MOLECULAR GENETIC ANALYSIS OF A CONSTITUTIONAL CHROMOSOME TRANSLOCATION IN A PATIENT WITH GANGLIONEUROBLASTOMA

ANTONY MICHALSKI

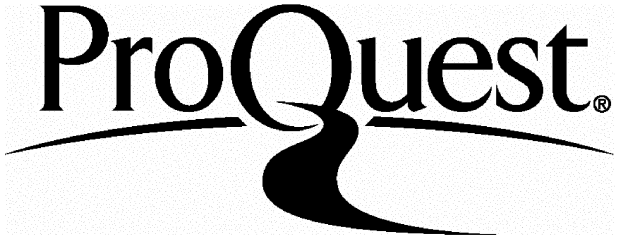
A thesis submitted for the degree of Doctor of Philosophy in the University of London

July 1994

The Imperial Cancer Research Fund Laboratory of Molecular Genetics

Department of Haematology and Oncology, Institute of Child Health

The University of London



ABSTRACT

Molecular genetic analysis of a constitutional chromosome translocation in a patient with ganglioneuroblastoma.

This study aimed to isolate and characterise the breakpoint junction fragment of a constitutional t(1;13)(q22;q12) from DG, a 5 year old boy with ganglioneuroblastoma, to see whether a gene critical to neural development has been disrupted.

Somatic cell hybrids were constructed by fusing the lymphoblastoid cell line from DG with mouse 3T3 cells. One subclone contained a derivative chromosome 1 but no other material from chromosomes 1 and 13.

Unique probes were isolated by Alu PCR cloning from a somatic cell hybrid containing 13pter to 13q14 as its only human component. Analysis of 180 clones yielded four probes which flanked the DG breakpoint. The probes were mapped by Southern hybridisation, PCR and FISH using cosmids isolated from a chromosome 13-specific cosmid library. All four probes were located in 13q12; 3'Alu66 mapped above and 3'Alu 78,71 and 62 mapped below, the DG breakpoint.

The four 3'Alu probes and an STS for the oncogene, FLT1, were used to screen a YAC library. The fourteen YACs identified with these markers were characterised by PFGE, hybridisation and PCR fingerprinting. PCR analysis using STS for the ends of the YAC identified by FLT1, suggested that this YAC was chimeric, which was confirmed by FISH. Cosmid c12.2, mapping 250 kb distal to 3'Alu66 by FISH, has replaced FLT1 as the closest marker distal to the DG breakpoint. Two YACs were identified with 3'Alu66. An STS from the left hand end of one of the YACs (a 1.3 Mb, non-chimeric YAC) was used to isolated five further YACs. The resulting contig of seven YACs will allow high definition mapping of 13q12 and may include clones that contain the DG breakpoint junction fragment and, hence, a gene important in Nb development.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr John Cowell, for his precision of thought and all the advice and encouragement he provided throughout the course of this thesis. John also made all the hybrids and performed their cytogenetic characterisation. I am grateful to Dr Finbarr Cotter for the help and advice on PCR and to Dr Helena Kempski and Ms Leslie Ann Hawthorn for the FISH analysis. Dr David Shapiro at St. Judes was of great assistance and without Dr Rosemary Shannon's original identification of the patient none of this work would have been performed. Much of the work would not have been possible without excellent backup from the Imperial Cancer Research Fund (ICRF) Oligonucleotide Synthesis department and the Genome Analysis Laboratory where Dr Mark Ross, in particular, was of great help.

I would like to thank Sheila Giles for all her help in the preparation of this and other manuscripts and for her cheerful acceptance of unreasonable deadlines. I am also grateful to the ICRF photographic department for their help with the figures. My colleagues in the laboratory deserve a special mention for their practical advice, explanations and companionship. The other members of the Department of Haematology and Oncology always provided both moral and financial support and shouldered an extra work load during the writing of this thesis.

I am indebted to the Medical Research Council for the Training Fellowship award as well as to The Neuroblastoma Society, the Research into Childhood Cancer fund and the Emma Killingback Fund for their support.

Finally, I would like to thank my family, especially Julie, for the tolerance, acceptance and understanding they have shown throughout the preparation of this manuscript.

To children with neuroblastoma and to their families

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND SYMBOLS:

Nucleic acids and their constituents

G, A, T, C, U	The bases guanine, thymine cytosine, uracil
GDP	Guanidine diphosphate
NTP (GTP, ATP etc)	The nucleoside triphosphates; guaninetriphosphate etc.
dNTP (dGTP etc)	The deoxynucleoside triphosphates
ddNTP (ddGTP etc)	The dideoxynucleoside triphosphates
DNA	Deoxyribonucleic acid
cDNA	Complementary deoxyribonucleic acid
RNA	Ribonucleic acid
mRNA	Messenger RNA
tRNA	Transfer RNA

Chromosomes

X, Y	The sex chromosomes
1 - 22	The autosomes
p	The short arm
q	The long arm

Reagents/chemicals/enzymes/media

β-ME	beta-mercaptoethanol
BSA	Bovine serum albumin
DMSO	Dimethyl sulphoxide
DNase	Deoxyribonuclease
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra-acetic acid
HEPES	N-2-hybroxyethylene piperazine
H_2O	Water
HCl	Hydrochloric acid
KCl	Potassium chloride
L-agar	Luria agar
LB	Luria broth
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
MOPS	3-(N-morpholino)-propane sulphonic acid
NaCl	Sodium chloride
NaOH	Sodium hydroxide
-OH	Hydroxyl group

^{32}P	Phosphorous-32
PBSA	Phosphate buffered saline
RNase	Ribonuclease
^{35}S	Sulphur 35
SDS	Sodium dodecyl sulphate (lauryl sulphate)
SSC	Sodium salt citrate
TAE	Tris acetate
TBE	Tris borate
TE	Tris-EDTA-buffer
TEMED	N,N,N',N'-tetramethyl ethylene diamine
Tris	Tri-hydroxymethyl-aminomethane

Units

bp	base pairs of nucleic acid
kb	kilobase pairs of nucleic acid
Mb	megabse pairs of nucleic acid
kD	Kilodalton
cpm	Counts per minute
rpm	Revolutions per minute
°C	Degrees centigrade
l/ml/µl	Litre/millilitre/microlitre
g/mg/µg	Gram/milligram/microgram
X^{-1}(eg ml^{-1})	Per X (eg per millilitre)
A/mA	Ampere/milliampere
M/mM	Molar/millimolar
Ci/mCi/µCi	Curie/millicurie/microcurie
(v/v)	Volume to volume ratio
(w/v)	Weight to volume ratio

Other

Nb	Neuroblastoma
WT	Wilms tumour
WT1	Wilms tumour 1 gene
BWS	Beckwith Wiedeman syndrome
Rb	Retinoblastoma
RB1	The retinoblastoma gene
pRb	Retinoblastoma protein
PCR	Polymerase Chain Reaction

RFLP	Restriction Fragment Length Polymorphism
YAC	Yeast artificial chromosome
PFGE	Pulsed field gel electrophoresis
SSCP	Single Strand Conformation Polymorphism
α	Alpha
β	Beta
γ	Gamma
λ	Lambda

INTRODUCTION

INTRODUCTION

1.0 OVERVIEW

The aim of this PhD thesis was to isolate and characterise the breakpoint junction of a constitutional t(1;13)(q22;q12) translocation from a patient with stage II ganglioneuroblastoma in order to ascertain if a gene vital in neural development had been disrupted by this genetic change. In order to justify this approach, it is essential to review the evidence for the importance of DNA damage in cancer formation and to introduce the concepts of oncogenes and tumour suppressor genes, concentrating on those changes found in neuroblastoma (Nb). Many molecular abnormalities have already been described in Nb and, although they explain some aspects of disease behaviour, the genetic changes responsible for Nb tumorigenesis remain unknown. The epidemiology and clinical behaviour of Nb allow predictions to be made as to the nature of the genetic changes responsible for the tumour phenotype and suggest that inactivation of a critical tumour suppressor gene may be important in Nb formation. There are some patients with a variety of genetic diseases where the analysis of constitutional chromosome translocations has established the location of genes responsible for the disease phenotype. Positional cloning and the investigation of candidate genes were the strategies available for the analysis of the t(1;13)(q22;q12) breakpoint.

1.1 DNA AND CANCER

Cancer is a genetic disease, at least at the cellular level since the progeny of tumour cells also produce tumours. The occurrence of hereditary predisposition to malignancy was recognised over a century ago (reviewed in Knudson 1977) and drew attention towards a link between genetics and tumour formation. More direct evidence for the importance of DNA damage in tumorigenesis came from the observation that carcinogens such as ultraviolet light, ionising radiation or chemicals (e.g. Nitrogen mustards) are also mutagens (Ames et al, 1973). Similarly, patients with defects of DNA repair (Xeroderma pigmentosa) (Scherly et al, 1993) were found to be at increased risk of developing malignancy. A direct link between genetics and cancer was first proposed in 1914 when Boveri, Professor of Zoology at the University of Wüzburg, wrote a small, 64 page book on the origin of malignant tumours in which he postulated that chromosome abnormalities were the cause of all malignant transformation. He suggested that random genetic events in a single cell caused chromosomal abnormalities which conferred a permanent growth advantage and gave rise to clonal expansion with millions of similarly altered cells forming a malignancy. Even though he was not able to visualise chromosomes at the time, this hypothesis has proved to be remarkably accurate. Much more is now known about the type of genetic

alterations associated with malignant transformation as well as the importance of the degree of differentiation of the cell in which the genetic change occurs.

Three types of genetic change exist:
(1): genome mutations which change the number of chromosomes and have their effect by altering gene dosage.
(2): chromosome mutations in which the structure of the chromosome is altered (e.g. deletion, translocation).
(3): gene mutations which, unlike the other two types, are invisible by light microscopy and are recognised by a phenotypic change and/or by direct genetic analysis.

Mutations can occur in germinal or somatic cells. Germinal mutations occur in the first or second meiotic division and are important in producing genetic variability within species. The mutation may render the germinal cell non-viable but, if not, will be present in every cell of the offspring produced and may produce its effects late on in life. If the change in phenotype produced by the mutation is beneficial, the mutation may well spread throughout the gene pool. 'Neutral' mutations, neither advantageous nor disadvantageous to the organism, lead to allelic variations. Indeed, many such genetic polymorphisms exist and are routinely used in genetic counselling and linkage analysis. If the mutation is disadvantageous, the offspring may die before reproductive age. Conversely, mutations occurring in somatic cells will be transmitted only to the direct descendants of the original cell, the number of which will be determined by the degree of differentiation of the cell (and hence its proliferative capacity) and any growth advantage the mutation confers.

Cancer arises from the abnormal and uncontrolled division of cells. Whereas normal cells in tissue culture will cease dividing when they make contact with other cells, malignant cells lose this 'contact inhibition' and form piled up foci of cells. Cancer cells differ from their normal counterparts in other ways, such as their ability to grow without an exogenous supply of growth factors. How DNA damage to normal cells resulted in the change to the malignant phenotype was unclear until the discovery of two major classes of genes, namely; oncogenes and tumour suppressor genes.

1.2 ONCOGENES

Evidence for specific genes having a role in carcinogenesis came from a study of the RNA tumour viruses in the 1970s. RNA viruses usually carry only three genes; one coding for the viral envelope (ENV), one for the group specific antigen (GAG) and a third for the enzyme reverse transcriptase (POL) which catalyses the rather heretical transcription of viral RNA to DNA allowing it to be incorporated into the host genome. RNA viruses associated with malignant transformation carry a fourth gene, called a viral oncogene (*v-onc*) capable of inducing tumours *in vivo* in animals and in

immortalised human cell lines *in vitro* (Varmus, 1984). Surprisingly, *v-onc* DNA sequences were not found to be unique to viruses, but shared broad homology with genes carried in mammalian and avian cells (Bishop, 1987). These cellular counterparts were termed cellular oncogenes (*c-onc*) or proto-oncogenes. However, although mammalian and avian cells carry these 'oncogenic' sequences they obviously do not all undergo malignant transformation. Fine structure analysis of the *v-oncs* of acutely transforming retroviruses showed that they were abnormal in some way. In some cases, the *v-oncs* were expressed as fusion proteins, containing both proto-oncogene and viral sequences, whereas in other cases the structure of the *v-onc* differed from that of its cellular counterpart (Varmus, 1984). The role of oncogenes as active instigators of malignant transformation was established by a number of lines of evidence. Infecting immortalised cells in culture with Rous sarcoma virus (RSV) was known to cause malignant transformation and the oncogene in RSV was called *v-src*. Temperature sensitive variants of RSV occur and their ability to cause transformation can be abolished by culturing infected cells at a non-permissive temperature for *v-src* expression, suggesting that the oncogene was acting dominantly to induce transformation (Martin, 1970). Further evidence that oncogenes, and not other viral DNA sequences, were responsible for malignant transformation came from experiments in which immortalised cells in tissue culture were transfected with DNA from human malignant tumours. When DNA from human bladder cancer cells was introduced into NIH 3T3 cells, for example, foci of transformed cells developed (Shih and Weinberg, 1982). Cells from these transformed foci were found to contain DNA sequences that shared homology with the *v-onc* of the Harvey murine sarcoma virus (H-*ras*) and Kirsten sarcoma virus (K-*ras*), both of which are retroviruses causing tumours in animals (Reddy et al, 1982, Tabin et al, 1982). The cellular homologues of these viral genes formed the *ras* family of oncogenes and the difference between the transforming oncogene and the normal proto-oncogene was due to a single amino acid substitution (Reddy et al, 1982, Santos et al, 1982, Tabin et al, 1982). The discovery that mutations in the *ras* gene could be induced by treatment of cells with carcinogens provided further evidence that oncogenes were directly involved in malignant transformation (Sukumar et al, 1983). Since then, many different oncogenes have been described in human malignancy and their involvement in malignant transformation is now beyond doubt.

While the evidence above is a compelling argument for the role of oncogenes in the malignant process, elucidation of the effects of *c-oncs* on cell growth and development and the discovery of mechanisms by which *c-onc* function can be come deranged added further weight to the argument.

1.2.1 *Oncogene function*

The hallmark of the malignant phenotype is the abnormal growth of cells. Hence, if abnormalities in the structure and function of oncogenes are important in producing the cancer phenotype, it would seem logical that proto-oncogenes are involved in the control of cellular growth and differentiation. The control of cell growth and development is a complex process which is poorly understood. Cells are under the influence of endogenously or exogenously produced growth factors which bind to growth factor receptors on the cell surface and changes in conformation of the receptor activate a variety of messages, including the membrane-bound tyrosine kinases. These signals are transduced by cytoplasmic tyrosine kinases and serine/threonine kinases with *ras*, *raf* and mitogen-activated protein (MAP) kinases having central roles in this process. Within the nucleus, nuclear binding proteins and transcriptional regulators further modulate this growth signal. Cellular oncogenes have been associated with each of the above stages which are shown schematically in figure 1.1. Each of these stages will now be discussed in more detail.

1.2.1.a *Growth factors*

Under normal conditions cell growth is controlled by environmental influences. Cell to cell contact usually results in decreased cell division but some cells must retain the capacity for rapid proliferation. For example, in circumstances such as wound repair, rapid cell growth is essential. Some extra cellular proteins, such as platelet derived growth factor (PDGF) can act as mitogens, stimulating cell division. PDGF is composed of two chains, A and B, and it was found that the product of the v-sis oncogene (isolated from the Simian sarcoma virus) had considerable homology to the B chain of PDGF (Doolittle et al, 1983, Waterfield et al, 1983). In some circumstances, such as rapid normal tissue growth or a malignancy, the cell produces its own growth factors, which act in an autocrine loop. In normal growth this stimulus can be overridden, but in malignant transformation it becomes independent of regulatory influences. Attaching the normal *c-sis* gene to a strong promoter results in transformation of cells in culture, but only of those cells that express PDGF receptors (Gazit et al, 1984). DNA sequences homologous to *sis* are expressed in some sarcomas and glioblastomas but not in normal human fibroblasts, and it may be that *sis* resulted in the malignant transformation of these tumour cells by mimicking the action of PDGF (Eva et al, 1982).

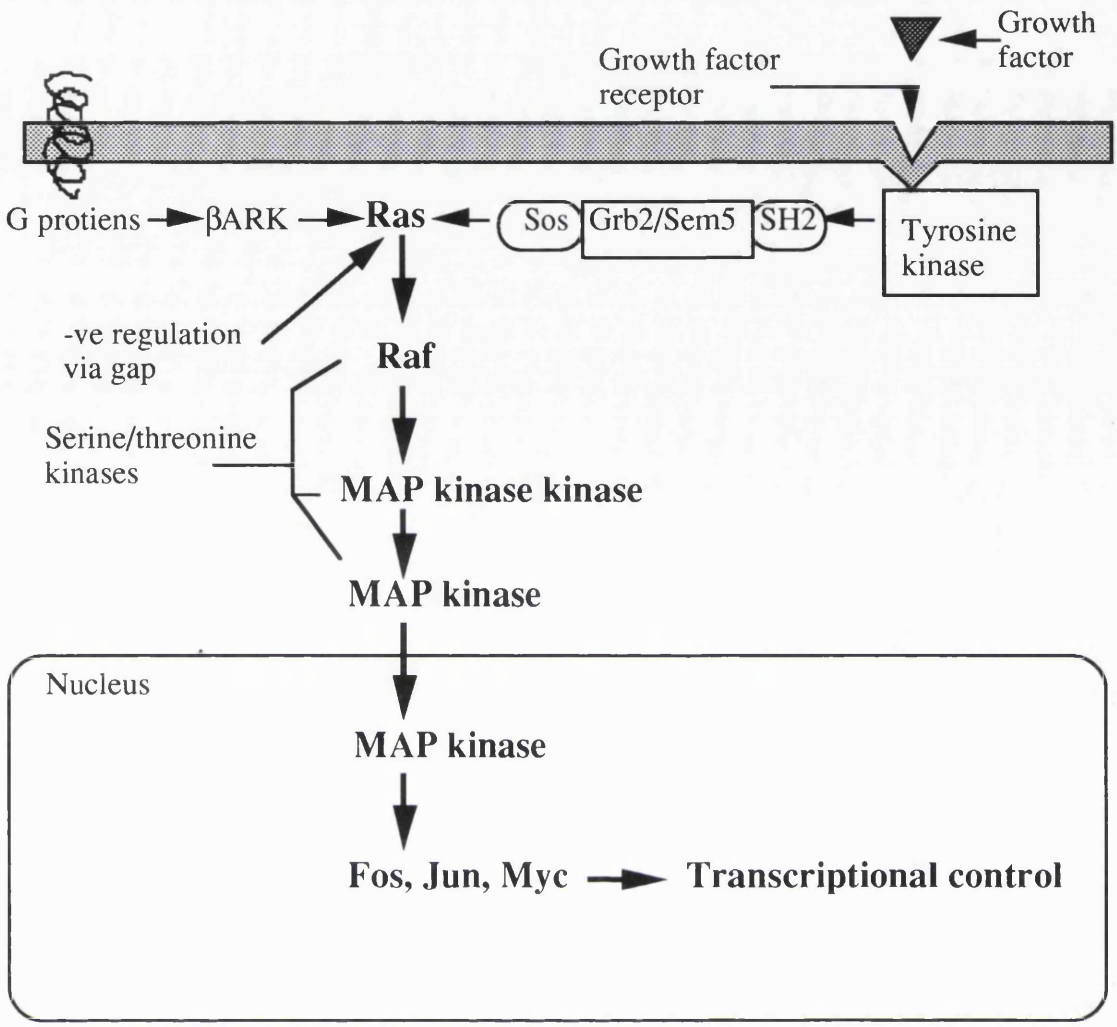


Figure 1.1 A diagrammatic representation of the transduction of a mitogenic stimulus at the cell surface into the increased transcription of genes involved in cell growth. A given growth factor binds to a membrane bound growth factor receptor and activates the associated tyrosine kinase receptor (shown on the right hand side of the diagram). Changes in the tyrosine kinase lead to the assembly of a complex (SH2, Grb2/Sem5, and Sos) that causes activation of Ras. Conversion of Ras from the GDP bound to the active GTP bound state can also be mediated by G proteins and related compounds (shown on the left hand side of the diagram). Activation of Ras leads to signal transduction via Raf (the product of an oncogene related to *ras*). Raf acts as a MAP kinase kinase kinase activating a series of serine/threonine kinases which result in the increased production of MAP kinase. Nuclear growth factors such as Fos and Myc then mediate the increased transcription of growth factors.

1.2.1.b *Growth Factor Receptors*

The protein products of several oncogenes act as part of cell surface receptors and, therefore, are involved in the transduction of the growth factor message outside the cell to an intracellular message, usually by activating a membrane bound tyrosine kinase. Small alterations in the configuration of the receptor can mimic the presence of a growth factor, activating tyrosine kinase even in the absence of exogenous growth factor and this appears to be the case with *v-erb*-B, the viral homologue of *c-erb*-B, which codes for the cellular receptor of epidermal growth factor (Downward et al, 1984). The receptor bound tyrosine kinase may itself be overactive as in the case of *v-fms* (the viral counterpart of *c-fms*) which codes for the colony stimulating factor CSF-1 receptor kinase in macrophages (Sherr et al, 1985). The normal CSF-1 receptor possesses tyrosine kinase activity and stimulates rapid cell division of macrophage precursors (Metcalf, 1985). The product of the v-*fms* oncogene differs from CSF-1 at its carboxy terminus and it is thought that this molecular change results in the mutant receptor being constitutively activated (Roussel et al, 1988).

1.2.1.c *Signal transduction from the membrane to the nucleus*

Signals received at the cell surface must be transduced in some way so that appropriate genes are transcribed in the nucleus. In the steroid receptor super-family this signal transduction is relatively straightforward. Steroid hormones are small, lipid soluble molecules that pass freely through the cell membrane without having to interact directly with membrane bound receptors. The steroid receptors exist in the cytoplasm complexed to various heat-shock proteins such as hsp90, hsp70 and hsp56. Binding of the hormone to the receptor results in dissociation from the heat-shock proteins and the receptor/hormone complex migrates to the nucleus where it binds to specific DNA sequences called hormone response elements (HRE) and directly initiates transcription (reviewed in Evans, 1988).

Most growth messages are, however, transduced in a much more complex manner and the oncogene, *ras*, was known to be involved in this process. The normal protein product of the *ras* gene (Ras) binds to guanine nucleotides. In its inactive state it is bound to GDP whereas when bound to GTP it becomes active in signalling. Normally the duration of activity of GTP bound Ras is short as both the intrinsic GTPase activity of the protein and extrinsic GTPases result in reversion to the inactive GDP bound form. Mutated forms of Ras have greatly reduced intrinsic GTPase activity and, therefore, continue to be active for longer. Ras activity is also controlled by GTPase activating proteins (gap) which activate cytosolic GTPase. Amongst the best studied of these are p120 Gap (Gibbs et al, 1984) and the gap related domain of neurofibromin, the product of the neurofibromatosis type 1 (NF1) gene (Xu et al, 1990). Patients with NF1 are predisposed to a variety of malignancies including tumours

derived from the Schwann cells that make up nerve sheaths. Malignant Schwann cell tumours from patients with NF1 (which must carry a NF1 gene defect) were found to have very high levels of GTP bound and, therefore, active Ras (Basu et al, 1992).

Mechanisms of transduction of message from membrane bound receptors to Ras have recently become clearer. At least two separate signal cascades are known to function via the activation of Ras. The first involves membrane bound tyrosine kinases. When bound to extracellular ligand the membrane bound tyrosine kinase receptors are phosphorylated and then interact with the Grb2/Sem5 family of adapter proteins either directly or via specific SH2 domain-containing adapter proteins (Rozakis-Adcock et al, 1992). The phosphorylation of the membrane bound tyrosine kinases is itself controlled by a diverse family of protein tyrosine phosphatases (Fischer et al, 1991). The Grb2/Sem5 proteins then interact with proteins that activate Ras by exchanging the inactive GDP for an active GTP molecule. The best known of these guanine-nucleotide-releasing factors is Sos, the protein product of the Son-of-sevenless gene involved in the optic development of the fruit fly, *Drosophila*. It may be that the Grb2/Sem5 proteins bind to Sos via specific SH3 domains which are known to be involved in signal transduction (reviewed in Blenis, 1993). The end result would be that phosphorylation of the membrane bound tyrosine kinase binds Sos to the cell membrane, holding it in close proximity to plasma membrane-associated Ras where it could promote Ras to exchange a GDP molecule for an 'active' GTP nucleotide. This cascade appears to be broadly conserved among species having been found to be important in the vulval development of nematodes, ocular development of *Drosophila* and in mammalian cells (Blenis, 1993).

The second signalling cascade which appears to function via activation of Ras is that associated with the G proteins which are involved in a variety of pathways including those responsible for phospholipase activity, modulation of ion channels and inhibition of adenylate cyclase. Some agonists of G-protein-coupled receptors such as thrombin and lysophosphatic acid act as mitogens, stimulating cell proliferation via a pertussis toxin (PTX)-sensitive signalling pathway (van Corven et al, 1993). The G-protein signalling pathway is conserved among species, from yeast to mammals, but how do the receptors couple to Ras? The receptors couple to trimeric GTPases consisting of α, β and γ subunits. In the yeast, Saccharomyces cerevisiae, the α subunit appears to regulate the ability of the $\beta\gamma$ dimer to activate the signalling cascade whereas in mammals it appears that signalling can be regulated by both α and $\beta\gamma$ subunits (reviewed in Cook & McCormick, 1994). The $\beta\gamma$-mediated signalling can be abolished by transfecting cells with plasmids expressing *ras*-inhibitory proteins suggesting that signal transduction via $\beta\gamma$ dimers was mediated through Ras (Crespo et al, 1994). The $\beta\gamma$ dimer is known to bind to the β-adrenergic receptor kinase, (βARK), a cytosolic enzyme involved in the adaptive response to hormonal stimulation. The $\beta\gamma$ dimers bind

at the carboxy terminus of βARK to a region which has homology to a variety of signals involved in cellular signalling, the Pleckstrin homology (PH) domain. Of note is the fact that mammalian nucleotide-exchange factors such as Sos also have PH domains. Hence, it may be that the βγ dimer assembles a signalling complex on the plasma membrane analogous to that composed of SH2, the Grb2/Sem5 and Sos. The effect of the βγ dimer, βARK complex would, therefore, be to activate Ras (reviewed in Cook & McCormick, 1994).

It appears, therefore, that a variety of signalling pathways have their effects via Ras but how does Ras mediate DNA transcription? Whereas the transduction of cell surface signals was effected by protein tyrosine kinases it was known that factors directly affecting transcription were serine/threonine kinases so somewhere there had to be a 'kinase switch'. Over the last two years the mechanisms of this switching from tyrosine kinase activation at the cell membrane to serine/threonine kinase activation in the cytoplasm have been described (reviewed in Egan & Weinberg, 1993, Blenis, 1993, and Ruderman, 1993). One approach taken to elucidate the nature of the control of the serine/threonine kinases was to start with a known phosphorylation event involved in the control of cell growth and work backwards to identify the signalling pathway. The phosphorylation of the S6 protein of the 40S ribosomal subunit was chosen for study and the kinase responsible was duly identified as RSK. The control of RSK activity was found to be via a family of serine/threonine kinases, mitogen activated protein kinases (MAP kinases), also called extracellular signal related kinases (ERK). MAP kinases themselves were found to be activated by serine/threonine kinases, termed MAP kinase kinases (or MEK in mammalian systems) and even MAP kinase kinases appeared to be controlled by serine/threonine phosphorylation (via MAP kinase kinase kinases). The link between activation of membrane bound tyrosine kinase receptors and the MAP kinase signalling pathway was found to be activation of Ras. Indeed, expression of oncogenic ras mutants in the absence of mitogenic stimuli leads to persistent MAP kinase activation, whereas the introduction of ras inhibitory proteins or over-expression of gap blocks the activation of the MAP kinase cascade. The protein product of *raf*, an oncogene known to mimic some of the effects of *ras*, forms bimolecular complexes with Ras and it is via Raf that Ras has its effect. Raf has been identified as having serine/threonine kinase activity and, in the majority of species studied, acts as a MAP kinase kinase kinase. In summary, therefore, signals from protein tyrosine kinase and G-protein associated cell surface receptors are transduced via Ras and then Raf through the MAP kinase pathway. It appears that signal transduction via Ras, Raf and MAP kinases is conserved through evolution from yeast to mammals with only minor modifications. In spite of the fact that the signalling pathway is conserved between species the terminology is not and many components of the MAP kinase signalling pathway have different names in yeast, *Drosophila* and

mammals. Although appealingly simple this scheme may not be the entire story. It appears that MAP kinase can phosphorylate Raf which could imply that a negative feedback loop controlling the effect of the MAP kinase cascade exists.

The original analysis of the MAP kinases concentrated on their effect in the transduction of mitogenic stimuli. Indeed, the phosphorylation of Jun and Fos is dependant on MAP kinase activity and the products of these two cellular oncogenes are known to be important in the production of the activator protein-1 (AP-1) transcription factor which modulates expression of genes involved in growth regulation, differentiation and neoplastic transformation (Bernstein et al, 1994). However, mitogenic signals are not the only ones to be transduced by MAP kinases. The growth of unfertilised eggs of vertebrates is arrested in metaphase by the activity of cytostatic factor (CSF) which is itself activated by MAP kinase (Haccard et al, 1994). The detailed interactions of the various components of the Ras/Raf/MAP kinase pathway remain unclear but their further analysis may well yield important insights into the precise mechanisms involved in malignant transformation.

1.2.1.d *Oncogenes acting in the nucleus*

Several oncogenes code for proteins located in the nucleus. The *myc* family of oncogenes has three members: *c-myc* (the cellular homologue of the *v-myc* oncogene found in several avian retroviruses), L-*myc* (associated with human small cell lung cancer) and N-*myc* (amplification of which is found in some Nb tumours). As amplification of N-*myc* is important in Nb progression (section 1.5.5.c) the structure and function of the *myc* oncogenes will be dealt with in detail. All three *myc* genes code for nuclear oncoproteins. *c-myc* encodes a nuclear protein with a short half life which is expressed when mitogens are present. Unlike other early response genes it continues to be expressed throughout the cell cycle as long as mitogens are present. If *c-myc* expression falls, cells go into G0 after the next cell division whereas if *c-myc* levels remain high, cells are committed to continue to divide. *c-myc* , therefore, encodes a repressor of genes that cause growth arrest (reviewed in Cole 1986). If *c-myc* is constitutively activated, and yet cells are grown in deficient medium, apoptosis (programmed cell death) occurs (Evan et al, 1992). Several observations have implicated the *myc* genes in cancer formation. The viral homologue of *c-myc* (*v-myc*) was identified as the oncogene of avian sarcoma virus (MC29) which causes cancer in chickens (Weiss et al, 1985) and there is evidence to suggest that *c-myc* can co-operate with *ras* to transform cultured cells (Land et al, 1983). Direct evidence for the role of *c-myc* in human malignancy came from the observation that *c-myc* DNA sequences are found to be amplified in cells from some types of human leukaemia (Collins and Groudine, 1982, Dalla-Favera et al, 1982). However, the most compelling evidence

for the involvement of *c-myc* in a human neoplasm came from the study of Burkitt's lymphoma. This high grade B-cell malignancy occurs in 2 forms: a slow growing jaw mass (high incidence in Africa called the endemic or African type) or a rapidly growing, abdominal lymphoid malignancy (non-African or sporadic type). In both types the *c-myc* gene on chromosome 8 is juxtaposed with the promoters of immunoglobulin genes, which are constitutively activated in B-cells. Translocations involve the immunoglobulin heavy chain on chromosome 14, the kappa light chain on chromosome 2 or the lambda light chain on chromosome 22 (reviewed in Croce CM, 1985). In cells carrying these translocations, the close proximity of the immunoglobulin promoter and enhancer sequences, combined with point mutations and rearrangements occurring in the *c-myc* gene, are thought to result in deregulation of *c-myc* (Spencer and Groudine, 1991). The phenotypic difference between the endemic and the sporadic types can be explained by the proximity of the immunoglobulin promoter to the *c-myc* gene. In the African type the promoter is some distance away, but in the non-African, more aggressive, tumours the translocation is actually within the first exon of the *c-myc* gene (Pelicci et al, 1986). The co-expression of *c-myc* and the oncogene BCL-2 (an inhibitor of apoptosis) results in particularly aggressive B-cell disease (Klein, 1991). Although the mechanisms by which *c-myc* exerts its effects are poorly understood, there is evidence that it is involved in transcriptional regulation. The end terminus of Myc (the protein product of *c-myc)* contains a region capable of activating transcription in mammalian cells (Cole et al, 1986). The structure of the Myc oncoproteins is similar to that of the basic-helix-loop-helix (bHLH) and basic-leucine zipper (bLZ) families of transcription factors (Penn et al, 1990). Both bLZ and bHLH families bind DNA as dimers and, therefore, in order for the basic regions of each monomer to make specific contact with DNA, the transcription factor must bind to itself (homodimerisation), or to a closely related molecule (heterodimerisation). Myc is unable to form homodimers but does complex with the related protein Max. Max is a bHLH- LZ protein which exists as four distinct isoforms due to alternative splicing. The functional significance of the different isoforms is not known. Max is able to form homodimers but preferentially complexes with Myc. Whereas levels of Myc fluctuate according to cellular growth signals, Max is a stable protein expressed at similar levels in both proliferating and quiescent cells. The concentration of the Myc/Max heterodimer is, therefore, determined by the rate of Myc synthesis (Berberich et al, 1992). There is evidence from work in yeast that Myc activates transcription through CACGTG and CACATG sequences in collaboration with Max (Crouch et al, 1993). The identity of the genes regulated by the Myc/Max complexes, however, remains unknown and the role of Max in transcriptional regulation is also poorly understood.

1.2.2 *Mechanisms of oncogene activation*

Transformation of cells by retroviruses may have led to the discovery of oncogenes but infection with retroviruses is not the cause of the majority of human malignancies. Abnormalities of the cellular counterparts of viral oncogenes, *c-oncs*, however, have been associated with tumour development. Three major mechanisms exist: gene amplification, increased transcription of a single copy gene and the production of a mutated protein product.

1.2.2.a *Gene Amplification*

DNA amplification is a common mechanism for increasing gene dosage. Gene amplification is not always pathological; for example, chorion genes are amplified as part of the normal pattern of development of the oocytes of Drosophila or Xenopus (Kafatos et al, 1985). However, there are now many examples of oncogene amplification in solid tumours with amplification of N-*ras, abl, myb, egfr, erb*B2 and of all three members of the *myc* family having been identified in tumours or cell lines (reviewed in Schwab & Amler, 1990). The amplified genes either insert themselves 'en bloc' into chromosomes forming homogeneously staining regions (HSRs) or coalesce to form double minutes (DMs) (Cowell, 1982). There is evidence that oncogene amplification has a role in the development or progression of malignancies. In the case of Nb, for example, patients with tumours in which amplification of N-*myc* has been identified have more aggressive disease (section 1.5.5.c). Similarly, amplification of *c-myc* in lung tumours is associated with a poor prognosis (Prins et al, 1993). More direct evidence for the role of amplified oncogenes in tumour formation comes from *in-vitro* experiments. Increasing the expression of N-*myc* by the use of an expression vector can result in the transformation of human cell lines in tissue culture (Schwab et al, 1985) and, as gene amplification results in increased gene expression, we can infer that amplified oncogenes are involved in tumorigenesis. Perhaps the most convincing data for the role of oncogene amplification comes from experiments in which antisense oligonucleotides were used to decrease oncogene expression. The promyelocytic cell line, HL60, overexpresses *c-myc*. The addition of a *c-myc* antisense oligonucleotide to HL60 cells in culture resulted in a decrease in proliferation and an increase in differentiation along the granulocytic line in both a sequence-specific and dose-dependant manner (Prins et al, 1993).

1.2.2.b *Increased transcription*

Increased transcription of a single copy gene arises when the oncogene is put under control of an active promoter. This is the case in the B-cell lymphomas (discussed in section 1.2.1d) where the *c-myc* oncogene is put under the transcriptional control of promoters for immunoglobulin loci. There is *in vitro* experimental data to

support this theory. *c-mos*, the cellular homologue of the *mos* oncogene from Maloney murine sarcoma virus (*m-msv*), is incapable of inducing transformation of cells alone. However, when a construct containing *c-mos* is linked to the *m-msv* long terminal repeat (which contains viral promoter sequences) was transfected into cells, transformed foci arose (Blair et al, 1981).

1.2.2.c *Mutated Protein Product*

Chromosome translocations can juxtapose gene sequences giving rise to fusion proteins with oncogenic activity. The classic example is the Philadelphia chromosome, characteristic of chronic myeloid leukaemias (CML), which results from a translocation between chromosomes 9 and 22 (Nowell and Hungerford, 1960). The t(9;22) fuses the *c-abl* (cellular homologue of the Abelson mouse leukaemia virus) oncogene with *bcr*, a gene known as the 'breakpoint cluster region' on chromosome 22 (Adams, 1985). The resulting fusion protein has a greatly increased tyrosine kinase activity (Adams, 1985). The t(9;22) translocation is not specific to CML and is found in around 10% of cases with acute lymphoblastic leukaemia (ALL). In CML the resulting fusion protein is 210 kD in size, whereas in ALL it is only 190 kD suggesting that the translocation breakpoint is different in the two diseases (Hermans, 1987).

In some diseases there is considerable heterogeneity in the translocations seen. For example, in infantile leukaemia a t(4;11) translocation is seen in 90% of cases, fusing the ALL-1 gene on 11q23 (a gene which has homology to the Drosophila trithorax gene (Djabali et al, 1992) to sequences on chromosome 4. However, many other translocations have been reported and, although the breakpoint on chromosome 11 (11q23, in the ALL-1 gene) remains constant, the position of the reciprocal translocation breakpoint varies.

The protein products of these translocations can retain the functional characteristics of both of the normal proteins from which they were derived. For example, acute promyelocytic leukaemia is associated with a t(15;17) translocation juxtaposing the retinoic acid receptor RARα on chromosome 17q21 with a novel transcription factor, PML on chromosome 15q23. The fusion protein has transcription factor activity but may also act as a block to retinoic acid produced differentiation. The latter effect can be overcome to some degree by huge doses of retinoic acid (Zelent, 1994).

Increased activity of a single oncogene, however, is not sufficient to induce the malignant transformation of an otherwise normal cell. Whilst it is true that *ras* can transform NIH 3T3 cells (Lowy et al, 1978) these cells already have an immortalised phenotype and can be considered pre-malignant. Normal cells can be transformed by *ras* but only when immortalised by *myc* (Land et al, 1983) which led to the concept that multiple oncogenes work together to transform cells (Weinberg, 1985).

1.3 TUMOUR SUPPRESSOR GENES

We have seen that oncogenes code for factors associated with cell growth and that they act dominantly; increased oncogenic effect resulting from gene amplification, increased gene expression or the increased activity of novel protein products. However, in the 1970s evidence was found for a second class of genes associated with malignancy but here it was loss of function, not increased effect, that resulted in tumour formation. Two lines of evidence were important in identifying this second class of genes namely, the statistical evaluation of heritable cancer syndromes and experiments in which hybrids were constructed between normal and malignant cells.

1.3.1 *Tumour/normal hybrids*

In 1969 Harris and Klein showed that fusion of malignant and normal cells *in-vitro* often resulted in the loss of the malignant phenotype. The hybrids were, however, genetically unstable, randomly losing chromosome material and, if the hybrids were grown through successive generations, some clones reverted back to the malignant phenotype. These experiments have been extended and the introduction of single human chromosomes or even micro-dissected regions of specific chromosomes into tumour cells also suppressed malignancy (Stanbridge, 1990). This suggested that normal cells carry genes that act to prevent malignancy and that their presence in the hybrids is capable of suppressing the transformed phenotype but, if these genes are lost through genetic instability, the malignant phenotype recurs. It appears, therefore, that these endogenous genes act recessively at the cellular level.

1.3.2 *Hereditary Cancer*

The recessive nature of these (tumour suppressing) genes did not fit well with the evidence from the eye tumour retinoblastoma (Rb). Rb is the commonest eye tumour of childhood and was known to be familial, the predisposition being transmitted as an autosomal dominant trait with variable penetrance (Vogel, 1979). In hereditary cases the tumour was diagnosed earlier and was more likely to be multiple and bilateral. This led Knudson to propose his now famous two-hit hypothesis (Knudson, 1971).

He postulated that Rb arose as a result of two genetic events (hits). In sporadic cases both hits had to occur as random genetic events in the same cell for the tumour to develop. In hereditary cases, however, one hit was a germline mutation, affecting every cell in the body, and all that was needed for Rb formation was a single genetic event in an appropriate cell. A single event in any retinal cell was much more likely to occur than two events in a single cell, hence the early onset and the greater likelihood of multiple tumours in familial cases. Two years later, Comings extended the concept by

suggesting that Rb may involve the loss of both alleles of a recessively acting gene that normally suppressed tumour formation (Comings, 1973). The existence of these genes, called 'tumour suppressors', 'anti-oncogenes' or 'recessive oncogenes', none of which accurately describes their normal role, is firmly established and several have now been cloned. The structure and function of three examples, RB1 (the retinoblastoma predisposition gene), WT1 (a gene predisposing to certain types of Wilms tumour) and p53 will be reviewed since they each have different mechanisms of action.

