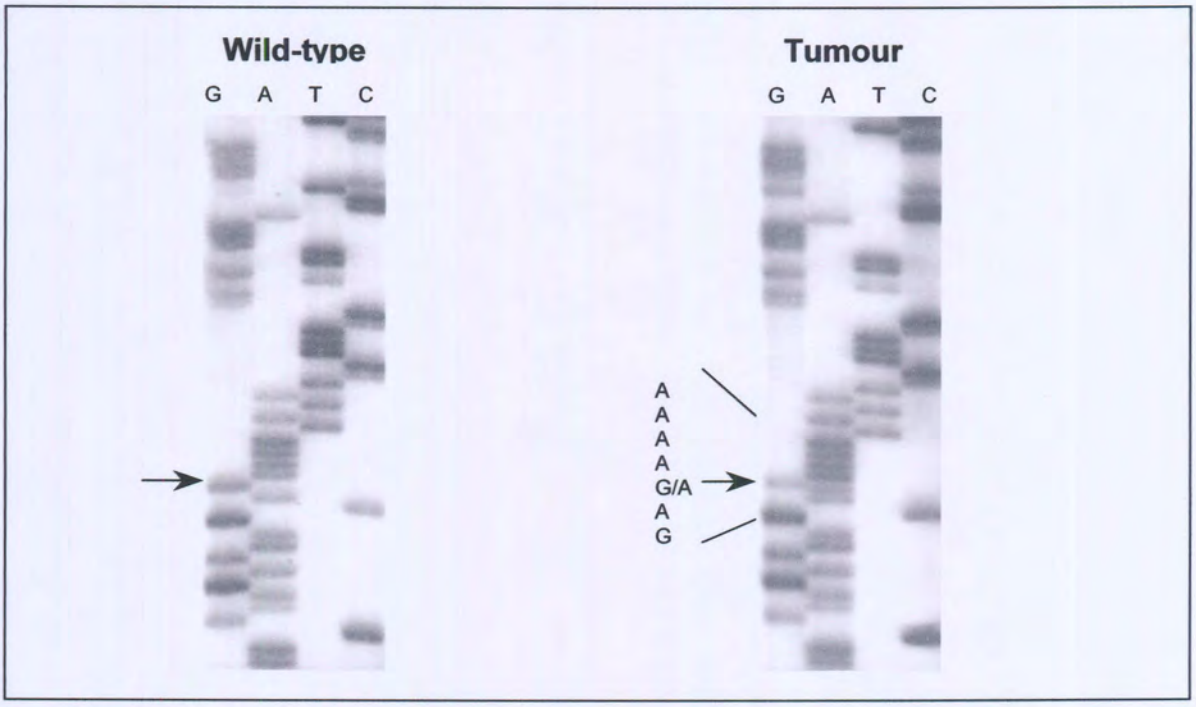
Arrows indicate the position of the mutation

Figure 4.13 Sequence analysis of the sense strand of exon 1 for the I88I mutation

Of the 208 patients 19 black individuals and only one white patient had this nucleotide substitution in codon 88. Of the 20 individuals carrying this polymorphism only three (CRC 7, 10, 46) were homozygous for the substitution. This implies that the adenine allele frequency of the patient population is 5,52% (23/416). When comparing the black and white patient groups, the adenine allele frequency differs significantly between the Caucasian (0.48%, 1/208) and black (10.58%, 22/208) populations (χ^2-test, p=0.0001).

Mutation Q123Q

Four white and 5 black patients had a heterozygous G→A transition in codon 123 (Figure 4.14). Thus there is a 2,16% (9/416) A-allele frequency in the patient population.



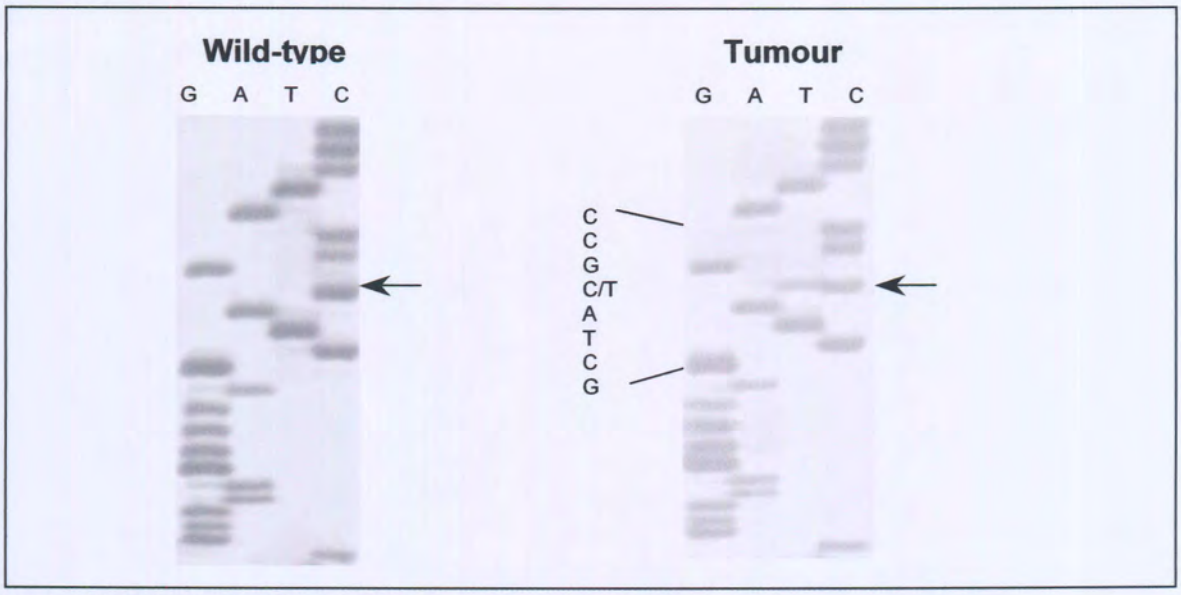
Arrows indicate the position of the mutation

Figure 4.14 Sequence analysis of the sense strand of exon 2 for the Q123Q mutation

Mutation T272T

A total of 18 black patients had the C→T transition, resulting in the synonymous T272T mutation (Figure 4.15).

All the patients were heterozygous for this change showing an 8.65% (18/208) frequency of the thymidine allele at nucleotide c.816 in the affected black patients. None of the white patients had this mutation.

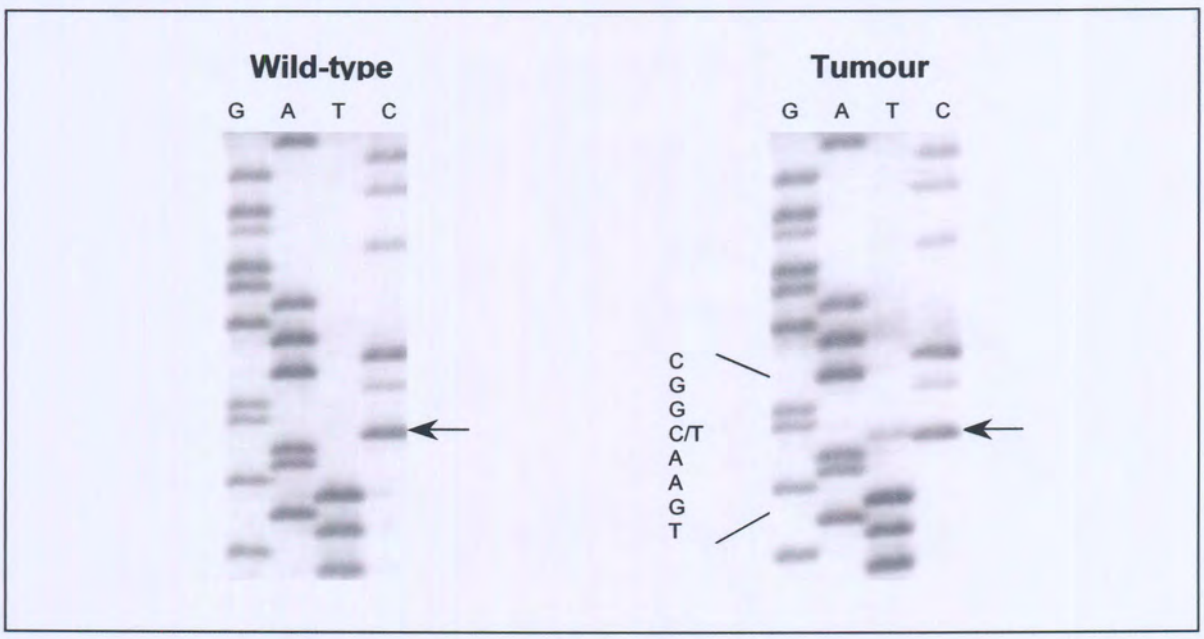


Arrows indicate the position of the mutation

Figure 4.15 Sequence analysis of the sense strand of exon 6 for T272T

Mutation N393N

Sequencing of exon 9 for patient CRW66 unveiled a single nucleotide substitution (C→T) in codon 393 of *STK11* (Figure 4.16). This patient was the only one out of 208 with this particular shift associated with this mutation. As normal tissue of this patient could not be analysed it was impossible to determine whether this mutation is somatic or germline in origin.



Arrows indicate the position of the mutation

Figure 4.16 Sequence analysis of the sense strand of exon 9 for N393N

Anatomic site of tumours with synonymous mutations

A total of 47 patients were found to have samesense mutations. The precise location of four tumours is unknown. The anatomical location of the other tumours with samesense mutations is given in figure 4.17.

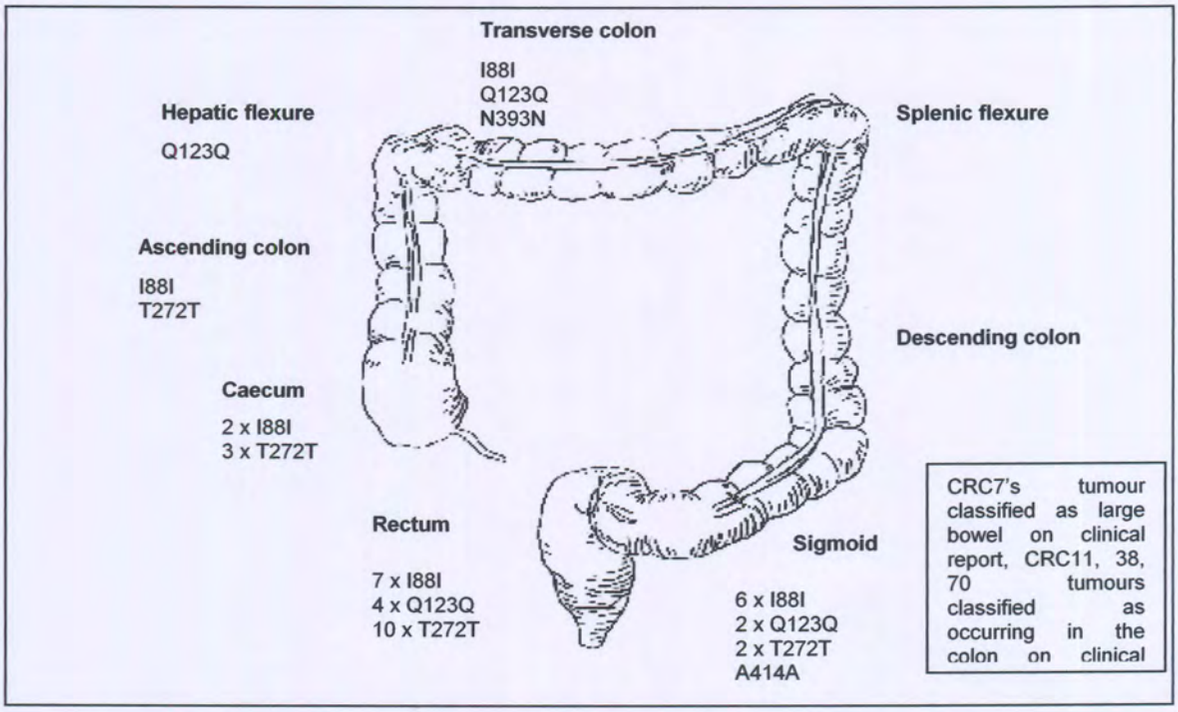


Figure 4.17 Location of tumours with samesense mutations in the colon

Twenty-one of the tumours with samesense mutations occurred in the rectum (21/97) 21.65% and 19.64% (11/56) eleven in the sigmoid colon. Altogether 11 tumours 25.58%, (11/43) were located on the right side of the colon. The remaining 1.98% (4/208) of the synonymous mutations belong to tumours of unknown colorectal origin.