1.3.3 *Retinoblastoma*

1.3.3.a *Evidence localising RB1 to 13q14*

The suggestion that mutation of a single gene is responsible for Rb promoted world-wide attempts to clone it. As RB1 was the first tumour suppressor gene to be cloned, a brief review of the approach used will be instructive. Rare patients with Rb also developed a range of other physical abnormalities and mental retardation. This complex phenotype was always associated with constitutional deletions of chromosome 13 and, as the number of deletions reported increased, a common region of overlap emerged in 13q14 (Francke and Kung, 1976, Yunis and Ramsay, 1978). Analysis of tumour karyotype also showed abnormalities of 13q14, but only in 10-20% of cases and the changes were seen against a background of other cytogenetic abnormalities (Benedict et al, 1983). Indeed, other changes such as extra copies of parts of 1q (Squire et al, 1985) or two extra copies of the short arm of chromosome 6, referred to as iso-chromosome 6p, were much more commonly seen than 13q14 abnormalities. Evidence implicating 13q14 as the location of the RB1 gene came initially for the study of a single patient. The child with retinoblastoma was found to have only 50% of the normal level of an enzyme, esterase D (ESD), in their blood cells (Sparkes et al, 1980). ESD had been shown to map to 13q14 (Van Heyningen et al, 1975) and exists as two polymorphic isoforms which can be distinguished by their electrophoretic patterns (Hopkinson et al, 1973). By studying large numbers of patients it was possible to ascertain whether Rb predisposition and ESD were genetically linked. Tight linkage between the two was demonstrated (Sparkes et al, 1980) placing the Rb predisposition gene in 13q14. The recessive nature of the Rb gene at the cellular level was suggested by the finding of a patient with Rb and only 50% of the normal level of ESD in their constitutional cells who had no cytogenetically visible deletion in 13q14 but had undetectable ESD activity in their tumour cells, suggesting loss of both copies of a part of 13q14 and, hence the Rb predisposition gene, had occurred. It is important to realise, however, that although ESD and the Rb predisposition gene are tightly linked, over 95% of patients with Rb have normal levels of ESD in their constitutional cells (Cowell et al, 1986). Were there two different genes involved in Rb, one on 13q14 for

the hereditary form and another, elsewhere for the sporadic type? Six patients were identified who were heterozygous for ESD isoforms in their blood cells. Some had unilateral tumours and no family history and others bilateral Rb and affected relatives. Although all the tumours had ESD activity, four had become homozygous for one of the two ESD isoforms implying that both sporadic and familial forms involved loss of alleles in 13q14 (Godbout et al, 1983). Similar observations were made using DNA rather than protein polymorphisms (Dryja et al, 1984). These DNA polymorphisms may be phenotypically silent but they alter the length of DNA fragments obtained when the sample is cut with restriction endonucleases and are, therefore, called restriction fragment length polymorphisms (RFLP) (Cavenee et al, 1984). Comparison of constitutional and tumour DNA using analyses for either ESD or RFLPs, showed that markers which were heterozygous in constitutional cells became homozygous in tumours in 70% of cases. This loss of heterozygosity (LOH) occurred by a number of different mechanisms. The whole of one chromosome 13 could be lost and the remaining, mutant, chromosome reduplicated. Non-disjunction of sister chromosomes during mitosis will result in two copies of the mutant chromosome in a daughter cell. In both mechanisms above, LOH will be found for markers all the way along chromosome 13. Mitotic recombination, and appropriate segregation at mitosis, will also give rise to a cell homozygous for the mutant predisposition gene but in this case LOH will be limited to the region distal to the recombination event. Point mutations and small interstitial deletions may not lead to LOH even of markers which closely flank the predisposition gene. Studies showed that mitotic recombination or non-disjunction events were the mechanism for loss of heterozygosity in 85% of cases of Rb in which it was found (Zhu et al, 1992).

1.3.3.b *The cloning of the RB1 gene*

Two approaches were used in the isolation of the Rb gene (RB1); chromosome walking from the adjacent ESD gene (Lee and Lee, 1986) and the random isolation of probes from a flow sorted chromosome 13 library (Lalande et al, 1984). Using the second approach, a probe that mapped to 13q14 was identified (Lalande et al, 1984). This probe was used to isolate a DNA sequence that was conserved between species and recognised a transcript from a foetal retinal cell line but failed to detect a transcript in Rb tumours (Friend et al, 1986). A cDNA clone (4.7R) identified structural abnormalities on Southern blotting in 15-40% of Rb tumours (Friend et al, 1987). When Northern blots were made from RNA derived from tumours, and probed with 4.7R it was found that, in 70% of Rb tumours, a full length transcript was apparently expressed. The finding that predisposing constitutional chromosome translocations involving breakpoints in 13q14 interrupted the coding sequence of the gene (Higgins et

al, 1989, Mitchell and Cowell, 1989) provided the final proof that RB1 was the retinoblastoma predisposition gene.

1.3.3.c *RB1 structure and Function*

The RB1 gene encompasses a huge 200 kb of genomic DNA and is composed of 27 exons most of which are less than 200 bp in length (McGee et al, 1989). RB1 contains a leucine zipper motif consisting of an arrangement of amino acids where every seventh one is a leucine residue. This structure is often found in transcription factors and appears to be important in facilitating dimerisation of proteins. The RB1 gene encodes an mRNA product 4757 bp long (Friend et al, 1987), and is translated to a protein (pRB) of 105-115 kD. pRB is a nuclear protein with a central role in cell division. It exists in unphosphorylated and phosphorylated forms, its phosphorylation being necessary for cells to progress from G1 to S phase of the cell cycle. Studies of the DNA tumour viruses demonstrated that the adenovirus virus E1A (Early region 1 A) protein formed a complex with pRB. pRB bound to those viral sequences known to be important for E1A oncogenic activity (Whyte et al, 1988). Proteins from other DNA tumour viruses, the large T-antigen of SV40 (DeCaprio et al, 1988) and the E7 protein of human papilloma virus (Dyson et al, 1989) acted in the same way preventing pRB from acting as a brake on DNA synthesis and cell division.

How does pRB exert its effect? It appears that transcription of a wide variety of cell cycle regulating genes, including *c-myc*, is controlled by the interaction of a cellular transcription factor E2F. E2F is thought to complex with the promoter regions of *c-myc* and other cell cycle genes initiating transcription of their mRNA. Unphosphorylated pRB complexes with E2F preventing it from activating genes responsible for cell division. Phosphorylation of pRB, or complex formation with the DNA virus oncoproteins, destroys its ability to bind E2F which can then stimulate cell growth unhindered (Nevins, 1992). Other proteins such as the Rb associated protein p107 and the cell cycle regulated cyclin A-CDK2 kinase complex also have a role.

pRB is central to the control of cell cycling, but does it have a role in differentiation? After all, histologically Rb tumours resemble undifferentiated, foetal retinal cells and a failure of normal retinal differentiation was suggested as a primary event in Rb development. Evidence for the importance of pRB in differentiation comes not from retinal cells but from muscle. In foetal muscle cells a protein, myo-D, is required for differentiation of primitive myoblasts into myotubes which lack the potential to divide. Myo-D interacts with unphosphorylated pRB both *in vivo* and *in vitro* through the same region that controls adenovirus E1A and SV40 large T-antigen interaction (Gu et al, 1993) In this way pRB appears to alter the ability of Myo-D homodimers, but not heterodimers, to bind DNA and cause muscle differentiation.

1.3.3.d *Tissue specificity of RB1*

As pRB is central to the control of cell division in all cells does this correlate with an increased risk of all malignancies in patients with familial Rb? Patients cured of Rb are indeed at increased risk of other malignancies and it is those patients who carry the germinal (familial) mutation and not the sporadic cases who are at risk. Treatment of patients with radiotherapy or chemotherapy with alkylating agents such as cyclophosphamide is associated with an increased risk of second malignancy. The most common second tumour is osteosarcoma with a two-hundred to five hundred fold increase of incidence for limb osteosarcomas and a risk at least twice that high for tumours within the radiation field (Hawkins et al, 1987). Other second tumours include soft tissue sarcomas, central nervous system tumours and basal cell carcinomas of the skin (Draper et al, 1986). How frequently individuals with genetic Rb develop any second tumour is unclear. Some groups estimate an 8% cumulative risk at 20 years post treatment for Rb (Hawkins et al, 1987) others as high as 90% at 30 years (Abramson et al, 1984). At a cellular level many tumours show mutations of RB1 or changes in its expression. Studies of breast and lung tumours, to which Rb patients do not appear to be predisposed, show LOH at loci on chromosome 13. In small cell lung carcinoma (SCLC) 91% of cases exhibit LOH for 13q (Yokota et al, 1987) . Furthermore, absence of expression of RB1 mRNA was found in 60% of SCLC cell lines, 18% of which were found to have structural DNA abnormalities in 13q14 (Harbour et al, 1988). However, these mutations are less likely to represent primary causative genetic events but rather yet another molecular change in the progression of the tumour. One could easily imagine that the loss of function of a gene that usually acts as a brake on cell division would give a tumour cell considerable growth advantage. Hence, clonal selection for cells that acquire mutations in RB1 may be seen in many tumours as they progress.

1.3.3.e *RB1 and patient management*

As Rb can be familial, the diagnosis of Rb in an individual child has far reaching consequences for the family as a whole. All other children in the family may need to be screened and advice given to other family members as to whether or not they are carriers of a predisposing mutation. The possibility of incomplete penetrance means that all children in the family (sibs and cousins) and all children of affected parents have to be considered at risk. Screening consists of frequent ocular examination under a general anaesthesia which has emotional and physical implications for the child and constitute a major clinical commitment for ophthalmology departments. Yet, at least half of the children screened would not have inherited the mutant RB1 allele and are at no more risk of developing the tumour than the general population. A means of identifying those at risk therefore would be a significant advantage in the clinical

management of this disease. Cloning the gene led to the identification of naturally occurring phenotypically insignificant polymorphisms within the gene which allow individual chromosomes to be identified (Wiggs et al, 1988) By comparing the patterns of inheritance of these markers with that of the disease phenotype it is possible to track the defective gene through generations allowing family members carrying the predisposing mutation to be identified. Using a panel of RFLPs it is possible to offer screening to 95% of families and prenatal detection is also possible (Onadim and Cowell, 1991) Only 10-20% of Rb patients have a family history of the disease, the majority presenting as apparently sporadic cases. In order to offer genetic counselling to the families of these patients, the causative mutation must be identified. It would be a huge undertaking to sequence the whole 200 kb genomic region for each patient and family member. However, each of the 27 exons can be amplified by the polymerase chain reaction (PCR) and the mutant exons, even those with a single base pair change, will show altered mobility on denaturing polyacrylamide gel electrophoresis; a technique known as single stand conformation polymorphism (SSCP) (Hogg et al, 1992) The rest of the family can then be screened for the presence of the abnormal exon.

1.3.4 WILMS TUMOUR

Rb provides the paradigm for understanding tumour suppressor genes with the tumour, which resembles its foetal tissue counterparts , arising from the loss of function of both copies of the RB1 gene. Wilms tumour (WT) is another embryonal tumour of childhood and, in the classical form, has triphasic histology: stromal, blastemal and epithelial elements are all seen. WT occurs in 1:10,000 normal children and has a peak incidence between 3 - 6 years of age. It is usually unilateral, but bilateral tumours occur in 6% of cases and 8% of unilaterally affected patients have more than 1 primary tumour in their involved kidneys. Bilateral cases also present earlier than unilateral WT and so it was not surprising that Knudson and Strong (1972) postulated a 2-mutation model for WT development. However, several lines of evidence suggested that the situation in WT was more complex than for Rb. The incidence of familial WT is low, 2.4% and, unlike Rb tumours in familial cases did not appear earlier than their sporadic counterparts as would be expected from a 2-hit model (Bonaiti-Pellie et al, 1992) . WT is associated with other malformations and four main complex phenotypes were identified; WAGR, Beckwith Wiedemann syndrome, Perlman syndrome and the Denys Drash syndrome.

In some patients WT co-existed with aniridia, genitourinary malformations and mental retardation (the WAGR phenotype) and these patients had constitutional interstitial deletions of 11p13 (Riccardi et al, 1978) suggesting that genes responsible for this collection of phenotypes were located there.

The second association was with Beckwith-Wiedemann syndrome (BWS). This syndrome is characterised by abnormalities of the growth of numerous organs (tongue, pancreatic islets, hemihypertrophy) and predisposes children to several embryonal tumours especially WT, though hepatoblastoma, adrenocortical carcinoma and rhabdomyosarcoma are also seen. Unlike WAGR patients, most patients with BWS have a normal karyotype although duplications of 11p15 have been reported and, in familial cases of BWS, there is close genetic linkage with genes coding for β–globin, insulin and Harvey-*ras*, all located in the 11p15 (Koufos et al, 1989). The two syndromes (WAGR and BWS) also differ in the type of intra-renal abnormalities that co-exist with WT. Metanephric blastemal tissue of the developing kidney that fails to differentiate persists as nephrogenic rests (NR). These foci of relatively undifferentiated tissue can be diffuse or multifocal and kidneys so affected are labelled as having nephroblastomatosis. NR are closely associated with WT, the two conditions often co-existing in the same kidney, and have been thought of as pre-malignant lesions, although, as 1:100 children have NR and only 1:10,000 develop WT, the majority of NR must not 'progress' (Beckwith et al, 1990). Two patterns of NR occur: intralobar NR (ILNR), which are thought to occur in early embryogenesis and perilobar NR (PLNR) which occur later. Patients with WAGR have the (earlier) ILNR whereas those with BWS have the (later) PLNR. Although there is no histological difference in the WT found in the two syndromes, tumours arise later in the BWS group (Slater and Mannens, 1992).

Two more syndromes that predispose to WT development have been described. Perlman syndrome is characterised by a high neonatal death rate, dysmorphic features, mental retardation and a high risk of developing WT, with most tumours presenting in the first year of life and often involving both kidneys. The syndrome is very rare, with only twelve recorded cases in the world literature, but it is distinct from BWS and may involve a different genetic locus , also on 11p15 (Grundy et al, 1992).

The Denys Drash syndrome is another rare condition in which a predisposition to WT is seen. Affected patients present with severe urogenital aberrations which result in heavy proteinuria and glomerular disease, abnormal genitalia (often sufficiently severe to cause pseudohermaphroditism) and prominent ILNR. Patients who survive the renal dysfunction are virtually certain to develop WT. The molecular basis of the predisposition to WT in Denys Drash syndrome has now been elucidated (reviewed in Hastie 1992).

Analysis of the rare familial Wilms showed that in the majority of cases WT predisposition was not linked to any markers on chromosome 11 (Grundy et al, 1988). So, unlike Rb, at least three genes seem to be important in WT development, located at 11p13, 11p15 and outside chromosome 11 altogether.

1.3.4.a *WT1 gene*

Cloning the WT1 gene from 11p13 started with definition of the smallest region of overlap of the deletions found in WAGR patients. A panel of somatic cell hybrids which carried various deletions and translocations of human 11p13 allowed localisation of probes within 11p13. Molecular cloning of the region between the markers which defined the smallest region of overlap identified a gene (WT1) spanning 50 kb of genomic DNA and encoding an mRNA 3 kb long (Call et al, 1990, Gessler et al, 1990). The WT1 gene product is located in the nucleus and has DNA binding motifs in the form of 4 carboxy-terminal zinc 'fingers' of the Cys2-His2 variety and a proline-glutamine rich amino terminal region. The zinc finger domains share homology with mammalian polypeptides involved in the early growth response (EGR and CR families) which mediate transition from G0 to G1 in the cell cycle (Rauscher et al, 1990). *In situ* hybridisation studies using WT1 mRNA probes showed WT1 expression in developing gonad and early epithelial structures within the developing kidney where it mediates the differentiation of metanephric mesenchyme to an epithelial phenotype (Pritchard-Jones et al, 1990). In tumours WT1 expression is found only in neoplastic structures whose normal counterparts also express the gene, with stromal predominant tumours showing very low levels and epithelial predominant tumours very high expression (Pritchard-Jones and Fleming, 1991). It may be that the failure of renal epithelial differentiation 'forces' some cells down the line of stromal differentiation, thereby explaining the persistence of stroma which does not express WT1 in tumours. The role of WT1 in disease is best shown in the Denys Drash syndrome . Mutations of one or other of the zinc fingers of WT1 have been observed in the constitutional cells of all ten patients studied by Pelletier et al (Pelletier et al, 1991). Nine of the ten patients had mutations in the third zinc finger and in the remaining case it was the second zinc finger that was affected. Abnormalities of WT1 in patients with the Denys Drash syndrome findings have been confirmed by other authors (Baird et al, 1992). However, in sporadic WT, although some tumours have been shown to have abnormalities in WT1 (Haber et al, 1990) the majority have normal WT1 structure, so other genetic changes must be important in the development of sporadic WT.

1.3.4.b *11p15 BWS and Genomic Imprinting*

Although only 30% of WT show allele loss for 11p, LOH is more common in 11p15 than 11p13 (Mannens et al, 1990, Wadey et al, 1990). One of the unexpected results of these studies was the observation that maternal alleles were lost preferentially (Pal et al, 1990, Schroeder et al, 1987). The activity of autosomal alleles is normally independent of their parental origin but when the sex of the parent from which the allele is inherited affects its expression the gene is said to be genomically imprinted (Feinberg, 1993). 11p15 has been shown to be important in WT with both paternal

uniparental disomy of 11p and paternal 11p15 duplications reported (reviewed in Slater and Mannens 1992). The imprinting, however, need not be permanently stamped on a somatic cell and it may be lost during malignant progression. Indeed, 'relaxation of imprinting' has been postulated as a mechanism of uncovering oncogenic mutations in a way analogous to LOH for non-imprinted alleles (reviewed in Rainier et al, 1993). The nature of the gene or genes in 11p15 involved in WT1 predisposition remain unclear.

1.3.5 *P53*

WT1 mutations are probably only one of a number of genetic changes required for WT formation and the effects of WT1 mutation seem to be specific for the urogenital system. This tissue specificity is also seen in Rb which is primarily involved in two malignancies (Rb and osteosarcoma) with only secondary involvement in other tumours. By contrast, the p53 oncogene is primarily involved in a large number of separate tumours derived from both embryonal and fully differentiated tissues. The nuclear phosphoprotein, p53, was originally thought to be a dominantly acting oncogene and the evidence for this contention seemed strong. It was originally discovered in extracts from transformed cells and was found to complex with the large T antigen of SV40 (Lane and Crawford, 1979). High levels of p53 were often found in DNA derived from tumour cells and the protein product was found to have a longer half life in transformed than in normal cells. Acting along with *ras*, a variety of genomic DNA clones of p53 were found to be capable of transforming rat embryo fibroblasts in culture (Parada et al, 1982). However, all of these transforming clones turned out to be in mutant forms of p53 (Hinds et al, 1989). The normal, wild type, p53 acted as a tumour suppressor in that it could inhibit the growth of malignant cells in culture (Baker et al, 1990) and decrease the ability of these cells to form tumours in animals (Chen et al, 1990). Mutations in p53 and often high levels of mutant gene expression were found in a very wide spectrum of tumours both embryonal malignancies such as alveolar rhabdomyosarcoma and other paediatric sarcomas and tumours derived from terminally differentiated cells (e.g. breast, lung). In the 1960s a cancer predisposition syndrome was identified which conferred the risk of not just a single tumour type as in Rb, but of a wide spectrum of tumours. Named after the authors of the paper describing it (Li and Fraumeni, 1969), the Li-Fraumeni syndrome (LFS) consisted of familial childhood sarcomas and a strikingly increased risk in the family of other neoplasms, particularly of breast carcinoma, brain tumours, osteosarcoma, leukaemia and adrenocortical carcinoma. Other frequently observed tumours were melanoma, gonadal germ cell tumours and carcinoma of the prostate pancreas and lung. In family members all of these tumour types developed at an early age and were often multiple. Sporadic cases of all these tumours were also shown to have abnormalities of the p53 gene. The pattern of early and multiple tumour formation in an affected family was reminiscent of

that of Rb, but RB1 mutations were not found to be responsible. The number of affected families with surviving members was too small for genetic linkage analysis to be performed so, a 'candidate gene' approach was taken and in 1990 Malkin et al showed that affected members of LFS families had constitutional abnormalities of p53. What was remarkable was that a single genetic defect predisposed to such a large variety of malignancies. Some tissue specificity exists however, in that both sporadic colon and breast carcinoma show p53 mutation. The incidence of colon and breast carcinomas is similar in the general population but in LFS breast carcinoma is very common indeed and colon carcinoma much less so, suggesting p53 to be a more critical step in the genesis of breast carcinoma than in bowel malignancy. How do p53 mutations lead to tumour formation? In the paradigm of Rb, mutations simply render alleles non-functional and both RB1 alleles have to be lost for a tumour to develop. This is not the case in p53 where the mutant product is often found at high levels due to its long half life. Cells transformed by *ras* and mutant p53 have been developed in which the amount of p53 can be controlled by an inducible promoter. Decreasing the synthesis of mutant p53 leads to a loss of the transformed phenotype. Conversely, increasing mutant p53 levels can cause tumour formation even in the presence of normal wild type p53. Two explanations have been suggested. In the first, wild-type p53 complexes with mutant p53 and is unable to exert its action as a tumour suppressor, a so-called 'dominant negative' effect. In the second, the mutant p53 is thought to acquire oncogenic potential of its own, a 'gain of function' mutation. These theories need not be mutually exclusive and the different modes of action may depend on the nature of the mutation that has occurred (reviewed in Levine 1991). The function of wild-type p53 can also be decreased by forming complexes with the protein products of the DNA tumour viruses or with amplified cellular genes such as MDM2. As MDM2 is amplified in a large proportion of human sarcomas, tumours known to be part of LFS, the inactivation of wild-type p53 by complexing with MDM2 may be an important mechanism in disease formation (Vogelstein and Kinzler, 1992).

The site and type of p53 mutation varies between tumours. Although the study of LFS families initially suggested a hot-spot between codons 245 and 258, mutations can occur in any of the codons which are conserved between species. Some striking correlations exist with tumour phenotype. For example, Chinese patients who develop liver tumour in which aflatoxin-B-1 and hepatitis-B are risk factors, show mutations at a single nucleotide pair of codon 249. The type of mutation also varies with transitions predominating in colon, brain and lymphoid malignancies, whereas lung and liver tumours exhibit transversions (Hollstein et al, 1991).

What does p53 actually do? Evidence now points to its role in programmed cell death (apoptosis) soon after cell injury. Apoptosis is needed in the course of the embryonic development of various organs and in the maturation of the immune system.

Two research groups have developed strains of mice in which p53 genes were disrupted in the germ line (p53 knockout mice). By selectively breeding mice which were heterozygous for the disrupted gene, offspring in which there was homozygous deletion of p53 could be obtained. The mice were developmentally and immunologically normal suggesting that apoptosis during their development had proceeded without hindrance in spite of the absence of one or both p53 genes. Both groups showed that thymocytes from these mice are very resistant to radiation, failing to undergo programmed cell death, mice carrying homozygous inactivation being even more resistant to the effects than those which were heterozygous p53 mutations. However, when phorbol ester was administered to the knockout mice to mimic the 'natural' signals responsible for programmed cell death of T-cells in the thymus, normal apoptosis was seen (Clarke et al, 1993, Lowe et al, 1993). The mechanism is unclear, but p53 may function as an inducer of apoptosis in cells that have undergone genetic damage and, in its absence, cells which have acquired oncogenic mutations survive and cancer formation occurs.

1.4 THE MULTI-STEP NATURE OF CANCER

We have seen that two classes of gene important to tumour development exist. Oncogenes, acting dominantly and tumour suppressor genes, the loss of function of which, results in malignant development. No single genetic change can induce tumour development, multiple steps are necessary. The number of genetic changes can be as little as two, as in the loss of function of both RB1 genes. Conversely, many steps may be necessary, as in colon cancer, where the loss of function of three separate tumour suppressors and the activation of the oncogene, *ras* , is necessary (Vogelstein et al, 1988). With this background in mind, let us now turn to neuroblastoma.

1.5 NEUROBLASTOMA

Neuroblastoma (Nb) is the commonest solid tumour of childhood, accounting for 8% of all paediatric malignancies. It is a tumour of the post-ganglionic sympathetic nervous system, occurring most frequently in the adrenal glands, or in collections of nerve cells on either side of the spine (paraspinal ganglia of sympathetic chain). It has a capacity to metastasise to bone, bone marrow, lymph nodes, skin and more rarely, brain. Most, but not all Nb tumours, produce excess catecholamines which account for the high blood pressure and irritability seen in patients with the disease. The excess catecholamines are excreted in the urine where their detection greatly aids the diagnosis of Nb.

1.5.1 *Epidemiology*

Every year 9 in every million children under the age of 15 will develop Nb (Young et al, 1986). Of these, 41% of patients will be below 2 years of age at presentation and 91% below the age of 10. Boys are more likely to develop Nb than girls, the ratio of males to females is 1.22:1 (Kinnier Wilson and Draper, 1974).

1.5.2 *Complex phenotypes and Nb*

The analysis of the complex phenotypes was instrumental in identifying the location of genes important in the development of Rb and Wilms' tumour. Do any such phenotypes exist for Nb?

Beckwith Wiedemann syndrome has been associated with a wide variety of embryonal tumours, with some groups putting the risk of a patient with BWS developing a malignancy at 7.5% though this figure is probably elevated due to a reporting bias (Wiedemann, 1983). The tumour most commonly seen is WT, with adrenocortical carcinoma, hepatoblastoma and rhabdomyosarcoma also reported (Sotelo et al, 1980). It is thought that these tumours arise from cells that have failed to differentiate fully, (embryonal rest cells as discussed under section 1.3.4). Only 6 neural crest tumours associated with BWS have been reported in the literature. 2 ganglioneuromas, which can arise from a neuroblastoma by differentiation (Wiedemann, 1983, Sirinelli et al, 1989), one ganglioneuroblastoma (Chitayat et al, 1990) and 3 neuroblastomas (Emery et al, 1983, Sirinelli et al, 1989). No abnormalities of constitutional karyotype were found in any of the cases, but high definition molecular mapping of the 11p15 region was not performed. It is uncertain whether this frequency of Nb is more than would be expected by chance alone.

The second candidate for a complex phenotype associated with Nb is the group of conditions that form the neurocristopathies. Neural crest cells arise from the epithelium at the site of the closure of the neural tube. Cells migrate from the neural

tube and differentiate into melanocytes, Schwann cells, peripheral nerves, sympathetic ganglia, cells of the adrenal medulla and the parasympathetic ganglia of the gut. The fate of individual neural crest cells is determined by their position in the neural crest, with different regions giving rise to different ganglia (Boland, 1974). The failure of migration of neural crest cells to the parasympathetic ganglia in the gut gives rise to a failure of normal gut motility called Hirschprung's disease. A gene for Hirschprung's disease has recently been mapped to the pericentric region of chromosome 10 (Angrist et al, 1993). Four cases of the co-existence of Hirschprung's disease and neuroblastoma have been reported as well as one patient (Chatten and Voorhess, 1967) who was found to have small foci of neuroblastoma in the adrenal gland (neuroblastoma *in situ,* which may be more akin to the nephrogenic rests of WT than a true malignancy). In three of the four case (reviewed by Michna et al, 1988) the tumour was a metastatic thoracic neuroblastoma. In the most recent case, the child had a multifocal ganglioneuroblastoma (Michna et al, 1988).

The patient reported by Chatten and Voorhess was part of a familial Nb aggregation, a sib dying of multifocal Nb at 2 years of age. One other familial aggregation has been reported (Clausen et al, 1989) in which cases of neuroblastoma, ganglioneuroma, neurofibromatosis type 1 (NF1), Hirschprung's disease and jaw winking (a condition due to abnormalities of cervical ganglia) have been reported in three generations of one Dutch family. No individual in this family had more than one of the above abnormalities and high resolution constitutional karyotype analysis of the propositus and her nuclear family were normal. Of particular interest is the fact that individuals with NF1, a well-characterised familial disorder, are not considered to be at increased risk of Nb. It is difficult to decide whether the diseases are all related, with a single genetic change being responsible for aberrant neural crest cell migration and increased predisposition to malignancy, or whether cells which migrate abnormally are subjected to more genetic change and therefore acquire more potentially oncogenic damage. Is it genetics or geography that is at fault? Either way no constitutional karyotypic abnormalities were found which could act as clues to the localisation of an important gene. A mouse model for neurocristopathy has been developed. A strain of transgenic mice was developed in which the SV40 large T-antigen and the prokaryotic lac Z gene were both under the control of myelin basic protein promoter. These mice developed Schwannomas and facial osteogenic sarcomas at 5 months of age. Both osteoclastic cells of the facial skeleton and Schwann cells are considered to be neural crest derivatives. Although interesting, its relevance to Nb is unclear especially as the mice did not develop Nb (Jensen et al, 1993).

1.5.3 *Familial Nb*

We have noted two families with more than one case of Nb in section 1.5.2. but how common is familial aggregation of neural crest tumours? Knudson and Strong (1972) reviewed the case records of the MD Anderson Hospital in Texas and supplemented their data with a review of the literature. They found 25 cases of Nb where a familial aggregation was present. The age of presentation and the number of primary tumours at diagnosis of this small group was compared with those of a much larger number (504) of unselected, presumably non-familial, cases. They found that in familial cases, 56% were diagnosed in the first year of life compared with 26% in the unselected group and, whereas cases with multiple primary tumours form only 5% of all Nb, they comprise 23% of familial cases. Familial Nb, therefore, appeared to present earlier and was more likely to be multiple which led to them proposing that Nb, like Rb, was due to two hits, one on each allele of a critical gene. The prediction was that 50% of the offspring of multiple tumour cases would be predisposed gene carriers and they calculated the penetrance of Nb to be 63%. They also deduced that 22% of all cases of Nb occur in gene carriers. Does more recent data support this contention? The last large literature review of familial neuroblastoma (Kushner et al, 1986) identified 23 separate families in whom Nb or ganglioneuroblastoma occurred in more than one family member. There were 55 patients in all and the relationships between affected subjects included siblings, half siblings, cousins, parent/child and monozygotic twins. The four monozygotic twins were interesting as in each twin pair both were diagnosed within two months of each other. The parent/child aggregations, all but one parent had localised stage I disease and in the one case where the father had an unresectable neuroblastoma, the behaviour of the tumour was unusual in that he was a long-term survivor after sub-total resection and low dose radiotherapy alone. As in the series reported by Knudson and Strong, the age at diagnosis and the number of primary tumours at diagnosis were calculated for familial and non-familial groups. In a series of sporadic cases of Nb the median age at diagnosis was 2.5 years with 25% diagnosed under the age of a year. For the familial group in the Kushner series the median age is 9 months and more than 60% of cases, were identified before the patient's first birthday. Multiple primary tumours were found in 23% of familial cases compared to 5% of sporadic Nb. This status certainly seems to support the Knudson and Strong hypothesis but what of the quoted genetic risks? The risks to offspring of non-familial cases of Nb certainly seems to be much less than the 5-8% that can be calculated from Knudson's data (reviewed in Kushner, 1986). However, within the families which had been identified with parent/child cases of Nb there is evidence of a transmissible mutation. Here, Knudson's suggestion is born out as if the mutation is constitutional, the risk of sibs being affected can be calculated at 30% (50% which is the chance of inheriting the parent's affected allele x 63% which is Knudson's estimate of the gene penetrance). If

the parent's mutation is acquired as a result of a post-meiotic event the risk of further children developing the disease would fall. The data are difficult to interpret for a number of reasons and have to be treated with caution. In particular, the numbers are very small and, unlike Rb, the affected children often die before reproductive age. Furthermore, the calculations are based on the incidence of clinically obvious disease in which cases of neuroblastoma *in situ* , which has an incidence at least 50 times greater than that of overt disease, are not included (Beckwith and Perrin, 1963). If these small rests of tissue do represent incipient tumour due to a germline mutation, not including them would seriously affect the calculation of the incidence of familial Nb. Unlike Rb, where RFLP analysis could identify asymptomatic family members carrying predisposing mutations, we have no genetic marker to follow in Nb. There has been a report (Rudolph et al, 1988) of a familial transmission of a fragile site at 1p13.1 (induced by culturing cells in folate deficient medium and treating them with aphidocholine). In this report a family with neural crest tumours was analysed and a patient with ganglioneuroblastoma and her mother, shared the same fragile site. A sib had died of congenital Nb but no lymphoblasts were available to test. The importance of this finding is unclear and it may represent a chance association rather than true genetic linkage.

We can only decide if a given case is caused by a hereditary predisposition to Nb by looking for evidence for regressed disease in family members, focusing particularly on the families of cases in multiple primary tumours. Robertson et al (1991) described a three-generation family, the four members of which developed single or multiple neural crest tumours. Their conclusions that careful family histories ought to be taken and DNA from family members stored for future analysis are very pertinent. It may be that familial Nb, like familial WT, represents a rare sub-group of tumours and it is possible that the causative mutation in familial cases is different to that of sporadic tumours.

1.5.4 *Clinical aspects*

A brief overview of the clinical aspects of Nb will allow a better understanding of the relationship between disease phenotype and tumour biology. In order to ensure that correct treatment is given to an individual child and that the results of treatment can be analysed, tumours are divided into clinical stages depending on the degree to which the tumour has spread. Up to 1988, the most usual system for staging was that developed by Evans (Evans et al, 1971). This staging system is shown diagrammatically in figure 1.2 and the key points are summarised below :

Stage I: small tumours, localised to the organ of origin.

Stage II: unilateral tumours, which had spread further than the organ in which they had originated, but which had been completely resected.

Stage III: tumours extending across the midline. There was uncertainty as to what constituted the midline as well as whether contralateral lymph node involvement with tumour constituted an extension of the tumour across the midline.

Stage IV: metastatic disease

Stage IVS: metastatic disease with a special pattern of spread.

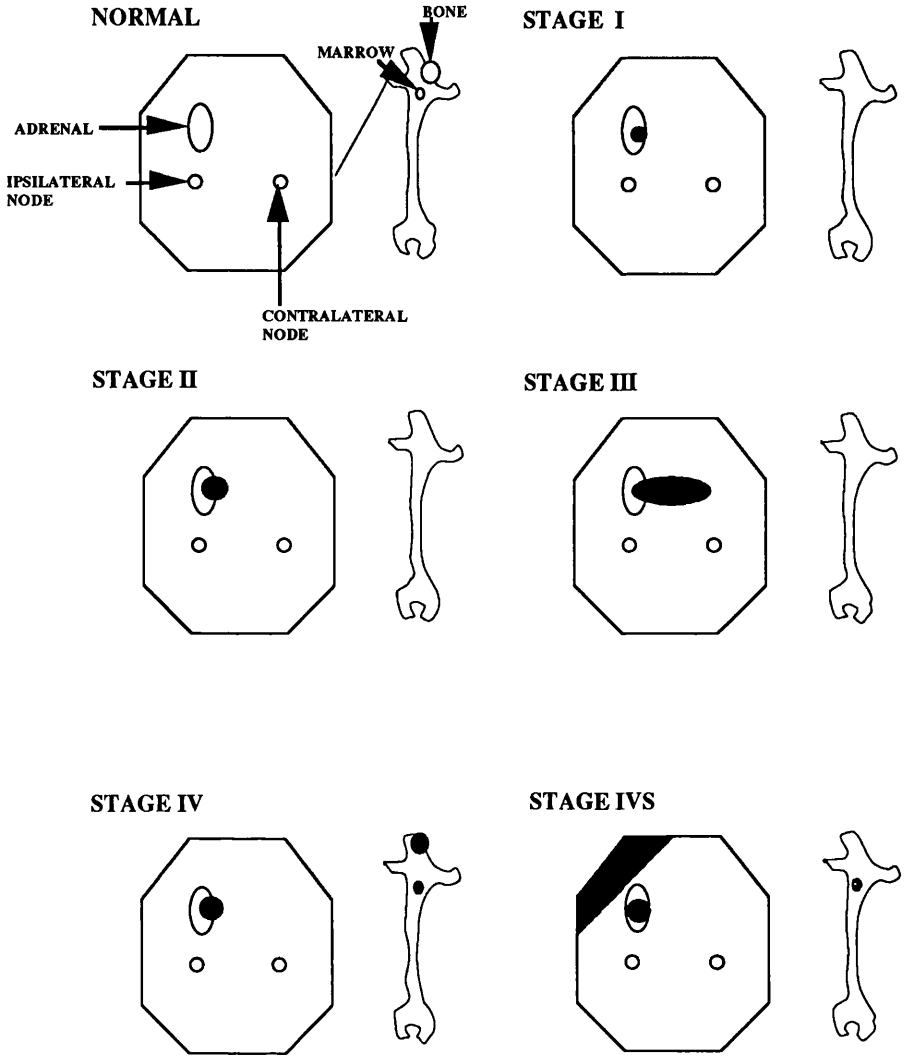


Figure 1.2. A diagrammatic representation of the Evans staging system for Nb. The location of potential sites of disease is given in the diagram entitled 'normal'. Sites affected with disease are shown in black, whereas the open (unfilled) sites are clear of tumour.

Stage IVS disease is particularly fascinating. Occurring mainly in children under 1 year of age it is defined as a small Stage I or II primary tumour with a particular pattern of metastatic disease. Metastatic deposits are found in the skin, presenting as bluish lumps, in the bone marrow, where they should form only a small proportion of nucleated cells and, most significantly, to liver, with huge, homogenous hepatic involvement. Indeed, it is the rapid, severe liver enlargement that causes the mortality in this stage, usually from respiratory failure by pushing the diaphragm up, thereby decreasing the space the lungs have to expand into. Here we have aggressive, fast growing, metastatic malignancy, and yet it has the capacity to resolve spontaneously without treatment (D'Angio et al, 1971, Evans et al, 1980). This observation led to the suggestion that stage IVS was simply a transient failure of differentiation of a neurocristopathy rather than a true malignancy (D'Angio et al, 1971). Some Stage IVS tumours, however, after the initial spontaneous disappearance, recur and go on to kill the patient (Nakagawara et al, 1990).

There is a broad correlation between stage and outcome in that patients with stage I and II tumours do well with minimal treatment and yet most patients greater than one year of age with stage IV tumours die in spite of intensive, multi-modality therapy (Evans et al, 1987). The huge discrepancy in survival between localised and disseminated tumours led the Japanese to set up a programme for mass screening for Nb by examining the urine of every infant at six months of age for the presence of elevated urinary catecholamines (Sawada et al, 1982). The rationale was that if Nb behaved like colon cancer, starting off as a small localised tumour then gradually growing and acquiring metastatic potential, identifying and removing it at an early stage would improve the chances of survival. Population screening was adopted and many infants identified as having sub-clinical Nb. These infants had low stage disease (Stage I, II and IVS) and the prognosis of patients with Nb (the number of survivors divided by the number of cases) improved (Sawada et al, 1987). However, there was concern that the incidence of Nb was rising and that screening identified patients with disease that may well have disappeared on its own with no treatment and without coming to medical attention. The inclusion of these patients would obviously increase the prognosis, but had screening actually reduced the death rate from neuroblastoma? Although there were claims that mass screening had been successful at reducing the death rate from Nb (Hanawa et al, 1990), Japanese population based statistics of the cause of death had not been very accurate before the start of the screening programme and it was difficult to be certain of the fall in mortality from Nb. A further worrying observation was that some cases, which had proved negative on mass screening, presented months later with aggressive Stage IV disease which went on to kill the

patient. Studies in Quebec and France are proceeding in order to attempt to ascertain whether mass screening does indeed have a role (Woods et al, 1992).

The point of developing a staging system was to be able to predict the outcome of a given patient. However, some stage II tumours (which would be treated with surgery alone) progressed rapidly and were fatal and conversely some stage IV tumours (which would be treated with aggressive chemotherapy and radical surgery) resolved spontaneously. This relative inability of the staging system to give an accurate prognosis constituted a great problem in designing rational management plans for individual children. As the system of staging did not accurately differentiate the groups of children who need markedly different types of therapy, a revised staging system was introduced (Brodeur et al, 1988). The revised system, which used Arabic rather than Roman numerals, is shown diagrammatically in figure 1.3 and summarised below, the major changes being in the definition of the old Stage II and III Nb.

Stage 1: localised tumour confined to the area of origin. Complete gross excision with or without microscopic residual disease. Identifiable ipsilateral and contralateral lymph nodes negative microscopically.

Stage 2a: Unilateral tumour with incomplete growth excision. Identifiable ipsilateral and contralateral lymph nodes negative microscopically.

Stage 2b: Unilateral tumour with complete or incomplete growth excision with identifiable disease in ipsilateral regional lymph nodes. Identifiable contralateral lymph nodes negative microscopically.

Stage 3: Tumour infiltrating across the mid-line with or without regional lymph node involvement. Or, unilateral tumour with contralateral regional lymph node involvement. Or, midline tumour with bilateral regional lymph node involvement.

Stage 4: Dissemination of tumour to distant lymph nodes, bone, bone marrow, liver and/or other organs (except as defined in Stage 4S).

Stage 4S : localised primary tumour as defined for stage 1 or 2 with dissemination limited to liver skin and/or bone marrow.

Although this new staging system helped reduce some of the anomalies created by the original Evans classification, it still proved to be inaccurate at predicting the prognosis for a small percentage of patients. Hence, interest developed in the analysis of other factors which could accurately forecast disease outcome. Research was divided into the analysis of clinical criteria and studies of the histopathology and molecular biology of tumour tissue.

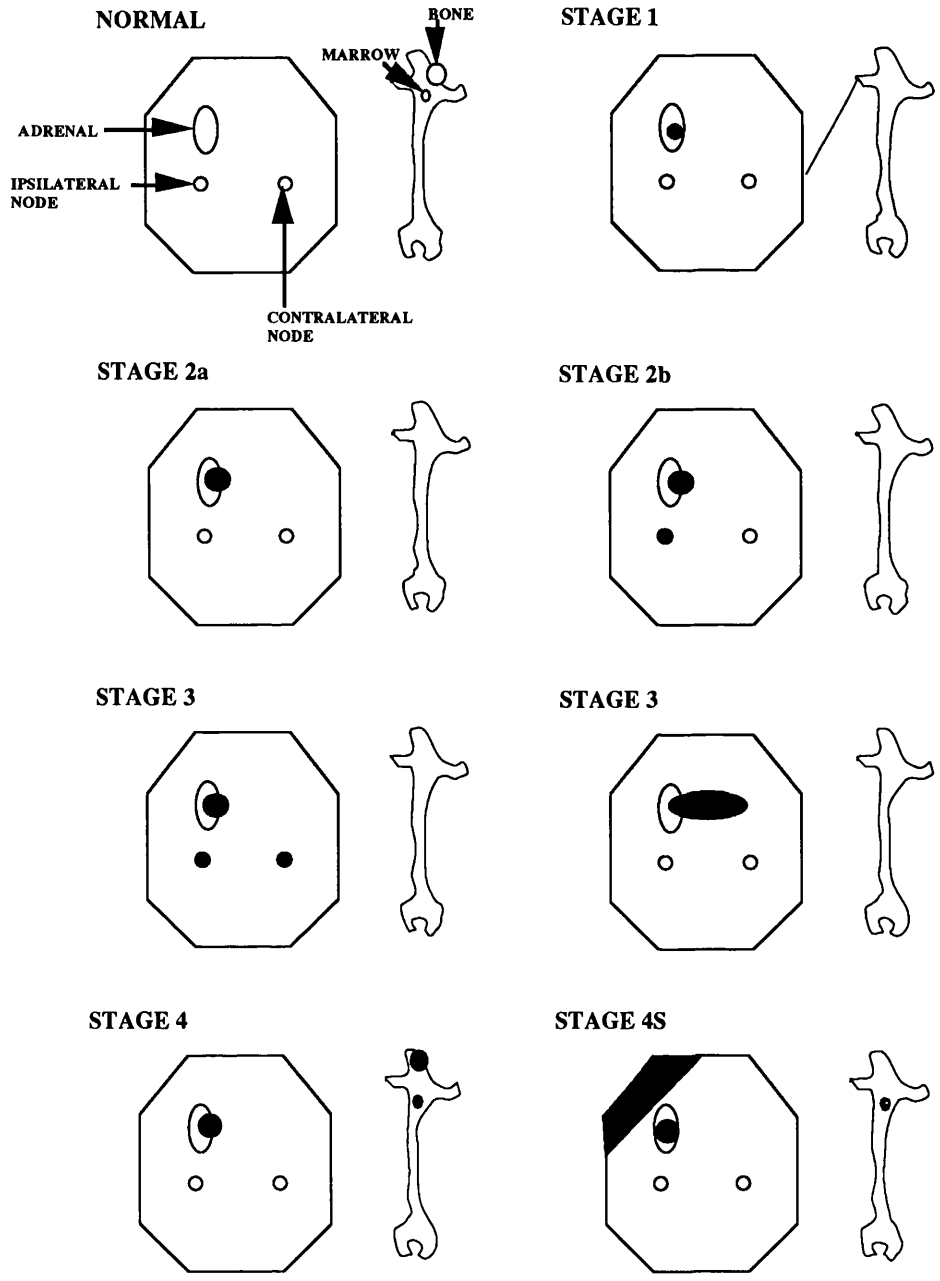


Figure 1.3. A diagrammatic representation of the INSS staging system for Nb. The potential sites of disease are shown in the diagram entitled 'normal'. The extent of disease in a given stage is shown by the black (filled) sites whereas unfilled sites are free of tumour. Compared with the Evans classification, this scheme subdivided stage 2 disease into 2a (without node involvement) and 2b (with node involvement). The definition of stage 3 disease is broader than in the Evans classification, also including unilateral tumours with contralateral nodal involvement.

The clinical criteria used were the patient's age and certain biochemical parameters in blood and urine. Patients under one year of age usually did well, with a much higher long-term survival than their older counterparts. This was partly due to the fact that infants have a higher incidence of early stage disease than older children (40% versus 20%) (Bernstein et al, 1992). However, even infants who have true metastatic Stage 4 disease seem to do better than older children with long-term survival of about 50% compared to a 6 year survival of 10-20% in older children (Bernstein et al, 1992). If the actual survival curves for infants with stage 4 disease are examined, two distinct populations are found. One population has a poor chance of survival, similar to that of older children, whereas the other group has an excellent prognosis (Finklestein et al, 1979). Molecular biology has identified factors that allow the two groups to be differentiated (section 1.5.5). The patient's sex makes a difference, because males have a greater likelihood of presenting with metastatic disease, as a group they have a poorer chance of survival (Pritchard et al, 1989).

Certain biochemical measurements correlate very well with survival. Thus, elevation of neurone-specific enolase or the muscle specific enzyme, lactate dehydrogenase is predictive of worse outcome. Similarly, elevation of ferritin, an iron binding protein and acute phase reactant, is also associated with poor prognosis. The precise type of catecholamine excreted in the urine also has been associated with differences in long-term survival. These factors, along with surgical stage have been combined in complex algorithms that allow patients to be stratified into 'risk groups' (Evans et al, 1987).

Pathological classifications of tumours have also been developed, which allow good and bad prognostic groups to be differentiated. The presence of neural differentiation and abundant stroma were classified as favourable features in the comprehensive but unwieldy system proposed by Shimada et al (1984). Simpler variations of the Shimada classification have been developed (Joshi et al, 1993). Patients have been stratified into 'risk groups' using algorithms that use both histology and biochemical criteria (Silber et al, 1991). Although many of these algorithms do allow more accurate prognosis to be given to an individual patient they do not increase our understanding of the basic biology of the disease. Studies of the cell and molecular biology of Nb, however, have given us clear insights into the phenotypic effects of certain genetic changes.

1.5.5 *Cell and molecular biology of Nb*

Clinical staging procedures do not always accurately predict an individual patient's prognosis and the treatment options in Nb vary widely (from minimal to very intensive). Hence, many of the analyses of the biology of Nb have attempted to identify molecular abnormalities which will allow definition of prognostic groups, though some groups have investigated genetic changes at a more fundamental level as well. The molecular abnormalities that have received most attention are abnormalities of the short arm of chromosome 1 (by karyotype analysis and LOH), changes in chromosome number, amplification of the oncogene N-*myc*, and more recently, defects in the nerve growth factor receptor pathway.

1.5.5.a *Abnormalities of the short arm of chromosome 1*

Since the original description by Brodeur et al (Brodeur et al, 1977) karyotype analyses have been performed on large numbers of tumour samples. On reviewing banded karyotypes, 70% of primary tumours and 85% of neuroblastoma cell lines were found to exhibit loss of genetic material from 1p (Brodeur et al, 1981, Gilbert et al, 1984). The size of the deletions identified varied, some tumours being monosomic for virtually all of 1p, others showing a more subtle cytogenetic change. It was hoped that identification of the smallest region of overlap of the deletions would allow localisation of a gene or genes important for the tumorigenesis of Nb (in a way similar to that employed for WT to define the 11p13 locus in WAGR patients with 11p13 deletions). However, as it is much easier to acquire metaphase spreads suitable for analysis and establish cell lines from aggressive tumours, the incidence of 1p deletion observed by Brodeur (Brodeur et al, 1981) would be skewed towards more advanced disease.

Karyotype analysis is a crude method of identifying deletions with only large defects (1 Mb or more) being visible, and, therefore, attempts were made to more accurately localise the extent of 1p loss using known molecular markers from 1p. Initially, two approaches were taken; in the first the extent of the 1p deletions in tumours was inferred by LOH studies using polymorphic markers (Fong et al, 1989) whereas in the second the deletion was mapped directly in somatic cell hybrids containing the derivative chromosome 1 (Hunt and Tereba, 1990, Ritke et al, 1989). Both strategies depended on the acquisition of molecular markers, whose position was known. The relative position of these molecular markers to each other has been determined by using somatic cell hybrid mapping panels or by genetic linkage in the large reference families collected by CEPH (Centre d'Etude du Polymorphism Humain). Consensus decisions as to the location of a particular probe were reported by the committee on the genetic constitution of chromosome 1 (HGM 12).

Were any known 'candidate' genes included in the consensus deletion? The genes encoding nerve growth factor B (located in band 1p13) and that for the proto-oncogene, *c-jun* (which maps to 1p22) were both more centromeric than the consensus deletion. L-*myc* is located in 1p32 and, although this locus was deleted in some tumours, the majority of deletions were telomeric to it (Hunt and Tereba, 1990). The consensus deletion was defined differently by different authors: 1p32 to 1pter (Hunt and Tereba, 1990) or in the interval between FGR and D1Z2 in 1p36 (Fong et al, 1989). Martinsson et al (1989), used microdissection to obtain DNA from the 1p35 to 1pter region. The resulting DNA was digested with EcoR1 and then cloned into a plasmid vector. The cloned fragments were then mapped using a somatic cell hybrid panel. Using these probes obtained by microdissection, LOH for 1p36 1.2 was identified in 8 of 9 (89%) stage III and IV neuroblastomas (Weith et al, 1989).

Molecular cloning of CpG islands from 1p36 has led to the isolation of novel human gene Heir-1, which encodes a protein with a helix loop helix (HLH) configuration (Ellmeier et al, 1992). The protein product is almost identical to the mouse HLH 1462 protein and shares homology with murine Id and Drosophila emcHLH proteins which inhibit differentiation of muscle and sensory organs respectively. Heir-1 is expressed at high levels in adult kidney, lung and adrenal medulla, but not in adult brain. 10/12 neuroblastoma cell lines were shown to express very low levels of mRNA for Heir-1 but two Nb cell lines which did not over-express the oncogene N-*myc* had high Heir-1 expression suggesting an inverse correlation between Heir-1 and N-*myc* expression in Nb (Ellmeier et al, 1992).

It may be that the deletions in 1p36 are not random but, as in Wilms tumour are governed by genomic imprinting. Using RFLP analysis Caron et al (1993) found that 28% of the tumours (15/53) showed allelic loss at 1p36 and, in 13/15 of these, the 1p36 allele loss was of maternal origin. The picture is not clear, however, with other groups (Cheng et al, 1993) reporting loss of paternal allele in 6/10 informative tumours consistent with a random distribution. Explanations for the discrepancy between the two reports could be due the patient selection and differences in the probes used to detect LOH.

What became apparent, however, was that 1p deletions were a characteristic of aggressive stage III and IV tumours. These high-grade malignancies were much easier to karyotype and derive cell lines from and this introduced a bias in the reported series (Brodeur et al, 1981). LOH studies performed on all Nb tumours showed the incidence of 1p deletion to be much lower (28%) than the 70% previously quoted (Fong et al, 1989). The presence of 1p abnormalities correlated well with advanced stage (Franke et al, 1986). Indeed, when stage II tumours with 1p deletion were identified they behaved aggressively and were nearly always fatal (Brodeur, 1989, Christiansen and Lampert, 1988). The discovery that the presence of 1p deletion in the tumour predicts

bad outcome more accurately than surgical stage led to the development of methods of assessment of 1p loss that use less tissue than Southern blotting and are more rapidly performed. PCR using 2 loci containing variable numbers of tandem repeats has been reported (Peter et al, 1992) and identified LOH 1p in 9/29 patients with Nb, the results of PCR correlating well with LOH assessment by Southern blotting. However, Nb tumours are not composed entirely of malignant cells; with normal stroma, reactive Schwann cells and lymphocyte infiltrate separating areas of malignant invasion. Many PCR methods use very small 'slivers' of tissue and it would be possible to select a region of the biopsy which did not contain a significant number of malignant cells and wrongly assume that no 1p deletion was present. Other groups have analysed genetic polymorphisms in genes known to map to 1p36 using the single strand conformational polymorphism (SSCP) technique. Allele loss was reported in 8 of the 39 tumours in which information could be obtained and one gene, tumour necrosis factor receptor 2 (TNFR2), was said to be located in the consistent region of overlap (White et al, 1993). Fluorescent in situ hybridisation (FISH) techniques using centromeric specific DNA probes have been used to detect numerical chromosome changes. By using probes from 1p36 and the centromere of chromosome 1 and comparing the number of centromere signals with that of the signals from 1p36, 1p36 deletions can be identified. Furthermore, as this technique can be performed on tumour imprints or paraffin sections of tissue, the results of FISH in a particular cell can be correlated directly with that cell's morphology ensuring, therefore, that it is only the tumour cells which are being analysed (Christiansen et al, 1992, Stock et al, 1993).

There is evidence that the gene or genes at 1p36 may not be entirely specific for Nb. Allelic loss for 1p35 to 1p36 has been described in a number of other tumour types which are completely unrelated to Nb. Examples of these include colon cancer (Leister et al, 1990), bladder cancer (Poddighe et al, 1992) and mixed alveolar and embryonal rhabdomyosarcoma where a t(1;13)(p36;q12) translocation was described in one tumour (Biegel et al, 1991). Indeed, in colon cancer cell lines with 1p deletions, reintroduction of the 1p34-1p36 region by microcell hybridisation led to suppression of tumorigenicity (Tanaka et al, 1993). The gene for familial malignant melanoma and its precursor lesion, dysplastic nevus (both neural crest derivatives), has been assigned to distal 1p and allele loss has been identified in primary tumours and cell lines (Dracopoli et al, 1989). Hepatomas also show LOH for 1p36 (Simon et al, 1991). Pheochromocytoma, an adrenal tumour which is often associated with a multiple endocrine neoplasia type 2 syndrome, shows allele loss on 1p, but more centromeric regions (1p22) are involved (Moley et al, 1992). There is evidence to suggest that 1p36 is a hot-spot for recombinant events and it also has been reported to be the site of preferential integration of certain serotypes of adenovirus (Romani et al, 1990). In summary, allelic loss for 1p36 is found in high stage, aggressive neuroblastomas as

well as being associated with the progression of other tumours. A candidate gene, Heir-1 has been isolated from the region and its expression is inversely related to that of the oncogene N-*myc*.

1.5.5.b *LOH of areas other than 1p*

The finding that LOH for 1p did not occur in the majority of Nb tumours led to the search for other areas of the genome that may harbour genes important in Nb development. To date, three other areas have been identified. LOH for 14q was found at the D14S1 locus in 6/12 (50%) patients, all but one of which were stage 3 and 4 tumours (Suzuki et al, 1989) and at a variety of loci on 14q in two other small studies (Srivatsan et al, 1993, Takayama et al, 1992). A larger study of 59 cases showed that LOH at four loci on chromosome 14 occurred in 23% of cases and was also associated with a poor prognosis. Tumours that showed LOH for markers from 14q infrequently had evidence of N-*myc* amplification (2/10) or 1p deletion (3/10). Indeed 7/10 had neither genetic change but still behaved in an aggressive fashion, suggesting that a gene important in the development of 'poor risk' Nb may be located on 14q (Fong et al, 1992).

The data for LOH on chromosomes 11 and 13 are less convincing. One study of 45 primary tumours reported LOH for all of chromosome 11 in four cases, 11p only in 1 case and 11q only in two cases (Srivatsan et al, 1993). The study by Fong et. al. (1992) found LOH at 11q loci to occur in less than 5% of the tumours they studied which cast doubt on the relevance of this event. Suzuki (1989) found LOH for 13q in 2 of 11 patients at the D13S3 locus (13q33-qter) but this finding has not been repeated by other groups.

1.5.5.c *N-myc amplification*

A substantial proportion of Nb tumours have cytogenetic evidence of gene amplification in the form of extra chromosomal double minute chromatin bodies (DMs) (Cowell, 1982). In tumour cell lines, chromosomally integrated homogeneously staining regions (HSRs), are seen in addition to DM's, but HSRs are only infrequently found in the metaphase spreads from the tumours themselves (Gilbert et al, 1984). The nature of the amplified genetic material became clear when an oncogene related to the *v-onc*, v-*myc* was found to be amplified in 8/9 cell lines tested (Schwab et al, 1983). The amplified oncogene was called N-*myc* (MYCN) the normal location of which is the distal short arm of chromosome 2; 2p24 (Schwab et al, 1984). The structure and mode of action of the *myc* genes is reviewed in section 1.2.1.d, but briefly they are bHLH genes acting as heterodimers and affecting transcription of cell cycle related genes (Ma et al, 1993). Is N-*myc* relevant to Nb development? The tumorigenic effect of H-*ras* in rat embryo cells is augmented by the introduction, in an appropriate

expression vector, of the N-*myc* gene under the control of a constitutive promoter (Schwab et al, 1985). Immortalised rodent and human cells can be transformed by the same mechanism (reviewed in Schwab and Amler 1990). In tissue culture, down-regulation of N-*myc* expression with retinoic acid (Thiele et al, 1985) or with antisense oligonucleotides to N-*myc* RNA (Negroni et al, 1994) leads to a decrease in colony formation. Hence, there is compelling evidence that N-*myc* is somehow involved in neuroblastoma tumorigenesis or its progression. Initial attempts to quantify the degree of N-*myc* amplification were performed by comparing the intensity of bands produced by Southern blotting and revealed that 38% of primary tumours had 3-300 fold amplification of N-*myc* per haploid genome (Brodeur and Seeger, 1986). How is it that gene amplification occurs?

Initial size fractionation experiments suggested that the amplified segment of DNA (amplicon) was 290-430 kb in size and that the amplicons were arranged as head to tail repeats (Schwab and Amler, 1990). More recent work with the SMS-KAN Nb cell line confirms the head to tail orientation of the N-myc repeats but estimates the size of the amplicon as being 1.2 Mb (Schneider et al, 1992). In cell culture, the precise head to tail arrangement was stable over multiple generations which contrasted with the finding that in amplified DNA from drug resistant cell lines, the amplicons were arranged head to head and that their structure was profoundly unstable (Schwab and Amler, 1990). The mechanism of gene amplification is also unclear and even in cell lines with N-*myc* amplification the endogenous gene still exists as a single copy in its usual location on chromosome 2. Replicated copies may form circular units which are then excised. Further, extra chromosomal, replication may then ensue and could explain the difference in size seen in the multiple DMs within a given cell (Schwab and Amler, 1990). Replication of N-*myc* alleles to produce gene amplification may not be a random process. In a recent study it was shown to be the paternal allele that was amplified in 12/13 cases (Cheng et al, 1993) suggesting that N-*myc* may be subject to genomic imprinting.

Amplification of N-*myc* in an Nb tumour is very strongly associated with rapid tumour progression and bad outcome (Seeger et al, 1985) The majority of tumours which exhibit N-*myc* amplification are Stage IV and would be expected to do badly, but even patients with Stage II disease, which normally has a 90% survival, were found to fare poorly if the tumours showed N-*myc* amplification. 1,200 patients with Nb were studied and, although 5-10% 'low stage' tumours were found to have N-*myc* amplification, the patients with these tumours showed rapid disease progression and died (Brodeur, 1989). Similar results were found in Stage IVS disease in which the majority of cases undergo spontaneous resolution of their tumours (unless the sheer tumour bulk kills the patient first). However, a small proportion of the patients had a fatal recurrence of the disease, even after spontaneous resolution, and it was in this

group that the tumours were found to have >10 copies of N-*myc* per haploid genome (Nakagawara et al, 1990). Was it just that N-*myc* amplification was an abnormality that was acquired during tumour progression? In a study of 60 patients, the number of copies of N-*myc* in the primary tumour, in metastatic deposits and in any subsequent recurrent disease was analysed. Within a given tumour, the degree of amplification of N-*myc* was found to be remarkably constant throughout the clinical course of the disease (Brodeur et al, 1987) suggesting that amplification was an intrinsic feature of these tumours. Caron et al (1993) reported a single stage 4S patient whose liver biopsy showed no amplification of N-*myc* at diagnosis and yet tissue from a lymph node recurrence 6 months later showed 40 fold amplification of N-*myc*. Sampling error in the liver biopsy could explain this finding, especially as no mention was made of any histological study in the biopsy specimen used for N-*myc* studies. This is the only case in the literature in which N-*myc* amplification is reported to have changed substantially between diagnosis and recurrence.

What of the 40% of patients with advanced disease who have single copy N-*myc* in their tumours, yet still go on to die? Although there is broad correlation between N-*myc* copy number and the level of *N-myc* mRNA expression, some tumours with single copies of N-*myc* do over express this gene (Wada et al, 1993). However, increased expression of N-*myc* in the absence of N-*myc* amplification does not predict bad outcome (reviewed in Brodeur, 1992). There has been a single report of an Nb cell line derived from a tumour which had been resected from a patient with Stage III disease where N-*myc* was not amplified but the protein product of N-*myc* appeared to have a prolonged half life (100 min versus 30 min) and this may have been important in determining the aggressiveness of the disease (Cohn et al, 1990). However, this was only a single case and is unlikely to be the biological explanation for the aggressive nature of Nb in all of the 40% of patients with advanced disease who do not have N-*myc* amplification. Thus, the presence of N-*myc* amplification in tumours identifies a group of patients who will do poorly, worse than their staging may suggest, but other tumours behave equally aggressively without increased activity of this oncogene. Are there other explanations? We know that the N-*myc* protein product (Myc) forms heterodimers with Max and changes in the production of Max, could, therefore, affect Myc activity. Indeed, when a mutated Max, which inhibited the sequence-specific DNA binding of Myc, was used in the rat Nb model system, there was reversal of the changes in gene expression that had been induced by N-myc (Billaud et al, 1993). Obviously, changes in Max need to be investigated further. Could it be that genes with oncogenic potential are amplified along with N-*myc* but in tumours with a single copy N-*myc* these other genes are amplified alone? This seems unlikely as evidence of gene amplification (DMs, HSRs) are not seen in the absence of N-*myc* amplification. Studies of the 'N-*myc* amplicon' reveal that the only CpG island is within the N-*myc*

gene itself, making it unlikely that other transcribed sequences are present (Hiemstra et al, 1994). The strong predictive value of N-*myc* amplification has led to its use in clinical decision making especially in Stage 4S patients, where its presence has been used to make the decision to give intensive treatment rather than no treatment at all.

Nb is often diagnosed on small biopsies of the primary tumour or on aspirates of bone marrow and the resulting specimens are too small for Southern blotting studies to be performed. PCR based techniques have been employed with the density of the N-*myc* PCR product being compared with the density of PCR product from a gene of known copy number. This use of non-amplified control genes reduces the effect of differential DNA loading or PCR kinetics on the final result. PCR based techniques appear to be accurate in identifying N-*myc* amplification of greater than 10 copies per haploid genome when compared with Southern blotting (Peter et al, 1992). However, as for 1p deletion, the patchy nature of Nb infiltration, results in the admixture of variable percentages of normal cells in the DNA used as a template in the PCR reaction and this could easily lead to false negative results, especially if very small amounts of tissue are used in the reaction. FISH techniques have been developed and allow semi-quantitative assessment of N-*myc* copy number with the advantage that Nb tumour cells can be morphologically identified and the results from these cells alone interpreted (Taylor et al, 1994).

N-*myc* amplification is not entirely specific to Nb tumours, and has also occasionally been found in small cell lung cancer, retinoblastoma and astrocytoma, all of which are derived from neuroectoderm (Schwab and Amler, 1990). Many studies have shown the coexistence of 1p deletion and N-*myc* amplification and the proposed mechanism of action of the Heir-1 gene at 1p36 suggests that the two events may be causally related (Ellmeier et al, 1992). However, opinions as to the presence of a direct relationship between the two vary. Fong et al (1989) analysed a total of 47 primary tumours and found a statistically significant correlation between the two molecular changes. Weith et al (1989) found that, of 9 Stage 3 and 4 tumours, 8 had 1p LOH but only 2 had N-*myc* amplification by Southern blotting. Other tumours have been identified with N-*myc* amplification but no 1p LOH, so it is unlikely that N-*myc* amplification is an event occurring only in tumours with 1p allelic loss.

1.5.5.d *Ploidy*

The DNA content of Nb tumour cells is an accurate indicator of prognosis. Initially, this was measured by flow cytometry and expressed as a ratio of cellular DNA content in the tumour and in normal cells. A DNA index of 1.0 represents a diploid tumour cell, 1.5 triploid and 2 tetraploid. The finding of near diploid DNA content is associated with poor response to treatment and early death, whereas patients with an increased DNA content (hyperploidy) do well. In infants under 12 months of age the

power of this assessment was striking, with 100% of patients with tumours showing hyperploid DNA content surviving compared with 20% of patients whose tumours had a diploid DNA content (Look et al, 1991). Assessment of surgical stage alone was unable to predict survival with this degree of accuracy. Ploidy did not give useful predictive information in two groups: infants with Stage DS (the equivalent of Stage 4S) and in children over 24 months of age with group D (the equivalent of Stage 4). Diploid tumours were associated with N-*myc* amplification and 1p deletion but the correlation was not absolute. Statistical analysis shows that, in different groups of children, there was variation in which genetic abnormality was most predictive of bad outcome (Bourhis et al, 1991, Look et al, 1991). The association of hyperploidy with favourable outcome has been seen in other childhood malignancies such as acute lymphoblastic leukaemia, embryonal rhabdomyosarcoma and medulloblastoma but not in patients with WT, alveolar rhabdomyosarcoma or osteosarcomas (reviewed in Look et al 1991) suggesting that changes in DNA content do not underlie universal mechanisms of tumour formation. As in N-*myc* amplification and 1p deletion, techniques have been developed to assess DNA content by FISH. In the assessment of ploidy probes specific to the centromere of chromosome 8 were used, and the number of 'signals' per tumour cell were used as the measure of DNA index (Taylor et al, 1994). By allowing morphological identification of cells this method allows tumour cells alone to be assessed which is advantageous. However, equating the number of centromeres of a particular chromosome with ploidy as assessed by DNA content may not be accurate in all circumstances (Stock et al, 1993).

1.5.5.e *Abnormalities of the nerve growth factor pathway*

As in most embryonal tumours of childhood, neuroblastoma cells morphologically resemble foetal neural crest cells. Neuroblastoma *in situ* is known to occur at 100 times the frequency of clinical Nb so that the vast majority of neuroblastoma *in situ* must differentiate into benign ganglioneuromas during normal growth and development. Unlike most embryonal tumours, however, Nb tumour cells can show differentiation either spontaneously or following treatment with cytotoxic drugs or differentiating agents such as retinoic acid. There is, therefore, evidence that Nb arises due to a block to normal differentiation of neural crest cells (Baker et al, 1989). Nerve growth factor (NGF) is known to be vital in inducing differentiation and promoting survival of developing neuroblasts, indeed it can even lead to differentiation of a subsection of Nb cells grown in tissue culture (reviewed in Baker et al 1989). Over the last few years there has been an increased interest in the pathways by which NGF activity is mediated. Two cell surface receptors have been described; $gp75^{NGFR}$ is a low affinity receptor coded by the LNGFR gene. A high affinity receptor $gp140^{TRK\text{-}A}$ (TrkA) with tyrosine kinase activity is encoded by the TRK-A gene (Hempstead et al,

1991). The two receptors work together to some extent with both being required for optimal function but NGF signal can be transduced by TrkA alone in some cultured cells (Kaplan et al, 1991). The TRK-A gene is expressed in foetal sympathetic nervous tissue and expression is maintained in adults. Transgenic mice, with defects of the tyrosine kinase portion of the TrkA receptor, have been developed. Most of these mice died within a month of birth and appeared to have virtual absence of the paraspinal ganglia of the sympathetic nervous system as well as sensory abnormalities (Smeyne et al, 1994). However, NGF is not the only neurotrophin; brain derived neurotrophic factor (BDNF), neurotrophin 3 (NT3) and neurotrophin 4 (NT4) have also been described and both NT3 and NT4 can bind to TrkA to some extent. Similarly, TrkA is not the only member of the family of neurotrophin receptor kinase which contains two other receptors; TrkB (Squinto et al, 1991) and TrkC (Lamballe et al, 1991). TrkB is the receptor for both NT4 and BDNF and TrkC appears to be specific for NT3. The functional significance of the three different members of the Trk receptor family is becoming clearer. Transgenic mice with homozygous deletions of TrkB have motor and sensory defects (Klein et al, 1993) and the TrkC null mutant mice have failure of development of neurones important in proprioception (Klein et al, 1994).

The fact that defects in TrkA have such a profound effect on the development of the sympathetic nervous system suggest that dysregulation of this receptor may be important in Nb development. However, the first tumour in which TrkA has been found to act as an oncogenic product was not Nb but colonic carcinoma where a recombination event between the TRK-A gene and HTMNM (a gene encoding human tropomyosin sequences) produced an oncogenic protein, presumably with aberrant tyrosine kinase activity (Radice et al, 1991). The TRK-A gene has been mapped to 1q32- 41 (Miozzo et al, 1990) and the HTMNM gene to 1q31 suggesting an interstitial rearrangement of 1q as the oncogenic event in this particular colonic tumour. In Nb no such rearrangement has been identified but TrkA expression has been found to be a powerful independent prognostic factor. Tumours that do not express TRK-A mRNA are high stage, often aggressive and ultimately fatal. N-*myc* amplification was strongly associated with reduction of TRK-A mRNA but TRK-A under expression was found to be an independent prognostic factor even when N-*myc* amplification was taken into account (Suzuki et al, 1993). In the same study expression of the low density receptor gp75NGFR did not correlate with TRK-A expression, histology, tumour stage or survival. Other neurotrophins have been implicated in Nb development. Whereas aggressive disease correlates with the lack of TRK-A expression, expression of TRK-B and BDNF are both found in immature Nb. Indeed, it may be that in Nb tumours which express TRK-B mRNA, BDNF acts in an autocrine fashion to promote survival of tumour cells (Nakagawara et al, 1994). Clearly, the neurotrophins and their receptors have a role in the pathogenesis of Nb and require further investigation.

1.5.5.f *Other molecular abnormalities identified in Nb*

I do not intend to comprehensively review the many other molecular abnormalities that have been infrequently identified in patients with Nb. However, certain findings deserve at least a mention. Neurofibromatosis type 1 (NF1) is an autosomal dominant condition characterised by cutaneous malformations and a predisposition to certain malignancies. The NF1 gene has been cloned and its protein product, neurofibromin, shares extensive homology with the mammalian *ras* GTPase activator protein, p120 ^{GAP} (reviewed in The et al). Neurofibromin is, therefore, implicated in the negative control of *ras*-mediated signal transduction. Although patients with NF1 are predisposed to a variety of tumours derived from neural tissues they do not have an increased risk of developing Nb. Indeed, there is no evidence that abnormalities of *ras* contribute to Nb development. Although a significant association was found between the expression of H-*ras* and the survival of Nb patients (without NF1), it was those patients whose tumours has increased expression of H-*ras* that survived and not the reverse (Tanaka et al, 1991).

Increased expression of other oncogenes, such as *ret* (Ikeda et al, 1990) and *nm23* (Leone et al, 1993) have been noted in Nb cell lines but it is probable that these molecular events are related to the progression of the cell lines under study rather than being central to Nb development.

1.5.5.g *Summary of the molecular pathology of Nb*

Colonic carcinoma has been the model for the interaction of multiple genes in the development and progression of malignant change with metastatic disease losing the function of three different tumour suppressors (DCC on 18q, p53 on 17p and fap on 5q) and showing activation of *ras* (Fearon and Vogelstein, 1990) Thus, a colonic adenoma will exhibit one or two molecular changes, a carcinoma three or more and metastatic disease the full spectrum of the 'colon cancer specific' oncogenic abnormalities. Progression of the disease is determined by a step-wise accumulation of genetic changes. However, in spite of the fact that many molecular events have been described in Nb tumours, Nb does not appear to act like this at all. N-*myc* amplification has been shown to be a stable feature of a given tumour and tumours without N-*myc* amplification do not acquire it as they progress (Brodeur et al, 1987). It appears that N-myc amplification, diploid chromosome number, TRK-A under expression and 1p deletion/LOH are closely correlated and together identify a subset of aggressive Nb. The association of LOH for 14q with these changes is less marked but considerable overlap still exists (Suzuki et al, 1989). However, about half of Nb tumours that go on to kill patients and all 'low grade' tumours are characterised simply by the absence of all these changes and there is no data to suggest a gradual, step-wise accumulation of these

events in Nb. Other genetic abnormalities must , therefore, exist and may be at least as important in the evolution of Nb as those noted to date.

1.6 CLUES TO THE LOCATION OF OTHER GENETIC CHANGES IN Nb.

1.6.1 *Tumour karyotype analyses*

As we have seen from the discussion of Rb and WT, clues as to the location of genetic changes important in the development of malignancy can be obtained from the analysis of complex phenotypes, tumour karyotype and LOH studies, linkage analysis and the investigation of patients with constitutional chromosome abnormalities.

Potential complex phenotypes that have been associated with Nb development are discussed in section 1.5.2. There is no compelling evidence that any of these associations represents a stable identifiable phenotype and constitutional karyotype analysis does not reveal any abnormalities as was seen for WT and Rb.

The analysis of tumour karyotypes is fraught with difficulties both in obtaining and interpreting the results. Karyotype analysis of solid tumours is technically very demanding and usually only successful in tumours with a high mitotic rate. Aggressive tumours are, therefore, over-represented in any analysis of this kind. The abnormalities seen represent the sum total of all the genetic events that have occurred in a given tumour cell, not only those responsible for the original malignant transformation but also those associated with subsequent disease progression. Tumour karyotypes are often hugely complex, with many identifiable abnormalities, and the identification of important genetic changes against a background of other abnormalities may be very difficult indeed. However, many studies of tumour karyotypes have been performed in Nb. 1p deletion was the most consistent change followed by trisomies for the long arms of chromosomes 1 and 17 (Gilbert et al, 1984).

1.6.2 *LOH studies*

LOH studies in Nb are described in sections 1.5.5.a and 1.5.5.b and, although 1p LOH predominates, low frequencies have been described for 14q, 11 and 13q33-qter.

1.6.3 *Genetic linkage analysis*

Close linkage of a known molecular marker to the Nb phenotype would provide good evidence as to the location of a gene responsible for the disease phenotype. However, as discussed in section 1.5.3, the number of familial aggregation of Nb is small and the majority consist of only two affected members in a family. DNA has only rarely been stored and because of the high mortality of Nb, it is not possible to follow up the majority of cases in order to obtain constitutional DNA samples. There is simply not enough material for informative linkage studies to be performed at the present time.

1.6.4 *Analysis of constitutional chromosome abnormalities*

In patients with genetic disease, the analysis of constitutional chromosome translocations has proved to be a fruitful way of identifying the position of genes important in determining the disease phenotype. This has been the case for Rb (Yunis and Ramsay, 1978), WT (Riccardi et al, 1978), NF1 (Fountain et al, 1989) and familial adenomatous polyposis (Varesco et al, 1989). These structural abnormalities are predominantly deletions, but where translocations are identified the breakpoints usually interrupt the predisposition gene (Higgins et al, 1989, Mitchell and Cowell, 1989). Whereas in deletions many genes may be lost, in translocations the breakpoint defines a discreet locus which can then be characterised. The finding of constitutional chromosome abnormalities in a patient with Nb may provide very valuable clues as to the location of genes important in Nb development. How many constitutional chromosome abnormalities have been reported in patients with Nb?

Two recent reports of constitutional changes involve 1p36. Laureys et al (1990) reported a 9 month old boy with a stage III Nb tumour who had a constitutional t(1;17)(p36;q12-21). Family history revealed that a child of the mother's niece had developed a thoracic Nb. In spite of his young age, the 9 month old boy developed progression of his tumour on therapy and died. Biegel et al (1993) noted an interstitial constitutional deletion (1p36.1-1p36.2) in a girl who had been originally investigated for developmental delay and dysmorphic features. At 5 months of age she was diagnosed as having Nb, and although the stage was not noted in the report the degree of bone marrow infiltration described is suggestive of stage IV disease. This child did well on chemotherapy. Thus, in both of the cases of Nb with constitutional chromosome abnormalities, the patients had high grade, stage 4 disease which occurred at a young age. It may be that analysis of these constitutional changes will allow

identification of the gene or genes in 1p36 that are responsible for the development of the subset of aggressive Nb characterised by 1p deletion. As noted in section 1.5.5.g, however, the majority of Nb are not characterised by 1p deletion and it may be that other genes will be important in the tumorigenesis of this group. Do other constitutional chromosome changes exist?

Moorhead and Evans (1980) reported constitutional translocations in 2/37 patients with Nb who had karyotype analysis. One child with a stage II Nb, had a constitutional translocation between the short arms of chromosome 4 and 7, but the relevance of the genetic change is unclear, as four unaffected family members were found to carry the same translocation. A patient with ganglioneuroblastoma carried a t(11;16)(q23;q24) as did the unaffected mother. More complex constitutional abnormalities have also been reported, one case of 2p trisomy and a familial t(2;16)(p13;p11) in 8 month old patient with stage IVS disease was reported by Nagano (1980). Hecht (1982) described a family with 4 cases of neuroblastoma in two generations. The propositus presented with stage IVS Nb and a paracentric inversion of the long arm of chromosome 11 (INV11) and a deletion of the short arm of chromosome 21 ($21p^-$) considered to be a normal cytogenetic marker. Her father carried the INV11, the breakpoint being 11q21 and 11q23. His daughter by another marriage died of Nb at 5 years of age and his chromosomally normal son went on to have a chromosomally normal child with congenital metastatic Nb (probably Stage IVS). The propositus's mother, who carried the $21p^-$ had a son by a previous marriage who died at 1 hour of age from pregnancy associated complication and at post-mortem had adrenal neuroblastoma. The disease does not appear to segregate with the genetic abnormalities noted in this family. Feingold et al (1971) reported a fatal Stage IV neuroblastoma in a morphologically normal 2.5 year old girl whose mother later gave birth to a female infant with multiple congenital abnormalities typical of trisomy 13 which was confirmed on chromosomal analysis. The infant died at 2 days of age and at post-mortem small foci of Nb were found in both adrenals (which could represent neuroblastoma *in situ).* The relevance of this finding to clinical Nb is unclear and chromosomes were not performed on the index case.

In section 2.0 a patient with ganglioneuroblastoma and a constitutional t(1;13)(q22;q12) is described. It appears, therefore, that this patient is the only case of a potentially predisposing constitutional chromosome abnormality occurring in a patient with low grade Nb and, as such, represents a unique resource for the investigation of genes vital in producing this phenotype.

1.7 STRATEGIES FOR CLONING GENES IMPORTANT IN Nb

Cloning a gene by 'classical genetic' techniques depended on a knowledge of the biochemistry of the defect that led to the development of the disease phenotype. The amino acid structure of the normal protein was then deduced by purification and chemical analysis of the product. Once the amino acid sequence was known the DNA sequence of the coding region of the gene could be deduced. This 'functional cloning' approach hinged on the mechanism of disease formation being known. For example, in the Li Fraumeni syndrome only the vaguest clues as to the mechanism of development of the disease phenotype were available (section 1.3.5). In families with Li Fraumeni syndrome the early age of onset of malignancy and the fact that multiple primary tumours that were observed were reminiscent of the pattern of disease seen in Rb. The RB1 gene had been cloned and was not found to be mutated in patients with the Li Fraumeni syndrome. Mutations in the p53 gene, the only other tumour suppressor gene known at the time, had been identified in sporadic cases of many of the tumours included in the Li Fraumeni phenotype. Hence, Malkin et al (1990) decided to analyse the role of the p53 gene in the Li Fraumeni syndrome and found that abnormalities of p53 were indeed responsible for the phenotype.

In the case of phenylketonuria there was much more accurate information as to the function of the gene responsible for the disease phenotype. Phenylketonuria was known to be due to deficiency of the enzyme phenylalanine hydroxylase. The enzyme had been purified from the livers of rodents, monkeys and humans and its structure and function were well described but the location of the gene was not known. The mRNA for phenylalanine hydroxylase was purified from rat liver by polysome immunoprecipitation (using antibodies to the protein) and cDNA probes were derived from the mRNA. The short cDNA probes were used to screen a cDNA library derived from human liver tissue and a full length transcript was identified (Kwok et al, 1985). The cloning of the phenylalanine hydroxylase gene and the identification of RFLPs within it allowed prenatal diagnosis and carrier detection in around 90% of affected families.