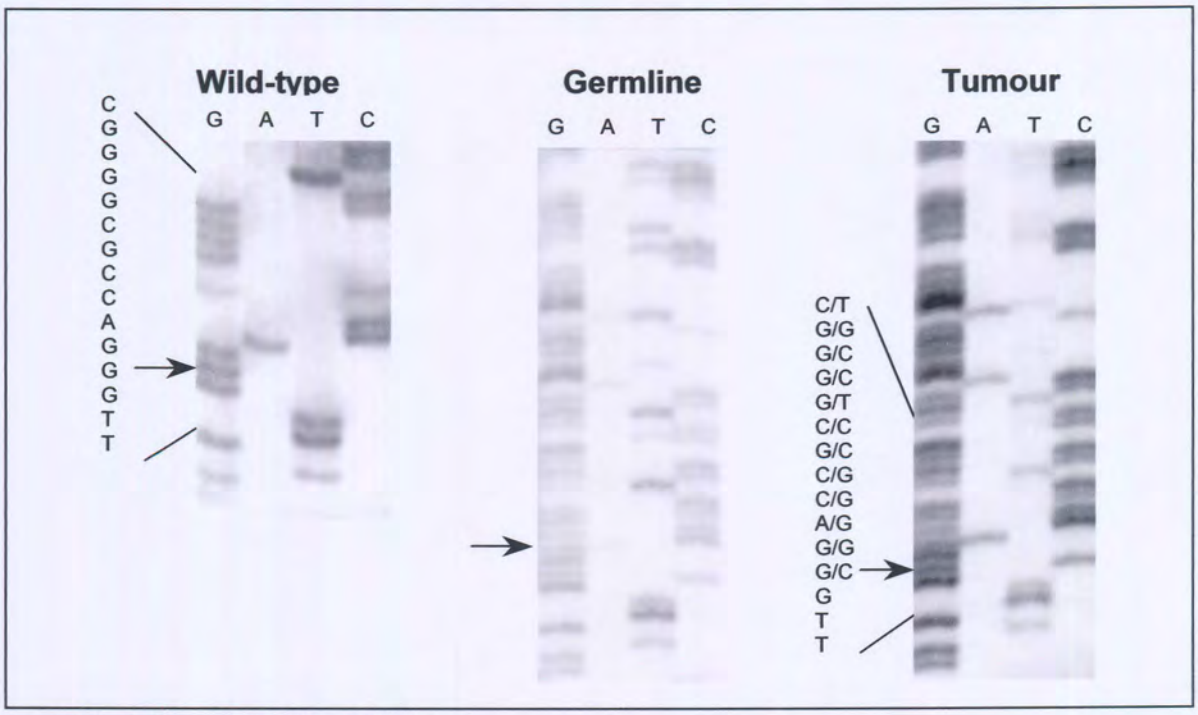
4.3.3 Intronic mutations

Mutations located within the introns of genes are currently described as being either splice site mutations or intronic variants. In this study 9 different mutations

were identified in the intronic regions of *STK11* of which eight are novel. The one previously reported mutation is in intron 2 (IVS 2+24 G→T). Four of the mutations occurred in intron 3, and were all transitions, G→A. Two mutations each were located in introns 2 and 6 and one in intron 5. Except for a deletion of six base pairs in exon 2 all other mutations were single nucleotide changes with 6 transition and 3 transversion mutations, mostly situated between 16 to 37 bases away from the exon boundaries.

Six base-pair deletion in intron 2 (IVS 2 + 19del6)

Apart from having the samesense mutation (Q123Q) in exon 2 CRC176 also has a novel 6 base-pair deletion in intron 2. Results obtained from sequencing are depicted in figure 4.18. From the sequence analysis the six deleted nucleotides can be identified as ggaccg starting at the nucleotide located 19 bases downstream from the 3'-end of exon 2. The individual is heterozygous for this deletion, which also occurs in the germline. Considering the location of this deletion is 19 bp downstream of exon 2 it is unlikely to be of any significance.

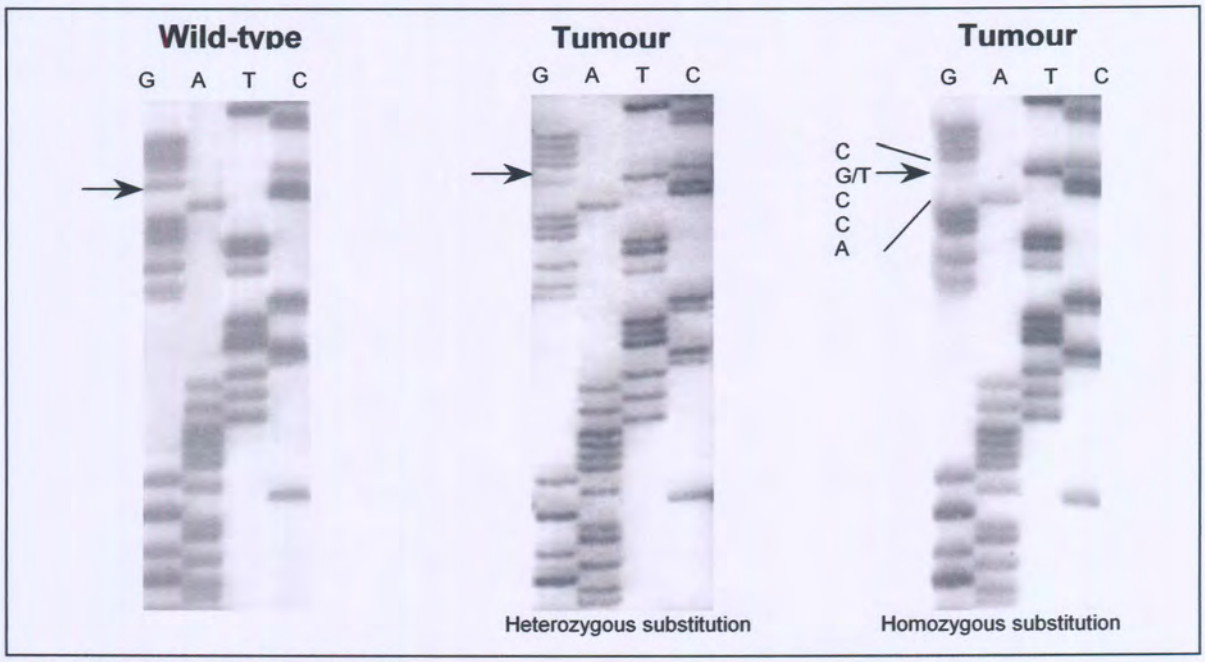


Arrows indicate the positions of the mutation

Figure 4.18 Sequence analysis of the sense strand of exon 2 for (IVS 2 + 19del6)

Of all the mutations identified in this study the single nucleotide substitution from guanine to thymidine located 24 bases into intron two (Figure 4.19) is the only previously reported mutation Avizienyte *et al.*, 1998; Resta *et al.*, 1998; Su *et al.*, 1999; Olschwang *et al.*, 2001, which is considered to be silent.

Twenty-seven black and no white patients were found to be homozygous for the mutation. A further seventeen white and twenty-two black affected individuals were heterozygous for the nucleotide substitution. Thus of 208 individuals screened, 66 (31.73%) have the G→T (IVS 2 + 24) transversion. Overall the T-allele frequency therefore is 22.35% (93/416) and the wild type G-allele 77.65% (323/416). Interestingly the T-allele frequency of the black patient group (76/208; 36,5%) is significantly increased (χ^2-test p=0.0001) compared to that of the Caucasian group (17/208; 8,2%)

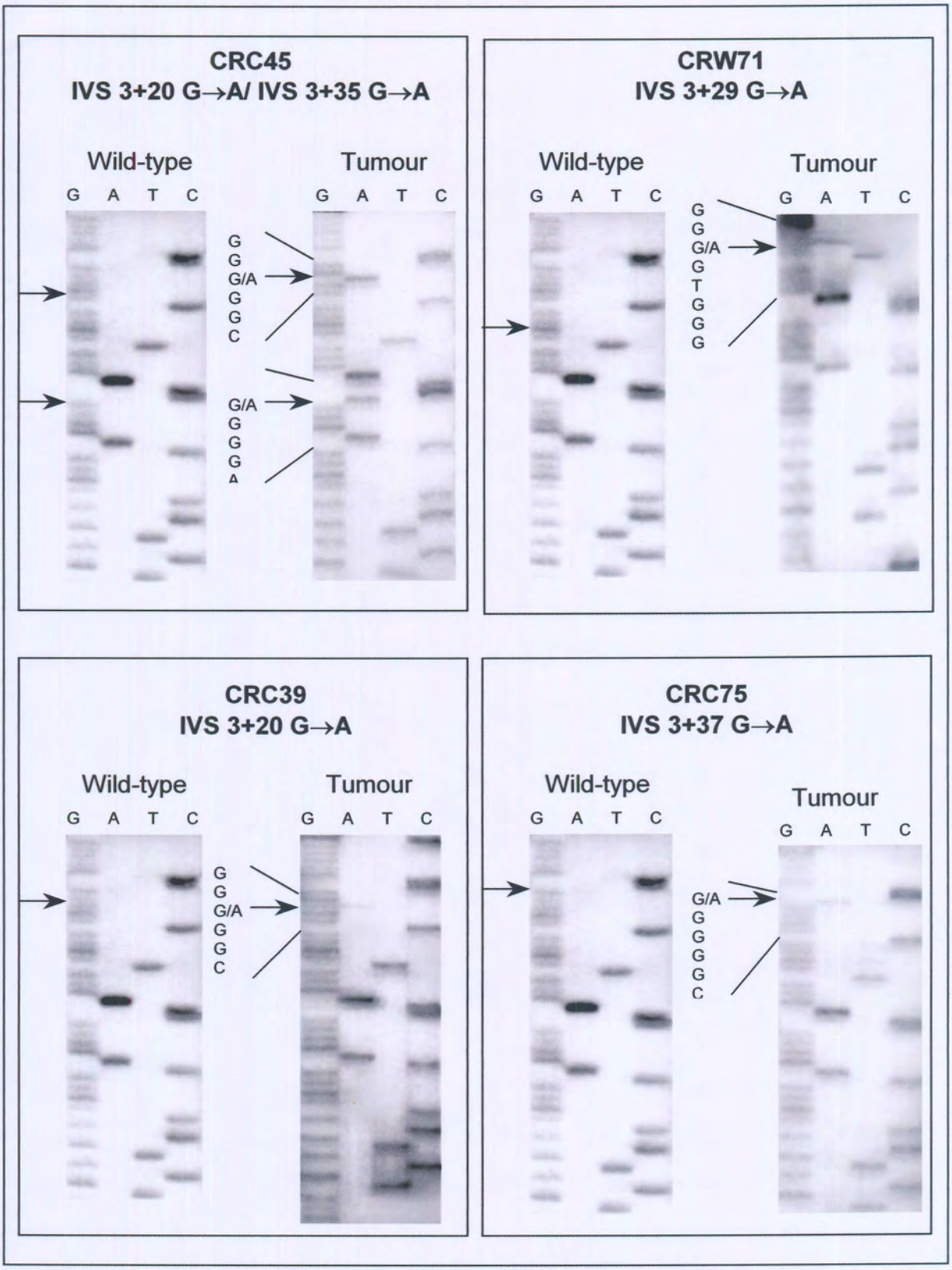


Arrows indicate the position of the mutation

Figure 4.19 Sequence analysis of the sense strand of exon 2 for G to T transversion in intron two (IVS 2+24 G→T)