Glucose-6-phosphate dehydrogenase is an enzyme important in determining NADPH levels in cells and deficiency of this enzyme leads to an autoimmune haemolytic anaemia. Over 300 variants of the protein exist and many are associated with decreased enzyme activity. It was hoped that the cloning of the gene would allow a greater understanding of the biological features of the variant proteins. The amino acid structure of a short segment of the enzyme (peptide KMMTKK) was known and from this the corresponding DNA structure could be inferred. The resulting probe (a 17 bp oligonucleotide) was used to screen three cDNA libraries and analysis of the resulting

cDNA clones allowed the entire coding region of the gene to be sequenced (Persico et al, 1986).

By contrast, the second approach to cloning a given gene presupposes no knowledge of the structure or function of the gene but relies instead on mapping the gene of interest to a precise location on a chromosome. Originally referred to as 'reverse genetics' this approach is now more correctly called positional cloning. Initial clues as to the location of the gene can come from linkage analysis using polymorphic probes, originally RFLP but now more often markers such as dinucleotide repeats. However, even in diseases where many informative meioses are available for study the gene of interest can only be assigned to a genetic interval of about 1 centimorgan (cM). Depending on the frequency of recombination events in a given area 1 cM corresponds to a physical distance of about 1 Mb. The availability of cytogenetic rearrangements provides good evidence for the site of the gene and most of the genes that have been mapped using positional cloning have been located at the site of chromosomal changes. Thereafter, the principle steps are to develop molecular markers which closely flank the gene, then, using these markers, to isolate the intervening DNA fragment from which the gene of interest can be cloned. Many genes have now been isolated by positional cloning using a variety of approaches selected examples of which are discussed below.

The cloning of the RB1 gene is often held up as the model for the positional cloning approach. The value of constitutional cytogenetic abnormalities in identifying the location of genes important in the disease phenotype was shown by the fact that the identification of deletions of 13q14 seen in a small number of patients with Rb led to the analysis of markers from this region. Further analysis, initially with protein polymorphisms (ESD) and then with RFLP, confirmed that the Rb phenotype was tightly linked to markers from 13q14, a finding that was also supported by LOH studies (section 1.3.3). However, the actual cloning of the RB1 gene did not occur via a step-wise analysis of the region of interest but rather by the chance identification of a clone isolated from a λ phage library made from DNA obtained from flow sorted copies of chromosome 13 (Lalande et al, 1984). Of 200 clones that did not hybridise to human DNA on primary screening 44 were analysed further and 18 turned out to act as unique sequences of which 13 were mapped to chromosome 13. When hybridised to DNA derived from a patient with Rb who had a deletion of 13q14 and to normal DNA two of these clones (H3-8 and H2-42) showed decreased signal intensity in the lane containing DNA derived from the patient with the deletion, suggesting that the probes were located in the region of the deletion. A 1 phage walk from pH3-8 allowed the identification of three further unique sequences (Dryja et al, 1986). Using these reagents, Friend et al (1986 and section 1.3.3b) identified a cDNA clone from the RB1 gene. Bearing in mind

that chromosome 13 has been estimated to contain 114 Mb of DNA (Morton 1991) it was amazingly lucky that of only 13 clones derived from the whole chromosome two were found to map within a small region of 13q14 and one from within the genomic sequence of the RB1 gene.

Not surprisingly, the positional cloning of other genes has not been blessed with such good fortune. The cloning of the dystrophin gene is a good example of the meticulous analysis necessary to accurately localise the position of a gene and underscores the value of rare cytogenetic abnormalities in these analyses. Duchenne muscular dystrophy (DMD) is an X-linked disorder characterised by progressive muscle degeneration leading to death in the second or third decade of life. Linkage analyses and cytologically detectable constitutional chromosome analyses suggested that the gene responsible for the DMD phenotype was located in Xp21. But more that simply acting as clues to the location of the gene, the cytogenetic abnormalities were central to the positional cloning strategy undertaken . A cytogenetically visible deletion of Xp21 was identified in a patient with three different X-linked diseases (DMD, retinitis pigmentosa and chronic granulomatous disease as well as the McLeod red blood cell phenotype). In order to derive novel probes from the region of the DMD gene, DNA from a lymphoblastoid cell line made from this patient was sonicated and mixed with DNA fragments obtained following Mbo 1 digestion of DNA from a human cell line containing four copies of the X chromosome. A 200 fold excess of sonicated DNA from the patient was used and, in order to maximise the opportunity for unique sequences to anneal, the phenol enhanced reassociation technique (PERT) was employed (Kunkel et al, 1985). Following annealing of the mixture, three different types of double stranded molecules were produced. The commonest species represented reassociated strands of the patient's DNA. Because the DNA had been randomly fragmented by sonication the ends of these reassociated molecules were composed of variable lengths of single stranded DNA and, hence, these molecules would not be easily cloned into a plasmid vector. The second species of double stranded DNA represented strands from the Mbo 1 digested cell line which had annealed to DNA strands from the patient and again the ends of the molecules would have random lengths of single stranded DNA making it difficult to clone them into plasmids. The third and rarest form of DNA molecule would be composed of homologous single strands of Mbo 1 cut DNA from the cell line the ends of which would allow them to be cloned into the Bam H1 site of a plasmid vector. Because the vast excess of DNA from the patient had 'mopped up' sequences common to both sources of DNA these clonable strands would be enriched for sequences from the deleted region of Xp21. The resulting clones were mapped on a panel of somatic cell hybrids containing various deletions of the X chromosome and seven clones which mapped to the deleted region of Xp21 were identified. Of these, one (pERT87 [DXS164]) was found to be absent in five of 57

unrelated males with DMD (Monaco et al, 1985). Independently, Ray et al (1985) isolated another probe from Xp21 which also detected deletions in DMD patients, again capitalising on the cytogenetic abnormality found in a single patient. Although X-linked, occasional females can develop a from of DMD and it was found that these women carry reciprocal translocations between the X chromosome and autosomes, the breakpoint on the X chromosome interrupting the gene for DMD. As, in this situation the normal X chromosome is preferentially inactivated, these patients go on to develop DMD. One woman was identified as having an t(X;21), the breakpoint in chromosome 21 being in 21p12, a region known to code for ribosomal RNA. As a result of the translocation, the DMD gene was now adjacent to a ribosomal genes, probes for which were available. By using these known ribosomal genes as probes a region of the DMD gene was identified and one clone (XJ1) was found to be closely linked with DMD in families and to detect deletions in patients.

In order to clone sequence from the coding region of the DMD gene a chromosome walk was initiated from the pERT87 locus. Chromosome walking involves using a known probe to screen genomic DNA libraries for other clones that contain not only sequences homologous to the probe but also sequences not contained in the original probe. When these 'secondary' clones are used as probes against genomic DNA libraries, further clones are identified thus extending the region of cloned DNA. Nine bi-directional walks from the pERT78 locus led to the cloning of 220 kb of genomic DNA from the region of the DMD gene. 50 unique subclones from this region were isolated and used to screen Southern blots derived from Hind III digested DNA from various species (monkey, rodent , chicken and cow). Only two of the 50 subclones showed this 'cross species conservation' and, as important genes are conserved in evolution, were thought to represent parts of the coding region of the DMD gene. One of the two clones also recognised an RNA transcript from foetal skeletal muscle and this clone was then used to screen a cDNA library derived from foetal skeletal muscle (Monaco et al, 1986). By screening further cDNA libraries the 14 kb full length transcript of the DMD gene (which was named dystrophin) was cloned (Koenig et al, 1987). Localising the cDNA clones on long range maps of the Xp21 region led to the characterisation of the genomic structure of the dystrophin gene which is composed of at least 65 exons stretching over a huge 2.5 Mb of DNA.

The presence of cytogenetic abnormalities was central to the identification for the genes important in the phenotypes of Rb and DMD. The difficulty of accurately localising the position of a gene of interest in the absence of cytogenetic changes is well illustrated by the cloning and characterisation of the cystic fibrosis gene, a process that was beautifully summarised in a series of articles in Science (Rommens et al, 1989; Riordan et al, 1989 and Kerem et al, 1989). Cystic fibrosis (CF) is the commonest

autosomal recessive disease in the caucasian population, with an incidence of 1 in 2000 live births and a carrier frequency of 1 in 20. There were, therefore, many families in whom genetic linkage analysis could be performed. Using a large number of polymorphic DNA probes the locus for CF was assigned to 7q31 and the closest flanking makers (MET and D7S8) were estimated as being 1.5 Mb apart. The isolation of the CF gene from this defined region proved to be more difficult than originally imagined. The first step was to develop a physical map of the region between the flanking markers and isolate further probes. To this end, a genomic library made from DNA sequences derived from flow sorted copies of chromosome 7 was examined. The position of individual markers relative to the flanking markers was established by mapping using panels of somatic cell hybrids containing various deletions of chromosome 7, by genetic linkage and by analysing the presence of pairs of markers on very long fragments of DNA which had been separated by pulsed field electrophoresis (PFGE). Using these methods, two further markers (D7S340 and D7S122) were isolated, both closely linked to CF and the order of markers was established as MET-D7S340-D7S122-D7S8. The physical distances between these probes were 500 kb, 10 kb and 980 kb respectively. Whereas in DMD it was relatively easy to orientate markers with respect to the DMD gene simply by their presence or absence in deletions from patients with the disease, the absence of cytogenetic rearrangements in CF presented real problems at this stage because there was no easy way of telling if any of these sequences came from the CF gene. The most successful approach was to laboriously clone the whole of the region most likely to contain the gene and then search for candidate gene sequences. Chromosome walking had been useful in cloning contiguous DNA sequences but such walks are often impaired by the presence of 'unclonable' regions of the genome. Hence, chromosome walking was supplemented with chromosome jumping in which large DNA fragments obtained by digestion with the restriction enzyme Mbo1 were separated by PFGE and ligated to a selectible marker (sup F; a tRNA suppressor gene) under conditions that favoured circularisation of the resulting product. When the circularised ligation products were digested with Eco R1, the small resulting fragments contained the selectible marker flanked by genomic DNA sequences that had been separated by over 50 kb. The products were then ligated into a phage vector and packaged *in vitro*. When plated on bacteria which lacked the sup F gene only those phage genomes which carried the sup F marker would replicate and form plaques. Screening this library with markers from the region of interest identified phage clones that contained not only the marker used as a probe but also a sequence from over 50 kb away. Further screening of genomic DNA libraries with the phage clones identified in this way allowed the isolation of markers from which bi-directional chromosome walks could be initiated. By localising clones derived from the jumps on the long range restriction map of the region, the direction of the jump relative to the

original markers could be ascertained. In total, seven jumps were made and 49 recombinant phage and cosmid clones were isolated from ten different genomic libraries. All the clones had to be localised on the long range map of the region to ensure that the cloned DNA sequences were colinear with the genomic sequence. This huge effort resulted in the cloning of 280 kb of contiguous DNA from the region of interest. The cloned region was further extended to more than 500 kb by the isolation of genomic DNA sequences using cDNA clones from the region of the CF gene. A variety of strategies were used to isolate potential candidate genes from the cloned DNA. The search for sequences that were conserved among species was technically the easiest approach but was limited by the presence of repetitive elements. Other methods included the identification of CpG islands, which are often found at the 5' end of genes, and the isolation of cDNA clones derived from the normal tissues which are affected in the CF disease phenotype. Eventually, a clone which cross hybridised with bovine, rodent and chicken DNA was identified and was found to contain a CpG island. A probe derived from this clone was used to screen a cDNA library derived from mRNA extracted from the sweat cells of an individual who did not have CF and a single positive clone was identified. Further screening of cDNA libraries led to the cloning of a 6 kb region that contained all 24 exons of a gene. But how was it decided that this gene really was the CF gene? Again, the absence of cytogenetic abnormalities was a problem as no abnormalities could be identified on Southern blots of DNA from CF patients. Sequencing the cDNA for the full length transcript and comparing the sequence with that obtained from patients with CF identified a three base pair deletion which resulted in the loss of a phenylalanine residue in codon 508. This mutation (ΔF_{508}) is the commonest one found in the CF population. The predicated structure of the CF protein (now called the cystic fibrosis transmembrane conductance regulator) was consistent with it being a transmembrane protein which fitted very well with its predicted role from physiological studies as a transmembrane ion channel.

The cloning of the genes for CF and DMD was a very laborious process. In each case the approach was to delineate the smallest possible region of interest from which individual single copy sequences were derived. Each of the single copy sequences was then tested independently for conservation among species and its homology to transcribed sequences. Once a transcribed sequence was identified it was used in further rounds of screening to derive the full length transcript. The characterisation of the genomic DNA from the gene of interest then depended on localising the full length transcript on a long range restriction map of the region of interest and cloning a large number of individual DNA fragments which together contained the entire sequence of the gene. Cloning very large DNA fragments from the region of interest and deriving transcribed sequences directly from them would simplify both the identification of

transcribed sequences from the gene of interest and the characterisation of the genomic structure of the gene. Whereas cosmid vectors allow the cloning of DNA fragments up to 50 kb long, the development of yeast artificial chromosome vectors (YACs) enabled fragments of 1 Mb to be cloned and recent libraries have inserts of up to 3 Mb in length. The power of YAC technology is well illustrated by the cloning of the gene for X-linked agammaglobulinaemia (XLA), a disease in which recurrent bacterial infections arise due to the deficiency in antibody production (Vetrie et al, 1993). The XLA locus had been mapped genetically to Xq21.3-22. With the identification of flanking markers (DXS442 and DXS101) the XLA locus was localised to a 2 cM interval and one polymorphic locus (DXS178) showed no recombination with the XLA phenotype. A long range map of the region had been constructed which established the order of 12 markers flanking the XLA locus and suggested the presence of a number of CpG islands in the region of the DXS178 locus. A contig of 68 YACs was established which spanned 6.5 Mb of genomic DNA from the region of interest. One YAC (y178-3) was gel purified and used to probe a cosmid library derived from a patient with a 49, XXXXX karyotype as described by Holland et al (1993). The 48 cosmids that were identified were localised into 12 regions of the long range map of the region. DNA from YAC y178-3 was immobilised on a hybridisation membrane and used to 'select' clones from a cDNA library derived from a B cell malignancy (see section 4.4.2 for details of cDNA selection). The resulting cDNA clones were enriched for the region of genomic DNA contained in YAC y178-3. One aliquot of the pool of enriched cDNA was sub-cloned and gridded out for further analysis whereas another was used to probe the 48 cosmids that had been identified with YAC y178-3. 13 of the 48 cosmids hybridised to the enriched cDNA and restriction fragments from these 13 were used to screen the gridded cDNA clones, identifying 104 clones which were arranged into 12 regions of Xq21 on the long range map. When representative cDNA clones from each of these regions were hybridised to Southern blots of DNA from XLA patients, cDNAs from two regions detected rearrangements. The three cosmids that spanned these two regions were hybridised to the gridded enriched cDNA clones and detected 25 positive clones. Hybridising two of these 25 cDNA clones to the original (un-enriched) cDNA library led to the identification of a 2.6 kb cDNA clone that hybridised strongly to all 25 clones indicating that all 25 were derived from a single transcript. The 2.6 kb transcript encodes a protein with extensive homology to intracellular (non-receptor) protein tyrosine kinases and has been called agammaglobulinaemia tyrosine kinase (atk). The beauty of the technique is that it makes the YAC do double duty; not only identifying expressed sequences from the region of interest but also identifying cosmids which allow the cDNA clones to be located on the long range map of the region, thereby reducing the total number of cDNA clones that need to be screened.

Although the techniques of positional cloning and the analysis of candidate genes have been presented as separate, recent advances in the mapping of the human genome will lead to the future amalgamation of the two approaches. The Human Genome Project has started to bear fruit and high resolution genetic maps based on sequences derived from short tandem repeats and other sequences have been constructed (Gyapay et al, 1994). YAC contigs spanning entire chromosomes have already been constructed for chromosome 21 and X allowing researchers access to cloned DNA fragments without the need for laborious chromosome walking and jumping and the coverage of the remaining chromosomes with YAC contigs is being extended (Cohen et al, 1993). In tandem with the improved maps of genomic DNA, several thousand cDNA clones have been partially sequenced and oligonucleotide primers have been designed. These oligonucleotide primers allow the intervening cDNA sequence to act as an 'expressed sequence tag' (EST), akin to the STS that are available for genomic DNA. Many of these sequences have been mapped and a database (dbEST) is now available (Boguski et al, 1993). Used together, the advances in positional mapping and the sequences of known cDNA clones may well speed up the process of gene identification. For example, the gene for autosomal dominant retinitis pigmentosa was found to be linked to markers on the long arm of chromosome 3. The gene for rhodopsin was also known to map to the region and a direct search for mutations in the rhodopsin gene found a C to A transversion in 17 of 148 patients but none of 102 unaffected individuals (Dryja et al, 1990). This has been called the 'positional candidate' approach and it is envisaged that when a given disease phenotype is mapped to a region of the genome, it will be possible to search for all the expressed sequences in that region. The sequence whose function most closely matches the potential function of the disease gene will then be selected as the candidate and investigated further (Ballabio, 1993). However, it has to be borne in mind that in 1989,when this project was started, the available maps of the genome were relatively crude, YAC technology was in its infancy and many of the techniques of genome analysis discussed above had not yet been reported.

1.8 AIMS OF THIS THESIS

This PhD study centres around a single patient with stage II ganglioneuroblastoma who was found to have a constitutional t(1;13)(q22;q12) translocation. The premise was that the balanced translocation predisposed him to developing his tumour (discussed in sections 1.6.4 and 4.1) and that the isolation and characterisation of the breakpoint junction fragment may lead to the cloning of a gene or genes important in the development of ganglioneuroblastoma. Initially, the sole information available as to the position of the breakpoint was from standard cytogenetics. The physical dimensions of the long and short arms of all the chromosomes have recently been published (Morton, 1991). The long arm of chromosome 13 is 98 Mb long and is 'composed' of 9 Giemsa positive and negative bands. The number of bands is, however, an inaccurate indication of the size of a given region as it depends of the size of the bands and the degree of resolution. At the 850 band resolution level, 13q12 represents approximately 10% of 13q and its size can, therefore, be estimated as 10Mb. Similarly, 1q is 135 Mb in length and is composed of 11 Giemsa positive and negative bands which vary significantly in size. 1q22 is one of the smaller bands and its size is between 9 and 12 Mb. Clearly, both of these bands were far too large to attempt any form of walking, even in YACs, and would have to be further subdivided into regions of more manageable size. The breakpoint junction fragment could be approached from the chromosome 1 or the chromosome 13 end and the choice of direction depended on the availability of reagents that allowed the region of the breakpoint to be characterised most accurately. Although there is some evidence that there may be a familial predisposition to Nb, the number of families is very small (section 1.5.3), far too small for any genetic linkage studies with polymorphic markers from 13q12 or 1q22 to be performed. Hence, localisation of the region of the breakpoint had to be performed using physical mapping information alone. The strategy I adopted (see section 3.0) was to use somatic cell hybrids to isolate one of the derivative chromosomes from all other chromosome 1 and 13 material to allow molecular mapping of the breakpoint. I then intended to isolated markers which closely flanked the breakpoint and go on to use these markers to institute a chromosome walk in YACs to clone the region of the breakpoint junction. Having isolated the region of interest on a single YAC I then intended to clone the breakpoint junction fragment .

MATERIALS AND METHODS

MATERIALS AND METHODS

2.0 CASE REPORT

Patient DG presented at five years of age with a two year history of a gradually enlarging left-sided neck mass but otherwise he had no clinical problems and was developmentally and morphologically normal. His parents and two sisters were well and there was no family history of malignant disease. Examination showed a 10 x 7cm neck mass but no other abnormalities. Magnetic resonance imaging (MRI) scanning showed the mass to be separate from the cervical spine and deep to the carotid vessels. Biopsy showed the mass to be a cervical ganglioneuroblastoma but his urinary catecholamine metabolite levels were normal and bone marrow aspirate and trephines were free of malignant cells, suggesting that the tumour was relatively benign. A well encapsulated ganglioneuroblastoma was removed surgically and he received no further treatment. The tumour was not sent for karyotype studies but cytogenetic analysis of his lymphocytes showed a constitutional t(1;13)(q22;q12) translocation; cytogenetic studies of his parents and two sibs were normal. Seven years later he is well with no tumour recurrence.

2.1 MOLECULAR TECHNOLOGY

All chemicals used in this project were analytical grade and were purchased from a variety of companies including Sigma, BDH Lab supplies, Bethesda Research Laboratories (BRL), Pharmacia and Boehringer Mannheim. Non radioactive dNTPs and random hexanucleotides (pD(N)6) were purchased from Pharmacia. Radio-labelled dNTPs were from Amersham International and were used in accordance with stringent safety regulations. The solutions and buffers were made in double distilled water (ddH_2O) or distilled water (dH_2O), obtained from a Milli Q 4 BOWL and Milli R04 system (Millipore Ltd.) respectively. When required, pH adjustments were made by the addition of an appropriate acid or alkali and checked with a pH meter (pH meter 240, Ciba Corning Diag. Ltd.). All solutions were either autoclaved or filter sterilised and stored according to manufacturers instructions. Pipette tips, eppendorf and other tubes were also autoclaved before use if they had not been gamma irradiated by the manufacturer. Centrifugation was either in a microfuge (MSE microcentaur) or in a Juoan CR412 centrifuge. Visualisation of DNA was by staining with ethidium bromide, an intercalating dye that fluoresces in ultraviolet (UV) light. Ethidium bromide stained gels were illuminated from below by a Chromatovue TM-20 transilluminator (Ultraviolet Products Inc., California, USA). Photographs of the gels were taken by either a Polaroid Cu5 hand camera (5 inch lens/CU5 88-48 hood

providing 0.42x magnification onto Polaroid 667 film ISO 3000) or by a camera with a 16 mm Cosmicar television lens connected to a video copy processor (Model K61S, Mitsubishi) producing images on thermal paper.

2.2 ISOLATION OF PROBE DNA

DNA probes are often cloned into plasmid vectors which confer antibiotic resistance upon bacteria. Transfecting bacteria with a plasmid and growing them in media containing antibiotic selection allows the production of large scale amounts of plasmid DNA containing the probe. The probe can be separated from the plasmid vector by digestion with restriction enzymes and gel electrophoresis. These steps will now be discussed in detail.

2.2.1 *Working with bacteria*

Escherichia coli strain JM83 (λ^-, *ara*, Δ(*lac-pro* AB), *rps*L, *thi*-1, ϕ80*lac*ZΔM15) (Pharmacia) was used in all manipulations involving bacteria except for high efficiency cloning where Maximum efficiency DH5α competent cells (BRL) were used. Bacteria were grown at 37°C in the appropriate culture media: liquid cultures in a rotary shaker (controlled environmental shaker, New Brunswick Scientific, Edison, NJ, USA) whereas cultures on agar plates were grown in a dry incubator (LEEC CV2 Nottingham UK). Media and stock solutions were made with ddH_2O and autoclaved at 121°C for 15 min to ensure sterility. The following media were used:

Luria Broth (LB): Difco Bactotryptone 15g/l, Difco Bacto Yeast Extract 5g/l, NaCl 5g/l.

Luria Agar (L-agar): as LB but with Difco Bacto Agar 14g/l.

L-agar was melted in a microwave oven as required. Transformed bacteria were grown on antibiotic selection with the addition of the appropriate antibiotic after the LB was cooled to 65°C to prevent inactivation of the antibiotic. For short term use (up to 4 weeks) transformed bacteria were maintained on L-agar plates at 4°C, for long term storage cell stocks were frozen at -70°C in LB containing 15% (v/v) glycerol.

2.2.2 *Plasmid Vectors*

Probes come in various plasmid vectors of which Bluescript SK (Stratagene, UK) was the most usual. Bluescript has an origin of replication, an ampicillin resistance gene, a cloning site capable of being cleaved by multiple restriction enzymes and a colour selection system in the form of a ß-galactosidase gene system. Bacteria grown in the presence of 50 μg/ml of X-gal will produce a blue colour if the galactosidase gene is intact but will remain white if the gene has been disrupted by the presence of a cloned DNA sequence.

2.2.3 *Transformation of bacteria*

Transformation of bacteria with plasmid can be done in two ways: by using an electric shock to introduce the plasmid (electroporation) or by rendering the bacterial cell wall more 'porous', allowing the introduction of the plasmid DNA. We chose to routinely use the latter method.

2.2.3.a *Making competent cells*

The composition of all solutions is given in table 2.1. Incubating bacterial cells in an ice cold solution containing rubidium chloride renders them more likely to be transformed by plasmid DNA but the mechanisms remain unclear. 100 µl of an overnight culture of E.coli JM83 is added to 7 ml of LB and grown at 37°C shaking continuously for 2 hours until cloudy. 1 ml of this culture is then added to 200 ml of LB in a sterile flask and grown until the optical density of a sample of the culture is 0.48 at a wavelength of A^{530} using a spectrophotometer zeroed against LB. The culture is cooled on ice for 15 min and centrifuged in a Jouan CR412 centrifuge at 2000 rpm, at 5°C, for 5 min. The supernatant is discarded and the cells resuspended in 60 ml TAB1. The cells are left on ice for 15 min and then pelleted by centrifugation as above. They are then resuspended in 8 ml TAB 2 and left on ice for 20 min after which they are kept at 4°C overnight, and then stored in 150 µl aliquots at -70°C .

Table 2.1

Solutions for preparation of competent cells

TAB 1:	10 mM Calcium chloride
	35 mM Sodium acetate
	15% (v/v) glycerol
	- pH 5.9 with acetic acid then add:
	100 mM Rubidium chloride
	50 mM Magnesium chloride.
TAB 2:	10 mM MOPS, pH 6.8 with HCl
	10 mM Rubidium Chloride
	75 mM Calcium chloride
	15% (v/v) glycerol

2.2.3.b *Transforming competent bacteria*

50 ng of plasmid DNA was added to ice cold competent cells, mixed gently with the pipette tip and left on ice for 45 min to allow the DNA to adhere to the bacterial cell wall. Bacteria were then heat shocked at 42°C for 2 min to allow entry of plasmid DNA into the cells. 1 ml of pre-warmed LB was added and the cells left at 37°C in a water bath for 30 min to allow expression of antibiotic resistance. Aliquots of the transformed cells (10 µl and 100 µl) were then spread on L-agar plates containing 50 µg/ml of ampicillin and left overnight at 37°C. The transformation efficiency was calculated as the number of bacterial colonies transformed by 1 µg of plasmid DNA and should average around $1x10^6$ colonies/µg.

2.2.4 ***Preparing plasmid DNA***

Small scale bacterial cultures were prepared in 10 ml of LB ampicillin in a Universal tube. For large scale preparations, bacteria were grown in 100 ml of LB/ampicillin in an autoclaved conical flask or a Schott bottle mounted securely in an orbital shaker. Care was taken to ensure adequate oxygenation by not overfilling the containers and using cotton wool bungs. Bacterial cultures were grown for no more than 16 hr as autolysis of bacteria with degradation of both plasmid and probe DNA may result. The composition of all solutions is given in table 2.2.

Table 2.2

Solutions for preparation of plasmid/cosmid DNA:	
Solution 1:	50 mM glucose 25 mM Tris-HCl, pH 8 10 mM EDTA, pH 8
Solution 2:	0.2 mM NaOH 1% SDS - made fresh for each use
Solution 3:	3 M Potassium acetate pH 5 with acetic acid

2.2.4.a *Preparation of plasmid DNA - small scale*

1.5 ml of the overnight culture was pipetted into an eppendorf tube and spun in a microfuge at 13,000 rpm for 3 min. The supernatant was discarded and the pellet was resuspended in 200 µl of Solution 1. A few grains of lysosyme were added and the tube left for 2 min at room temperature. The cells were lysed by adding 400 µl of Solution 2, rendering the solution transparent. The bacterial DNA and protein was precipitated by adding 200 µl of Solution 3 and leaving the tube on ice for 20 min. Centrifugation at 13,000 rpm for 10 min allowed separation of the plasmid containing supernatant, and the precipitate, containing the protein and bacterial DNA, was discarded. Pipetting off the supernatant was done with care, as the precipitate is electrostatic and readily sticks to the pipette tip. Small scale plasmid DNA preparations were only used to confirm the presence of a sequence of interest and were not routinely treated with RNase. Plasmid DNA was precipitated by adding 0.6 x vol. iso-propyl alcohol and leaving on ice for 20 min. Centrifugation at 13,000 rpm for 10 min pelleted the plasmid DNA which was washed in 70% ethanol, air-dried and resuspended in 50 µl of TE.

2.2.4.b *Large Scale plasmid preparation*

The principles are the same but on larger volumes. A 100 ml culture was divided into 2 x 50 ml Falcon tubes and centrifuged at 4,000 rpm for 10 min. The supernatant was discarded and both precipitates resuspended in a total of 10 ml of Solution 1. 20 ml of Solution 2 and 10 ml of Solution 3 were added as above and, following DNA precipitation and washing, the plasmid DNA was resuspended in 5 ml of TE. 25 µl of 10 mg/ml RNase were added and the DNA incubated at 37°C for 40 min. Extraction with equal volumes of equilibrated phenol and phenol/chloroform (50:50) removed the remaining protein and inactivated the RNase. At each stage the sample was mixed thoroughly, then spun to separate the aqueous and solvent layers. The upper aqueous layer containing the DNA was transferred to a fresh eppendorf tube leaving the protein behind as a coagulated layer at the interface. A final extraction with an equal volume of chloroform alone removed the remaining phenol which would interfere with subsequent enzyme digestion of the plasmid. The DNA was precipitated in 2 x vol. ethanol and 1 x vol. of 3M Sodium acetate, washed in 70% ethanol and resuspended in 500 µl of TE. The concentration of the final DNA solution was ascertained by comparative DNA fluorescence as described in section 2.4.7.b.

2.2.5 *Generating insert*

10 μg of plasmid DNA was incubated with 20 units of a restriction enzyme appropriate to the cloning site into which the probe had been inserted, along with the restriction enzyme buffer, for 2 hr. Digestion was confirmed by agarose gel electrophoresis using 2 μl of the digest and, if fully digested, the whole sample was separated by gel electrophoresis, along with molecular weight markers. Two bands resulted, a linearised plasmid and the insert. The gel was placed on a cleaned UV transilluminator, the insert band cut out with a scalpel and placed in a 1.5 ml eppendorf tube. The insert DNA was separated from the agarose using commercially available kits (Geneclean II, Biotech Inc, La Jolla, USA.) according to manufacturer's instructions. The concentration of insert DNA was ascertained using comparative DNA fluorescence (section 2.4.7.b).

2.3 TISSUE CULTURE

All hybrids and cell lines were grown in appropriate media at 37°C in 7% CO_2 and all manipulations performed in a laminar flow hood under strict aseptic conditions. Hybrids were constructed by Dr. J. Cowell as described in Cowell (1992).

2.3.1 *Lymphoblastoid cell lines:*

Aliquots of lymphoblastoid cell lines were added to 30 ml of Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% foetal calf serum (fcs) and 2 mM glutamine. Suspensions were split 1:3 once the medium became orange, by pouring 20 ml of fresh medium into 2 new flasks and adding 10 ml of the original culture. 20 ml of fresh medium was then added to the original flask resulting in all 3 flasks containing 10 ml of cells and growth factors secreted into the original medium and 20 ml of fresh medium.

2.3.2 *Adherent cell lines:*

Adherent cell lines were grown in E4 medium supplemented with 10% fcs and 2 mM glutamine until just confluent. The medium was then removed and the cells washed in phosphate buffered saline (PBS). Cells were removed from the flasks in which they had been grown by adding 1 ml of trypsin, swirling to ensure complete coverage of cells and leaving for 5 min. The cells were washed off the flask with 5 ml medium, which inactivates the trypsin, and gently pelleted by centrifugation in a Jouan BR3.11 centrifuge (1200 rpm, 5 min). Cells were resuspended in medium and split 1:3 into new flasks.

2.4 PREPARATION OF GENOMIC DNA

The aim of this procedure is to produce DNA of high molecular weight in high concentration and free from contamination with RNA, protein, phenol and salt. Care was taken at each stage to avoid excessive shaking as shearing forces fracture DNA resulting in its degradation. The composition of the solutions used in the preparation of genomic DNA is given in table 2.3.

Table 2.3

Solutions for preparation of DNA

Lysis Buffer:	100 mM NaCl 100 mM Tris-HCl, pH 8 10 mM EDTA, pH 8
Equilibrated Phenol:	400 ml phenol -equilibrate with: 400 ml 1 M Tris-HCl, pH 8 then: 400 ml 0.1 M Tris-HCl, pH 8 0.4% (w/v) dihydroxyquinoline
RNase:	10 mg/ml in ddH_20 - boil for 10 min to inactivate residual DNase.
Proteinase K:	50 mg/ml in ddH_2O -incubate 65°C 1 hr to remove DNase
TE:	10 mM Tris-HCl, pH 8 1 mM EDTA, pH 8

2.4.1 *Harvesting of Cells*

Lymphoblastoid cells grown in suspension were removed from the growth medium by pipetting and centrifuged at 1,200 rpm for 5 min. at 20°C. The supernatant was discarded and the cells were resuspended in 10 ml of lysis buffer. Once fully resuspended the cells were lysed by adding SDS to 0.1% (w/v). The addition of SDS before complete resuspension resulted in clumping of cells, and reduction of DNA yield.

Adherent cells grown in flasks were harvested by pouring off the medium and washed in phosphate buffered saline (PBS). 3 ml of lysis buffer plus 0.1% SDS were added to each flask which was gently rotated to ensure the lysis solution was in contact with all of the cells. The cells were left to lyse for 10 min and the lysate removed from the flask by gentle rotation to 'strip' the lysed cells off the base.

Tumour samples were first deep-frozen by immersion in liquid nitrogen and then ground to a fine dust with a mortar and pestle. The ground tissue was resuspended in lysis buffer to which SDS 0.1% was added.

2.4.2 *Proteinase K from a stock solution of 10 mg/ml*

Proteinase K was added to the DNA lysate at a working concentration of 0.1 mg/ml and incubated at 37°C overnight or at 65°C for 1 hour, the former being more efficient at removing protein from the DNA solution.

2.4.3 *RNase from a working stock of 10mg/ml*

RNase was added to the lysate at a working concentrate of 50 μg/ml and incubated at 37°C for 40 min.

2.4.4 *Phenol/chloroform extraction of DNA*

A volume of equilibrated phenol equal to that of the lysate was added to the DNA and gently mixed for 2 - 3 min. The phenol was separated from the aqueous DNA by centrifugation at 4,000 rpm for 20 min in a Jouan CR412 centrifuge. The clear DNA solution formed the upper layer and the coagulated protein was at the phenol/DNA interface. The DNA was removed using a wide bore pipette and the phenol discarded. This procedure was repeated using 50:50 phenol and chloroform together (the addition of the chloroform increased the definition of the DNA/phenol interface). To remove all traces of phenol, which would otherwise interfere with subsequent DNA analysis, the DNA solution was finally extracted with an equal volume of chloroform alone.

2.4.5 *Ethanol Precipitation*

DNA was precipitated by adding 2 x vol. of ethanol and 1/10 vol. 3 M Sodium acetate. Cooling on ice for 20 min resulted in DNA precipitation, with the high molecular weight DNA forming lumps which could be 'spooled out' using a sealed glass Pasteur pipette. The DNA was washed in 70% ethanol to remove the salt, air-dried and then resuspended in the minimum possible volume of TE. This resulted in a solution containing at least 500 μg of DNA per ml.

2.4.6 *Preparation of DNA from cell lines for use in PCR reactions*

Screening large numbers of hybrid subclones for the presence or absence of a given sequence is easier by PCR than by Southern blotting and also requires much less DNA. The following method of preparing DNA to act as the template in PCR reactions can be performed much more rapidly than utilising the steps noted in sections 2.4.2 to 2.4.5. Cell lines are harvested from a single 9cm petri dish as described in section 2.4.1 and then transferred to an Eppendorf tube. The cells are spun in a microfuge for 30 sec, gently resuspended in 1 ml PBS and then spun once more. The supernatant is discarded and the final pellet of cells is resuspended in 25 µl of lysis mix, the composition of which is given in table 2.4.

Table 2.4

Cell lysis mix for rapid preparation of DNA for PCR

10mM Tris HCl pH 8.3
1.25 mM $MgCl_2$
50 mM KCl
0.01% gelatine (w/v)
0.45% Tween 20
0.45% NP40

2.4.7 *DNA quantitation*

Two methods were used: spectrophotometry and comparative analysis of fluorescence.

2.4.7.a *Spectrophotometry*

DNA concentration is determined by the absorbance of light through a sample of DNA which was diluted 1:20 in TE. The spectrophotometer was zeroed against an absorbance matched cuvette containing TE alone. Readings were taken at wavelengths of A^{260} and A^{280}. Nucleic acids alone absorb light at A^{260} whereas at A^{280} light is also absorbed by protein and other contaminants. This absorbance of light (the optical density (OD)) is related to the concentration of DNA (an OD at A^{260} of 1 is 50 µg/ml of DNA) The ratio of absorbance of 260/280 should be between 1.7 and 1.9 if the DNA is free from contaminating proteins and this measure of DNA 'cleanliness ' is the major advantage of this method. The disadvantage is that the DNA often settles out in

the photocell, making readings unstable and the method also requires relatively large volumes of DNA.

2.4.7.b *Comparative analysis of DNA fluorescence*

This method compares the fluorescence of an aliquot of DNA in ethidium bromide with that of DNA of known concentration. 1 µl of the test DNA sample was added to 10 µl of 2 µg/ml of ethidium bromide placed on transparent parafilm and mixed with a pipette tip. 1 µl aliquots of DNA standards of known concentration (5 ng/µl to 80 ng/µl) were added to 10 µl of 2 µg/ml ethidium bromide as above. The parafilm was placed on an ultraviolet light transilluminator and photographed. The DNA standard which compared most closely in fluorescence intensity to the test DNA is taken as the concentration of the sample. Dilution of the sample may be required to ensure that it falls within the scale of the panel of control DNAs. Although only semi-quantitative, this method has the advantage of ease, speed and uses much less DNA than spectrophotometry. This was the method of choice for DNA probes isolated by the Geneclean technique.

2.5 SOUTHERN BLOTTING

2.5.1 *Restriction Enzyme Digestion*

Restriction endonucleases are enzymes produced by bacteria to restrict the growth of other bacteria by cutting their DNA at specific sites, identified by a precise sequence of oligonucleotides. Their names come from the bacteria from which they were derived (e.g. Eco R1 is *Escherichia coli* R enzyme 1). They are used to cleave DNA in a reproducible way dependant on the presence of the recognition sequences and appropriate incubation conditions. 5 units of enzyme were used for each µg of genomic DNA to be digested, for plasmid and cosmid preparations this was reduced to 2 units/µg.

2.5.2 *Gel electrophoresis*

Electrophoresis through an agarose gel separates DNA fragments according to their size with the smaller ones migrating more rapidly than the larger. The rate of migration is determined by the size of the DNA fragment, the percentage of agarose in the gel and the voltage applied.

Agarose (ultrapure electrophoresis grade, Gibco-BRL) was melted in 1 x TAE buffer at the required concentration in a microwave oven and cooled to 65°C in a waterbath. Gels were cast in a 20 cm x 20 cm perspex former with the size of the perspex combs determining the capacity of each well. The percentage of agarose was

chosen so as to optimise the separation of the fragments of interest. For fragments 1.5 kb and smaller 2% gels were used, whereas for larger fragments 1% gels were used. All gels were run in a horizontal gel tank containing 1 litre of 1 x TAE buffer with ethidium bromide at 0.5 μg/ml . Loading buffer was added in a 1:10 dilution to all samples to increase their density, ensuring the DNA stayed in the wells. The composition of all the buffers used in gel electrophoresis is given in table 2.5. Molecular weight markers were loaded alongside the samples so that the size of the DNA of interest could be accurately determined. The choice of the molecular weight marker depended on the expected size of the DNA of interest, the larger fragments were more easily assessed using HHR markers. These consisted of a mixture of equal amounts of two digestions of lambda DNA: one digested with HindIII the other with HindIII then Eco R1 (Gibco BRL). Smaller fragments (less than 2 kb) were sized using 1 kb ladders (Gibco BRL). Examples of the electrophoretic mobility of the markers used are shown in figure 2.1 and 2.2.

'Checker gels' or electrophoresis of PCR products were run at 200 volts (Kingshill 15AOIC stabilised power supply), but if more accurate size separation was required, as for Southern blotting, the gels were run at 35 volts for 16 hours (overnight).

Separation of DNA products was visualised by placing the gel on an ultraviolet light transilluminator and the gel photographed with a ruler laid alongside so that the position of the band on the gel could be correlated with its molecular weight shown on the markers.