Mutations in intron 3

Downstream from the 3'-end of exon 3 there are a series of CpG islands. In intron 3, four guanine residues were substituted by adenine. Patient CRC45 and CRC39 had the mutation IVS 3+20G->A and patient CRC45 had an additional mutation IVS 3+35G->A. CRC39 was heterozygous for the mutation, CRC45 on the other hand was homozygous for both mutations. IVS 3+29 was a heterozygous change in patient CRW71. The furthest mutation into the intronic region of exon 3 was found in patient CRC75, who also harbours V150M. This mutation is IVS 3+37 G→A. Sequence analysis of the mutations are represented in figure 4.20.

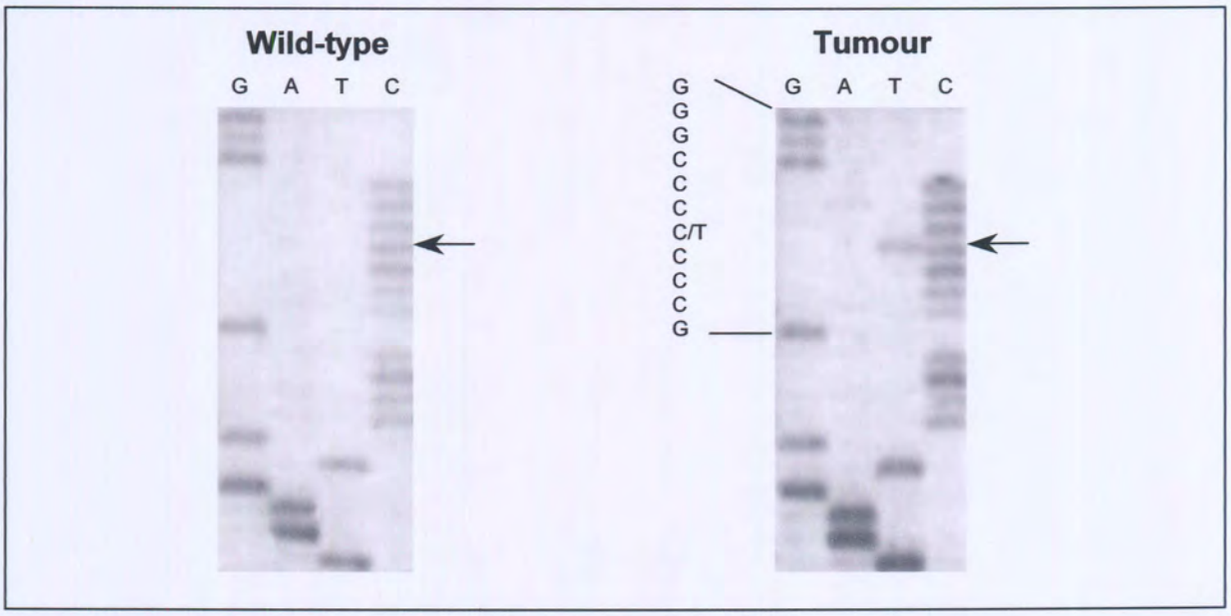


Arrow indicates the position of the mutation

Figure 4.20 Sequence analysis of the sense strand of intron 3 in CRC45, CRW71, CRC39 and CRC75

Cytosine to thymidine substitution in intron 5

A white patient (CRW118) had a single nucleotide substitution, 16 bases into the intronic region of exon 5. A representation of wild-type versus the mutant sequence is given in figure 4.21. This patient is heterozygous for IVS5+16C->T.

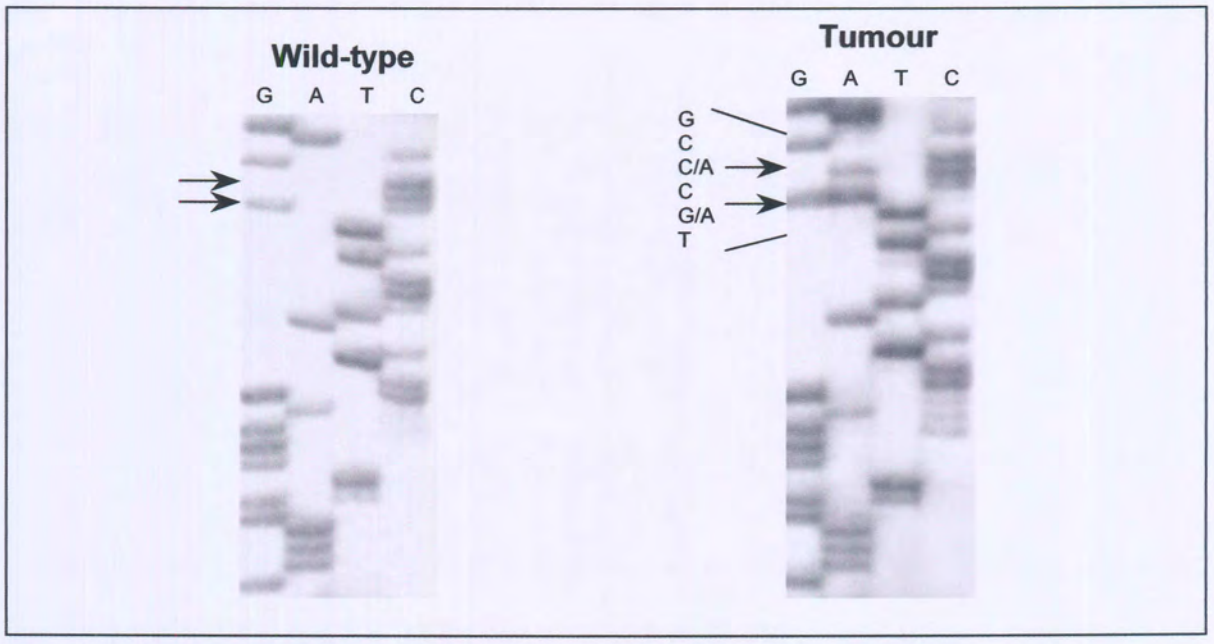


Arrow indicates the position of the mutation

Figure 4.21 Sequence analysis of the sense strand of CRW118 of exon 5

Two novel mutations in intron 6

Two white individuals (CRW19, CRW111) had two identical mutations located in intron 6 (Figure 4.22). The first mutation a G→A is located 20 bases downstream (IVS6+20G->A) and the second mutation a C→A is located a further 2 bases downstream (IVS6+22C->A). from exon 6. Both individuals were heterozygous for the two mutations



Arrows indicate the of the mutation

Figure 4.22 Sequence analysis of the sense strand of intron 6 in CRW19 and CRW111

Anatomical site of tumours with intronic mutations

A total of 79 intronic mutations were identified in this study of which 63 (79,74%) occurred in tumours in the left side of the colon. Ten mutations were identified in right-sided tumours (12,65%). The remaining 7,61% of the intronic mutations was found in tumours without a clear description of the colonic location.

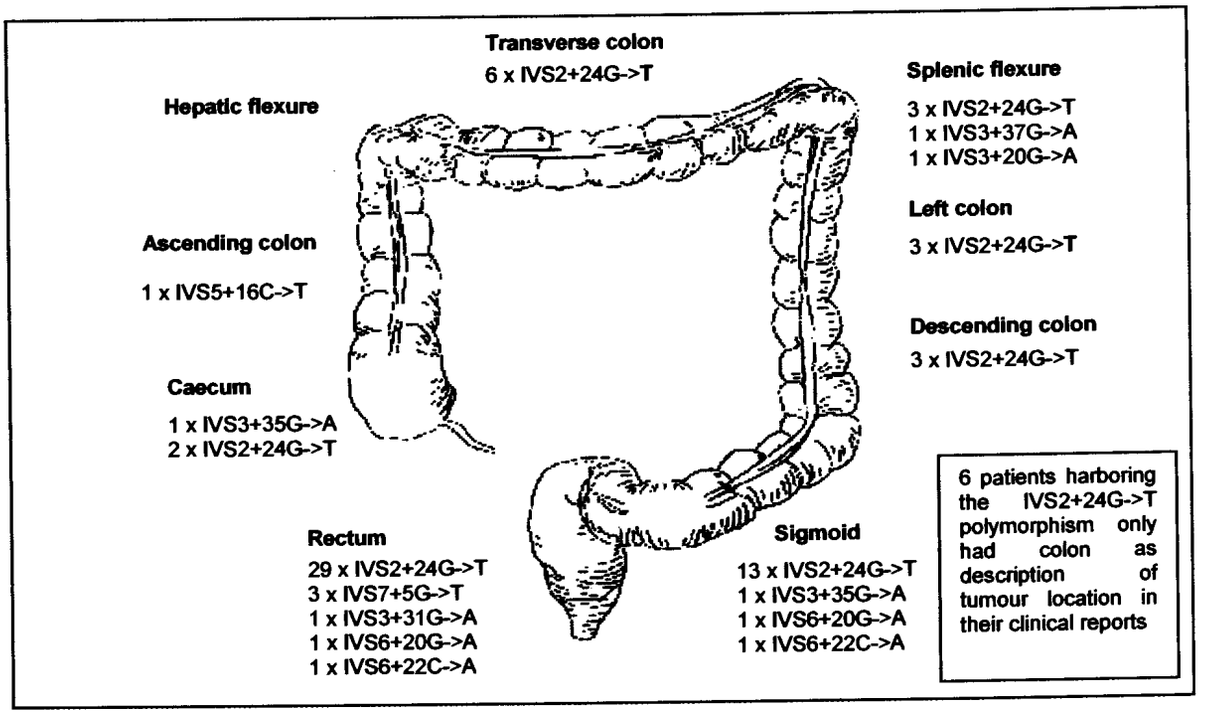


Figure 4.23 Location of tumours with intronic mutations in the colon

4.4 SUMMARY OF ALL MUTATIONS

In total 73 tumours from black patients and 25 tumours from white patients were found to have mutations in the *STK11* gene. This observation clearly indicates a significantly higher mutation rate of *STK11* in the black population group compared to the white patients (χ^2-test p=0.0001). Five missense, 3 synonymous and 5 intronic mutations were detected in the 73 tumours from black patients, whereas only three missense, 3 synonymous and 5 intronic mutations were detected in the 25 tumours from Caucasian patients. These mutations and anatomic site of the tumours are summarised in table 4.2.

Table 4.3 Mutations and the anatomical site of the tumours

Patient number	Mutation	Anatomical site
CRC1	I88I	Caecum
CRC2	I88I	Rectum
CRC7	I88I	Large bowel
CRC9	T272T	Rectum
CRC10	I88I	Descending sigmoid
CRC11	T272T/ F12I	Colon
CRC12	K83N	Caecum
CRC16	I88I	Ascending colon
CRC17	T272T	Caecum
CRC24	IVS2+24G->T	Sigmoid
CRC27	Q123Q	Sigmoid
CRC28	T272T	Caecum
CRC31	IVS2+24G->T	Left colon
CRC32	IVS2+24G->T	Left colon
CRC34	T272T	Rectum
CRC38	T272T/ Q123Q/ IVS2+24G->T	Colon
CRC39	T272T/ IVS3+35G->A	Caecum
CRC41	I88I/ IVS2+24G->T	Transverse colon
CRC43	IVS2+24G->T	Transverse colon
CRC45	I88I/ T272T/ IVS3+20G->A/ IVS3+35G->A	Sigmoid
CRC46	I88I/ IVS2+24G->T	Sigmoid
CRC47	IVS2+24G->T	Colon
CRC52	T272T	Ascending colon
CRC59	I88I/ IVS2+24G->T	Sigmoid
CRC62	IVS2+24G->T	Sigmoid
CRC63	IVS2+24G->T	Rectum
CRC66	IVS2+24G->T	Rectum
CRC68	IVS2+24G->T	Rectum
CRC69	IVS2+24G->T	Rectum
CRC70	I88I/ IVS2+24G->T	Colon
CRC71	IVS2+24G->T	Transverse colon
CRC73	IVS2+24G->T	Transverse colon
CRC75	V150M/ IVS2+24G->T/ IVS3+37G->A	Splenic flexure

CRC76	IVS2+24G->T	Left colon
CRC78	IVS2+24G->T	Descending colon
CRC80	IVS2+24G->T	Splenic flexure
CRC82	I88I/ IVS2+24G->T	Sigmoid
CRC84	IVS2+24G->T	Sigmoid
CRC86	IVS2+24G->T	Sigmoid
CRC87	T272T	Rectum
CRC88	I88I/ IVS2+24G->T	Rectum
CRC90	IVS2+24G->T	Rectum
CRC91	T272T/ IVS2+24G->T	Rectum
CRC93	IVS2+24G->T	Rectum
CRC98	IVS2+24G->T	Rectum
CRC123	I88I/ I88S/ IVS2+24G->T	Rectum
CRC100	IVS2+24G->T	Rectum
CRC124	IVS2+24G->T	Descending colon
CRC127	IVS2+24G->T	Sigmoid
CRC128	I88I	Caecum
CRC144	T272T/ Q123Q/ D207N	Sigmoid
CRC147	I88I	Rectum
CRC149	Q123Q	Hepatic flexure
CRC152	IVS2+24G->T	Caecum
CRC170	IVS2+24G->T	Rectum
CRC171	IVS2+24G->T	Rectum
CRC173	T272T/ IVS2+24G->T	Rectum
CRC174	IVS2+24G->T	Rectum
CRC175	T272T	Rectum
CRC176	T272T/Q123Q	Rectum
CRC178	I88I/ IVS2+24G->T	Rectum
CRC179	IVS2+24G->T	Rectum
CRC180	I88I	Distal rectum
CRC182	T272T/ IVS2+24G->T	Rectum
CRC183	I88I	Rectum
CRC185	T272T/ IVS2+24G->T	Rectum
CRC186	IVS2+24G->T	Colon
CRC190	IVS2+24G->T	Sigmoid
CRC191	IVS2+24G->T	Rectum
CRC192	IVS2+24G->T	Rectum
CRC194	IVS2+24G->T	Rectum

CRC195	IVS2+24G->T	Rectum
CRC196	T272T	Rectum
CRW3	IVS2+24G->T	Rectum
CRW16	IVS2+24G->T	Colon
CRW19	IVS6+20G->A/ IVS6+22C->A	Sigmoid
CRW24	I88I/ IVS2+24G->T	Sigmoid
CRW27	IVS2+24G->T	Rectum
CRW28	IVS2+24G->T	Colon
CRW29	IVS2+24G->T	Sigmoid
CRW36	IVS2+24G->T	Rectum
CRW48	IVS2+24G->T	Caecum
CRW63	IVS2+24G->T	Splenic flexure
CRW66	IVS2+24G->T/N393N	Transverse colon
CRW69	IVS2+24G->T	Rectum
CRW71	IVS3+31G->A	Rectum
CRW82	IVS2+24G->T	Sigmoid
CRW83	E16K	Rectum
CRW97	IVS2+24G->T	Descending colon
CRW108	Q123Q/ IVS2+24G->T	Transverse colon
CRW110	Q123Q	Rectum
CRW111	IVS2+24G->T/ IVS6+22C->A	Rectum
CRW115	A389V/ P411S/ A414A	Sigmoid
CRW118	IVS5+16C->T	Ascending colon
CRW119	IVS2+24G->T	Rectum
CRW121	Q123Q	Rectum
CRW129	Q123Q	Rectum
CRW136	IVS2+24G->T	Sigmoid

CHAPTER 5
CONCLUSIONS

Colorectal cancer (CRC) is one of the most common forms of cancer in Western nations, it is however uncommon in sub-Saharan Africa. In South Africa there is an approximate ten-fold lower incidence of CRC in black patients compared to Caucasian patients. This could be due to differences in lifestyles and environment that exist between the various population groups. Underlying molecular events could also account for the difference in susceptibility to colorectal cancer.

To determine the possible involvement of the *STK11* gene, tumours from 104 black and 104 Caucasian patients were screened for mutations using exon-by-exon PCR SSCP analysis. In total 73 tumours from black patients and 25 tumours from Caucasian patients were found to have sequence variants in the *STK11* gene. Five missense, 3 synonymous and 5 intronic mutations were detected in the tumours from black patients, whereas only three missense, 3 synonymous and 5 intronic mutations were detected in the tumours from Caucasian patients. One Caucasian patient's tumour had two missense mutations and may therefore be a compound heterozygote. Whether the synonymous or intronic changes affect the function of STK11 has yet to be determined, but this appears unlikely.

In total 8 novel missense mutations, one of which was germline (in a black patient), were identified in seven tumours (~3,4% 7/208) from 5 black and two Caucasian patients. Thus 4.8% (5/104) of tumours from black patients and 1.9% (2/104) of tumours from Caucasian patients harbour *STK11* missense mutations. Although this is a 2.5-fold difference it is considered not significant (χ^2-test p=0.442). The effect of the missense mutations on the protein function can be evaluated by applying the criteria of Cotton and Scriver (1998), i.e. on the basis of segregation with the disease, which is not possible in this study, conservation across species, population frequency and predicted functional significance. Applying these criteria it would appear that four could possibly have pathogenic significance (E16K, K83N, I88S/R and D207N).

The present study is the sixth to suggest that somatic mutation of the *STK11* gene in sporadic colorectal cancer of Caucasians is an infrequent event (Avizienyte *et al.*, 1998; Forster *et al* 2000; Resta *et al.*, 1998; Launonen *et al.*, 2000; Wang *et al.*, 1998). Contradictory results were obtained by Dong *et al.* (1998) in a study on Korean CRC patients. They reported a high frequency (30.43%, 7/23) of mutations occurring solely in left-sided tumours. In the present study of black patients 4,8% (5/104) of the tumours were mutated, these include 4 somatic (4/104, ~ 3,8%) and one germline (1/104, 0,96%) mutation. This is only the second study of a non-Western population to show somatic mutations in sporadic cases of CRC. Compared to the study of Dong *et al* (1998), the present study shows an eight fold lower incidence of somatic mutations, however as this is a much larger study (104 patients), it may be a truer reflection of somatic mutations in CRCs of a non-Western population. Furthermore with regard to the anatomic site of tumours with somatic missense mutations, the present study found that for black patients 7,69% (2/26) of the left-sided tumours, 2% (1/50) of rectal tumours and 4,54% (1/22) of right-sided tumours harboured mutations. Thus the frequency of missense mutations of left-sided CRC tumours compared to right-sided tumours was not significantly elevated ($\chi 2$ –test , 1df, $p = 0,881$) in the black population.

Peutz-Jeghers Syndrome was diagnosed in 0.96% (1/104) of black patients via a germline mutation. Using this data the frequency of PJS associated with CRC in the general black population can be estimated. According to the National Cancer Registry (Sitas *et al.*, 1998), the lifetime risk for CRC in black South Africans is 0.22% (0.25% male; 0.2% female). Thus the risk of PJS-associated CRC in black South Africans is 0.96% of 0.22% or 1 in 47500 subjects.

In summary, this study represents the first investigation into the role of the *STK11* gene in putative sporadic cases of CRC from both black and Caucasian South African patients. The study shows that mutations of the *STK11* gene are infrequent in sporadic CRCs in the South African Caucasian population, and more frequent in the South African black population. This may be a reflection of the differences in lifestyle and incidence of CRC in the different populations.