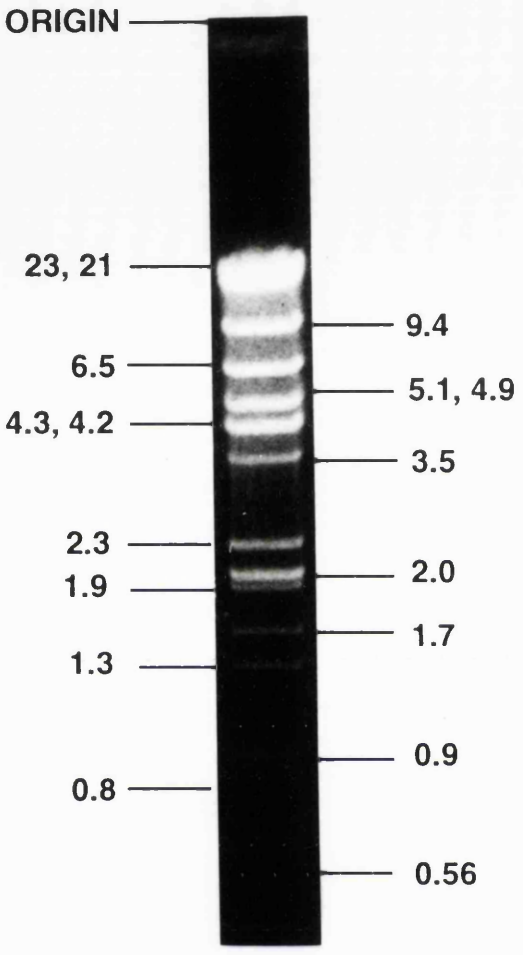


Figure 2.1. A photograph showing the λ HHR markers separated by electrophoresis on a 1% agarose gel. The sizes of the bands are given in kb.

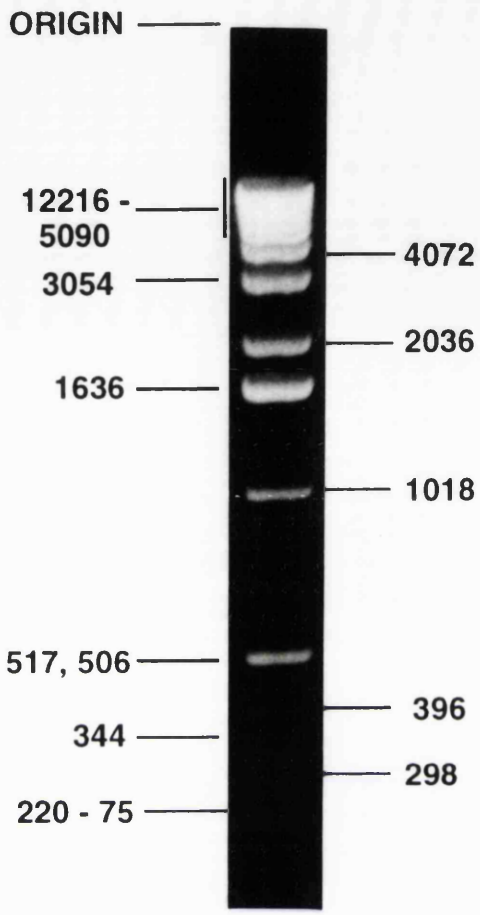


Figure 2.2. A photograph showing the 1 kb DNA ladder (Gibco BRL) following electrophoresis on a 2% agarose gel. The sizes of the bands are given in bp. It is not possible to distinguish any if the 8 bands between 12216 and 5090 bp in length and they appear as a broad smear (uppermost 'band').

2.5.3 *Southern Blotting of agarose gels*

Agarose gels are fragile and difficult to store for long periods so the DNA separated by electrophoresis is routinely transferred to a nylon-based membrane (filter) for further analysis by the technique of Southern transfer (Southern, 1975). The aim is to use capillary action to transfer as much of the DNA as possible, in single-stranded form, from the gel to the filter and fix it there so that the membrane acts as a usable copy of the original gel. The DNA is rendered single stranded by floating it in a denaturing solution of 1.5M NaCl; 0.5M NaOH for 45 min and the gel is then neutralised (0.5M Tris Base; 0.3M sodium citrate; 3m NaCl pH 5.5 with conc. HCl) for 1 hr in the same way. For blotting, a 2 cm high platform was placed in a tray and covered with 2 layers of 3MM Whatman paper which overlapped the platform and touched the bottom of the tray forming a wick. The tray was filled with 20 x SSC to just below the level of the platform. The gel was washed for 30 sec in 2 x SSC, carefully inverted, placed on a blotting platform and overlaid with the hybridisation membrane (Hybond N, Amersham, UK) which had been prewetted in 2 x SSC. Saran Wrap was used to cover the areas of exposed 3MM paper surrounding the gel to reduce evaporation of the 20 x SSC during the procedure and ensure that the capillary action is through the gel only. 3 pieces of 3MM Whatman paper, cut to the same size as the membrane and also presoaked in 2 x SSC, were placed over the Hybond N and air bubbles extruded by rolling a cylindrical plastic tube over the top. A stack of paper towels, approximately 10 cm thick, was placed over the Whatman paper and a weight, usually a metal tray, placed on top of the towels to ensure adequate and even contact of all of the layers. Blotting was performed overnight, after which the position of the wells on the gel, which should now be very thin, was marked on the membrane. The membrane was washed briefly in 2 x SSC, blotted dry and baked at 80°C for 2 hr to covalently link the DNA to the membrane. Storage was at room temperature, sandwiched between pieces of 3MM Whatman paper. Table 2.5 shows the composition of all solutions used in gel electrophoresis and Southern blotting.

Table 2.5

Solutions for gel electrophoresis and Southern blotting

Gel electrophoresis:

TAE:	0.04 M Tris acetate
	0.002 M EDTA
	- 1 litre of 50x stock is:
	242 g Tris base
	57.1 ml glacial acetic acid
	100 ml 0.5 M EDTA, pH 8
TBE:	0.089 M Tris base
	0.089 M Boric acid
	0.002 M EDTA, pH 8
	- 1 litre of 10x stock is:
	108 g Tris base
	55 g Boric acid
	9.3 g EDTA
Loading Buffer:	50% (v/v) glycerol
	0.42% (w/v) bromophenol blue.

Southern Blotting:

Denaturing solution:	1.5 M NaCl
	0.5 M NaOH
Neutralising solution:	0.5 M Tris base
	0.3 M Sodium citrate
	3 M NaCl
	- pH 5.5 with conc. HCl
20x SSC:	3 M NaCl
	0.3 M Sodium citrate

2.5.4 *Hybridisation*

The technique of Southern hybridisation involves radioactively labelling a DNA probe, denaturing it to render it single stranded, and applying it to a membrane to which single stranded DNA has been fixed. If a sequence complementary to the DNA probe exists on the membrane, the probe will hybridise to it and be recognised as a black band on X-ray film laid on top of the hybridised membrane (autoradiography). The composition of all the buffers used is shown in table 2.6.

2.5.4.a *Oligolabelling*

The aim is to generate a single-stranded copy of the probe labelled with a substance that allows detection of the probe once it has hybridised to homologous sequences immobilised on the membrane. Although detection systems based on colour change, or light production exist the most sensitive method remains labelling the probe with radioactivity. Oligolabelling results in exact copies of the single stranded probe being made using random hexanucleotides, non-radioactive deoxyribonucleotides dATP, dGTP and dTTP and radioactive dCTP, and the reaction is catalysed by the Klenow fragment enzyme as described in Feinberg & Vogelstein (1983).

70 ng of probe DNA was made up to a total volume of 21 µl with ddH_2O, boiled for 3 min to denature the DNA and snap cooled on ice. 3 µl of 10x oligolabelling buffer (OLB) and 1.5 µl of 10 mg/ml bovine serum albumin (BSA) were added followed by 3 µl radio-phosphorous labelled deoxycytidinetriphosphate (^{32}p dCTP, ›5000 Ci/mmol). 1.5 µl of Klenow DNA polymerase (NBL) was added to the reaction mix which was left at room temperature in a perspex box for 4 - 16 hours. To increase the specificity of the hybridisation, the labelled probe was separated from the un-incorporated ^{32}p dCTP by elution through a sephadex G-50 column. Sephadex G50 was autoclaved with sufficient TE to ensure full adsorbtion of fluid to the Sephadex beads. A ball of polymer wool was inserted into a 145mm glass Pasteur pipette and Sephadex suspension added until the Pasteur pipette was full. TE drips out of the column leaving compacted Sephadex beads. The column was clamped vertically on a retort stand and a rack containing nine eppendorf tubes positioned below. The 30 µl of labelled probe mix was added to the column followed by 150 µl aliquots of TE. The first four 'drip through' aliquots were collected in a single eppendorf but, thereafter, each 150 µl aliquot was dripped into a separate tube until all nine tubes had been used. Tube 1 contained 600 µl, tubes 2 to 9 150 µl. Each tube was placed 20 cm from a collimated Geiger Muller monitor (Mini-L. Mini Instruments Ltd., Essex, UK). The eluted probe formed the first 'peak' of radioactivity, normally tubes 4-6, after which the radioactive signal declined until tube 8 and 9 in which the second peak represented un-incorporated nucleotides.

The specific activity of the probe was calculated from a 5 µl sample using a Bioscan QC 2000 counter (V.A.Howe Co. Ltd., London, UK). The specific activity was usually in the region of 2 x 10^7 cpm /µg of probe DNA but this depended on the quality of the DNA and the age of the radionucleotides used.

Table 2.6

Solutions used in Southern hybridisation:

10x Oligolabelling buffer:

Solution A: 1.25 M Tris-HCl, pH 8
125 mM Magnesium chloride
225 mM ß-mercaptoethanol
0.5 mM of dGTP, dATP, dTTP

Solution B: 2 M Hepes, pH 6 with NaOH

Solution C: 90 OD A^{260} units/ml random hexanucleotides in TE

-Mix solutions A, B, C in the ratio of 2:5:3

Hybridisation buffers:

Denhart's: 4x SSC
10x Denhart's solution
0.1% SDS
0.05 mg/ml boiled, sonicated salmon sperm
-100x Denhart's solution is :
2% (w/v) bovine serum albumin
2% (w/v) Ficoll
2% (w/v) polyvinylpyrrolidone

Church: 7% SDS
0.5 M NaPi, pH 7.2
1 mM EDTA
1% Bovine serum albumin
- 1 M NaPi is:
71 g Na_2HPO_4 in 1 l ddH_2O
-pH 7.2 with phosphoric acid

2.5.4.b *Hybridisation*

The aim of the hybridisation technique is to:

- prepare the membrane so as to minimise the non-specific binding of the radioactive probe.
- To add single stranded probe DNA
- To perform the hybridisation reaction in such a way as to optimise the specificity of annealing between probe and membrane bound homologue without decreasing sensitivity. This could be achieved by altering the type of hybridisation fluid used and introducing competitive binding of any repetitive DNA sequences in the probe.

Hybridisation solution needs to contain a wetting agent, a salt concentration which allows annealing of homologous sequences with little non-specific binding and a 'blocking agent' to prevent physical trapping of radioactive probe in the irregularities in the membrane. The most frequently used hybridisation solution was Church buffer, which produced autoradiographs with little background signal. However, some probes produced poor signal intensity with Church buffer and for these the less stringent Denhart's solution was used. Both solutions were used for hybridisation at 65°C. It is important that the membrane is in contact with freely circulating hybridisation fluid at a strictly controlled temperature and I found that placing the membrane in rotating cylindrical glass bottles in a hybridisation oven (Hybaid, UK) to be the most efficient method. Membranes were prewetted in 2 x SSC if Denhart's hybridisation solution was used or 0.5 M NaPi for use with Church buffer. The membrane was placed on a nylon mesh in the appropriate solution and, if more than one membrane was to be hybridised with the same probe, a further nylon mesh was inserted between each membrane to ensure adequate circulation of hybridisation solution. The membranes were rolled up, placed in a pre-warmed glass hybridisation bottle and 20 ml of 2 x SSC or 0.5 ml NaPi at 65°C was added. The hybridisation bottle was then rolled in the direction which caused the membrane to adhere to the sides of the bottle. The solution was poured off, hybridisation solution was added (minimum 15 ml) and the bottle placed in the rotisserie of the preheated hybridisation oven in such a way that the direction of rotation of the rotisserie did not cause the membranes to roll up again and lose contact with the hybridisation fluid.

The membranes were 'prehybridised' (to block any physical trapping of probe) for 2 hours if Church buffer was used or for 16 hours in Denhart's was used. Only in the case of Denhart's solution was the hybridisation solution replaced with a fresh aliquot after prehybridisation.

2.5.4.c *Addition of Probe*

The separated probe was boiled for 3 min, snap cooled on ice and added to the hybridisation fluid in the cylindrical glass bottle, aiming to achieve a specific activity of 1 x 10^6 cpm per ml of hybridisation solution. Hybridisation was for 16 hr at 65°C. However, if the probe was known to contain repetitive elements, their effect could be reduced by competitive binding with Cot-1 DNA (total human DNA enriched for repetitive sequences, Gibco BRL). The exact amount of competition can be calculated as described by Sealey et al (Sealey et al, 1985) but in most cases successful results were obtained by adding 5 µl of 10mg/ml Cot-1 DNA, 50 µl of 20xSSC, 25 µl of ddH_20 and 20 µl of 0.1% (w/v) SDS to 100 µl of the separated probe, boiling for 3 min and then allowing it to anneal at 65°C for 1 hr. The probe was then added to the hybridisation fluid in the bottle without further boiling. Competition could also be applied at the oligolabelling stage, and the addition of 100 µg/ml of sheared human placental DNA to the hybridisation fluid further reduced the effect of repetitive elements in the probe.

2.5.4.d *Washing*

Non-specific binding of un-incorporated nucleotides and probe DNA occurred in spite of adequate prehybridisation. Washing aimed to displace the non-specific hybridisation but leave the probe hybridised to homologous sequences on the membrane. The stringency of washing is determined by salt concentration and temperature: the lower the salt concentration and the higher the temperature, the more stringent the wash. Reducing salt concentration is a greater change in stringency than increasing the temperature. The radioactive hybridisation solution was poured off and the remaining traces removed by swilling the bottle out with 2 x SSC 0.1 % SDS for Denhart's, 100 mM NaPi 0.1% SDS for Church buffer. If the washing conditions for the probe were known, the correct washes were performed in the hybridisation bottle, if not, sequential increases in stringency were applied. Deciding whether to increase the stringency of the wash was based on checking the filter with a Geiger Muller monitor (Mini-L, Mini Instruments Ltd., Essex, UK). High generalised activity indicated the need for more stringent washing , or the presence of repetitive sequences in the probe.

2.5.4.e *Autoradiography*

Once washed, filters were blotted with 3MM Whatman paper to remove surface fluid and sandwiched between pieces of Saran wrap. They were then placed in light proof cassettes and exposed to XAR-5 autoradiography film (Kodak) with intensifying screens (Kronex Quanta 3, DuPont) at -70°C for 16 hr and the film was then developed. The position of a black band on autoradiography denoted the presence of a sequence

homologous to the probe at the corresponding position on the filter. If the band appeared pale, the exposure time was increased. To allow the membranes to be re-used more quickly than if the radioactivity band was allowed to decay naturally, the radioactive probe was stripped off by immersing the membranes in boiling 1% SDS, 10 mM Tris HCl pH 8 and leaving to cool to room temperature with vigorous agitation.

2.6 POLYMERASE CHAIN REACTION

Polymerase chain reaction (PCR) allows selective amplification of short DNA sequences between selected oligonucleotide primers from DNA samples used as templates (Saiki (1985), Mullis (1987)). Although there are many variations, each PCR cycle has the following stages: denaturation of template DNA, annealing of oligonucleotide primers to template sequences, extension of the sequence from one of the oligonucleotide primer pair in the 5' and 3' direction until it reaches the other primer, resulting in the production of 2 strands homologous to the template. The number of template sequences is, therefore, doubled with each cycle. 30 cycles of PCR would result in 2^{30} ($1x\ 10^9$) fold amplification of the original DNA template, assuming complete fidelity of the reaction. Very small amounts of template DNA (10-100 ng) can be used and the technique is less sensitive to DNA degradation than Southern hybridisation. However, minute amounts of contamination give rise to false positive results. Extreme caution must be exercised to prevent this occurring and a negative control (containing no template DNA) was always included in the experiment. The method is described in detail below:

2.6.1 *PCR Premix*

Reactions were carried out in a total volume of 50µl and required:

- DNA template (100 ng)
- a pair of oligonucleotide primers specific for the sequence to be investigated (2 µM)
- deoxyribonucleotide triphosphates (dATP, dTTP, dGTP, dCTP 200 µM each)
- appropriate buffer (5 µl of 1:10 Taq polymerase buffer) [500 mM KCl, 100 mM Tris HCl, 0.1% gelatine w/v, 1% Triton X w/v, 15 mM $MgCl_2$.] .
- 50 µl Mineral oil was added to prevent evaporation during the reaction
- a heat stable DNA polymerase (Promega, thermus aquaticus [Taq], 1 unit).

2.6.2 *Method*

All reactions were performed using an automated Techne 2 thermal cycler by the following method. An initial denaturation step of 10 minutes at 94° C was

performed and the temperature then dropped to 10° C above the annealing temperature. 1µl of enzyme mix (1 µl x 10 Taq buffer, 2 µl Taq polymerase, 7µl ddH2O) was added to each tube. The subsequent cycles proceeded as follows.

Denaturation at 94°C normally for 30 sec but longer if the sequence to be amplified was greater than 500 bp in length.

Annealing at a temperature determined by the melting temperature of the oligonucleotide primers being used, and for a time determined by the length of the expected sequence.

Extension at 72°C, the time being determined by the length of the expected product.

Number of cycles The amount of product doubles at each cycle but above 30 cycles the fidelity of the reaction can decrease as secondary non-specific amplification occurs. For screening for the presence or absence of a particular DNA sequence identified by the oligonucleotide primers 30 cycles was sufficient, but in circumstances where maximum product yield was necessary, up to 40 were used.

The precise reaction conditions for each set of oligonucleotide primers used is given in appendix 1.

5 - 10 µl of the resulting product was separated by electrophoresis on a 2% agarose gel with 1 kb molecular weight markers in TAE buffer including ethidium bromide and viewed on a UV light source.

2.6.3 *PCR cloning*

Using primers derived from the Alu human repetitive sequence, PCR will amplify short segments of DNA in between adjacent Alu sequences. These can be cloned into a plasmid vector, used to transform bacteria and then characterised. By choosing the DNA template and Alu primers with care, a library of inter-Alu sequences specific for a given chromosomal region of interest can be constructed (Cotter et al, 1990, Nelson et al, 1989).

2.6.3.a *Alu-PCR reaction*

2 µM of 3' Alu or 5' Alu primers were added to a 50 µl reaction PCR mix as above and 100 ng of DNA template added. The reaction conditions and primer sequences are given in appendix 1.

2.6.3.b *Cloning of PCR products*

The PCR products were purified on Sephadex G 50 spin columns. A ball of polymer wool was placed in the neck of a 1 ml plastic syringe and Sephadex G50 in TE was added. A 1.5 ml eppendorf tube without a lid was placed in a 15 ml plastic Falcon tube to collect eluent from the column. The syringe was inserted into the 15 ml Falcon

tube ensuring that the tip rested in the eppendorf, and the top of the Falcon tube screwed on. The 'spin column' was centrifuged in a Beckman J2.21 at 1000 rpm for 5 min to pack the Sephadex tightly and the eppendorf containing the excess TE replaced. The PCR product was added and the centrifugation step repeated. 100 µl of TE was added and the tube centrifuged a third time. The eluents from the latter two steps were pooled and the DNA precipitated by adding 2.5 x vol. ethanol and 0.1 x vol. 3M sodium acetate and cooling to -20°C for 6 hours. The precipitated DNA was pelleted by centrifugation in a microfuge at 13000 rpm for 20 min and the supernatant discarded. The pellet was washed in 70% ethanol to remove excess salt and the centrifugation step repeated. The supernatant was again discarded and the precipitate air dried and resuspended in 10 µl of TE. The DNA concentration was measured by comparative DNA fluorescence.

Bluescript SK vector (Stratagene) was linearised by cutting with Eco RV, which produced a 'blunt-ended' fragment, and analysed by electrophoresis on a 1% agarose gel. The linearised plasmid, a single band of 3 kb, was cut out of the gel and the DNA eluted by the Geneclean method described in section 2.2.5. The concentration of cut plasmid vector was also determined by comparative DNA fluorescence.

In order to calculate the molar ratios of insert (Alu PCR product) to vector (cut plasmid) the 'average' size of the Alu PCR product was taken as being 500 bp. The size of the cut vector was known and the molar ratios could be calculated as in the example below:

Alu PCR product 100 ng/µl , average size 500 bp
vector 100 ng/µl, size 3 kb.
Therefore an insert:vector ratio of 1:1 by volume would be 6:1 by molar concentration.

Aliquots of 100 ng (1µl) of Alu PCR product were ligated into EcoRV cut Bluescript at a variety of different insert to vector ratios (1:1 to 8:1) of which 8:1 gave the highest transformation efficiency. Ligations were performed at 14° C for 16 hr in a final volume of 20µl with Polyethylene glycol (PEG) 8000 (Sigma) at 5% concentration, 200 units of T4 DNA ligase and 2 µl of 10 x ligase buffer containing rATP. The ligation mix was diluted 1:5 with TE and 5 µl of the diluted reaction mix was used to transform 100 µl of maximum efficiency DH5α competent cells (BRL) according to manufacturer's instructions. Transformed cells were plated out onto LB agar plates containing ampicillin (100 µg/ml) and X-gal (50 µg/ml). Colonies with destruction of ß-galactosidase gene (white colonies) were gridded out onto LB-ampicillin plates and analysed further.

2.6.3.c *Identification of repetitive sequences by in-situ colony hybridisation*

Exact replicas of the plates on which the white colonies had been gridded were made and the colonies were grown at 37° C until 2 mm across. The plate lids were removed and the plates left at 4°C for 30 min to ensure that the colonies stood 'proud' of the agar. A sheet of Hybond N membrane was gently laid on top of the colonies and pinholes made through the filter into the gel to allow accurate orientation. After 5 min the filter was carefully removed with the colonies stuck to it. The membrane was floated, colony side up, on denaturing solution for 5 min and then transferred to neutralising solution for 5 min ensuring that the membrane did not become submerged which could result in the colonies floating off. Placing the membrane on dry 3MM Whatman paper ensured that DNA from the colonies was transferred to the membrane by capillary action. Filters were then baked at 80°C for 2 hr. 70 ng of sheared total genomic human DNA (average size 500 bp) was oligolabelled and hybridised to the colony filters as described in section 2.5.4.a-c. Church buffer was used as the hybridisation solution and washing was to 5 mM NaPi 0.1% SDS at 65°C for 30 min. Autoradiography was performed overnight without an intensifying screen. Dense black dots on the autoradiograph denoted the presence of repetitive sequences in the colonies at an identical position on the plate.

2.6.3.d *Amplification and Characterisation of Alu PCR plasmid inserts directly from colonies*

Cells from each colony of interest were transferred with a sterile toothpick to the PCR mixture described above and were amplified using the original 3' or 5' Alu primers. A 5 µl aliquot of the reaction mixture was run on a 2% agarose gel to confirm the size and concentration of PCR product. Simply touching the surface of the bacterial colony with the toothpick transferred sufficient DNA for PCR analysis. Indeed, transferring too much of the colony to the PCR mix inhibited the PCR reaction. If restriction enzyme digestion of the product was necessary, the appropriate enzyme was added directly to the PCR reaction mixture and incubated for 2 hours at a temperature appropriate for the enzyme. The PCR buffer acted as the buffer for the restriction enzyme with no modification. To purify the Alu PCR plasmid inserts, 1 µg of the product was run out on a 2% agarose gel, excised and extracted using Geneclean according to manufacturer's instructions.

2.6.3.e *Hybridisation of Alu PCR probes to somatic cell hybrid mapping panels*

Samples of 20µg of normal human mouse 3T3, Chinese hamster and DNA from somatic cell hybrids were digested with EcoR1 or Hind III and separated by gel electrophoresis for 16 hours at 30 volts in 0.8% agarose gel. The gel was then denatured, neutralised and the DNA transferred to Hybond N membranes as noted in

section 2.5.3. 70 ng of purified Alu PCR probe was oligolabelled by the random primer method and the filters were hybridised in Denhart's solution at 65°C in a Hybaid oven. Washing was to 2xSSC 0.1% SDS at 65°C for 30 min (section 2.5.4).

2.7 YEAST ARTIFICIAL CHROMOSOME ANALYSIS

Yeast artificial chromosome (YAC) technology allows the cloning of extremely large DNA inserts, up to 2 Mb in size, in a vector that enables them to function as independent chromosomes in the yeast host (AB1380 strain of Saccharomyces cerevisiae) and be recognised by the presence of selectible marker; usually the ability to grow in restricted medium. Using rare cutting restriction enzymes, total genomic DNA has been digested into large fragments which have been cloned into a YAC vector and used to transform yeast host cells. The resulting transformants have been gridded out and act as YAC libraries of large human DNA sequences e.g. the ICRF Human Genome Reference Laboratory (Lehrach and et al, 1990). However, the insert DNA often undergoes homologous recombination and, therefore, genetic rearrangement, especially the development of chimerism, is common. The aim of screening YAC libraries is to identify YACs containing the sequence of interest, ascertain their size, and characterise the DNA insert. Screening can be performed by hybridisation or PCR. Once obtained, YAC clones are grown on selected media (deficient in uracil in the case of the ICRF YAC library). For pulsed field gel electrophoresis (PFGE) the YAC DNA needs to be prepared in such a way as to ensure that the chromosomes stay intact. The shearing forces produced by the preparation of DNA in solution would fracture the YACs so all stages are performed with the YAC containing yeast cells embedded in agarose plugs. For standard enzyme digestion and PCR analysis, DNA of 20 kb is long enough and DNA can be prepared in solution. The cloned insert can be characterised in a number of ways including the rescuing of the ends of the insert and Alu fingerprinting. These steps are now described in detail.

2.7.1 *Screening YAC libraries*

2.7.1.a *Hybridisation*

In the ICRF library 20,000 YAC clones are gridded robotically onto 20 cm x 20 cm hybridisation membranes in ordered arrays and the DNA from the clones denatured, neutralised and fixed to the membrane. 70 ng of probe DNA is oligolabelled by the random probing method and hybridised to the YAC library membrane in Church buffer at 65°C for 16 hours. Washing is to 100 mM NaPi/ 0.1% SDS at 65°C for 30 min for cDNA probes, and 50 mM NaPi 0.1% SDS 65°C for 30 min for genomic probes. In order that the location of the positive clones can be assessed relative to the filter as a

whole, yeast host DNA is oligolabelled using ^{35}S dATP. 70 ng of sheared AB1380 DNA is oligolabelled by the random priming method outlined in section 2.5.4. except that the oligolabelling buffer is made up with dCTP, dGTP, and dTTP (i.e. lacking dATP). Instead of ^{32}p dCTP, the yeast host is labelled with 3 µl of ^{35}S dATP. The rest of the oligolabelling procedure was identical to that described for 'standard' DNA probes (section 2.5.4.a) but the radioactive signal strength of the ^{35}S labelled probes had to be assessed using a specially collimated Geiger Muller monitor. The ^{35}S dATP labelled yeast host hybridises to all the colonies on the filter but less intensely than the ^{32}p probe, allowing the intense probe signal to be localised on the background of weaker yeast host signal (figure 3.23). The co-ordinates for the positive clone are then accurately ascertained.

2.7.1.b *PCR*

Screening the ICRF YAC library by PCR was a three stage process. In this library 20,000 YAC clones are divided into 41 primary pools. Each primary pool is composed of DNA from all the colonies on 4 microtitre plates. Each microtitre plate has been gridded into rows and columns (12 x 8) and all the colonies in each row and column made into a DNA pool. The first step is to screen the 41 primary pools and identify any positives. Each primary pool contains DNA from almost 400 individual YACs so the amount of PCR product obtained from screening the primary pools is often very low and sometimes below the limits of detection by UV transillumination of an ethidium bromide stained gel. If there appear to be no positive YAC pools following UV transillumination, but the positive control in the PCR reaction has produced a good signal, the positive YAC pools can be visualised by Southern blotting the gel and probing it with an ^{32}P labelled aliquot of the positive control. This 'self to self' hybridisation increases the sensitivity of detection of the positives. The primary pools that had been identified by PCR screening are made up of DNA pools from 4 separate microtitre plates which were then screened individually to ascertain which contained the colony of interest. Once a positive microtitre plate had been identified by PCR, screening DNA pools made from the individual rows and columns allowed identification of the YAC containing the sequence of interest.

2.7.2 ***Growing YACs***

Yeast colonies identified by screening arrive as stab cultures in agar. They are streaked out on uracil-deficient medium which acts as the selective pressure for the retention of YACs in the yeast host. The composition of the media and other solutions used in YAC work are given in table 2.7. The plates are incubated at 30°C (**not** 37°C) for 48 hours until single colonies of 2 mm diameter appear. Single colonies are picked and added to 15 ml YAC broth and incubated at 30°C overnight with shaking at 200

rpm. 5 ml of the resulting culture was centrifuged at 2000 rpm for 10 min in a Jouan CR412 centrifuge, resuspended in 2 ml of YAC broth plus 15% glycerol, and kept at -70°C as a long-term stock. 1 ml is added to 100 ml YAC broth in an autoclaved conical flask and incubated overnight at 30°C with shaking at 200 rpm.
This culture will go to make solution DNA. The remaining 9 ml is used to make agar plugs of YAC DNA for pulsed field gel electrophoresis.

Table 2.7

Solutions used in yeast artificial chromosome work

Uracil deficient YAC broth:

A: 28 g Casamino acids
110 mg Tyrosine
40 g glucose
200 mg Adenine hemisulphate
-make up to 1860 ml with ddH_2O and autoclave in 400 ml aliquots

B: 33.3 g/500 ml Yeast Nitrogen base without amino acids
-filter sterilise

C: Tryptophan 1 mg/ml
-filter sterilise

-Mix 400 ml of A, 40 ml of B and 8 ml of C.
For the culture of the uracil dependant yeast host add:
10x Uracil stock: 200 µg uracil/litre -filter sterilise

Guanidinium hydrochloride for preparation of DNA in solution:
4.5 M Guanidinium hydrochloride
0.1 M EDTA
0.15 M NaCl
0.05% Sarcosyl, pH 8

Yeast lysis buffer for agarose plugs:
1% Lithium dodecyl sulphate
0.1 M EDTA
10 mM Tris-HCl, pH 8

2.7.3 *Making DNA*

Unlike bacteria, yeast are resistant to lysis by SDS until their outer coats have been digested with lyticase. The cells are then fragile and have to be handled with care as both shearing forces and osmotic differences may well rupture them, leading to reduction of DNA yield. The morphology of the cells changes to spheres, a shape that allows the maximum volume of cytoplasm to be contained within a given surface area of cell membrane. The cells are then called spheroplasts.

2.7.3.a *Making solution DNA*

Yeast cells are pelleted by centrifugation at 3000 rpm for 10 min in a Jouan CR412 centrifuge at 20°C and resuspended in 5 ml of 0.9 M sorbitol, 20 mM EDTA, 14 mM ß-mercaptoethanol (ß-ME). The sorbitol ensures osmotic stability of the cells after spheroplasting and the EDTA chelates magnesium and therefore inhibits autolysis of YAC DNA. 20 µl of 10 mg/ml lyticase (Sigma) resuspended in 0.05 M potassium phosphate buffer, is added and the yeast cells incubated at 37° C with gentle shaking. The cells are pelleted by centrifugation at 1,200 rpm and resuspended in 5 ml of guanidinium hydrochloride solution. Lysis is ensured by heating to 65°C for 10 min. On cooling to room temperature the solution is transparent and often viscous. The nucleic acids are precipitated by adding an equal volume of ethanol and incubated on ice for 30 min. Centrifugation at 2,500 rpm for 10 min pellets the nucleic acids which are resuspended in 2 ml of TE. RNase (10 mg/ml) stock is added to a working concentration of 100 µg/ml and incubated for 30 min at 37°C. Proteinase K is added to 200 µg/ml and the YAC DNA incubated at 65°C for 1 hour. The lysate is very proteinatious and DNA is purified by 3 phenol/chloroform extractions followed by an extraction with chloroform alone. The DNA is precipitated by adding 2.5 x volume ethanol, 0.1 x 3M sodium acetate and incubated on ice for 30 min. The DNA is pelleted, the ethanol removed, the pellet washed in 70% ethanol, air dried and resuspended in 250 µl of TE.

2.7.3.b *Making Yeast DNA in agarose plugs*

In order to prevent degradation of yeast chromosomes all steps need to be performed with meticulous attention to detail and with no gaps in between them. A 9 ml culture of yeast cells in YAC broth, grown as above, was centrifuged at 1,500 rpm in a Jouan CR412 centrifuge for 10 min. The supernatant was discarded and the cells washed once in 50 mM EDTA, pH 8 to precipitate magnesium, thereby preventing endouclease digestion. The cells were pelleted again as above and resuspended in 400µl 0.9 M Sorbitol, 20 mM EDTA, 14 mM ß ME and 1 mg/ml lyticase. This was added to 500 µl of previously prepared 2% low gelling temperature (LGT) agarose in

0.9M sorbitol, 20mM EDTA (to which 14 mM ß ME had been added once it had cooled to 45°C). The two solutions were mixed by gentle swirling, then kept at 45°C and dispensed in 100 µl aliquots into slots of a standard plug mould. The agarose plugs were allowed to set on ice for 20 min and then extruded into a 15 ml Falcon tube containing 5 ml of 0.9 M sorbitol, 20 mM EDTA, 14 mM ß ME, 10 mM Tris-HCl, 1 mg/ml lyticase. This was incubated at 37°C for 2 hr in order that the yeast cells became spheroplasted. The yeast cells were lysed by replacing the solution with 5 ml of filter sterilised lysis solution and incubating at 37°C for 60 min, after which the plugs should appear less opaque than before lysis. The solution was poured off, another 5 ml of yeast lysis solution added, and the plugs incubated overnight. The solution was then replaced with another 5 ml yeast lysis solution in which the plugs were stored at room temperature. The plugs were used directly for PFGE, but for enzyme digestion, were washed in TE for 2x 30 min at 50°C and then in TE for 2 x 30 min at room temperature.

2.7.4 *Characterising YACs*

The presence or absence of DNA sequences in the YACs was ascertained by either Southern hybridisation as described in section 2.5. or PCR. DNA for the PCR template came from solution DNA, from colonies (touched with a sterile toothpick which was stirred in the PCR premix), or from YAC plugs (by adding a small sliver of a yeast plug to the PCR mix). Methods used to characterise YACs are described in sections 2.7.4.a and b.

2.7.4.a *PFGE to ascertain the size of YACs*

Pulse field gel electrophoresis separates very large DNA fragments by using an electric current which alternates in two different directions (Schwartz and Cantor, 1984). 3 litres of 0.5 x TBE are placed in a pulse field electrophoresis tank (2015 Pulsaphor, LKB Bromma) and cooled to 10°C (2219, Multitemp II Thermostatic Circulator). A 1% agarose gel was made in 0.5X TBE without ethidium bromide and set in a rubber surround on a perspex gel tray into which locating holes had been drilled. A perspex comb was located at one end of the gel on small plastic supports. Once set, the gel was loaded with the yeast plugs, each well being large enough to take half a yeast plug. YACs were always loaded in duplicate and PFGE markers (Pharmacia) included on the gel. Figure 2.3 shows the sizes of the PFGE markers. Electrophoresis was at 150 volts (Electrophoresis Power Supply, EPS 500/400, Pharmacia) switch time 80 sec for 48 hr (LKB Bromma, Pulsaphor 2015) at 10°C. The gel was then cut off its supporting tray, stained for 20 min with gentle agitation in 100 ml of running buffer to which 100µl of 10 mg/ml ethidium bromide had been added and photographed over a UV transilluminator with a ruler alongside. Ethidium bromide was removed by placing the gel in 200 ml of running buffer for 20 min and then the gel was acid depurinated by gentle agitation in 250 ml dH_2O, 5 ml concentrated HCl for 20 min. The gel was denatured, neutralised and Southern blotted as described in section 2.5. Probing the blot with sheared total human DNA allowed the size and number of YACs per plug to be determined.

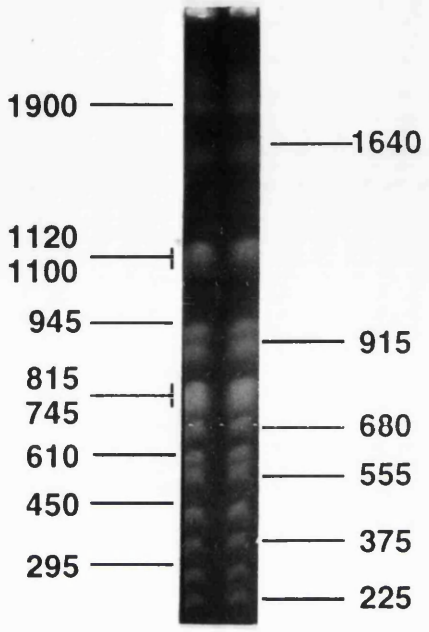


Figure 2.3. A photograph showing the sizes, in kb, of the bands of the Pharmacia PFGE markers. A 1% agarose gel was used and run at 10°C in 0.5x TBE. The running conditions were 200 V with a switch time of 70 sec for 15 hr, then a further 12 hr with a switch time of 120 sec.

2.7.4.b *Alu fingerprint of YACs*

Using primers derived from the human specific repetitive sequence, Alu, PCR allowed amplification of DNA fragments in between Alu repeats. Products were derived from the human inserts only and not from yeast host sequences. The pattern of bands derived from different YACs was used to determine whether the YACs contained identical human sequences. The individual bands could also be sub-cloned and used as probes. The following Alu primers were used: ALE 1, ALE 3, Alu IV, BK33 and 3' Alu; the reaction conditions and their oligonucleotide sequences are found in Appendix 1.

2.7.5 ***End Rescue of YACs***

Rescuing sequences from either end of the human insert in a YAC greatly facilitates YAC walking. Two main methods were used:

2.7.5.a *Alu/YAC vector PCR*

The aim of this procedure is to amplify the DNA sequence in-between the first Alu repeat in the human insert and specific regions close to the cloning site in the pYAC4 vector on either side of the cloning site. For each YAC three PCR reactions were performed. In the first tube YAC template DNA was amplified using one of the Alu primers mentioned above. In the second tube the same YAC template DNA and Alu primer were used but primers for the left-hand end of the pYAC4 sequence (1091) were also added. A third tube consisted of YAC template DNA, Alu primer and primers for the right hand end of the pYAC4 vector (HYAC-C). The products of the three PCR reactions were loaded next to each other on a 2% agarose gel and separated by electrophoresis. The banding pattern obtained by amplification of YAC DNA with Alu primers alone was compared with that obtained by Alu primers and either of the end sequences. Any extra bands produced by the addition of 1091 or HYAC-C primers represented human sequences in between the YAC end primer (1091 or HYAC-C) and the first Alu sequence in the human insert. These DNA fragments were excised and eluted from the gel using the Gene Clean technique. The reaction conditions for the PCR were as described for the specific Alu primer and are found in Appendix I. The success of the method depends on their being an Alu sequence within approximately 3 kb of the end of the human insert (the size limit of the PCR reaction).

2.7.5.b *Vectorette end rescue*

This technique (Riley et al, 1990) consists of digesting the YAC with a frequently cutting restriction enzyme and ligating an oligonucleotide cassette onto both ends of each of the resulting fragments. The two strands of DNA making up the

cassette are homologous at the ends of the cassette but non-complementary in the middle and for this reason the construct is often called a 'bubble'.

Using primers from within each end of the YAC vector and a primer for the bubble enabled us to rescue only the fragments at each end of the human insert, i.e. those containing the YAC end attached to part of the vector. The rescued ends were then digested with Eco RI, to remove YAC vector sequences and used as probes for hybridisation. To generate sequence tagged sites (STSs) from the YAC ends the rescued fragments were sequenced as described in section 2.8.

In order for the bubble to anneal to the digested YAC fragments the end has to be complementary to the cut site of the restriction enzyme used in the digest. The 'bottom' of the bubble was always the same but the 'top' varied according to the enzyme used. Rsa 1 (blunt end cloning) and Hinf I (sticky end cloning) were the enzymes most commonly used. To obtain annealed bubble for ligation, equimolar amounts of universal bubble bottom and restriction enzyme specific bubble top were heated to 65° for 5 min, and annealed by cooling over 2 hr to 37°C, then frozen in aliquots at -20°C. 1 μg of YAC DNA was digested with Rsa 1 or Hinf I and the completeness of digestion confirmed by gel electrophoresis. The restriction enzyme was inactivated by heating at 65°C for 10 min and 1 μl of annealed bubble was then added together with 1μl of $2x10^6$units/ml T4 DNA ligase (NBL, UK) and an appropriate amount of 10x ligase buffer containing rATP. Ligation was performed at 14°C overnight. The reaction solution was diluted 1:5 with ddH_2O and 4 μl aliquots were used for subsequent PCR analysis.