APPENDIX A

Previously reported mutations located within the *STK11* gene

Exon	Nucleotide change	Predicted effect	Sample site	Reference
?	Genomic rearrangement	Aberrant/absent protein	PJS[a]	Ylikorkala *et al.*, 1999
1	Homozygous del	absence	Biliary tumour	Su *et al.*, 1999
1	insGC	Fs51X	PJS	Nakagawa *et al.*, 1998
1	del29	Fs152X	PJS	Hemminki *et al.*, 1998
1	c.108C->A	Y36X	Pancreatic tumour	Su *et al.*, 1999
1	c.117insGC	Fs	PJS	Gruber *et al.*, 1998
1	c.125del12	Fs	PJS	Kruse *et al.*, 1999
1	c.144G->T	K48N	Unaffected individual	Resta *et al.*, 1998
1	c.145T->G	Y49D	Malignant melanoma	Rowan *et al.*, 1999
1	c.151del18ins6	delLMGD	PJS	Mehenni *et al.*, 1998
1	c.153insGG	Fs64X	PJS	Westerman *et al.*, 1999a
1	c.153del3	52del	PJS	Olschwang *et al.*, 2001
1	c.157insG	Fs162X	PJS	Trojan *et al.*, 1999
1	c.158delA	Fs63X	PJS	Ylikorkala *et al.*, 1999
1	c.165del10	Fs158X	PJS[a]	Westerman *et al.*, 1999a
1	c.169insG	Fs162X	PJS	Hemminki *et al.*, 1998
1	c.169delG	Fs57	PJS	Olschwang *et al.*, 2001
1	c.169G->T	E57X	PJS	Hemminki *et al.*, 1998
1	c.180C->G	Y60X	PJS	Hemminki *et al.*, 1998; Olschwang *et al.*, 2001
1	c.180C->A	Y60X	PJS	Wang *et al.*, 1999
1	c.197insT	Fs162X	PJS	Westerman *et al.*, 1999b
1	c.200T->C	L67P	PJS	Hemminki *et al.*, 1998
1	c.200T->G	L67R	PJS	Olschwang *et al.*,2001
1	c.208G->T	E70X	PJS	Hemminki *et al.*, 1998
1	c.211delG	Fs50X	PJS	Miyaki *et al.*, 2000
1	c.250A->T	K84X	PJS	Hemminki *et al.*, 1998; Gruber *et al.*, 1998; Ylikorkala *et al.*, 1999

1	c.256C->T	R86X	PJS	Ylikorkala *et al.*, 1999
1-9	Homozygous del	Absence of product	Pancreatic tumour	Su *et al.*, 1999
2	c.290ins131	Fs	PJS	Abed *et al.*, 2001
2	c.291G->T	Del ex 2	PJS	Miyaki *et al.*, 2000
2	c.297C->T	Q100X	PJS	Westerman *et al.*, 1999a
2	c.321del6	107H-K-N to H	PJS	Wang *et al.*, 1999
2	c.323A->G	K108R	PJS	Wang *et al.*, 1999
2	c.350ins4	Fs	PJS	Olschwang *et al.*,2001
2	c.354C->A	Y118X	PJS	Olschwang *et al.*,2001
2	c.357C->T	N119N	Malignant melanoma	Guldberg *et al.*, 1999
2-3	c.192del174	Truncated (in frame product of 376 amino acids	PJS	Hemminki *et al.*, 1998
3	c.396C->A	C132X	PJS	Olschwang *et al.*, 2001
3	c.403G->C	G135R	Malignant melanoma	Rowan *et al.*, 1999
3	c.406del9	Q-E-M-L to L	PJS	Ylikorkala *et al.*, 1999
3	c.407T->G	M136R	PJS	Olschwang *et al.*, 2001
3	c.418delC	Fs160X	PJS	Nakagawa *et al.*, 1998; Gruber *et al.*, 1998
3	c.454C->T	Q152X	PJS	Ylikorkala *et al.*, 1999
3	c.463insG	Fs162X or altered splicing	PJS	Westerman *et al.*, 1999a
4	c.470T->C	F157S	PJS	Ylikorkala *et al.*, 1999
4	c.474del7	Fs	PJS	Kruse *et al.*, 1999
4	c.484G->A	D162N	PJS	Westerman *et al.*, 1999a
4	c.488G->A	G163D	PJS, testicular cancer	Avizienyte *et al.*, 1998; Westerman *et al.*, 1999a
4	c.490C->A	L164M	PJS	Westerman *et al.*, 1999a
4	c.508C->T	E170X	Malignant melanoma	Guldberg *et al.*, 1999; Abed *et al.*, 2001
4	c.511G->A	G171S	Transverse colon Rectum	Dong *et al.*, 1998
4	c.516insT	Fs	PJS	Kruse *et al.*, 1999
4	c.526G->A	D176N	PJS	Mehenni *et al.*, 1998
4	c.528delC	Fs286X	PJS	Wang *et al.*, 1999
4	c.536del6	delKD	PJS	Resta *et al.*, 1998

4	c.541A->T	N181Y	PJS	Ylikorkala *et al.*, 1999
4	c.545T->C	L182P	PJS	Olschwang *et al.*, 2001
4	c.567ins25	Fs264X	PJS	Miyaki *et al.*, 2000
4	c.574insA	Fs 73 new amino acids after K191	PJS	Mehenni *et al.*, 1998
4	c.580G->A	D194N	PJS	Westerman *et al.*, 1999a
4	c.580G->T	D194Y	Malignant melanoma	Guldberg *et al.*, 1999
4	c.581A->T	D194V	Lung adenocarcinoma	Avizienyte *et al.*, 1999
4	c.595G->A	E199K	Rectum	Dong *et al.*, 1998
4	c.597insIVS4	isoform	PJS	Abed *et al.*, 2001
4-7	Inv/del	Codons 156-307 deleted	PJS	Jenne *et al.*, 1998
5	c.604del20	Fs258X	PJS	Ylikorkala *et al.*, 1999
5	c.622G->A	D208N	Sigmoid	Dong *et al.*, 1998
5	c.636del5	Fs	PJS	Olschwang *et al.*, 2001
5	c.644G->A	G215D	Sigmoid	Dong *et al.*, 1998
5	c.644insG	Fs264X	PJS	Miyaki *et al.*, 2000
5	c.650insC	Fs264X	PJS	Miyaki *et al.*, 2000
5	c.650delC	Fs	Pancreatic tumour	Su *et al.*, 1999
5	c.658C->T	Q220X	PJS	Gruber *et al.*, 1998; Kruse *et al.*, 1999; Ylikorkala *et al.*, 1999
5	c.666delC	Fs286X	PJS	Miyaki *et al.*, 2000; Olschwang *et al.*, 2001
5	c.694T->C	S232P	PJS	Yoon *et al.*, 2000
5	c.716delGGTC	Fs285X	PJS	Jenne *et al.*, 1998
5	c.724G->T	G242W	PJS	Olschwang *et al.*, 2001
5	c.725G->A	G242V	PJS	Olschwang *et al.*, 2001
5	c.733delC	Fs286X	PJS	Nakagawa *et al.*, 1998
6	c.738C->G	Y246X	PJS	Nakagawa *et al.*, 1998
6	c.744delC	Fs	PJS	Olschwang *et al.*,2001
6	c.747del14	Fs260X	PJS	Ylikorkala *et al.*, 1999
6	c.751G->A	G251S	PJS	Resta *et al.*, 1998
6	c.753delT	Fs	PJS	Olschwang *et al.*, 2001
6	c.766G->T	E256X	PJS	Connolly *et al.*, 2000
6	c.767A->C	E256A	PJS	Yoon *et al.*, 2000
6	c.773del8	Fs262X	PJS	Ylikorkala *et al.*, 1999

6	c.777C->A	Y253X	PJS	Jenne *et al.*, 1998
6	c.787delTTGT	Fs286X	PJS	Miyaki *et al.*, 2000; Resta *et al.*, 1998
6	c.789del4	Fs	PJS	Olschwang *et al.*, 2001
6	c.790del4	Fs285X	PJS	Ylikorkala *et al.*, 1999
6	c.796delAAC	Y-N-I to Y-I	PJS	Nakagawa *et al.*, 1998
6	c.831del2	Fs283X	PJS	Hemminki *et al.*, 1998
6	c.841delC	Fs	PJS	Olschwang *et al.*, 2001
6	c.842C->T	P281L	Rectum; Ovarian carcinoma	Dong *et al.*, 1998; Nishioka *et al.*, 1999
6	c.842delC	Fs283X	PJS	Wang *et al.*, 1999; Ylikorkala *et al.*, 1999
6	c.842delC	Fs286X	Sigmoid; Gastrointestinal cancer	Dong *et al.*, 1998; Nakagawa *et al.*, 1998
6	c.842insC	Fs284X	PJS	Nakagawa *et al.*, 1998; Gruber *et al.*, 1998
6	c.843delG	Fs286X	PJS	Jenne *et al.*, 1998
6	c.844insC	Fs284X	Prostate, colon, pancreatic	Boardman *et al.*, 2000
7	c.890G->A	R297K	PJS	Westerman *et al.*, 1999a
7	c.891C->T	A297S	Colon, endometrial	Boardman *et al.*, 2000
7	c.903delG	Fs 33 new amino acids after R301	PJS	Mehenni *et al.*, 1998
7	c.907del9	I-R-Q-H to N	PJS	Hemminki *et al.*, 1998; Olschwang *et al.*, 2001
7	c.909delC	Fs316X	PJS	Westerman *et al.*, 1999a
7	c.910C->T	R304W	PJS	Resta *et al.*, 1998
7	c.914delA	Fs335X	PJS	Ylikorkala *et al.*, 1999
7	c.916C->T	H272Y	Thyroid, cecal	Boardman *et al.*, 2000
8	del188	Fs404X	PJS	Hemminki *et al.*, 1998
8	c.923G->A	W308X	PJS	Ylikorkala *et al.*, 1999
8	c.924G->T	W308C	PJS	Mehenni *et al.*, 1998
8	c.936delA	Fs	Pancreatic tumour	Su *et al.*, 1999
8	c.941C->A	P314H	Colon cancer tumour	Resta *et al.*, 1998
8	c.957delA958G->T	Fs335X	Cervical adenocarcinoma	Avizienyte *et al.*, 1999
8	c.957del2insT	Fs	PJS	Olschwang *et al.*, 2001

8	c.971C->T	P324L	PJS	Yoon *et al.*, 2000
8	c.972del5	Fs324X	PJS	Boardman *et al.*, 2000
8	c.984C->T	T328T	Malignant melanoma	Guldberg *et al.*, 1999
8	c.989del9	Fs	PJS	Gruber *et al.*, 1998
8	c.989insC	Fs359X	PJS	Westerman *et al.*, 1999a
8	c.1024insG	Fs342X	PJS	Yoon *et al.*, 2000
8	c.1062C->G	F354L	Rectum; Left sided colorectal carcinoma	Dong *et al.*, 1998; Launonen *et al.*, 2000
8	c.1100C->T	T367M	Sigmoid	Dong *et al.*, 1998
9	c.1246A->T	K416X	PJS	Wang *et al.*, 1999

a) PJS indicates that blood samples of Peutz-Jeghers syndrome patient was used for the mutation screening.
b) Nucleotide positions given refer to STK11 cDNA sequence (Genbank U63333)

Mutations identified in the introns of the *STK11* gene

Exon	Nucleotide change	Predicted effect	Sample site	Reference
1	IVS1-2A->G	Exon skipping	PJS	Westerman *et al.*, 1999a; Abed *et al.*,, 2001
1	IVS1-1G->C	Exon skipping	PJS	Westerman *et al.*, 1999a
1	IVS1-1C->G	Aberrant splicing?	PJS	Nakagawa *et al.*, 1998
1	IVS1+1G->A	Aberrant splicing	PJS	Olschwang *et al.*, 2001
1	IVS1-1A->C	Aberrant splicing?	PJS	Yoon *et al.*, 2000
1	IVS1+36G->T	Silent	Pancreatic ; Colorectal tumours	Su *et al.*, 1999; Avizienyte *et al.*, 1998; Olschwang *et al.*, 2001
2	taatgaat> gtaatgc	Silent	colon cancer	Wang *et al.*, 1998
2	IVS2+24G->T	Silent	Colorectal; Pancreatic tumours; PJS	Avizienyte *et al.*, 1998; Resta *et al.*, 1998; Su *et al.*, 1999; Olschwang *et al.*, 2001
2	IVS2-49G->A	Silent	Colorectal carcinomas	Avizienyte *et al.*, 1998; Resta *et al.*, 1998
3	IVS3-51T->C	Silent	Colorectal; Lung tumour	Avizienyte *et al.*, 1998; Avizienyte *et al.*, 1999; Olschwang *et al.*, 2001
3	IVS3-1G->A	Aberrant splicing/ exon 4 skipping	PJS	Jenne *et al.*, 1998
3	IVS3-2A->G	Aberrant splicing	PJS	Olschwang *et al.*, 2001
3	IVS3+16insGGG	Silent	Pancreatic tumour	Su *et al.*, 1999
3	IVS3+47insGGG	Silent	Pancreatic tumour	Su *et al.*, 1999