Probes for Southern hybridisation were generated by PCR amplification of the YAC end clone followed by digestion with Eco R1 to recover the 'human' insert. The method was as follows: 1 μg each of universal bubble primer (224) and YAC end primer (1091 or HYAC-C) were used in a standard PCR mix. A denaturing step at 94°C for 10 min was performed and 1 unit of Taq polymerase was then added. The conditions for PCR were : 1.5mM magnesium, 94°C for 1 min, 67°C for 2 min, 72°C for 2 min, 35 cycles. If the resulting YAC end clones were under 2 kb in length, the times of the annealing and extension steps were reduced in future reactions. The YAC end clone consisted of the end of the human insert and a short stretch of YAC vector sequence (figure 3.29a). The YAC vector was cleaved off the human insert by adding Eco R1 to the PCR reaction mix and incubation at 37°C for 2 hr, the Taq buffer acting as the buffer for the restriction enzyme with no modification. The Eco R1 cut PCR mix was separated by electrophoresis on a 2% agarose gel (figure 3.29b). Two bands were seen: the human sequence and the vector fragment (240 bp for HYAC-C, 160 bp for 1091). The band corresponding to the end of the human DNA carried in the YAC

vector was then excised and the DNA eluted using the Geneclean technique according to manufacturer's instructions.

Single stranded sequencing of YAC ends was performed as described in section 2.8. The primer for the bubble (224) was biotinylated and the DNA strands of the resulting PCR product separated by alkali denaturation. The biotinylated strand was 'captured' by adsorbtion onto streptavidin coated magnetic beads which were removed from the solution with a magnet. The non biotinylated strand remained in the NaOH solution from which it was then precipitated. Both strands could then be sequenced using primers close to the cloning site as described in section 2.8

2.8 SEQUENCING

Sequencing was performed by the di-deoxy chain termination method (Sanger et al, 1977) directly from PCR products after separation of the individual strands of DNA. The composition of all solutions used is given in table 2.8.

2.8.1 *Separation of DNA strands*

For separation of the DNA strands, one of the primers used in the PCR reaction was biotinylated allowing immobilisation of the DNA on streptavidin coated magnetic beads (Dynal). The PCR reaction was as described above except that the amount of primer used was decreased to 25 to 50 pM to avoid saturation of the streptavidin coated beads with excess biotinylated primer. 30 µl of Dynabeads were washed twice in 60 µl of TES, all separations being carried out on a magnetic particle concentrator (MPC, Dynal) and resuspended in 30µl of TES. Products of the original PCR reaction were separated from the mineral oil overlying them with a pipette and then incubated with the 30 µl of resuspended Dynabeads for 5 min at room temperature to allow annealing of the double stranded, biotinylated PCR product to the streptavidin-coated beads. The tube was then placed in the MPC, one side of which is magnetic, and the beads, with the double stranded DNA attached, migrated to the side of the tube. The supernatant was gently removed with a pipette and discarded. The double stranded DNA was denatured by resuspending the beads in 100 µl of freshly made 0.15M NaOH at room temperature for 5 min. Both single strands of DNA could then be isolated, the biotinylated strand bound to the beads and the non-biotinylated strand remaining in the NaOH. DNA from the non-biotinylated strand was precipitated by the addition of 0.1 x vol. 3 M Sodium acetate and 2.5 x vol. 100% ethanol followed by incubation at -70°C overnight. Precipitated DNA was removed by centrifugation in a microfuge at 13,000 rpm for 20 min, washed in 70% ethanol, air-dried and resuspended in 7 µl ddH_2O. The

biotinylated DNA, immobilised on the beads, was washed in 100 µl of TES, once in 100 µl ddH_2O, resuspended in 14 µl ddH_2O and stored at -20°C.

2.8.2 *Di-deoxy sequencing*

Both DNA strands were sequenced using a version 2 Sequenase kit (USB). The enzyme, sequenase, is a genetic variant of the bacteriophage T7 DNA polymerase generated by *in vitro* genetic manipulation (Tabor and Richardson, 1987). 0.5 pM of the appropriate primer was annealed to 7 µl of template DNA, still attached to the biotinylated beads, with 2µl of 5 x sequenase reaction buffer by heating to 65°C for 2 min and cooling to room temperature over a 30 min period. DNA was then labelled with α^{35}SdATP (1000 Ci/mmol) and then the extension of DNA synthesis terminated by incubation with each of the 4 di-deoxynucleotides (ddNTPs) which lack the 3' OH group necessary for chain elongation. The resulting products would each be terminated at different lengths and, as all would have incorporated radio labelled dATP, their position could be identified by autoradiography. The method was as follows:

The following were added to the annealed template primer :1 µl of 0.1M DTT, 2 µl of 1:15 diluted labelling mix, 0.5 µl (5 µCi) of α^{35}SdATP and 2 µl of sequenase diluted 1:8 in enzyme dilution buffer. The reaction mix was incubated for 3 min at room temperature and then 3.5 µl aliquots were added to 0.5 ml eppendorf tubes containing 2.5 µl of each of the four ddNTPs (ddGTP, ddTTP, ddCTP, ddATP), incubation being for 5 min at 37°C. The reaction was terminated by the addition of 4 µl of stop solution. The resulting products were then denatured and separated by polyacrylamide gel electrophoresis.

Table 2.8

Solutions used in the separation of DNA strands and di-deoxy-sequencing

TES (STE):	10 mM Tris-HCl 1 mM EDTA 0.1 M NaCl
5x sequenase buffer (USB):	200 mM Tris-HCl, pH 7.5 100 mM Magnesium chloride 250 mM NaCl
5x labelling mix (USB):	7.5 µM each of dGTP, dCTP, dTTP
Enzyme dilution buffer (USB):	10 mM Tris-HCl, pH 8 5 mM DTT 0.5 mg/ml Bovine serum albumin
Stop buffer (USB):	95% formamide 20 mM EDTA 0.05% Bromophenol Blue 0.05% Xylene cyanol

2.8.3 *Polyacrylamide gel electrophoresis*

Polyacrylamide gels were cast between two glass plates, one of which was shorter (33.3 x 39.4 cm) than the other (33.3 x 42 cm). Before assembly, both plates were washed with detergent (Fairy Liquid) and warm water and rinsed in ddH_2O. One side of the shorter plate was siliconised by wiping the surface with Sigmacote (Sigma) then drying the plate in a fume cupboard. When dry, the siliconised surface was again rinsed in ddH_2O. Plastic spacers, 42 cm long, 1 cm wide and 0.25 mm deep, were placed on either side of the larger plate and two additional short spacers (2 cm) were placed on top of the long spacers at the bottom of the plate, increasing the final distance between the plates at the bottom to 0.5 mm. The shorter glass plate was then placed, face down, on top of the larger plate with the siliconised surface adjacent to the larger plate. The edges of the glass plate were sealed with 33 mm wide PVC electrical tape. Polyacrylamide was polymerised from acrylamide during gel casting in a radical reaction, starting with Ammonium persulfate (Sigma) and TEMED (N,N,N',N'-tetra methyl ethylene diamine, BRL). 6% gels were made by adding 27 ml of 40% (w/v) acrylamide stock (acrylamide: Bis-acrylamide, 19:1, NBL) to 75.8 g ultra pure urea (Sigma), 18 ml of 10 x TBE and ddH_2O were added to make a final volume of 180 ml. The solution was warmed to 30°C and 180 µl of TEMED and 216 µl of freshly made 25% ammonium persulfate added to start polymerisation. The gel mixture was introduced between the glass plates with a syringe and, when the gap between the plates had been filled, a plastic comb inserted in between the plates to a depth of 5 mm whilst the gel was still liquid. The plates were clamped at the top and sides by bulldog clamps and left horizontally in a tray for polymerisation to occur. Once the gel was set, it could be stored overnight by laying soft paper towels soaked in 1xTBE over the comb and covering the towels with Saran wrap to prevent desiccation of the gel. In order to use gel, the clamps and tape were removed, the glass plates were washed and the comb was also removed. The gel was placed in a vertical gel electrophoresis tank (BRL/Life Technologies Model 2) with the shorter glass plate innermost and clamped in position. One litre of 1 x TBE running buffer was added, 500 ml to the upper and 500 ml to the lower reservoir chambers. Unpolymerised acrylamide and urea crystals were removed by injecting 1 x TBE in between the wells of the gel at least twice. This prevented fragmentation of the bands obtained on the resulting sequencing gel. The gel was then connected to a suitable power supply (LKB Bromma 300Xi, Biorad) and preheated by running at 60 W constant power for 30 min, prior to loading the samples. The wells were once again cleaned by injection of 1 xTBE and the samples loaded with duckbill pipette tips (Anachem) having been denatured for 3 min at 90°C.

There is no need to remove the Dynabeads prior to loading the biotinylated strand. Electrophoresis was at 60 W constant power for 2-4 hr, following which the buffer was drained off and the gel assembly withdrawn from the tank. The shorter siliconised plate was prised off and the gel, adherent to the lower, larger plate, was placed in a large plastic tray and fixed in 2 l of 10% methanol:10% glacial acetic acid for 40 min in a fume hood. The gel was transferred to 3MM Whatman paper, covered in Saran Wrap and dried for 90 min in a gel drier (model 583 Biorad), attached to a vacuum pump (Genevac CVP 50). The gels were overlaid with autoradiography film (Kodak, XAR5) inside light proof cassettes for 16 hr at room temperature without an intensifying screen and the sequence ladder read on a light box.

2.9 λ PHAGE ANALYSIS

λ phage are bacterial viruses that have been altered in vitro to allow cloning of larger DNA fragments than is possible in plasmid vectors. Commercially available chromosome specific libraries (ATCC, USA) are available in this vector. All amplification and screening of λ phage has to be done in bacterial cultures capable of being infected by the viruses and E coli, strain LE392 were used. The composition of all solutions used in phage analysis is shown in table 2.9.

Table 2.9

Solutions used in λ phage work

20% Maltose:	Maltose 20 g ddH_2O to 100 ml - filter sterilise (autoclaving will caramelise the sugar) - add 1 ml to each 100 ml medium
SM:	Per litre: 5.8 g NaCl 2 g $MgSO_4$ 7 H_2O 1 M Tris-HCl, pH 7.7 2% gelatine - autoclave

2.9.1 *Preparation of LE392*

LE392 cells were streaked out from a stab of frozen stock onto LB agar with no antibiotic selection and incubated at 37°C with shaking overnight. Single colonies of LE392 were added to 5 ml of LB in a 15 ml universal container and placed in a shaking incubator at 37°C for 8 hr. 3 ml of this culture (now thick with LE392) was added to 50ml LB + 0.2% maltose and incubated with shaking at 37°C overnight. Cells were pelleted by centrifugation in a Jouan CR412 centrifuge at 4000 rpm for 10 min at room temperature and resuspended in 0.4x the original volume (20 ml) of 0.01 M magnesium sulphate. Further 0.01 M magnesium sulphate was added until the optical density at A^{600} was 2 OD units, corresponding to $1.6x10^9$ cells/ml. Cells were stored in aliquots at 4°C.

2.9.2 *Calculating the titre of phage libraries*

Plates were made by pouring cooled, molten LB agar into 9 cm Petri dishes under sterile conditions in a fume hood . This 'bottom' agar formed a support for the bacterial lawn and phage. 7g of electrophoretic grade agarose was added to 1 litre of LB, autoclaved and cooled to 45°C; this would form the 'top' agarose eventually containing bacteria and phage.

10 µl of phage stock was added to 990 µl SM (10^{-2} dilution) and further serial 10x dilutions were made giving dilutions of 10^{-3} to 10^{-8} of the original stock. 100µl of the dilutions above were added to 200 µl of LE392 cells and incubated for 15 min at 37 °C. 3.5 ml of top agarose was added to the phage and bacteria, mixed and rapidly spread onto pre-warmed bottom agar plates. Once set, the plates were incubated up-side down at 37°C overnight and the number of clear plaques on the bacterial lawn counted. The titre of phage per µl of stock could then be calculated.

2.9.3 *Screening phage libraries*

The phage library was plated out at a density that gave clear plaques in the bacterial lawn 2 mm in diameter after overnight incubation at 37°C. Circular discs of hybridisation membrane (Hybond N) were laid on top of the Petri dish and pinholes were made in them and through the agar to ensure that the orientation of the filter relative to the plate was known. The filter was removed after 2 min using forceps and floated on denaturing solution for 3 min, neutralising solution for 3 min, 2x SSC for 3 min and the DNA transferred to the membrane by lying the filter, phage side up, on 3MM Whatman paper as described in section 2.6.3. Filters were baked for 2 hr at 80°C and then hybridised with sheared total human DNA as described in section 2.6.3. The hybridisation fluid used was Church buffer and washing was to 0.05 M NaPi, 0.1 % SDS at 65°C for 30 min and autoradiography was performed overnight. The developed autoradiograph was placed on a light box and the plate from which the filter had been

made placed on top, the pin holes ensuring accurate orientation. Black dots on the autoradiograph indicated colonies containing repetitive human sequences. Colonies with no corresponding signal either contained no human insert or contained unique sequences and these were picked for further analysis. A sterile pipette tip was snipped off so that the diameter of the end was slightly greater than that of the plaque. The Gilson pipette was set to 30 µl and the plunger depressed. The pipette tip was located above the plaque and stabbed through both top and bottom agar leaving the plaque inside the tip. The plunger was released and the vacuum broke the plaque off its attachment to the plate, ensuring it stayed in the pipette tip. The plaque was extruded into 250 µl SM containing 50 µl of chloroform to inhibit bacterial growth, eluted for 5 hr at room temperature and then stored at 4°C.

2.9.4 *Amplifying phage DNA*

50 µl of plug eluent was incubated with 100 µl of LE392 at 37°C for 20 min then added to 3.5 ml top agarose and spread onto a pre-warmed bottom agar plate which was incubated overnight at 37°C. Confluent lysis of the bacterial lawn should occur and the phage were then eluted with 4 ml of SM with gentle shaking for 4 hr. 2 hr before the end of the elution, 0.5 ml of LE392 cells were added to 45 ml LB with 0.2% maltose and incubated at 37°C with shaking until the OD A^{600} was 1.0 (usually 2 hr). The 4 ml of eluent was then added to the 50 ml of LE392 cells and incubated at 37°C with shaking until flocculation indicative of bacterial lysis occurred. This usually took 2 hr. 0.5 ml of chloroform was added to lyse the remaining intact bacteria and then the chloroform was removed. 2.9 g NaCl and PEG to 10% were added, the tube inverted gently to dissolve the salt, and stored for 16 hr at 4°C to precipitate the phage. Centrifugation pelleted the phage, which were resuspended in 0.5 ml SM and vortexed with an equal volume of chloroform. Once the chloroform had been removed, SDS was added to 1% to ensure lysis of the phage. RNase was added to a working concentration of 100 µg/ml for 30 min at 37°C and then inactivated by the addition of proteinase K to a concentration of 200 µg/ml and incubated at 65°C for 1 hr. The phage DNA was phenol/chloroform extracted, ethanol precipitated and resuspended in TE.

2.10 COSMID ANALYSIS

Cosmid vectors are hybrids derived from plasmids and λ phage, that are able to accept approximately 45 kb of foreign DNA, three times more than that of phage vectors. A chromosome 13 specific cosmid library has been developed from flow sorted chromosomes by the ICRF Genome Reference Laboratory (Lehrach and et al, 1990)and was screened by filter hybridisation in a way identical to that used for YACs (section 2.7.1.a). Cosmids were grown in LB with 30 μg/ml kanamycin as the antibiotic selection and extraction of DNA was identical to that described for plasmids (section 2.2.4). Cosmids were used as complete probes for in-situ hybridisation or were used to generate unique probes. To derive unique sequences cosmids were digested with Eco R1, and the fragments separated by gel electrophoresis on a 1% agarose gel, photographed and Southern blotted. Probing the resulting filter with sheared total human DNA identified bands containing repetitive sequences. Bands present on the photograph of the gel but absent on the autoradiograph were likely to function of single copy DNA sequences when used as probes. A repeat electrophoresis of the Eco R1 cut cosmid DNA was performed, the 'unique' bands excised and the DNA eluted using the Geneclean technique.

RESULTS

RESULTS

3.0 OVERVIEW

The aim of this thesis was to isolate and characterise the breakpoint junction fragment of a constitutional t(1;13)(q22;q12) translocation identified in a patient (DG) with stage II ganglioneuroblastoma. The strategy was as follows:

- Isolate one of the derivative chromosomes from all other chromosome 1 and 13 material to allow the breakpoint to be mapped by constructing somatic cell hybrids between the lymphoblastoid cell line from DG and mouse 3T3 cells.
- Develop molecular markers which closely flank the DG breakpoint.
- To use these markers to isolate YACs.
- To construct a YAC contig around the breakpoint in an attempt to isolate the breakpoint junction fragment in a single YAC.

The results presented here extend those found in the three publications bound at the end of the thesis. Appendix 1 gives the sequences, conditions and the sizes of the PCR products expected of all the STSs used in this thesis. All the data on probes used in Southern hybridisation is given in Appendix 2.

3.1 ISOLATION OF THE DERIVATIVE CHROMOSOME 1 IN SOMATIC CELL HYBRIDS.

3.1.1 General considerations

Fusing rodent and human cells results in the production of somatic cell hybrids which randomly lose human chromosome material. Each hybrid clone may, therefore, carry a different complement of human chromosomes. After fusion, it is important to be able to select against the unfused 'parental' cells, both of human and rodent origin. Human cells are at least 100 times more sensitive to the sodium-potassium pump poison, ouabain, than rodent cells. As ouabain 'resistance' is a dominant trait, mouse/human hybrids will have the same resistance as the parental rodent cells. Growing the hybrids in ouabain containing medium will, therefore, select against unfused human cells without affecting the hybrids. Some regions of the human genome carry genes , missing from the mouse cells, which code for enzymes allowing the hybrid cells to grow on media that would not sustain the growth of normal rodent cells, thereby selecting against unfused rodent cells. Growing the hybrids in HAT

(hypoxanthine, aminopterin and thymidine) medium is the classic example of such a selection system. DNA precursors are normally synthesised from glutamine and dUMP, but if this pathway is blocked by aminopterin then human cells can still survive if they are supplied with hypoxanthine and thymidine by generating DNA precursors via a salvage pathway. The presence of the enzyme hypoxanthine phospho-ribosyltransferase (HPRT), on the human X chromosome, allows the conversion of hypoxanthine to DNA precursors, thymidine kinase (TK), on chromosome 17, allows the use of thymidine. The rodent cells used to make the hybrids were TK deficient and, therefore, growing the hybrids in HAT medium selected against unfused rodent cells.

HAT selection will not, however, influence whether chromosomes 1 or 13 are retained so once the rodent cells have died off, hybrids were then grown on HAT free medium. Hybrid clones were not kept growing for long periods of time as the chances of random loss of the chromosome of interest would have been increased. Sufficient cells were grown for cytogenetic analysis and DNA preparation and then the clones were kept as frozen stocks. Whenever a frozen stock was re-grown the presence of the chromosome of interest was confirmed by Southern blotting or PCR before any further studies were performed.

3.1.2 Construction and initial analysis of hybrids

A lymphoblastoid cell line was made from patient DG by infecting isolated white cells with Epstein-Barr virus (EBV), which immortalises B cells. The resulting cell line was maintained in RPMI-1640 medium supplemented with 10% foetal calf serum and 2 mM glutamine. This line was then fused with mouse 3T3 cells using polyethylene glycol (PEG molecular weight 1000) as described in Cowell (Cowell, 1992). The hybrid cells were cultured in HAT medium for four to six weeks (ouabain was added to the medium for the first two weeks) and then individual colonies were isolated by ring cloning. The resulting clones were assigned numbers, all pre-fixed by the letters DGF (DG fusion).

36 hybrid clones were grown in bulk and DNA prepared from them. Molecular analysis for the presence of one or other of the two derivative chromosomes was carried out by Southern hybridisation using three different probes: a probe for the MYCL locus in 1p22 (pL*myc* 10), a probe for the renin (REN) locus in 1q41 (pren) and Rb 3.8 (from the RB1 locus at 13q14). The normal G-banded copies of chromosomes 1 and 13 and the two derivative chromosomes taken from karyotype analysis of the DG lymphoblastoid cell line are shown in figure 3.1.a. The derivative chromosome 1 [der (1)] comprised 1pter-1q22::13q12-13qter. The derivative chromosome 13 [der(13)] is much smaller, containing 13pter-13q12::1q22-1qter. If a given hybrid contained the der(1) and no other chromosome 1 or 13 material, probes for MYCL and RB1 would be present but REN would be absent, whereas if the der(13) was exclusively present the

opposite would be found. In fact, all 36 colonies showed the presence of all three probes indicating that none carried either derivative chromosome as the sole material from chromosomes 1 and 13.

Cytogenetic analysis of the hybrid colonies (Cowell, pers comm) showed that DGF27 carried the der(1) chromosome though the normal copy of chromosome 1 was also present in a sub-set of cells. Single cells from DGF27 were sub-cloned and the resulting colonies were grown in bulk for DNA and chromosome analysis. One of these single cell clones, DGF27C11 (clone 11), was shown to hybridise with probes for MYCL but not REN indicating that the der(1) but not the normal copy of chromosome 1 was present. A probe for the polymorphic urinary mucin (MUC) locus in 1q24 (pMUC10), was also absent from DGF27C11, confirming the results obtained with REN (figure 3.2). A karyotype analysis of DGF27C11 confirmed the presence of the der(1) and the absence of the normal copy of chromosome 1 in this hybrid (figure 3.1.b). In situations were the amount of target DNA is low, Southern hybridisation can yield false negative results and excluding the presence of the normal copy of chromosome 1 on the basis of the absence of bands for REN and MUC was risky. PCR is at least a thousand times more sensitive than Southern blotting and so DNA from DGF27C11 was used as a template in PCR reactions using oligonucleotide primers for nerve growth factor B (NGFB, 1p12) and spectrin (SPECT, 1q22). The presence of NGFB and the absence of SPECT (figure 3.3) provided conclusive evidence for the absence of a normal copy of chromosome 1 in DGF27C11.

The probe for RB1 was present in DGF27C11 as would be expected if this hybrid contains the der(1) chromosome. However, proving that the normal copy of chromosome 13 was absent was much more difficult than for chromosome 1. Repeated cytogenetic analysis did not reveal the presence of the normal chromosome 13 in DGF27C11. In order to try and confirm that the normal copy of chromosome 13 was absent from DGF27C11, the hybrid was probed with the two most centromeric chromosome 13 markers available at this time, D13S6 (pHU10), and D13S1 (p7F12). Both markers were present in DGF27C11 (figure 3.2), indicating that either the DG breakpoint was more centromeric than both markers, or that the normal copy of chromosome 13 was present. Until such time as probes from 13p or proximal 13q12 became available we had to trust the cytogenetic analysis and assume that the normal chromosome 13 was absent from DGF27C11.

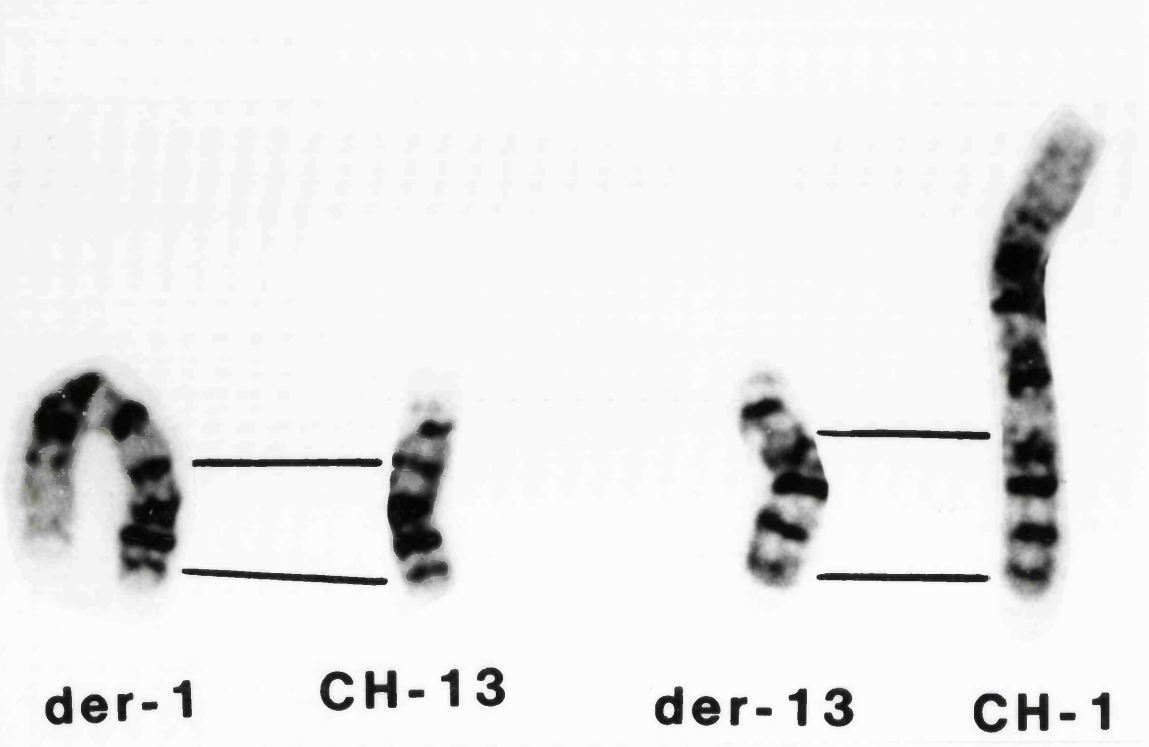


Figure 3.1.a. Trypsin-Giemsa banded metaphase chromosomes from the patient, DG, showing both derivative chromosomes and the normal copies of chromosomes 1 and 13.

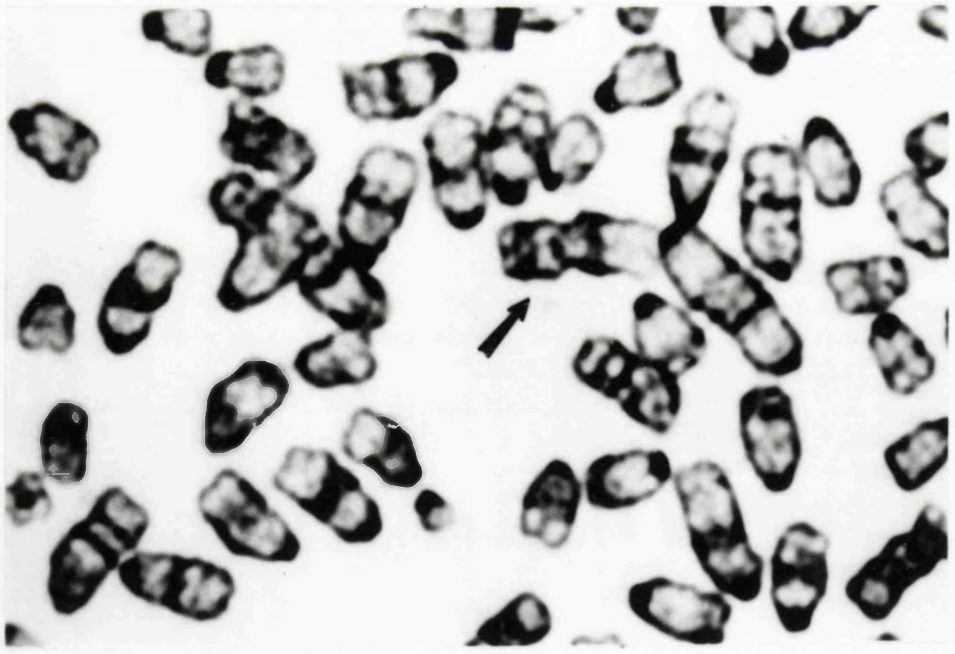


Figure 3.1.b Partial G-banded metaphase spread from the hybrid DGF27C11, showing the derivative (1) chromosome, indicated by the arrow.

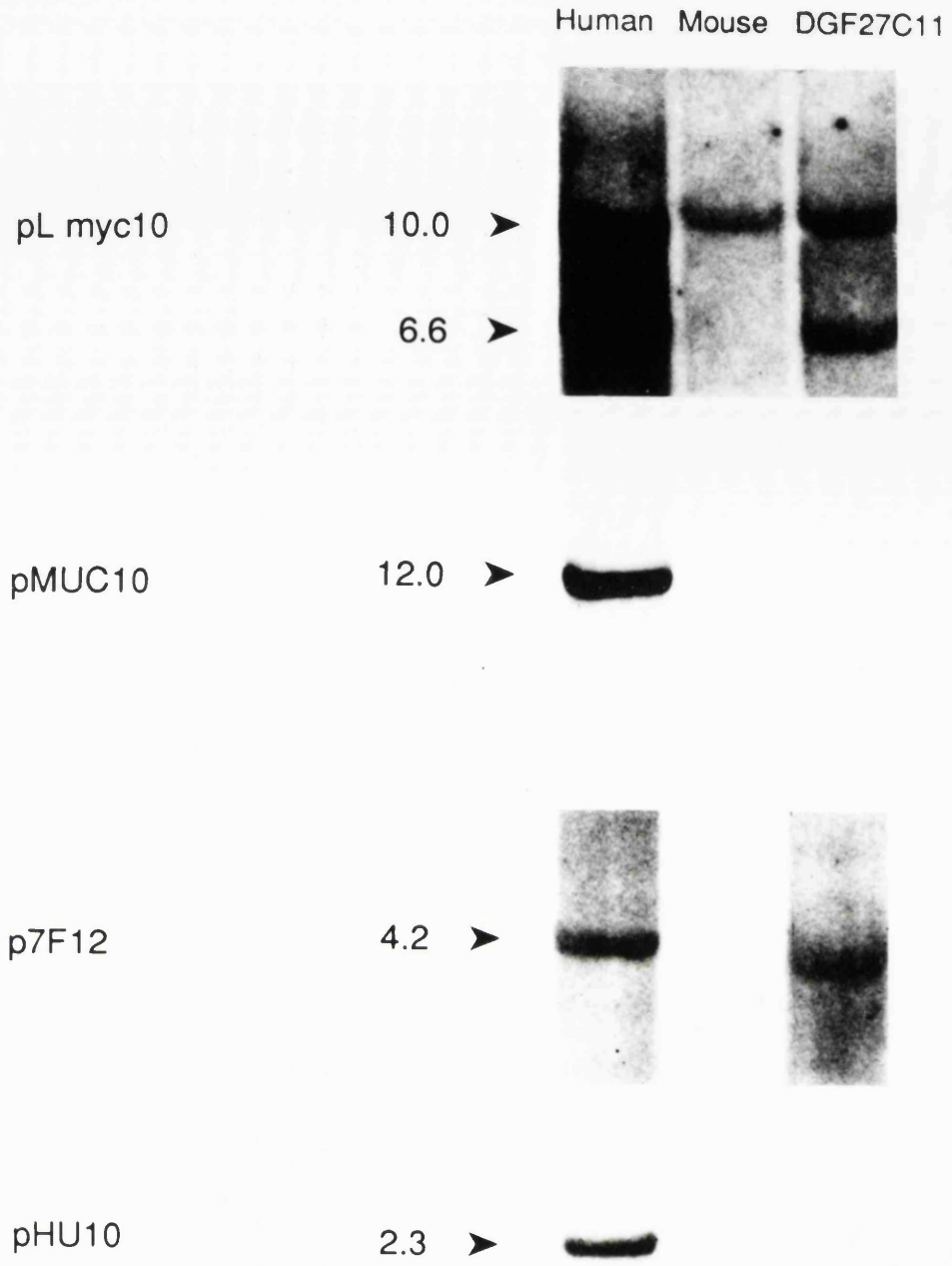


Figure 3.2. Southern blot analysis of hybrid DGF27C11. 4 blots are shown, in the upper 2 the DNA has been digested with Eco R1, in the lower 2 with Hind III. The sizes of the hybridising bands are shown on the left for each probe. The probe pL-myc10 from 1p32 is present in DGF27C11 but pMUC10, from 1q22, is not. The human specific band for pL-myc10 is 6.6 kb long but cross hybridisation with a 10 kb product is seen in the mouse. On chromosome 13, both p7F12 and pHU10 (the most centromeric probes available in 1990) are present in DGF27C11, confirming the presence of 13q12 in this hybrid. However, it is not possible to exclude the presence of the normal copy of chromosome 13 from this data.

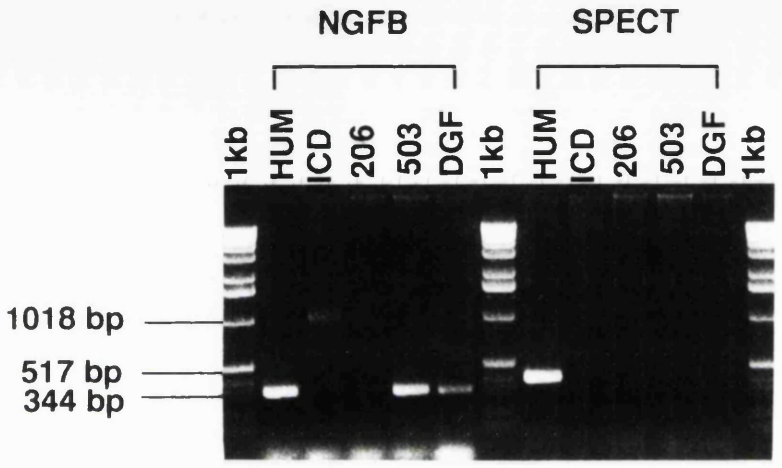


Figure 3.3 Characterising hybrid DGF27C11 using PCR. Oligonucleotide primers for nerve growth factor B (NGFB), located in 1p13, and spectrin (SPECT), located in 1q22 were used. Both of these genes were present in human genomic DNA (HUM) NGFB as a 325 bp band and SPECT as a 405 bp band. However, only NGFB was present in DGF27C11, confirming the exclusive presence of the der (1) in this hybrid. As expected, neither gene is present in ICD which contains only 13q21-pter, nor in 206 which contains only 1q but with a deletion at 1q21. Hybrid 503 (a hybrid with a derivative chromosome from a t(1;X) containing 1p but not 1q) was positive with primers for NGFB alone.

3.2 ANALYSIS OF THE DG BREAKPOINT WITH KNOWN MARKERS

3.2.1 *Somatic cell hybrid mapping panels*

The aim of mapping the breakpoint in DGF27C11 was to obtain markers that flanked the breakpoint closely enough to allow cloning of the breakpoint junction fragment using a bi-directional YAC walk. Although the exclusive presence of the der(1) chromosome in DGF27C11 enabled markers to be positioned above or below the breakpoint, this single hybrid did not allow the chromosomal location of individual probes to be determined. This was important as, for example, a probe that mapped above the breakpoint but was several chromosomal bands away would obviously not be useful for YAC walking.

In 1990 high definition mapping of cosmids by FISH was not available in the host laboratory and, therefore, probes were mapped on panels of somatic cell hybrids. Mapping panels were available for both chromosomes 1 and 13, each containing hybrids with various deletions of the respective chromosome. The panel for chromosome 1 is shown in figure 3.4 with the available molecular markers placed in the intervals between the deletions. The breakpoint in hybrid N-OH1 (Lamb et al, 1989) was in 1p36, far too great a distance away from the DG breakpoint to be useful. The closest breakpoint distal to the DGF translocation was that of Lix7 (1q24-1qter), a hybrid kindly provided by Dr. H. Middleton-Price. The closest proximal breakpoint was in GOS139A (1p22-1pter), a hybrid described by Cowell and Mitchell (1989). The 1p22-1q22 region could not be further sub-divided and represents a large region within which it would be impossible to accurately localise markers proximal to the DG breakpoint.

However, during the search for the RB1 gene the host laboratory had generated a somatic cell hybrid mapping panel for chromosome 13 which allowed 13q14-13q12 to be divided into 5 separate regions (fig 3.5). Hybrid ICD was described by Scheffer (1986), RHF14 by Mitchell (1991) and the other four hybrids by Cowell and Mitchell (1989). This panel enabled the relative position of probes from the region of interest to be determined. The most centromeric breakpoint was the proximal one in KBF11, a hybrid described by Cowell and Mitchell in which 13q12-13q14 has been deleted. Probes pHU10 and p7F12 were absent in this hybrid and present in DGF27C11, placing the DGF breakpoint within the deletion in KBF11 (but above the loci for pHU10 and p7F12) or even centromeric to the proximal breakpoint in KBF11. The chromosome 13 mapping panel allowed localisation of probes mapping below the DG breakpoint to within a sub-region of 13q12 and those above DG to 13q12-13pter.

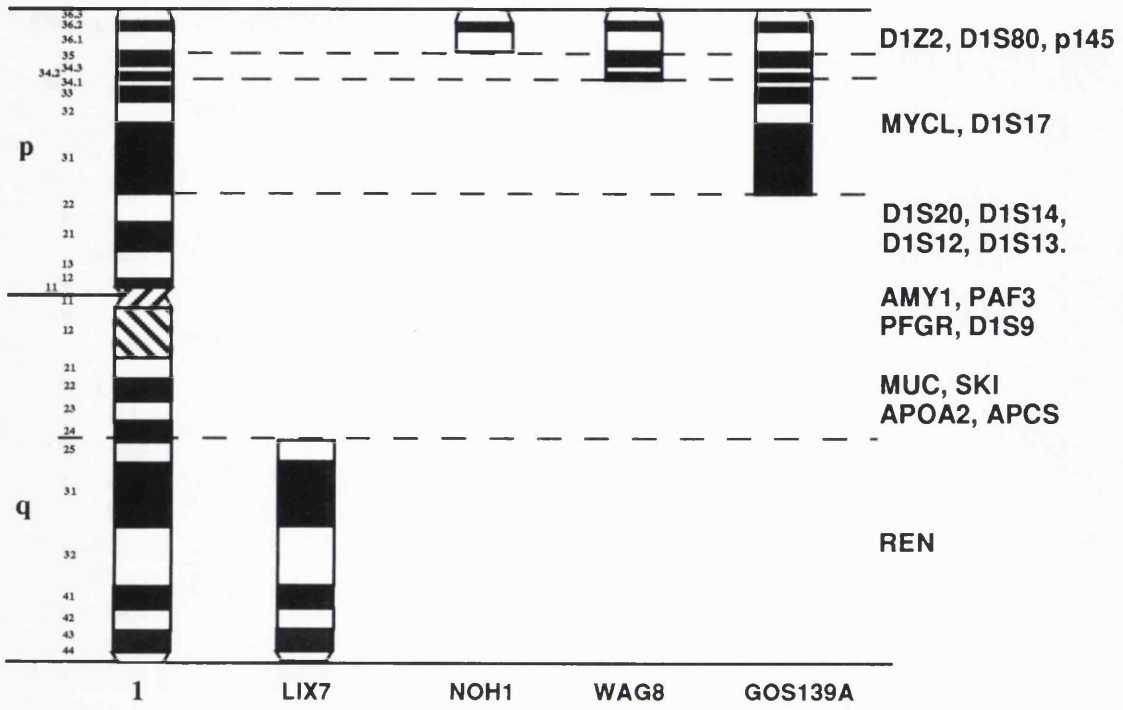


Figure 3.4. An ideogram of the chromosome 1-specific somatic cell hybrid mapping panel available at the onset of this thesis. The names of the various hybrids are given below. The regions of chromosome 1 contained within each hybrid is indicated diagramatically although other human chromosomes were also present. LIX7 was provided by Dr.H.Middleton Price, N-OH1 was described in Lamb et al, (1989), WAG8 by Dr. Y. Boyd and GOS139A was described by Cowell and Mitchell (1989). The map position of the probes shown on the right hand side of the diagram were determined by Southern hybridisation.

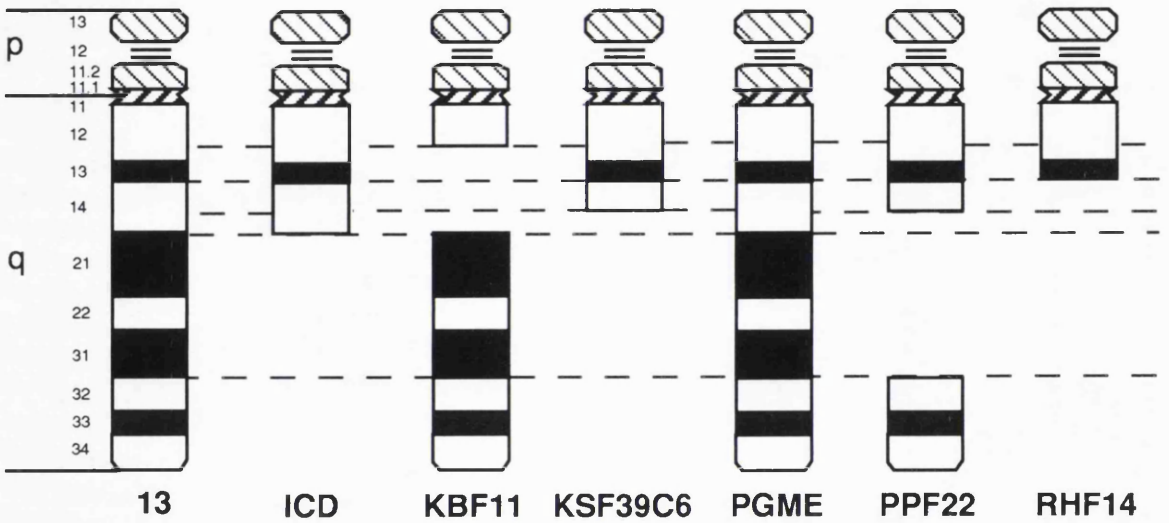


Figure 3.5. An ideogram of a panel of chromosome 13-specific somatic cell hybrids available for mapping the position of probes on chromosome 13. Hybrid ICD was characterised by Scheffer et al (1986), RHF14 by Mitchell et al (1991) and the other four hybrids by Cowell and Mitchell(1989).

Hence, the accuracy of localisation of probes around the chromosome 13 breakpoint was much greater than that obtainable for the chromosome 1 breakpoint. There was no good evidence to suggest that the putative Nb tumour suppressor gene was more likely to map to chromosome 1q22 than to 13q12. Indeed, the evidence implicating genes on chromosome 1 in Nb development is restricted to the abnormalities of 1p36 that have been observed in advanced stage tumours (section 1.5.5.a). Also, the breakpoint could be approached from either chromosome as, once the junction fragment was cloned, sequences from both chromosomes would be available. It was decided, therefore, to approach the breakpoint from the chromosome 13 direction.

3.2.2 *Mapping with known markers*

The location, relative to the DG breakpoint, of the known molecular markers available at the time is shown in figure 3.6. It can be seen that the DG breakpoint allowed the position of existing 1q probes relative to each other to be accurately established showing SKI, to be more centromeric than probes for APOA2, APCS and MUC (Michalski et al, 1992). Of importance is the fact that probe p145, for which LOH was shown in 90% of stage III and IV Nb tumours (Weith et al, 1989), is present in the hybrid excluding the possibility that a cytogenetically unrecognised deletion in 1p36, rather than the translocation, was the cause of the tumour formation in DG.

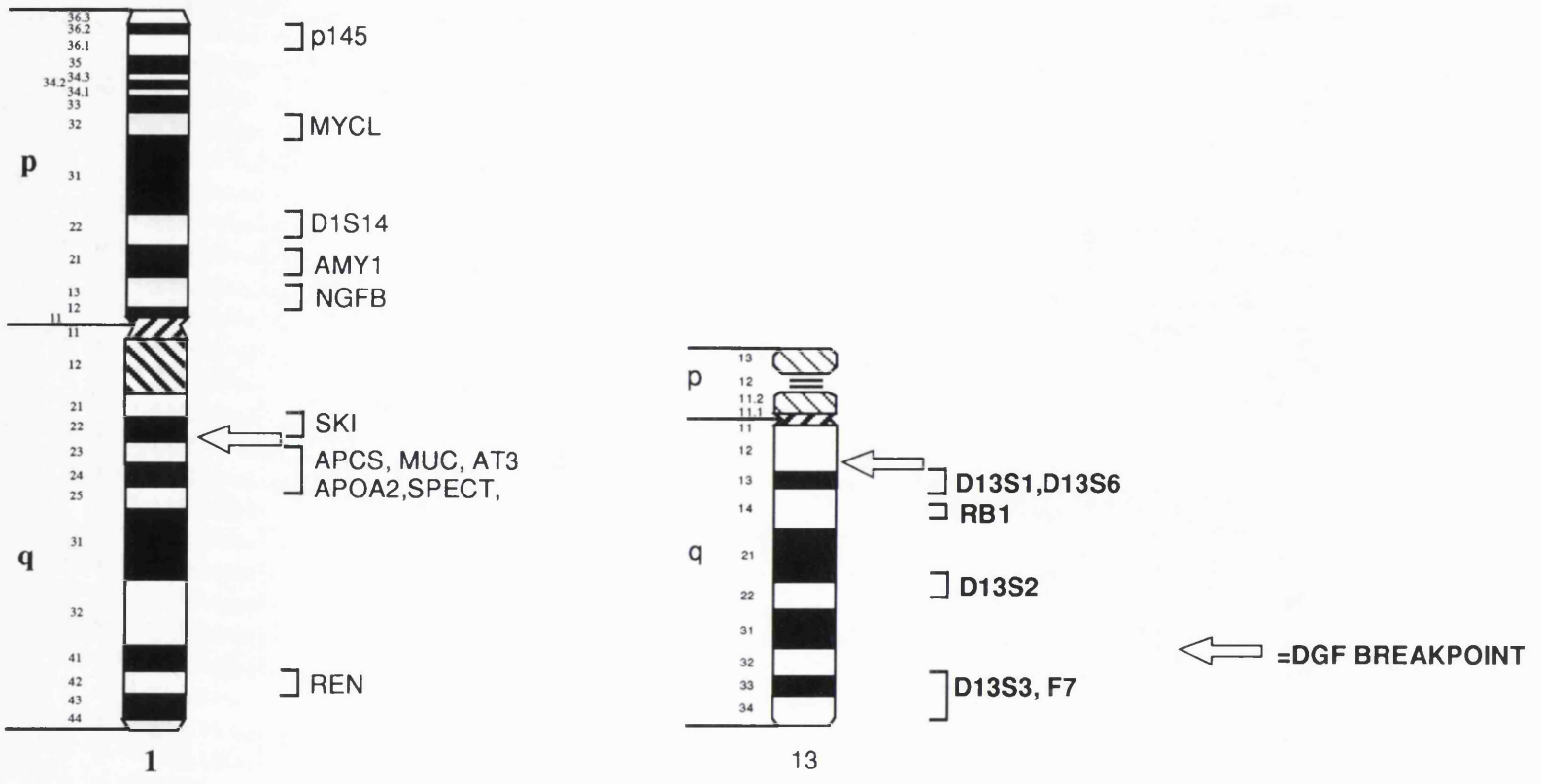


Figure 3.6. Ideograms of the normal copies of chromosomes 1 and 13 showing the map location of the probes used to characterise hybrid DGF27C11. The sites of the breakpoints involved in the reciprocal translocation in patient DG are indicated by the open arrows.

3.3 DEVELOPING NEW MOLECULAR MARKERS

3.3.1 Screening a chromosome 13 specific λ phage library

The most centromeric chromosome 13 probes (pHU10, p7F12) both mapped below the breakpoint in DGF27C11 and below the proximal breakpoint in KBF11. Therefore, in order to acquire markers that closely flanked the DG breakpoint, new probes had to be developed. The initial approach was to isolate and map unique sequences from a chromosome 13 specific λ phage library available from Lawrence Livermore laboratories as described in section 2.9.

The library was plated out using 100 µl of a 1:1000 dilution of the original stock, and 50 µl of LE392 cells which gave plaques of 2 mm diameter after overnight incubation at 37 °C. $3.8x10^5$ colonies per ml of original library stock were obtained which correlated well with the figure of $5x10^5$ quoted in the supplier's literature. Around 200 phage clones were present on each 9 cm diameter Petri dish and colonies from eleven Petri dishes (2,200 clones) were analysed further. Circular disks of hybridisation membrane were laid over the Petri dishes and their orientation on the agar determined by piercing the filter and the underlying agar with a pin. When the filter disks were removed from the surface of the agar, virus in the plaques stuck to the hybridisation membranes. The filters were then denatured, neutralised and baked. In order to determine which colonies carried repetitive human sequences, the disks were probed with sheared, sonicated total human DNA and autoradiography performed. Phage which contain repetitive human DNA sequences would hybridise strongly to the probe whereas those clones which contain unique sequences, or no human DNA at all, would not do so.

Further analysis of 24 phage clones was performed. Of these, 18 plaques had produced no signal on autoradiography, 3 plaques had given pale signal and the last 3 had yielded strong signal on hybridisation with total human DNA. Phage colonies were grown as described in section 2.9 and DNA successfully extracted from 19 of the 24 original clones. The DNA was digested with Eco R1 and the digests separated by agarose gel electrophoresis. Only six of the 19 phage clones had human DNA inserts, a non-recombination rate of 68% as compared to 14% quoted in the supplier's literature. Of the six colonies with inserts, three were from clones which had not produced any signal on hybridisation with total human DNA. Two of the six were from phage clones that had given pale signal on the autoradiograph and the last clone had hybridised strongly to the human DNA probe. All six of these inserts were excised from the gel and the DNA extracted. When used as probes for Southern hybridisation all six acted as highly repetitive sequences and were, therefore, useless for mapping.

On the basis of this data it appeared that λ phage library screening was unlikely to be a time effective method for generating flanking markers for the DG breakpoint. Even if unique probes were isolated, they could come from anywhere along chromosome 13 and the vast majority would not be useful as markers for the breakpoint under study. At this time the technique of Alu-PCR cloning became available and offered the possibility of deriving human specific DNA probes from the region of interest.

3.3.2 *Alu PCR cloning*

3.3.2.a *Generating inter-Alu sequences by PCR*

Alu is the commonest family of short interspersed repetitive elements (SINES) in humans with $9x10^5$ copies per haploid genome (discussed in section 4.2.2). Oligonucleotide primers have been developed for the 3' and 5' ends of the Alu repeat and can be used in PCR reactions to amplify human DNA sequences between Alu repeats as shown diagrammatically in figure 3.7. The products of Alu-PCR amplification can then be cloned and characterised for use as probes. The ideal Alu primers would be entirely human specific and yet still have homology to as broad a range of the family of human Alu sequences as possible. For theoretical reasons discussed in section 4.2.3, Alu 8559 was chosen as the 3' Alu primer and Alu 5684 as the primer for the 5' end of the Alu sequence. If primers to only the 3' end of the Alu repeat are used, human DNA sequences between Alu repeats whose 3' ends point towards one another will be amplified. If only 5' Alu primers are used, a different set of human sequences, those between Alu repeats whose 5' ends point towards each other, will be amplified. If both primers are used together, then the PCR product will contain both of the sets of human sequences above as well as those between Alu repeats whose 3' and 5' ends point towards each other, although the resulting PCR product may well contain a high percentage of repetitive sequences (section 4.2.5.).

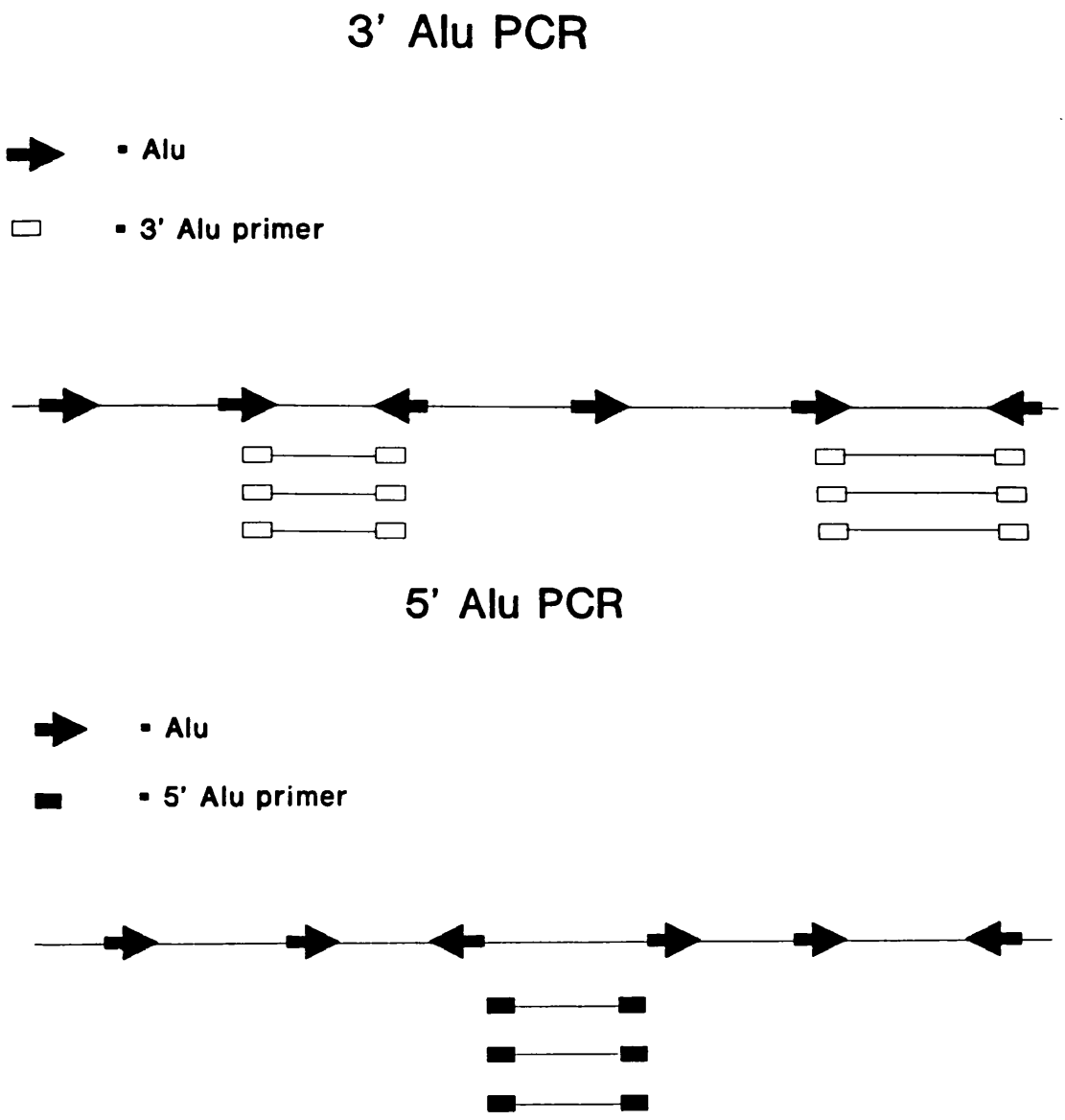


Figure 3.7. Schematic diagram describing the principle of PCR amplification of genomic DNA sequences using 3' and 5' Alu primers. The thin, horizontal line represents a strand of genomic DNA and the arrows represent Alu sequences in various orientations along its length. The open rectangles indicate the position of a primer for the 3' end of the Alu repeat (homologous to the arrow tips in the diagram) and the filled rectangles represent the 5' Alu primer (homologous to the arrow shafts). Thus, PCR using a given primer amplifies genomic DNA sequences which lie between Alu repeats which point towards each other. To develop a 3' Alu library only the 3' Alu primer is used and for a 5' Alu library only 5' Alu primers are utilised.

3.3.2.b *Selection of template DNA for Alu PCR*

The beauty of the technique is that by careful selection of the DNA used as a template in the PCR reaction only sequences from the region of interest will be amplified. ICD is a hamster/human somatic cell hybrid developed by Scheffer (1986), which contains 13pter-13q14 as its only human component. Molecular characterisation of ICD showed that the locus for RB1 is present, placing the breakpoint in distal 13q14 or 13q21.

Inter-Alu sequences were amplified by PCR from mouse 3T3, Chinese hamster, human, and ICD template DNA. The PCR reactions were performed with either the Alu 8559 (3' Alu), or the Alu 5684 (5'Alu) primers. A comparison of the PCR reactions is shown in figure 3.8. Using the Alu IV primer no sequences were amplified from mouse or Chinese hamster DNA. So many different sequences were amplified from genomic human DNA that, when the products of the reaction were separated by gel electrophoresis, a smear was seen. As ICD only contains a portion of a single human chromosome, PCR amplification using the human-specific Alu 8559 primer gave a multi-band pattern, with fragments between 1.1 kb and 300 bp visible on the gel. As no rodent sequences were amplified by the Alu 8559 primer, all the bands seen in ICD must be of human origin.

By contrast, the Alu 5684 primer was far less species specific, producing smears of amplified products from both mouse and Chinese hamster DNA. Incrementally raising the annealing temperature of the reaction from 58 °C to 65°C and then to 72°C decreased the number of different fragments amplified from rodent DNA but also reduced the number of bands in human and ICD (figure 3.9). At 72°C no rodent bands are seen but only one band was produced from ICD and subsequent analysis of this fragment (section 3.3.4) showed it to act as a high grade repeat.

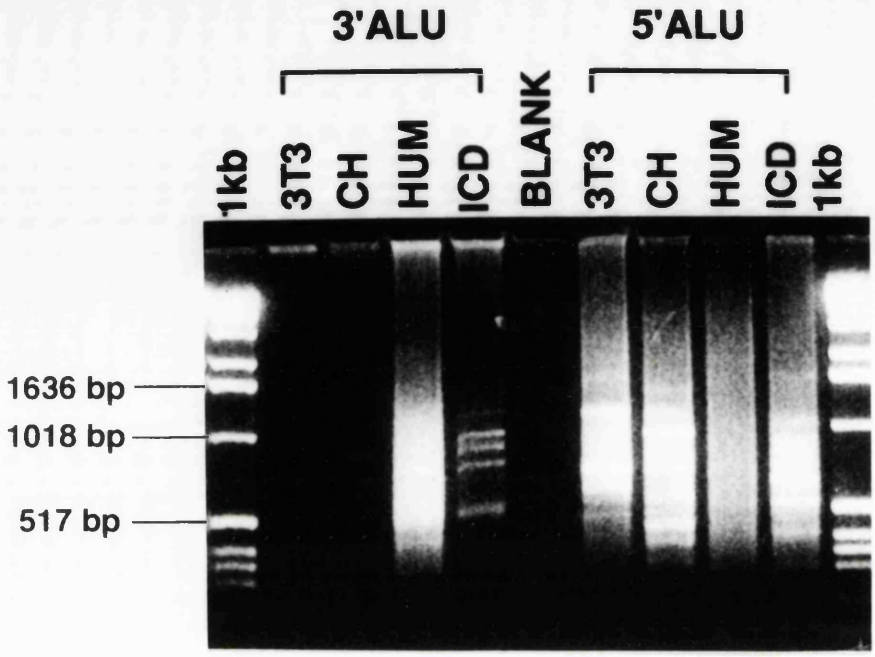


Figure 3.8. Comparison of the species specificity of different Alu primers when used at an annealing temperature of 58°C. When 3' Alu primers are used (left) no PCR product is derived from rodent DNA (3T3 is mouse and CH is Chinese hamster DNA). A smear is produced with human genomic DNA and several bands in ICD, which contains only the top half of chromosome 13. The bands in ICD must, therefore, be derived from human DNA. On the other hand, when 5' Alu primers are used the amplification appears much less species specific, producing smears not only in human DNA but also in rodent DNA (right).

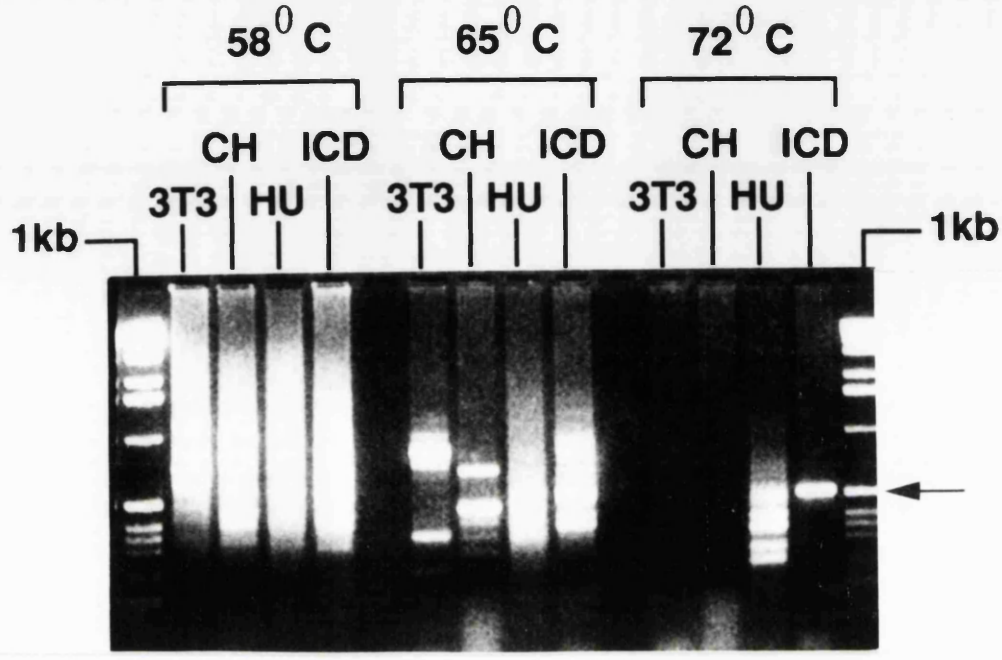


Figure 3.9. Amplification using 5' Alu primers can be made more species specific by increasing the annealing temperature of the PCR reaction. At 58°C a smear is seen in human (HUM) and rodent lanes (3T3 and CH) as well as in the mouse/human hybrid, ICD. Increasing the annealing temperature to 65°C still produces a smear in human DNA but many fewer rodent sequences are amplified and a discreet banding pattern is seen. A further temperature increase to 72°C eliminates all rodent signals but also reduces the number of bands from human DNA and only a single band is seen in ICD (arrow).

3.3.2.c *Cloning Alu-PCR products: general considerations*

Cloning the products of Alu-PCR amplification enables them to be separated from each other for individual analysis and stored for future use. A diagrammatic representation of Alu-PCR cloning is shown in figure 3.10. Two separate Alu-PCR libraries were made using ICD as the template; one from the PCR reaction products of amplification using the 3' Alu primer (Alu 8559) and one from the PCR reaction using 5' Alu primer (Alu 5684). The reaction products were purified on Sephadex columns and cloned into the 'blunt-ended' Eco RV site of Bluescript plasmid DNA as described in section 2.6.3. The ligated plasmid was then used to transform DH5α maximum efficiency competent cells.

Transformed cells were grown on LB agar containing ampicillin selection to ensure retention of the plasmid in the bacteria during progressive growth and to eliminate bacteria not containing a plasmid. Only single, distinct, colonies were picked for further analysis.

Some bacteria may have been transformed with plasmids that had re-circularised without containing any cloned DNA fragments. The addition of X-gal to the medium allowed these colonies to be identified as the absence of insert in the plasmid would have led to reconstitution of the LacZα gene and the resulting bacterial colonies would appear blue, whereas bacteria transformed with plasmids containing insert would form white colonies. As described in section 2.6.3, white colonies were then analysed further and the cloned Alu-PCR product was recovered either by extraction of plasmid DNA from large scale cultures, or by re-amplifying the insert using Alu-PCR.

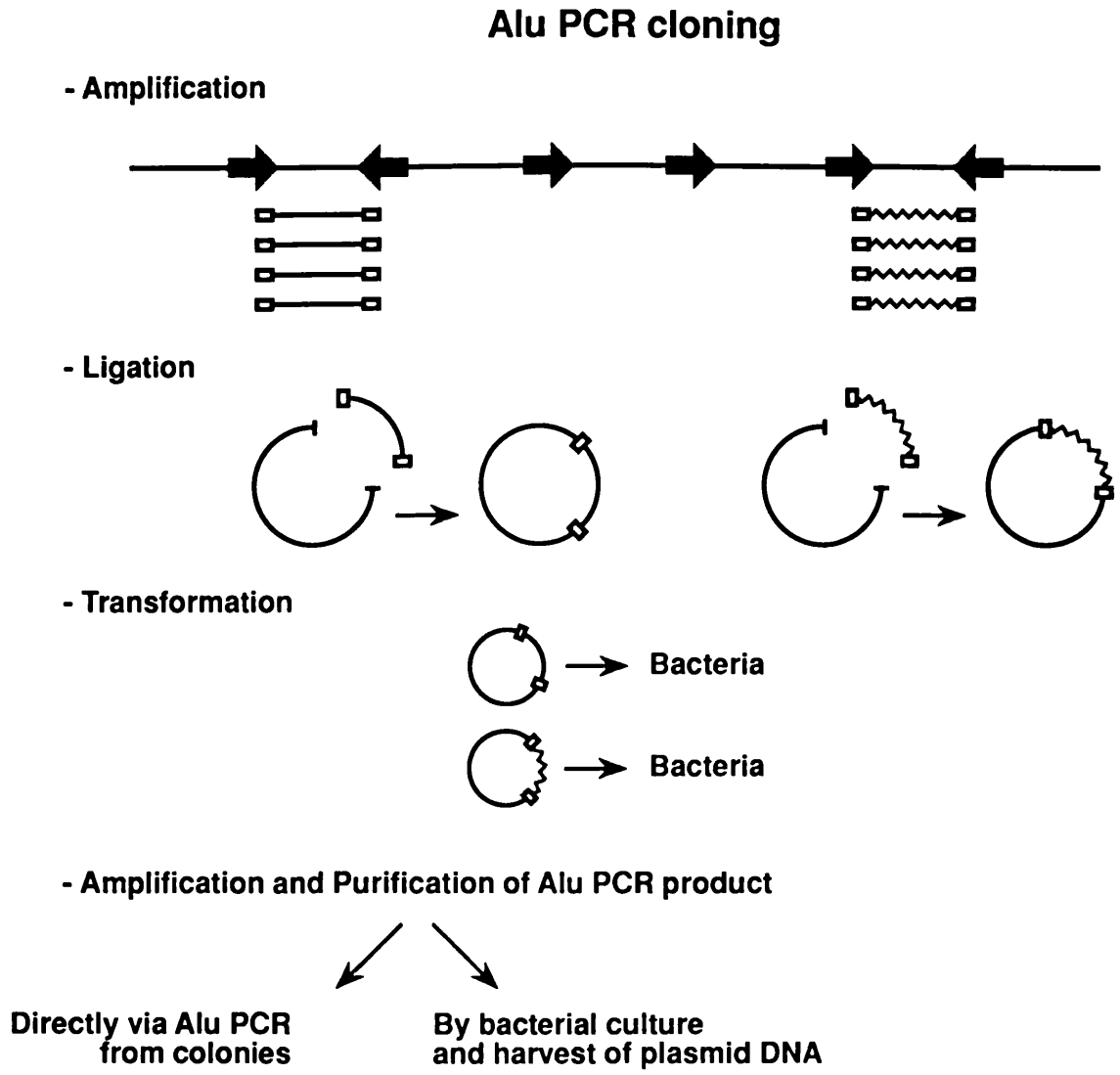


Figure 3.10. A diagram of the Alu PCR cloning procedure using 3' Alu primers. Inter-Alu sequences are amplified using single primers, then ligated into the Eco RV site of the Bluescript plasmid. The ligated plasmids are used to transform DH5α maximum efficiency competent cells and amplification of insert DNA is achieved either by a repeat round of PCR using 3'Alu primers or by growing the bacterial colonies in LBamp and preparing plasmid DNA as described in section 2.2.4.

3.3.3 *Characterising the 3' Alu-PCR library made from ICD*

3.3.3.a *Identification of unique sequences*

As discussed in section 4.2.6, blunt-end cloning is very inefficient. For the library made using the 3' Alu primer (Alu 8559), the transformation efficiency was only $2x10^5$ colonies per μg of plasmid DNA compared to $5x10^8$ colonies per μg of the control plasmid included with the DH5α cells. 185 white colonies were recovered and gridded out for further analysis. Cloned sequences were amplified directly from bacterial colonies by PCR using Alu 8559 primers and visualised after agarose gel electrophoresis as described in section 2.6.3.d. Of the 185 colonies analysed, 63 did not yield a product on two separate occasions. To determine whether this was due to some inadequacy of the PCR protocol to amplify inserts rather than their absence, standard minipreps of plasmid DNA were made from 12 of these 63 colonies. Digestion with Eco RV demonstrated the absence of insert DNA and these colonies were discarded. The 122 colonies containing insert DNA were gridded out onto microtitre plates and transferred to hybridisation membranes which were denatured, neutralised and baked as described in section 2.6.3.b. Probing with sheared, sonicated total human DNA showed strong hybridisation for 22 of the 122 colonies indicating that they contained repetitive sequences.

Large scale plasmid DNA preparation were made for 57 of the remaining 100 colonies (section 2.2.4.b). The inserts were still usually obtained by PCR; diluted plasmid DNA was used as the template for the PCR reaction obviating the need to keep large numbers of bacterial cultures growing indefinitely. The inserts were then used as probes for Southern hybridisation against the mapping panel shown in figure 3.5. No DNA competition was used yet 30 of these 57 inserts acted as unique sequences on Southern hybridisation. This was surprising as each probe had 30 bp of the Alu repeat sequence, derived from the Alu 8559 primer, on either end. None of the probes gave any signal in rodent DNA, so all the bands seen in the somatic cell hybrids used must have been derived from the human DNA they carried. However, it quickly became apparent that certain patterns of hybridisation were found repeatedly, suggesting that the clones from which the probes were derived contained the same cloned DNA inserts.

3.3.3.b *Detecting clones containing identical sequences*

It was important to identify colonies which contained the same cloned sequence to ascertain how many different probes had been developed. Sequencing all the probes would allow accurate comparison of individual clones but is too time consuming for initial screening. Similarly, comparison of the size of bands produced on Southern hybridisation to human DNA cut with different restriction enzymes would provide convincing indirect evidence for the similarity of two clones but, again, the method is

labour intensive. In order to rapidly ascertain whether clones were, in fact, identical two methods were employed. The first depended on hybridisation, the second on restriction enzyme fingerprinting. The relative strengths of these methods are discussed in section 4.2.7.

In the first method, inserts for each clone which produced a similar pattern on Southern hybridisation were generated by Alu-PCR and then loaded in adjacent lanes of an agarose gel. Following electrophoresis the gel was denatured, neutralised and Southern blotted. One of these inserts was then used as a probe and PCR products to which it hybridised were assumed to contain the same sequence. The autoradiograph produced by probing 15 Alu-PCR products with one Alu-PCR clone (3' Alu 71) is shown in figure 3.11.a. Although the probe recognised itself more strongly than any of the other sequences, there was significant cross hybridisation with other inserts, even those of different size. Hence, although this method was occasionally successful, some probes hybridised so strongly to many different bands that it was impossible to decide which should be coded as being identical.

The results of second method, restriction enzyme 'fingerprinting', are shown in figure 3.11.b. In this experiment, 11 probes were compared. The probes shown ranged in size between 250 bp (3' Alu 169) and 900 bp (3' Alu 78). On the basis of size alone, many of the probes could be confidently identified as being different. However, assuming probes were identical simply because they were the same size was not justified. The 'structure' of identical sized probes could be compared further by digesting them with frequently cutting restriction enzymes and comparing the pattern, or 'fingerprint', of bands produced. For example, probes 3' Alu 90, 71 and 53 all appeared to be the same size initially, but 3' Alu 71 gave a markedly different band pattern following digestion with Taq 1 and Rsa 1. Restriction enzyme fingerprinting gave more accurate results than hybridisation and, as it was also quicker, it became the method of choice for comparing different clones. Only 57 of the 100 clones with inserts that had not hybridised to human DNA on colony screening had been assessed by hybridisation. Inserts from the remaining 43 clones were assessed by restriction enzyme fingerprinting but no new probes were identified.

Occasionally, doubt remained as to whether two clones were identical. For example, 3'Alu 51, 66 and 77 were all 850 bp long. Following digestion with Rsa 1 and Taq 1, the fingerprints of 3' Alu 77 and 51 were identical, but the bands produced from 3' Alu 66 were slightly smaller (figure 3.11b). All three inserts were used as hybridisation probes to blots of genomic human DNA and produced identical sized bands on autoradiography (22 kb on a blot made from Eco R1 digested DNA), which suggested they were identical. When 3' Alu 66 and 77 were sequenced, they were found to be identical clones (figure 3.18). Clearly, it is important not to 'over interpret' minor differences in fingerprint patterns. Similarly, it is important to look at the sum

total of all the data obtained about a particular pair of probes and not be misled by a single, spurious, result. Clones 3' Alu 62 and 27 were identical in size and gave identical restriction enzyme fingerprints. They both produced 7 kb bands when used as probes for hybridisation to Eco R1 digested genomic human DNA. However, 3' Alu 62 appeared to be absent from DGF27C11 by hybridisation whereas 3' Alu 27 was present (figure 3.15). They were thought to be different probes until sequence analysis again proved them to be identical (figure 3.18b). When an STS was made from the sequence, 3' Alu 62 was found to be present in DGF27C11 (figure 3.21) by PCR and its absence from DGF27C11 by hybridisation was a false negative finding.

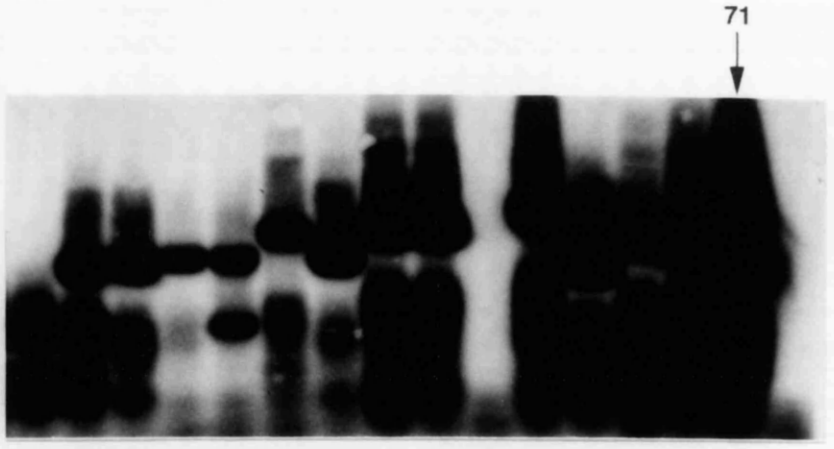


Figure 3.11.a A Southern blot of various 3' Alu PCR products probed with 3' Alu 71. Although the signal in the lane containing probe 71 is the strongest (arrowed), there is cross hybridisation with many of the other probes (even those known to be of different size) making this an inadequate method for deciding if individual sequences are identical.

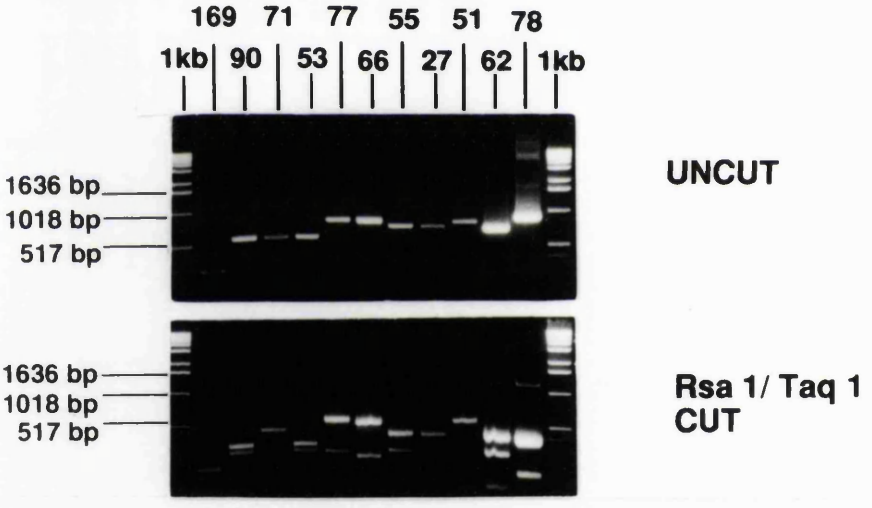


Figure 3.11.b. Comparison of the 3' Alu PCR products that had been found to be unique by Southern hybridisation. In the gel shown above the sizes of the individual 3' Alu clones (identified by their number) are demonstrated. Several appear to be the same size. When these PCR products are digested with Rsa 1/Taq 1 bands of different size are seen, and the pattern of bands produced for each probe can be compared (restriction enzyme fingerprinting). For example, whereas probes 90, 71 and 53 are all the same size uncut, probe 71 has a different fingerprint following enzyme digestion. Probes 66, 77 and 51, which are the same size, produce very similar fingerprints but both of the bands produced by enzyme digestion of 66 are slightly smaller than those of 77 and 51. Probes 55, 27 and 62 are identical both uncut and cut.

3.3.3.c *Attempts to generate unique sub-clones from repetitive probes.*

It was becoming increasingly obvious that, in spite of screening 185 colonies, only six different clones had been identified; 3' Alu 169, 90, 78, 71 66 and 62. In an attempt to increase the number of different probes generated for the 13q14-13pter region, I attempted to derive unique fragments from some of the probes that had acted as low grade repeats. To this end, three probes (3' Alu 61, 79 and 80) were digested with Alu 1 or Taq 1. When the resulting fragments were size separated by agarose gel electrophoresis, they were found to be similar in size to one another. The original intent had been to probe the gel with total human DNA. Fragments that did not hybridise to human DNA may represent unique sequences and would be excised from a second gel (made from the same digest) and used as probes. However, the fact that the fragments were so close together on the gel would have made it difficult to decide which band to excise and purify. Hence, 7 of the fragments that were generated were excised, gene-cleaned and used as probes in Southern hybridisation. All gave repetitive signals both in total human and hybrid (but not rodent) DNA. It appeared that the sub-cloning of fragments of repetitive probes would be time consuming, and offered no guarantees of generating further useful probes. In order to generate further probes from the 13q14-13pter region a second Alu-PCR library was made, this time using Alu 5684, a primer to the 5' end of the Alu repeat sequence.

3.3.4 *Characterising the 5' Alu libraries made from ICD*

PCR amplification of ICD using 5' Alu primers should yield a completely different set of human sequences to that produced by 3' Alu-PCR. I decided to make a 5' Alu-PCR library from ICD DNA using the Alu 5684 primer using a method identical to that for the production of the 3' Alu-PCR library. The problem was, as shown in figure 3.8, that primers to the 5' end of the Alu sequence also amplified rodent sequences. Increasing the annealing temperature of the PCR reaction could completely suppress the amplification of rodent sequences but also drastically reduced the number of human sequences amplified (figure 3.9). As we were trying to maximise the number of probes from the region of interest it seemed counter-productive to deliberately restrict the number of human sequences cloned by using a high annealing temperature. Hence, the library was made at an annealing temperature of 57°C. Figure 3.13 shows the products of the PCR amplification of ICD DNA using the Alu 5684 primer at 57°C. A multi-band pattern is seen, though many of the bands were expected to be of hamster origin. The products of the PCR reaction were purified on Sephadex columns and then ligated into Eco RV cut Bluescript plasmid vector at an insert to vector ratio of 8 to 1 (identical to that used for the 3' Alu library). The ligation mix was then used to transform DH5α competent cells which were grown on LB agar containing ampicillin and X-gal.

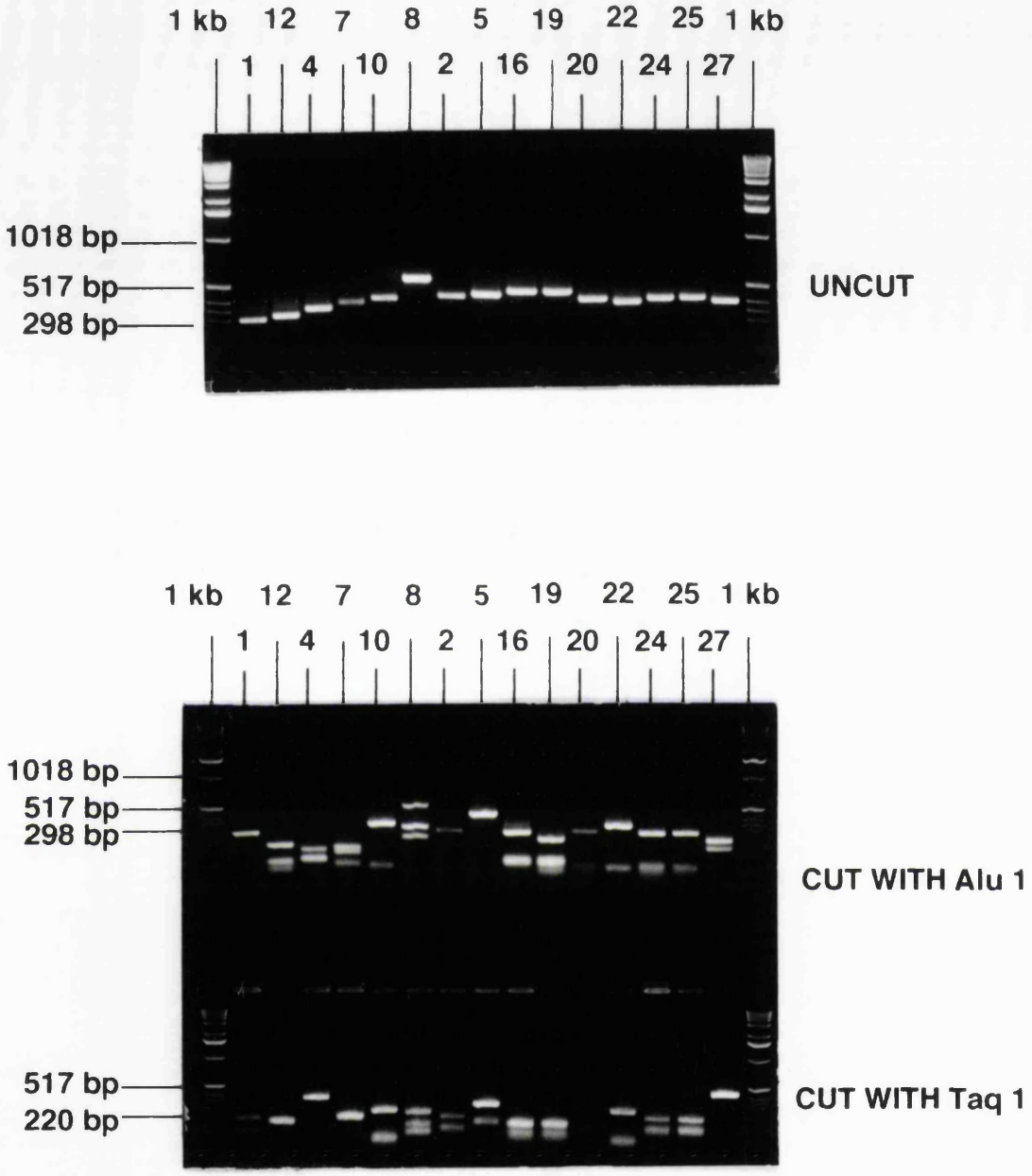


Figure 3.12. Comparison of 15 different 5' Alu-PCR clones, both uncut (above) and cut with either Alu 1 or Taq 1 (below). On the basis of the sizes of the bands produced after digestion, clones 2, 20, 24 and 25 were thought to be identical, as were clones 4 and 27, 10 and 22. Therefore, only 10 different DNA fragments had been cloned.