3	IVS3+49ins7	Silent	PJS	Olschwang *et al.*, 2001
3	IVS3+49G->C	Silent	Malignant melanoma	Rowan *et al.*, 1999
3	IVS3+50G->C	Silent	Malignant melanoma	Rowan *et al.*, 1999
4	IVS4-2A->T	Aberrant splicing?	PJS	Olschwang *et al.*, 2001
4	IVS4+31del14	Intron 4 retention	PJS	Abed *et al.*, 2001
4	delG	Silent	Colorectal carcinomas	Avizienyte *et al.*, 1998
5	IVS5-1G->A	Aberrant splicing?	PJS	Olschwang *et al.*, 2001
5	IVS5-6A->G	Silent	Pancreatic tumour	Su *et al.*, 1999
5	IVS5-51C->T	Silent	Pancreatic tumour; PJS	Resta *et al.*, 1998; Su *et al.*, 1999
5	IVS5+1G->A	Aberrant splicing	PJS	Olschwang *et al.*, 2001
5	IVS5+2insT	Aberrant splicing?	PJS	Nakagawa *et al.*, 1998
5	IVS5-52G->A	Silent	PJS	Resta *et al.*, 1998
6	IVS6-5del36	Fs or altered splicing	PJS	Westerman *et al.*, 1999
6	IVS6-8C->T	Silent	PJS	Olschwang *et al.*, 2001
6	IVS6+3G->C	Altered splicing	PJS	Boardman *et al.*, 2000
7	IVS6del52	Abnormal splicing/ truncated protein	PJS	Mehenni *et al.*, 1998
7	IVS7+7G->C	Silent	Colorectal; Pancreatic tumours	Avizienyte *et al.*, 1998; Su *et al.*, 1999; Olschwang *et al.*, 2001
7	IVS7+27C->T	Silent	PJS	Resta *et al.*, 1998
7	IVS7-1G->C	Silent	PJS	Ylikorkala *et al.*, 1999
8	IVS8+20C->T	Silent	PJS	Resta *et al.*, 1998
9	3'UTR+129C->T	Silent	PJS	Olschwang *et al.*, 2001
9	3'UTR+259G->T	Silent	PJS	Olschwang *et al.*, 2001
9	3'UTR+506G->A	Silent	PJS	Olschwang *et al.*, 2001
9	3'UTR+633C->T	Silent	PJS	Olschwang *et al.*, 2001
9	3'UTR+677T->A	Silent	PJS	Olschwang *et al.*, 2001

APPENDIX B

Black colorectal cancer patients

	Patient	Age	Tumour stage	Cancer site	Left or right
1	CRC 1	53m	Duke C	Caecum	R
2	CRC 2	48f		Rectum	L
3	CRC 3	72f	Duke B	Left colon	L
4	CRC 7	73m	Duke C	Large Bowel	U
5	CRC 8	51m		Rectum + polyp	L
6	CRC 9	Um		Rectum	L
7	CRC 10	66m		Descending sigmoid	L
8	CRC 11	72m	Duke B	Colon	U
9	CRC 12	51m	Duke B	Caecum	R
10	CRC 16	74m	Duke B2	Ascending colon	R
11	CRC 17	56m		Caecum	R
12	CRC 23	39f		Caecum	R
13	CRC 24	51f	Duke C	Sigmoid	L
14	CRC 25	32f	Duke C	Caecum	R
15	CRC 27	72m	Duke C	Sigmoid	L
16	CRC 28	43m		Caecum	R
17	CRC 29	52f		Rectum	L
18	CRC 30	16f	Duke C	Hepatic flexure	R
19	CRC 31	74f	Duke C	Left colon	L
20	CRC 32	68f	Duke B	Left colon	L
21	CRC 34	66m	Duke C	Rectum	L
22	CRC 36	27f	Duke C	Ascending colon	R
23	CRC 37	46f	Duke C	Sigmoid	L
24	CRC 38	72m		Colon	U
25	CRC 39	22m		Caecum	R
26	CRC 40	62m	Duke C	Descending	L
27	CRC 41	51f	Duke C	Transverse colon	R
28	CRC 43	76f		Transverse colon	R
29	CRC 45	31m	Duke B	Sigmoid	L
30	CRC 46	60f	Duke B	Sigmoid	L
31	CRC 47	40f	Duke B	Colon	U
32	CRC 52	52m	Gr II, Duke C	Ascending colon	R
33	CRC 53	63f	Gr II, Duke D	Caecum	R
34	CRC 59	65f	Gr II, Duke D	Sigmoid	L
35	CRC 62	28f	Gr II, Duke B2	Sigmoid	L
36	CRC 63	19m	Duke C	Rectum	L
37	CRC 65	70f	Gr II, Duke A	Rectum	L
38	CRC 66	63f	Gr II	Rectum	L
39	CRC 67	30f	Gr II	Rectum	L
40	CRC 68	70f	Gr II, Duke C	Rectum	L
41	CRC 69	50m	Gr II, Duke C	Rectum	L
42	CRC 70	67m		Colon	U
43	CRC 71	67m	Gr II, Duke B2	Transverse	R
44	CRC 73	29f	Gr III, Duke C	Transverse	R
45	CRC 75	62f	Gr II, Duke B2	Splenic flexure	L
46	CRC 76	66m	Gr II, Duke C	Left colon	L
47	CRC 78	24m		Descending colon	L
48	CRC 79	44f	Gr II, Duke A	Descending colon	L
49	CRC 80	44m	Gr II, Duke B2	Splenic flexure	L
50	CRC 82	60f	Gr II	Sigmoid	L

51	CRC 83	68f	Gr III, Duke C	Sigmoid	L
52	CRC 84	37f	Gr II	Sigmoid	L
53	CRC 86	64m	Gr II, Duke B2	Rectum	L
54	CRC 87	44m	Gr II, Duke C	Rectum	L
55	CRC 88	56m	Gr II, Duke C	Rectum	L
56	CRC 90	82f	Gr II, Duke C	Rectum	L
57	CRC 91	41m	Gr II	Rectum	L
58	CRC 93	80m	Gr II, Duke B	Rectum	L
59	CRC 95	49m	Gr II	Rectum	L
60	CRC 96	52f	Gr II, Duke B2	Rectum	L
61	CRC 97	50f	Gr II	Rectum	L
62	CRC 98	32f	Gr II	Rectum	L
63	CRC 99	65m	Gr II, Duke C	Rectum	L
64	CRC 100	64m	Gr II	Rectum	L
65	CRC 123	42f	Gr II	Rectum	L
66	CRC 124	72m	Gr II, Duke C	Descending colon	L
67	CRC 125	50m	Gr II, Duke D	Ascending colon	R
68	CRC 126	Uf	Gr II, Duke B2	Descending colon	L
69	CRC 127	53f	Gr II, Duke D	Sigmoid	L
70	CRC 128	34m	Gr II, Duke B	Caecum	R
71	CRC 144	46f	Gr II, Duke C	Sigmoid	L
72	CRC 146	49m	Gr II, Duke B1	Rectum	L
73	CRC 147	48f	Gr II, Duke C	Rectum	L
74	CRC 149	52m	Gr II, Duke B2	Hepatic flexure	R
75	CRC 150	39m	Gr II	Rectum	L
76	CRC 151	64f	Gr II, Duke B2	Caecum	R
77	CRC 152	83f	Gr II, Duke B1	Caecum	R
78	CRC 167	54		Rectum	L
79	CRC 168	55		Ascending colon	R
80	CRC 170	21f		Rectum	L
81	CRC 171	58f		Rectum	L
82	CRC 172	46m		Rectum	L
83	CRC 173	64m		Rectum	L
84	CRC 174	74f		Rectum	L
85	CRC 175	58m		Rectum	L
86	CRC 176	79f		Rectum	L
87	CRC 177	83f		Rectum	L
88	CRC 178	21f		Rectum	L
89	CRC 179	67f		Rectum	L
90	CRC 180	48f		Distal rectum	L
91	CRC 181	50m		Rectum	L
92	CRC 182	65m		Rectum	L
93	CRC 183	59f		Hipoplastic polyp Rectum?	L
94	CRC 184	21m		Rectum	L
95	CRC 185	51f		Rectum	L
96	CRC 186	67m		Colon	U
97	CRC 187	46m		Rectum	L
98	CRC 190	53f		Sigmoid	L
99	CRC 191	22f		Rectum	L
100	CRC 192	41m		Rectum	L
101	CRC 193	72f		Rectum	L
102	CRC 194	49f		Rectum	L
103	CRC 195	31		Rectum	L
104	CRC 196	45m		Rectum	L

Caucasian colorectal cancer patients

	Patient	Age	Tumour stage	Cancer site	Left or right
1	CRW 2	75f	Duke C	Colon	U
2	CRW 3	19m		Rectum-anal	L
3	CRW 16	34f	Gr II, Duke D	Colon	U
4	CRW 17	38f		Sigmoid	L
5	CRW 18	45m	Gr II	Rectum	L
6	CRW 19	52m	Gr II, Duke D	Sigmoid	L
7	CRW 21	52m	Gr II, Duke C	Rectum	L
8	CRW 22	52m	Gr II, Duke B2	Sigmoid	L
9	CRW 23	23f	Gr II, Duke C	Splenic Flexure	L
10	CRW 24	44m	Gr II, Duke C	Recto-sigmoid	L
11	CRW 25	72f	Gr II, Duke B	Transverse	R
12	CRW 26	57f	Gr II, Duke D	Left colon	L
13	CRW 27	82f	Gr II, Duke C	Rectum	L
14	CRW 28	71f	Gr II, Duke C	Colon	U
15	CRW 29	69m	Ge II, Duke B2	Sigmoid	L
16	CRW 34	51m	Gr II, Duke B2	Caecum	R
17	CRW 35	43m	Gr II	Rectum	L
18	CRW 36	55f	Gr II, Duke C	Rectum	L
19	CRW 37	60m	Gr II, Duke C	Splenic flexure	L
20	CRW 38	74f	Gr II, Duke C	Right colon	R
21	CRW 39	76m	Gr II, Duke B2	Rectum	L
22	CRW 40	80m	Gr III, Duke B	Sigmoid	L
23	CRW 41	66f	Gr II, Duke C	Sigmoid	L
24	CRW 43	68f	Gr II, Duke B1	Rectum	L
25	CRW 44	73m	Gr II	Recto-sigmoid	L
26	CRW 45	80f	Gr II, Duke B1	Rectum	L
27	CRW 46	37f	Gr II, Duke D	Descending colon	L
28	CRW 48	73f	Gr II, Duke B2	Caecum	R
29	CRW 50	70f	Gr II, Duke D	Sigmoid	L
30	CRW 51	56m	Duke B	Rectum	L
31	CRW 52	74f	Gr II, Duke C	Rectum	L
32	CRW 53	62m	Gr III, Duke C	Rectum and anus	L
33	CRW 54	79m	Gr II, Duke C	Rectum	L
34	CRW 55	77f	Gr II, Duke C	Sigmoid	L
35	CRW 59	64f	Gr II	Rectum	L
36	CRW 61	68m	Gr II, Duke C	Rectum	L
37	CRW 62	76m	Gr II, Duke B	Transverse colon	R
38	CRW 63	58f	Gr II, Duke B2	Splenic flexure	L
39	CRW 64	75f	Gr II, Duke B2	Sigmoid	L
40	CRW 65	51f	Gr II, Duke D	Sigmoid	L
41	CRW 66	79f	Gr II, Duke C	Transverse colon	R
42	CRW 67	73m	Gr II, Duke C	Sigmoid	L
43	CRW 68	53f	Gr II, Duke B1	Sigmoid	L
44	CRW 69	63m	Gr II, Duke D	Rectum	L
45	CRW 70	77f	Gr II	Rectum	L
46	CRW 71	88f	Gr II, Duke B1	Rectum	L
47	CRW 73	64f	Gr II, Duke C	Transverse colon	R
48	CRW 74	Um	Gr II, Duke C	Rectum	L
49	CRW 75	62f	Gr II	Transverse colon	R
50	CRW 76	77m	Gr II, Duke B1	Caecum	R
51	CRW 77	58f	Gr II, Duke C	Rectum	L
52	CRW 79	75m		Sigmoid	L
53	CRW 80	80f	Gr II	Rectum	L

54	CRW 81	74f	Gr II	Caecum	R
55	CRW 82	41f	Gr II	Sigmoid	L
56	CRW 83	68m	Gr II	Rectum	L
57	CRW 84	74f	Gr II	Sigmoid	L
58	CRW 85	79f	Gr III, Duke C	Colon	U
59	CRW 86	77f	Gr II, Duke B2	Sigmoid	L
60	CRW 87	68f	Gr II, Duke B2	Splenic Flexure	L
61	CRW 88	80f	Duke A	Rectum	L
62	CRW 89	75f	Gr II, Duke B2	Rectum	L
63	CRW 90	60f	Gr II, Duke B	Rectum	L
64	CRW 91	61f		Caecum	R
65	CRW 93	67m	Gr II, Duke B2	Rectum	L
66	CRW 94	61f	Gr II, Duke C	Caecum	R
67	CRW 95	72m	Gr II, Duke D	Caecum	R
68	CRW 97	82f	Duke B	Descending + rectum	L
69	CRW 98	76f	Gr II, Duke C	Caecum	R
70	CRW 100	64m	Gr II, Duke A	Left colon	L
71	CRW 101	72f	Gr III, Duke D	Sigmoid	L
72	CRW 102	70f		Rectum	L
73	CRW 103	68f		Sigmoid	L
74	CRW 104	70f	Gr II, Duke C	Sigmoid	L
75	CRW 105	36f		Colon	U
76	CRW 106	56m	Gr II, Duke C	Caecum	R
77	CRW 107	72m	Gr II, Duke B2	Transverse colon	R
78	CRW 108	67f	GR II, Duke B1	Transverse colon	R
79	CRW 109	75f	Gr II, Duke B	Splenic flexure	L
80	CRW 110	74f	Gr II, Duke C	Rectum	L
81	CRW 111	43f	Gr II	Rectum	L
82	CRW 112	78f	Gr II, Duke C2	Hepatic flexure	R
83	CRW 113	45f	Gr II	Colon	U
84	CRW 114	68f	Gr II	Recto-sigmoid	L
85	CRW 115	80m	Gr II, Duke B1	Sigmoid	L
86	CRW 116	68f	Gr II, Duke B2	Sigmoid	L
87	CRW 117	65m	Gr II	Caecum	R
88	CRW 118	67f	Gr II, Duke C	Ascending colon	R
89	CRW 119	57f	Gr II	Rectum	L
90	CRW 120	83f	Gr II, Duke B1	Rectum	L
91	CRW 121	59m	Gr II, Duke C	Rectum	L
92	CRW 122	77f	Gr II, Duke C	Rectum	L
93	CRW 124	71m		Rectum	L
94	CRW 125	67f	Gr II, Duke C	Rectum	L
95	CRW 126	69f	Gr II, Duke B1	Rectum	L
96	CRW 127	49m		Rectum	L
97	CRW 128	76f	Gr II, Duke C	Rectum	L
98	CRW 129	59f	Gr II, Duke B2	Rectum	L
99	CRW 130	58f	Gr II, Duke B	Rectum	L
100	CRW 132	73f	Gr II	Rectum	L
101	CRW 133	66f	Gr II, Duke C	Rectum	L
102	CRW 134	68f	Gr II, Duke B1	Descending colon	L
103	CRW 135	69f	Gr II, Duke C	Ascending colon	R
104	CRW 136	84m	Gr II, Duke B2	Sigmoid	L

APPENDIX C

Tumour stage classification in colorectal cancer

In most cancer types tumour size is the prognostic parameter. In colorectal cancer however tumour size doesn't always correlate with prognosis. Instead the extend of tumour infiltration and the presence of lymph node metastasis are the main factors that determine the prognosis. Taking this into consideration the first attempt to stage colorectal carcinoma was made in 1926. This staging system consisted of three stages named A,B and C. Cuthbert Dukes later modified this staging system, which is now recognised as the basis for staging of colorectal cancer. An addition was made to Dukes' classification to form a fourth stage, yet again this stage was further subdivided 1980. The complete staging therefore contains stages A, B, C1, C2, D1 and D2.

Dukes' classification

According to Dukes' classification stage A is infiltration of the tumour into the muscularis propria, stage B invasion occurs into the perirectal tissue. Stages C and D are both subdivided into C1, C2, D1 and D2. Stage C1 involves adjacent lymph node metastasis and C2 apical lymph node metastasis. D1 refers to "fixed" cancers that have invaded adjacent organs and D2 is associated with the presence of distant metastasis.

Broders' classification

This classification system grades the degree of histological differentiation of the tumour. Where grade I is well differentiated, grade II moderately diffirentiated, grade III poorly differentiated tumours and grade IV totally anaplastic tumours.

REFERENCES

Abed AA, Günther K, Kraus C, Hohenberger W, Ballhausen W. Mutations at the RNA level of the STK11/LKB1 gene in Peutz-Jeghers syndrome reveals complex splicing abnormalities and a novel mRNA isoform (STK11 c.597^598insIVS4). *Human Mutation* 2001; 18:397-410.

Antonarakis SE and the Nomenclature Working Group. Recommendations for a nomenclatur system for human gene mutations. *Human Mutation* 1998; 11:1-3

Antonarakis SE, Krawczak M, Cooper DN. The Nature and Mechanisms of Human Gene Mutation. In The Metabolic and Molecular Basis of Inherited Disease. Editors; Scriver CR, Beaudet AL, Valle D, Sly WS, Vogelstein B, Childs B *et al.* McGrawHill 2002; http://www.MMBIDonline.com.

Ausubel F.M, R. Brent, R.E. Kingston, D.D. Moore, J.G. Siedman, J.A. Smith, K. Struhl. 1993 Current protocols in molecular biology. John Wiley and sons inc. USA.

Avizienyte E, Roth S, Loukola A, Hemminki A, Lothe RA, Stenwig AE, *et al.* Somatic mutations in *LKB1* are rare in sporadic colorectal and testicular tumors. *Cancer Research* 1998; 58:2087-2090.

Avizienyte E, Loukola A, Roth S, Hemminki A, Tarkkanen M, Salovaara, *et al.* *LKB1* somatic mutations in sporadic tumours. *American Journal of Pathology* 1999; 154(3):677-681.

Bishop DT, Hall NR. The genetics of colorectal cancer. *European Journal of Cancer* 1994; 30(13):1946-1956.

Bishop JM. Molecular themes in oncogenesis. *Cell* 1991; 64:235-248

Boardman LA, Couch FJ, Burgart LJ, Schwartz D, Berry R, McDonnell SK, *et al.* Genetic heterogeneity in Peutz-Jeghers syndrome. *Human Mutation* 2000; 16:23-30

Broders AC. The grading of carcinoma. *Minesota Medicine* 1925; 8:726

Campbell MK. Biochemistry. Saunders College publishing, Orlando, Florida, 1995.

Cartegni L, Chew SL, Krainer A. Listening to silence and understanding nonsense exonic mutations that affect splicing. *Nature Reviews Genetics* 2002; 3:285-298.

Churchman M, Dowling B, Tomlinson IP. Identification of a novel mRNA species of the LKB1/STK11 Peutz-Jeghers serine/threonine kinase. *DNA Sequence* 1999; 10;255-261.

Collins SP, Reoma JL, Gamm DM, Uhler MD. LKB1, a novel serine/threonine protein kinase and potential tumour suppressor, is phosphorylated by cAMP-dependant protein kinase (PKA) and prenylated *in vivo*. *Biochemical Journal* 2000; 345:673-680.

Connolly DC, Katabuchi H, Cliby WA, Cho KR. Somatic mutations in the STK11/LKB1 gene are uncommon in rare gynecological tumour types associated with Peutz-Jegher's syndrome. *American Journal of Pathology* 2000; 156(1):339-345.

Cooper DN, Krawczak M. Human gene mutation. BIOS Scientific Publishers. Oxford, UK, 1993.

Cotton RG, Scriver CR. Proof of "disease causing" mutation. *Human Mutation* 1998; 12:1-3

Dipple KM, McCabe ERB. Phenotyps of patients with "Simple" mendelian disorders are complex traits: thresholds, modifiers, and systems dynamics. *American Journal of Human Genetics* 2000; 66:1729-1735.

Dong SM, Kim KM, Kim SY, Shin MS, Na EY, Lee SH, *et al.* Frequent somatic mutations in serine threonine kinase 11/ Peutz-Jeghers syndrome gene in left-sided colon cancer. *Cancer Research* 1998; 58:3787-3790.

Eeles RA, Stamps AC. Polymerase Chain reaction (PCR): The technique and its applications. R.G. Landes Company. 1993.

Forster LF, Defres S, Goudie DR, Baty DU, Carey FA. An investigation of the Peutz-Jeghers gene (LKB1) in sporadic breast and colon cancers. *Journal of Clinical Pathology* 2000; 53(10):791-793.

Gardiello FM, Welsh SB, Hamilton SR, Offerhaus GJA, Gittelsohn AM, Booker SV, *et al.* Increased risk of cancer in the Peutz-Jeghers syndrome. *The New England journal of Medicine* 1987; 316(24):1511-1514.

Gruber SB, Entius MM, Petersen GM, Laken SJ, Longo PA, Boyer R, *et al.* Pathogenesis of adenocarcinoma in Peutz-Jeghers syndrome. *Cancer Research* 1998; 58:5267-5270.

Guldberg P, thor Straten P, Ahrenkiel V, Seremet T, Kirklin AF, Zeuthen J. Somatic mutation of the Peutz-Jeghers syndrome gene, *LKB1/STK11*, in malignant melanoma. *Oncogene* 1999; 18:1777-1780.

Hanks SA, Hunter T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *The FASEB Journal* 1995; 9:576-596.

Hemminki A, Markie D, Tomlinson I, Avizienyte E, Roth S, Loukola A, *et al.* A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. *Nature* 1998; 391(8):184-187.

Hemminki A. The molecular basis and clinical aspects of Peutz-Jeghers syndrome. *Cellular and Molecular Life Sciences* 1999; 55:735-750.

Hodgson SV, Maher ER. A practical guide to human cancer genetics. Cambridge University press. Cambridge, 1993.

Jenne DE, Reimann H, Nezu J, Friedel W, Loff S, Jeschke R, *et al.* Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nature Genetics* 1998; 18:38-43.

Karuman P, Gozani O, Ozde RD *et al.* The Peutz-Jeghers gene product LKB1 is a mediator of p53-dependent cell death. *Molecular Cell* 2001; 7:1307-1319.

Kinzler KW, Vogelstein B. Landscaping the cancer terrain. *Science* 1998; 280:1036-1037.

Kruse R, Uhlhaas S, Lamberti C, Keller KM, Jackish C, Steinhard J, *et al.* Peutz-Jeghers syndrome: Four novel inactivating germline mutations in the *STK11* gene. *Human Mutation* 1999; 13(3):257-258.

Latchman D. Gene regulation a eukaryotic perspective, Chapman and Hall. London, 1997.

Liu H-X, Cartegni L, Zhang MQ, Krainer AR. A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA 1* and other genes. *Nature Genetics* 2001; 27:55-58.

Luuko K, Ylikorkala A, Tiainen M, Mäkelä TP. Expression of LKB1 and PTEN tumor suppressor genes during mouse embryonic development. *Mechanisms of development*; 83:187-190.

Launonen V, Avizienyte E, Loukola A, Laiho P, Salovaara R, Jarvinen H, *et al.* No evidence of Peutz-Jeghers syndrome gene *LKB1* involvement in left-sided colorectal carcinomas. *Cancer Research* 2000; 60:546-548.

Macleod K. Tumour suppressor genes. *Current Opinion in Genetics and Development* 2000; 10:81-93.

Mader S. Biology of genetics and inheritance. Wm. C. Brown Publishers. Dubuke, Iowa, 1993.

Maquat LE. The power of point mutations. *Nature Genetics* 2001; 27: 5-6.

Marignani PA, Kanai F, Carpenter CL. LKB1 associates with brg1 and is necessary for brg1-induced growth arrest. *Journal of Biological Chemistry* 2001; 276:32415-32418.

Marth G, Yeh R, Minton M, Donaldson R, Li Q, Duan S *et al.* Single nucleotide polymorphisms in the public domain: how useful are they? *Nature Genetics* 2001; 27:371-372.

McGarrity TJ, Kulin HE, Zaino RJ. Peutz-Jeghers syndrome. *The American Journal of Gastroenterology* 2000; 95(3):596-604

McGuffin LJ, Bryson K, Jones DT. PSIPRED protein prediction programme. UCL Bioinformatics Unit, Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK. http://bioinf.cs.ucl.ac.uk/psiform.

Mehenni H, Blouin J-L, Radhakrishna U, Bhardwaj SS, Bhardwaj K, Dixit VB. *et al.* Peutz-Jeghers syndrome: confirmation of linkage to chromosome 19p13.3 and identification of a potential second locus, on 19q13.4. *American Journal of Human Genetics* 1997; 61:1327-1334.

Miyaki M, Iijima T, Hosono K, Ishii R, Yasuno M, Mori T. *et.al.* Somatic mutations of *LKB1* and *β-Catenin* genes in gastrointestinal polyps from patients with Peutz-Jeghers syndrome. *Cancer Research* 2000; 60:6311-6313.

Nakagawa H, Koyama K, Miyoshi Y, Ando H, Baba S, Watatani M, *et al.* Nine novel germline mutations of *STK11* in ten families with Peutz-Jeghers syndrome. *Human Genetics* 1998; 103:168-172.

Nakagawa H, Koyama K, Nakamori S, Kameyama M, Imaoka S, Monden M, *et al.* Frameshift mutation of the *STK11* gene in a sporadic gastrointestinal cancer with microsatellite instability. *Japanese Journal of Cancer Research* 1999; 90:6633-6637.

Nathanson KL, Weber BL. 'Other' breast cancer susceptibility genes: searching for more holy grail. *Human Molecular Genetics* 2001; 10(7):715-720.

Nezu J-I, Oku A, Shimane M. Loss of cytoplasmic retention ability of mutant LKB1 found in Peutz-Jeghers syndrome. *Biochemical and Biophysical Research Communications* 1999; 261:750-755.

Nishioka Y, Kobayashi K, Sagae S, Sugimura M, Ishioka S, Nagata M, *et al.* Mutational analysis of *STK11* gene in ovarian carcinomas. *Japanese Journal of Cancer Research* 1999; 90:6629-6632.

Oxford concise colour medical dictionary. 2nd ed. Oxford New York:Oxford University Press, 1998.

Olschwang S, Boisson C, Thomas G. Peutz-Jeghers families unlinked to STK11/LKB1 gene mutations are highly predisposed to primitive biliary adenocarcinoma. *Journal of Medical Genetics* 2001; 38:356-360

Rebbeck TR, Kantoff PW, Krithivas K, Neuhausen S, Blackwood MA, Godwin AK, *et al.* Modifications of BRCA1-associated breast cancer risk by polymorphic androgen-receptor CAG repeat. *American Journal of Human Genetics* 1999; 64:1371-1377.

Resta N, Simone C, Mareni C, Montera M, Gentile M, Susca F, *et al. STK11* mutations in Peutz-Jeghers syndrome and sporadic colon cancer. *Cancer Research* 1998; 58:4799-4801.

Rowan A, Bataille V, MacKie R, Healy E, Bicknell D, Bodmer W, *et al.* Somatic mutations in the Peutz-Jeghers (*LKB1/STK11*) gene in sporadic malignant melanomas. *The Journal of Investigative Dermatology* 1999; 112(4):509-511.

Rustgi AK. Hereditary gastrointestinal polyposis and nonpolyposis syndromes. *The New England Journal of Medicine* 1994; 331(25):1694-1702.

Sandler RS. Epidemiology and risk factors for colorectal cancer. *Gastroenterology clinics of North America* 1996; 25(4):717-735.

Sapkota GP, Kieloch A, Lizcano JM, Lain S, Arthur JSC, Williams MR *et al.* Phosphorylation of the protein kinase mutated in Peutz-Jeghers cancer syndrome, LKB1/STK11, at Ser^{431} by $p90^{RSK}$ and cAMP-dependent Protein Kinase, but not its farnesylation at Cys^{433}, is essential for LKB1 to suppress cell growth. *Journal of Biological Chemistry* 2001; 276(22):19469-19482.

Sitas F, Madhoo J, Wessie J. Incidence of histologically diagnosed cancer in South Africa, 1993-1995. *National Cancer Registry of South Africa* 1998; 16.

Smith DP, Spicer J, Smith A, Swift S, Ashworth A. The mouse Peutz-Jeghers syndrome gene *Lkb1* encodes a nuclear protein kinase. *Human Molecular Genetics* 1999; 8(8):1479-1485.

Smith DP, Rayter SI, Niederlander C, Spicer J, Jones CM, Ashworth A. LIP1 , a cytoplasmic protein functionally linked to the Peutz-Jeghers syndrome kinase LKB1. *Human Molecular Genetics* 2001; 10(25):2869-2877.

Spigelman AD, Murday V, Phillips RKS. Cancer and the Peutz-Jeghers syndrome. *Gut* 1989; 30:1588-1590.

Strachan T and Read AP. Human molecular genetics. Bios Scientific Publishers, Oxford, 1997.

Stratakis CA, Kirrschner LS, Taymans SE, Tomlinson IPM, Marsh DJ, Torpy DJ, *et al.* Carney complex, Peutz-Jeghers syndrome, Cowden Disease, and Bannayan-Zonana syndrome share cutanous and endocrine manifestarions, but not genetic loci. *Journal of Clinical Endocrinology and Metabolism* 1998; 83(8):2972-2976.

Su GH, Hruban RH, Bansal RK, Bova GS, Tang DJ, Shekher MC, *et al.* Germline and somatic mutations of the *STK11/LKB1* Peutz-Jeghers gene in pancreatic and biliary cancers. *American Journal of Pathology* 1999; 154(6):1835-1840.

Tiainen M, Vaahtomeri K, Ylikorkala A, Mäkelä TP. Growth arrest by the LKB1 tumour suppressor: induction of $p21^{WAF1/CIP1}$. *Human Molecular Genetics* 2002; 11(13):1497-1504

Tiainen M, Ylikorkala A, Mäkelä TP. Growth suppression by LKB1 is mediated by a G1 cell cycle arrest. *Proceedings of the National Academy of Science. USA* 1999; 96:9248-9251.

Toribara NW, Sleisenger MH. Screening for colorectal cancer. *The New England Journal of medicine* 1995; 332(13):861-867.

Trejo-Becerril C, Sarmiento RG, Abad MM, Ichaso N, Delgado R, Cruz JJ, et. al. Immunohistochemical expression of p53 in breast carcinoma is associated with the intron 1 *Bgl*II polymorphism of the p53 gene. *Mutation Research 2000*; 452:231-236.

Trojan J, Brieger A, Raedle J, Roth WK, Stefan Zeuzem. Peutz-Jeghers syndrome: molecular analysis of a three-generation kindred with a novel defect in the serine threonine kinase gene *STK11*. *American Journal of Gastroenterology* 1999; 94(1):257-261.

Van Rensburg EJ. The genetic basis of cancer. *Specialist Medicine* 1997;46-53.

Wang ZJ,Taylor F, Churchman M, Norbury G and Tomlinson I. Genetic pathways of colorectal carcinogenesis rarely involve the PTEN and LKB1 genes outside the inherited hamartoma syndromes. *American Journal of Pathology* 1998; 153(2):363-366.

Wang Z-J, Churchman M, Avizienyte E, McKeown C, Davies S, Evans DGR, *et al.* Germline mutations of the LKB1 *(STK11)* gene in Peutz-Jeghers patients. *Journal of Medical Genetics* 1999; 36:365-368.

Wanebo HJ. Colorectal cancer. Mosby. St. Lois, Missouri, 1993.

Westerman AM, Entius MM, Boor PPC, Koole R, de Baar E, Offerhaus GJA, *et al.* Novel mutations in the *LKB1/STK11* gene in Dutch Peutz-Jeghers families. *Human Mutation* 1999a; 13:476-481.

Westerman AM, Entius MM, de Baar E, Boor PPC, Koole R, van Velthuysen MLF, *et al.* Peutz-Jeghers syndrome: 78-year follow-up of the original family. *The Lancet* 1999b; 353:1211-1215.

Ylikorkala A, Avizienyte E, Tomlinson IPM, Tiainen M, Roth S, Loukola A, *et al.* Mutations and impaired function of *LKB1* in familial and non-familial Peutz-Jeghers syndrome and a sporadic testicular cancer. *Human Molecular Genetics* 1999; 8(1):45-51.

Yoo LI, Chung DC, Yuan J. LKB1-A master tumour suppressor of the small intestine and beyond. *Nature Reviews Cancer* 2002; 2:529-535

Yoon K-A, Ku J-L, Choi HS, Heo SC, Jeong S-Y, Park YJ, *et al.* Germline mutations of the *STK11* gene in Korean Peutz-Jeghers syndrome patients. *British Journal of Cancer* 2000; 82(8):1403-1406