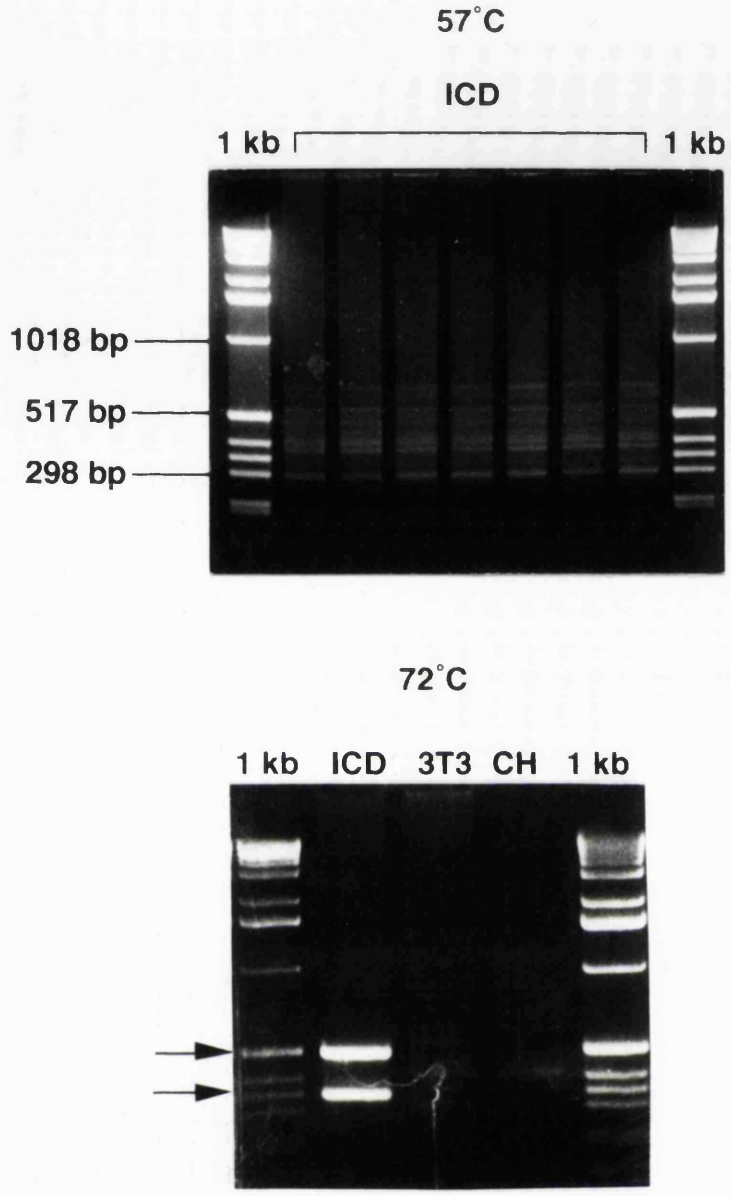


Figure 3.13. These photographs show the PCR products obtained from ICD using 5' Alu primers at two different temperatures. The 5' Alu library was made from the bands produced at 57°C (upper photo) in which a multi-band pattern was seen, although it was known that many of these bands were rodent in origin (see figure 3.8). At 72°C only two bands are seen in ICD (arrowed), with no product visible in mouse 3T3 or Chinese hamster (CH) as shown on the lower photograph.

In this experiment, the transformation efficiency was much higher than that obtained using the 3' Alu-PCR library; $3x10^8$ colonies per μg of plasmid DNA as compared with $2x10^5$. A total of 989 colonies grew but only 36 (3.6%) were white indicating an extremely high frequency of plasmids re-linearising without containing DNA inserts. The reasons for the difference in results between the 3' Alu and 5' Alu libraries is discussed in section 4.2.6. 30 of the white colonies were analysed further and 15 were shown to have no insert by PCR. The frequency of non-recombination in these white colonies was, therefore, 50% compared with 34% in the 3' Alu-PCR library.

From the analysis of the 3' Alu PCR library it was clear that the optimum method of assessing the number of different probes that had been obtained was comparison of the sizes of the inserts, both uncut and after restriction enzyme fingerprinting. The inserts in the 3' Alu library had been digested with Rsa 1 and then with Taq 1. However, a collaborating group was deriving inter-Alu sequences from a similar (though larger) region and assessing them by digestion with Alu 1 or Taq 1. In order to make comparisons between sequences easier, we chose to use the same enzymes here. The upper photograph of figure 3.12 shows the inserts from each of the 15 clones size separated by gel electrophoresis. The lower photograph in figure 3.12 shows the sizes of the fragments obtained when the inserts were cut with Alu 1 or Taq 1. The panel of 15 clones was composed of only 10 different sequences.

All 10 sequences were used as probes for Southern hybridisation and all gave repetitive signals in the total genomic human lane. One probe identified a low level repeat in human and could potentially have been useful if used with stringent Cot-1 DNA competition. However, whereas the repetitive 3'Alu -PCR probes never gave any signal in the mouse lane, all the 5' Alu probes did making them useless for mapping using DGF27C11, a human/mouse hybrid. If the signal that had been produced in the mouse lane had been a single band it may have suggested that the inserts were conserved sequences and, therefore, potentially interesting. However, all the probes produced smears or multi-band patterns in rodent DNA suggesting all were repetitive sequences.

In order to try and develop at least some usable probes with 5' Alu primers, the annealing temperature of the PCR reaction was increased to 72°C at which there was no amplification of mouse or Chinese hamster sequences. Only 2 bands were produced from ICD, both of which were presumably, human specific (figure 3.13). The products of 5' Alu-PCR were all similar in size and the 2 visible bands could have been made up of several bands. The bands were, therefore, separated as far as possible by prolonged electrophoresis on a 2.5% agarose gel. Thin slivers of the gel were then excised as shown in figure 3.14a; 2 above the upper band, 2 from the upper band, 2 between the

two bands, 2 from the lower band and 2 below the lower band A small segment of each was used as the DNA template in a PCR reaction using Alu 5684 primers at an annealing temperature of 58°C. The resulting PCR products are shown in figure 3.14.b and contain 3 bands, an extra one being visible just above the lower band in lane 5. When a filter made from the gel shown in figure 3.14b was probed with total human DNA all three different sized bands gave strong signal indicating that they were repetitive. Hence, in spite of using two different annealing temperatures, no unique probes had been generated from 13q14-13pter using primers to the 5' end of the Alu repeat.

The approach described above may appear to have been a considerable effort for very little gain but this must be viewed in the context of the very small number of probes available from 13q12 at the time. In the end, eight new probes were generated which formed the basis of the future mapping studies.

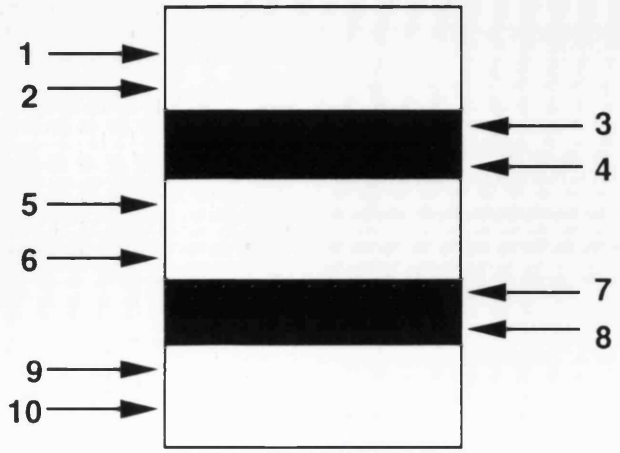


Figure 3.14.a. Diagrammatic representation of the bands (shown in figure 3.13.) produced using the 5' Alu PCR primer at 72°C and ICD as the DNA template. The position of the 10 thin sections of the gel that were cut out for futher analysis is shown by the arrows.

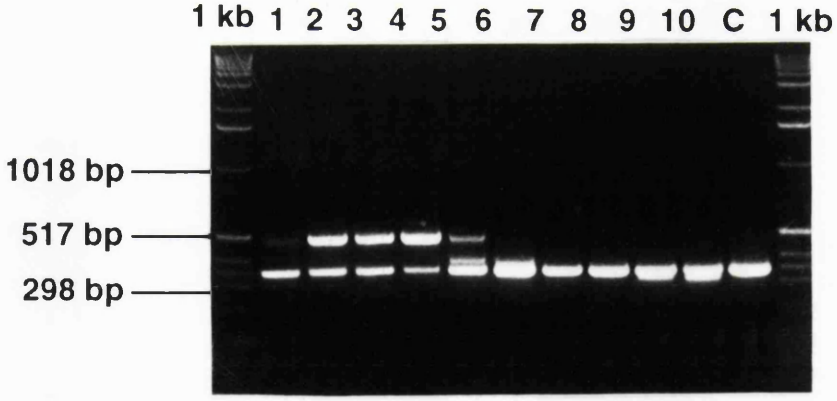


Figure 3.14.b. When the DNA from the 10 gel sections shown in figure 3.14.a. was amplified using 5' Alu primer at 58°C only three different size bands are seen. In order to check the reaction had worked an aliquot of 5' Alu 12 from the 5' Alu PCR library was used as a positive control (C).

3.3.5 *Mapping of unique 3' Alu-PCR probes*

Despite the considerable effort that had been expended in the generation and characterisation of inter-Alu probes, it appeared that only eight different unique probes had been isolated from the region of interest (the mapping experiments were performed before it became apparent that 3' Alu 66 and 77 were identical as were 3' Alu 62 and 27 as discussed in section 3.3.3.a), all from the 3' Alu-PCR library.

The location of all eight probes was determined by mapping them on the panel of somatic cell hybrids shown in figure 3.4. and to DGF27C11, the hybrid containing the der (1) chromosome from DG. All eight probes were present in all of the hybrids including KBF11, placing them in 13q12. The chromosome 13 deletion in hybrid KBF11 extends to the middle of 13q12 and, until the characterisation of the DG breakpoint, the proximal breakpoint in KBF11 was the most proximal one reported on chromosome 13. Hence all eight probes were more proximal then the previously reported most centromeric markers, pHU10 and p7F12.

The results obtained when these eight probes were hybridised to Southern blots of Eco R1 digested DNA from DGF27C11 are shown in figure 3.15. Five probes (3' Alu 77, 66, 78, 71, and 62) were not present in DGF27C11 and 3 probes (3' Alu 27, 169 and 90) mapped between the DGF27C11 breakpoint and the proximal breakpoint in KBF11 as shown in figure 3.15. The size of each 3' Alu probe, its presence or absence in DGF27C11 and the band pattern produced on hybridisation to Eco R1 digested human DNA, are shown in table 3.1. Although 3' Alu 27 and 62 had given identical patterns on fingerprinting and sizes of bands on Southern hybridisation, they appeared to map on different sides of the DG breakpoint and were assumed to contain different sequences. Hence, even though only a small number of probes had been generated from 13q14-13pter using Alu PCR cloning, we had derived markers that flanked the DG breakpoint more closely than any others reported.

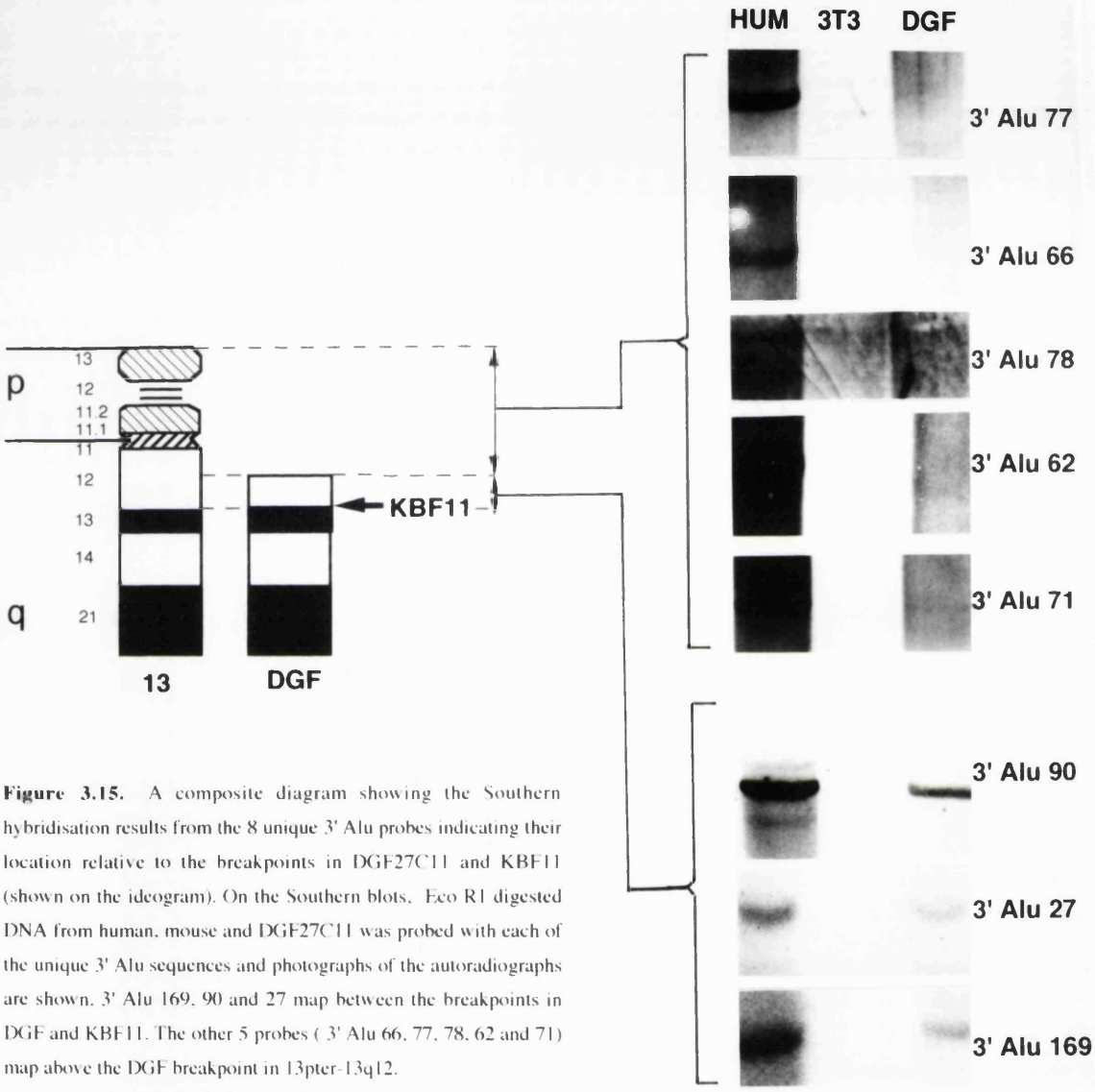


Figure 3.15. A composite diagram showing the Southern hybridisation results from the 8 unique 3' Alu probes indicating their location relative to the breakpoints in DGF27C11 and KBF11 (shown on the ideogram). On the Southern blots, Eco R1 digested DNA from human, mouse and DGF27C11 was probed with each of the unique 3' Alu sequences and photographs of the autoradiographs are shown. 3' Alu 169, 90 and 27 map between the breakpoints in DGF and KBF11. The other 5 probes (3' Alu 66, 77, 78, 62 and 71) map above the DGF breakpoint in 13pter-13q12.

Table 3.1

3' Alu probe	Size (bp)	Size of band on Eco R1 blot (kb)	Presence in DGF27C11 by hybridisation.
66	850	22	No
77	850	22	No
78	900	22	No
71	550	7	No
62	750	7	No
27	750	7	Yes
90	550	22&20 doublet	Yes
169	340	22	Yes

Details of the size, the band pattern on hybridisation to Eco R1 digested human DNA and the presence in DGF27C11 of each unique 3' Alu-PCR generated probe as assessed by Southern hybridisation.

Although it appeared that closely flanking molecular markers had been developed two major problems still remained. Firstly, the map position of the 3' Alu probes had been determined by the presence or absence of bands on Southern hybridisation, which is known to give false negative results, especially when the target DNA forms a small proportion of the DNA in the sample.

Secondly, the probes absent from DGF27C11 could map anywhere within the 13q12-13pter region and could, therefore, be far too far from the DG breakpoint to act as useful starts for a YAC walk. Two strategies were undertaken to determine the position of the probes which mapped above DGF27C11. The probes were used to isolate cosmids from the chromosome 13 specific cosmid library available in the ICRF, and the cosmids were then used as probes for FISH to normal metaphase spreads. The position of the hybridisation of the cosmids on normal copies of chromosome 13 could be then determined. The second approach was to try to acquire a more detailed somatic cell hybrid mapping panel for 13q12. In addition, sequencing the probes and developing oligonucleotide primers for the ends would allow mapping by PCR which is at least 1000 fold more sensitive than Southern hybridisation.

3.4 MAPPING USING COSMIDS AND FISH

In-situ hybridisation of fluorescence-labelled cosmids to normal metaphase spreads will allow their position on chromosome 13 to be visualised. The 3' Alu probes were all under 1 kb in length and successful FISH experiments depend on the probes being at least 5 to 10 kb long. Hence, the 3' Alu sequences had to be used to identify longer stretches of genomic DNA that could be used as probes for FISH. Each of the eight 3' Alu probes that were thought to be unique by Southern hybridisation was used to screen a chromosome 13 specific cosmid library. The ICRF library consisted of 20,000 cosmid clones, in the Lorist 4 vector. The 20,000 clones were gridded in arrays of 6x6 on a 20 cm x20 cm hybridisation membrane (Lehrach and et al, 1990). Each filter contained approximately five genome equivalents of chromosome 13. In order to be able to accurately identify the position of a potential positive clone on the filter, the membranes were co-hybridised with Lorist 4 arms which had been radio-labelled with ^{35}S dATP as described in section 2.10. The labelled cosmid vector gave faint signals from each of the colonies on the membrane. Figure 3.16.a shows that, when the 3' Alu probes were labelled with ^{32}P dCTP, they gave dense black dots that were easily distinguished from the background of ^{35}S signal.

For each probe, a series of potential positive clones were identified and then a two stage procedure was used to select those clones which actually contained the correct sequence used in the original hybridisation. Firstly, stab cultures of each of the potential positive clones were obtained and colony purified by streaking out on LB agar plates with kanamycin selection. Individual colonies were grown in arrays on LB agar with kanamycin and transferred to hybridisation membranes as described in section 2.9.4. The membranes were probed with the appropriate 3' Alu clone, those cosmid colonies that yielded strong hybridisation signal were analysed further. Cosmid DNA was prepared from individual colonies using the standard mini-prep procedure (section 2.2.4a). The resulting cosmid DNA was then digested with Eco R1 and separated by electrophoresis on a 1% agarose gel. The gel was blotted and probed with the 3' Alu sequence that had been used to identify the cosmid.

Figure 3.16 shows the typical results obtained by screening the library, in this case with 3' Alu 77. Figure 3.16.a is a photograph of a section of the library filter showing the positive clones showing up against a fainter background of host signal. The Eco R1 digest of the 7 potential positive clones identified is shown in figure 3.16.b. and it can be seen that a multi-band pattern is obtained. Although each of the cosmids shares at least one band of identical size with each of the other cosmids, the overall fingerprints of the seven cosmids are very different. Cosmids 2 and 6 in figure 3.16.b share four bands of identical size and differ by only one band implying that they may contain very similar human inserts. Indeed, as can be seen in figure 3.16.c, of these 7 clones, only cosmids 2 and 6 gave any signal when hybridised with 3' Alu 77.

The FISH data was kindly obtained by two co-workers in the ICH laboratory; Ms. L.A. Hawthorn and Ms. H. Kempski and confirmed by Dr. D Shapiro from St. Jude's Hospital, Memphis, USA. Table 3.2 summarises the results of screening the chromosome 13-specific cosmid library with the eight markers derived by Alu-PCR.

Table 3.2

3' Alu Probe	N° positives on library screen	N° positives after 2° screen	Identification number	Location by FISH
77	12	2	C108FO419	13q12
66	same as 77			
71	14	9	C108F0380	13q12-q14
78	18	8	C108E055	poor signal
90	17	3	C108H0710	many centromeres
169	212	not analysed		
62	5	5	C108A0723	13q12-q21
27	same as 62			

Details of the results obtained when the chromosome 13 specific cosmid library was screened with the eight 3' Alu unique probes. The number of potential positive clones identified by Southern hybridisation is noted as is the number that contained the sequence of interest after two rounds of further analysis as described in section 3.4. The identification number of the clone used as a probe for FISH is given and its chromosomal location by FISH is noted.

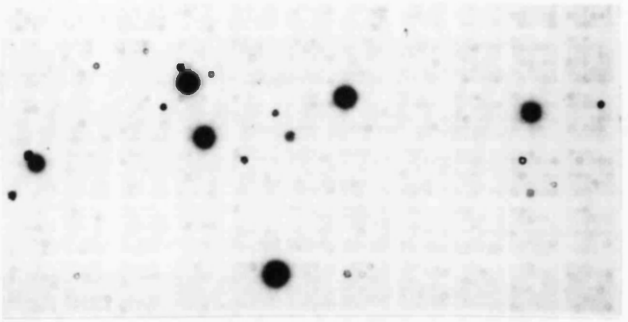


Figure 3.16.a. An autoradiograph showing the results from a typical screening of the chromosome 13-specific cosmid library. Individual cosmids are arranged in 6x6 grids which have been highlighted using ^{35}S labelled Lawrist arms. Superimposed on this weak signal are the six high intensity positive clones following hybridisation 3' Alu 77 labelled with ^{32}P.

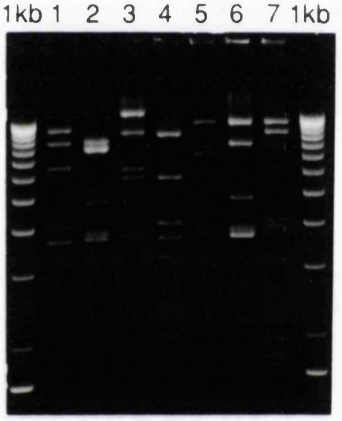


Figure 3.16.b. An Eco R1 digest of 7 cosmids identified by 3' Alu 77 showing multiple bands from each cosmid. Although several cosmids share identical sized bands, the overall 'fingerprints' of the seven cosmids are all different.

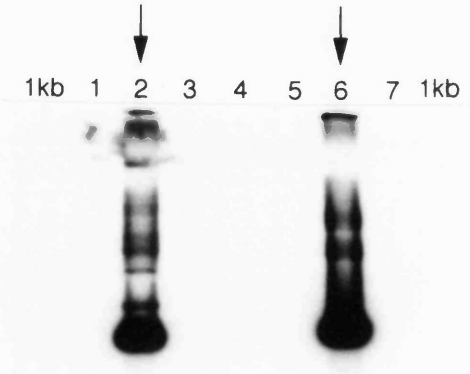


Figure 3.16.c. A Southern blot made from the gel shown in figure 3.16.b. was probed with 3' Alu 77. Only 2 cosmids (arrowed) actually hybridise with 3' Alu 77, the other 5 represent errors in the identification or picking of the cosmids.

Analysis of the data shown in table 3.2 allowed several important conclusions to be drawn. Firstly, FISH analysis allowed the accuracy of mapping data obtained from the screening of a panel of somatic cell hybrids using hybridisation to be checked. Five of the eight probes (3' Alu 77, 71, 66, 62 and 27) were found to map in 13q12, as suggested by previous mapping data (figure 3.15). The fact that these five probes all mapped to the region of interest and not to 13p implies that they are likely to closely flank the DG breakpoint. However, 3' Alu 90 appeared to hybridise to the centromeres of many different human chromosomes and not to 13q12. The presence of this probe in DGF27C11 by hybridisation was not due to its homology to sequences in 13q12-13qter but to sequences in the centromere of chromosome 1. In this way the data obtained by FISH had allowed more accurate localisation of the probe than was possible by mapping.

Secondly, the pattern of cosmid hybridisation proved to be a useful way of comparing probes. 3'Alu probes that identified identical cosmid clones were very likely to contain the same sequence and, therefore, this data supported the results noted in section 3.3.3b which suggest that 3' Alu 66 and 77 are identical as are 3' Alu 27 and 62.

Thirdly, the number of cosmid clones identified can be a sensitive way of deciding if a given clone is indeed functioning as a unique sequence. The cosmid library has a five times representation of chromosome 13 and, therefore, about five cosmids should be identified with each probe. 3' Alu 169, which identified 212 clones, functions as a low grade repetitive sequence and is not suitable for mapping.

Mapping by FISH was not, however, an unqualified success. 3' Alu 78 did not appear to be an obviously repetitive probe by FISH but the signal obtained was weak and it was not possible to determine the precise location of this clone. Section 4.2.8 discusses the potential reasons for the inability of FISH to accurately map 3' Alu 78 as well as the strengths and weaknesses of the technique.

3.5 DEVELOPING STS FOR THE 3' ALU PROBES

3.5.1 *General considerations*

Southern hybridisation can give false negative results when the molar ratio of target DNA in the sample is low. Screening by hybridisation also requires the preparation of relatively large amounts of DNA (15 μg per lane) and is intolerant of even small amounts of DNA degradation. The procedure is, therefore, time consuming and cumbersome for screening large numbers of samples. By contrast, screening using PCR can be performed directly from bacterial or yeast colonies, uses minute amounts of DNA and is both quicker and more sensitive than Southern hybridisation. However, in order to use PCR in screening protocols, it is necessary to develop oligonucleotide primers specific for both ends of the probe. These primers identify an operationally unique DNA sequence and constitute a sequence tagged site (STS) (Green and Olson, 1990). In order to develop oligonucleotide primers sequence data has to be obtained for both ends of each probe. As the 3' Alu sequences had been cloned into the Bluescript plasmid vector they could be amplified using primers for the T3 and T7 promoter regions. Biotinylation of only one of these primers meant that the amplified DNA strands could be separated, as discussed in section 2.8. The region of the Bluescript plasmid containing the T3 and T7 promoters and the polylinker is shown diagrammatically in figure 3.17a, both before and after a human insert has been cloned. The sequences of the primers for the T3 and T7 promoters, as well as that of the polylinker region are shown in figure 3.17b. It is clear from these diagrams that, if T3 and T7 primers were used to sequence the separated strands of DNA, sequence would be derived from the promoter, the polylinker region of the plasmid and the Alu primer sequence before any sequence could be obtained from the probe of interest. However, if Alu primers, which are internal to the promoter and polylinker regions, are used for sequencing, then sequence can be obtained from the probe much more easily.

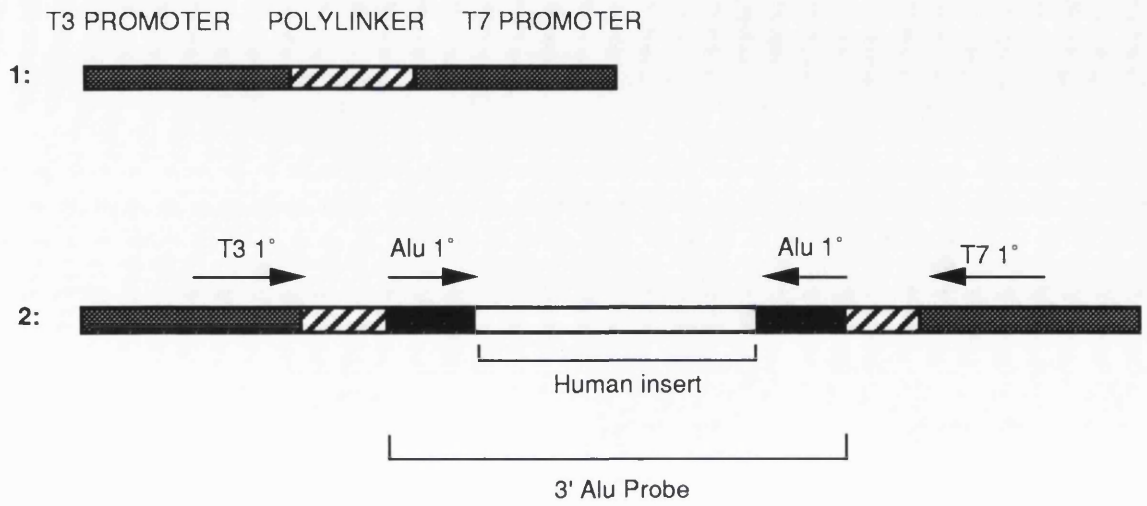


Figure 3.17.a. In (1) the relative positions of the T3 and T7 promoters and the polylinker in the Bluescript vector are shown. Following successful cloning of an inter-Alu fragment the organisation of the cloning site is shown in (2). The Alu sequences at either end of the probe are shown in black, the human insert in white. The position and orientation of the T3, T7 and Alu primers is noted.

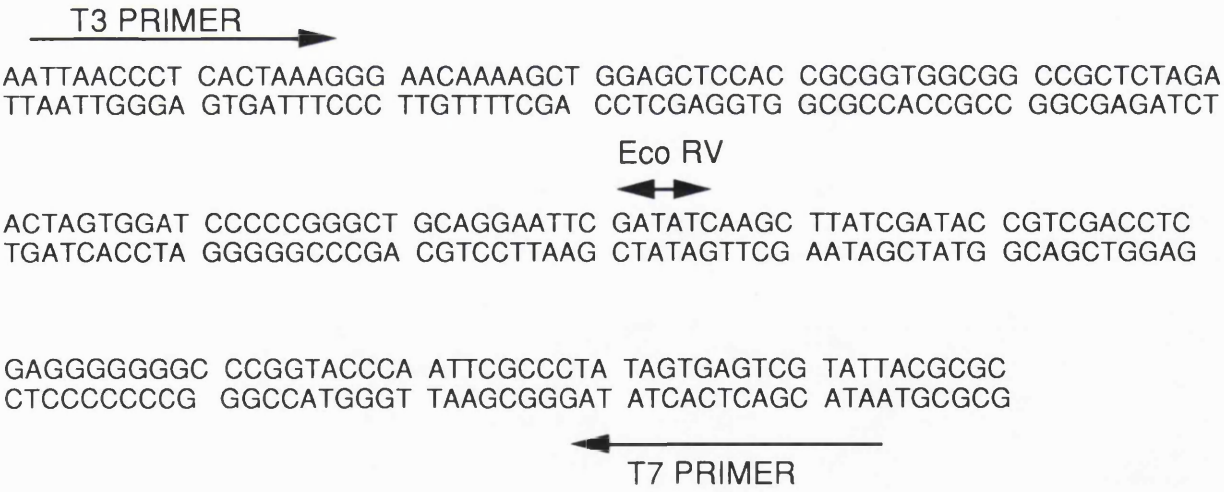


Figure 3.17.b. The sequence of the pBluescript II SK vector showing the position of the T3 and T7 primers relative to the Eco RV cloning site.

3.5.2 *Sequencing the 3' Alu probes*

Using the methods described in section 2.8, sequence data was obtained from all eight 3' Alu probes. Initially, biotinylated T3 primers and non-biotinylated T7 primers were used but, although good sequence data was obtained from the biotinylated strand, the signal from the non-biotinylated strand was often very weak. If the signal was too weak to read with confidence, the experiment was repeated using biotinylated T7 and non-biotinylated T3 primers, making the previously weak 'negative' strand the 'positive' strand, thereby improving signal strength.

The type of comb used to make the sequencing gel made a difference to the ease of interpretation of the results. When combs which produced wells 5 mm wide were used, the individual lanes were well separated which minimised the 'bleeding' of one lane into another. This type of comb was unrivalled for ease of reading of an individual sequence. However, because the wide combs spread the sample over a larger area, the bands appeared pale if the activity of the radio-labelled sample was low. By increasing the exposure time to 14 days, readable sequence could be obtained from most reactions but the length of exposure was often a hindrance to the timely development of STS. Also, the width of the comb reduced the number of samples that could be run on a given gel and if comparison of the sequences of many different samples was required the smaller 'shark's tooth' gel was used. Examples of sequences obtained with both types of comb are shown in figure 3.18.

All eight sequences showed a long poly (dA) tract at both ends of the clone adjacent to the Alu repeats. These A- rich region were usually so long that reading the sequence distal to them was difficult, even when the polyacrylamide gels were run for a longer time. The poor signal resolution was compounded by the frequent band compression seen on the sequencing gels at the distal end of the poly (dA) tract. Examples of these problems are shown in figure 3.18a, in which the long poly (dA) tracts and band compression identified in the sequence from the T7 end of 3' Alu 77 are arrowed. Similar findings are seen in figure 3.18b which shows sequences from the T3 end of 3' Alu 66, 62 and 27 run on a gel which had been made using a shark's tooth comb. Changing the dideoxy-nucleotides used from dG to dI and adding Mn buffer at 1:10 dilution did not reduce the band compression seen. As a result, very little unique sequence was obtained from most of the 3' Alu clones and this limited the choice of sequence from which to design primers.

One of the advantages of sequencing the 3' Alu probes was that it was possible to decide for certain which clones contained identical inserts. As shown in figure 3.18b, the sequences for 3' Alu 27, 55 and 62 were identical supporting the data obtained from restriction enzyme fingerprinting (section 3.3.3a and figure 3.11b). By comparing the sequences in figures 3.18a and b, it is clear that the T7 end of 3' Alu 77 and the T3 end

of 3' Alu 66 are identical (as are the T7 end of 3' Alu 66 and the T3 end of 3' Alu 77, data not shown) proving that these two probes contain the same sequence which has been inserted into the plasmid in opposite orientations. Hence, we now had irrefutable evidence that only six different 3' Alu sequences had been cloned; 3' Alu 169, 90, 78, 71, 66 and 62.

3.5.3 *Designing oligonucleotide primers for the 3' Alu probes*

In spite of the difficulties in obtaining good quality sequence data (section 3.5.2), it was still possible to design oligonucleotide primers for all 6 different probes. The aim of designing oligonucleotide primers was to define STS that were both specific and easy to use. The specificity of a given PCR reaction can be increased by raising the annealing temperature (Saiki et al, 1988). The annealing temperature is determined by the melting temperature (Tm) of the oligonucleotide primers which can be calculated from the formula Tm = 69.3+ 0.41(G+C%) - 650/L (where L is the length of the oligonucleotide primer). From this formula it can be seen that the higher the G+C ratio and the longer the primers, the higher the Tm will be. Primers were not designed from regions with an AT content of more than 70% as their resulting Tm would have been so low that spurious amplification of non-specific products may have occurred. Hence, none of the primers were designed from sequences in the poly (dA) tracts. Regions with too high a GC content were also unsuitable for primer design. Here, the problem is that secondary structure formation can occur in the template or the primers, resulting in reduced product yield or non-specific amplification (Innis, 1990). The Tm of each oligonucleotide in a given primer pair had to be similar and so care was taken to ensure a similar GC content in each. When this was not possible, the length of one the primers was altered so that its Tm equalled that of its partner. The sequences of all the primers used in this thesis are noted in appendix 1 and it can be seen that most were of the order of 20 bp long.

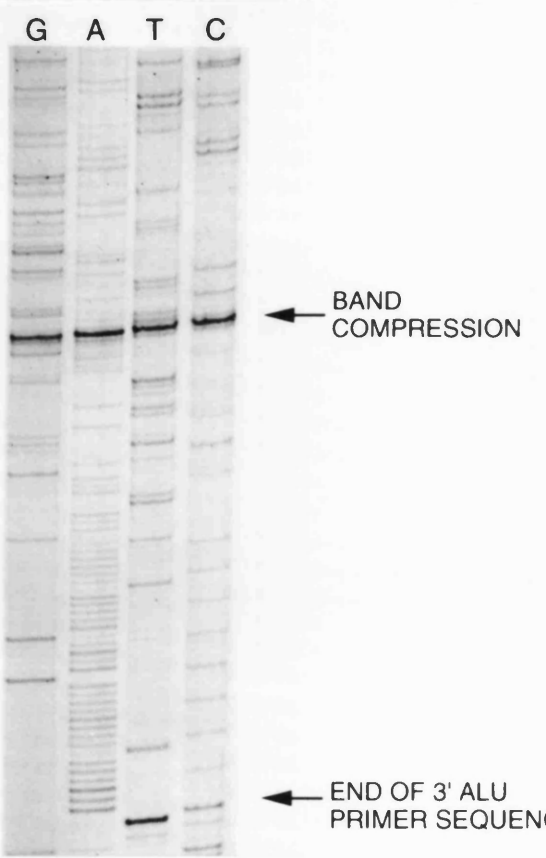


Figure 3.18 a. A portion of the DNA sequence obtained from 3' Alu 77 using the T7 primer. The lower part of the gel (up to the arrow) corresponds to the end of the Alu primer used to generate the 3' Alu library. Above the arrow is the sequence of 3' Alu 77 showing a long A-rich tract followed by a dense band compression.

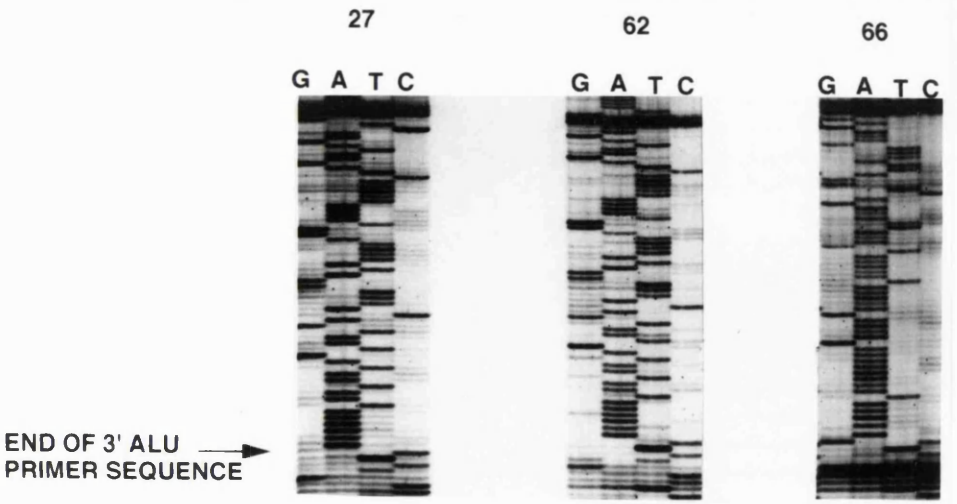


Figure 3.18.b. A 'shark's tooth' gel showing the sequence obtained from 3' Alu 27, 62 and 66 using the T3 primer. The lower part of the gel (up to the arrow) represents the end of the sequence of the Alu primer used to generate the 3' Alu library. Above the arrow the band pattern for 3'Alu 27 and 62 remains identical throughout, whereas that for 3'Alu 66 is different. Note that the sequence of 3' Alu 66 using the T3 primer is the same as that of 3' Alu 77 using the T7 primer (shown in figure 3.18.a.) demonstrating that these two probes are identical but have been inserted into the plasmid in different orientations.

3.5.4 *Optimising PCR conditions for STS for 3' Alu unique probes.*

The ideal PCR reaction would specifically amplify a single product in sufficient quantities to be visualised by agarose gel electrophoresis. In order for this aim to be achieved, the conditions of the PCR reaction need to be optimised.

As mentioned in section 3.5.3, the annealing temperature is an important determinant of the specificity of the PCR reaction. The annealing temperature was predicted from the Tm of the primers used. In practice, optimum specificity of amplification was most often obtained within 3°C of the Tm. Small differences in annealing temperature, however, could make a big difference to the specificity of the PCR reaction. Figure 3.19, for example, shows the products of two PCR reactions using 3' Alu 66 primers, in one reaction the annealing temperature was 57°C and in the second reaction it was 60°C (the Tm of these primers was 62°C). At 57°C there is a dominant 550 bp band but many other bands are seen even in lanes from which the dominant band is missing. Raising the annealing temperature just 3°C to 60°C results in the production of just the single, dominant, band.

The magnesium concentration of the PCR mix also has a profound impact on the PCR reaction. A higher magnesium concentration results in less specific annealing of the primers. The usual concentration employed was 1.5 mM but some primers required the use of 2.0 mM in order to give an adequate amount of product. An example of this is shown in figure 3.20, where no PCR product was seen using 1.5 mM magnesium but a good yield was produced simply by increasing the concentration to 2.0 mM.

Increasing the number of amplification cycles resulted in more product yield but could also reduce the specificity of the reaction, especially if over 30 cycles were performed. For example, at an annealing temperature of 60°C 3' Alu 66 primers gave a single band when 30 cycles are performed, but after 35 cycles a multi-band pattern resulted (the product looking identical to that obtained using the lower annealing temperature of 57°C in figure 3.19).

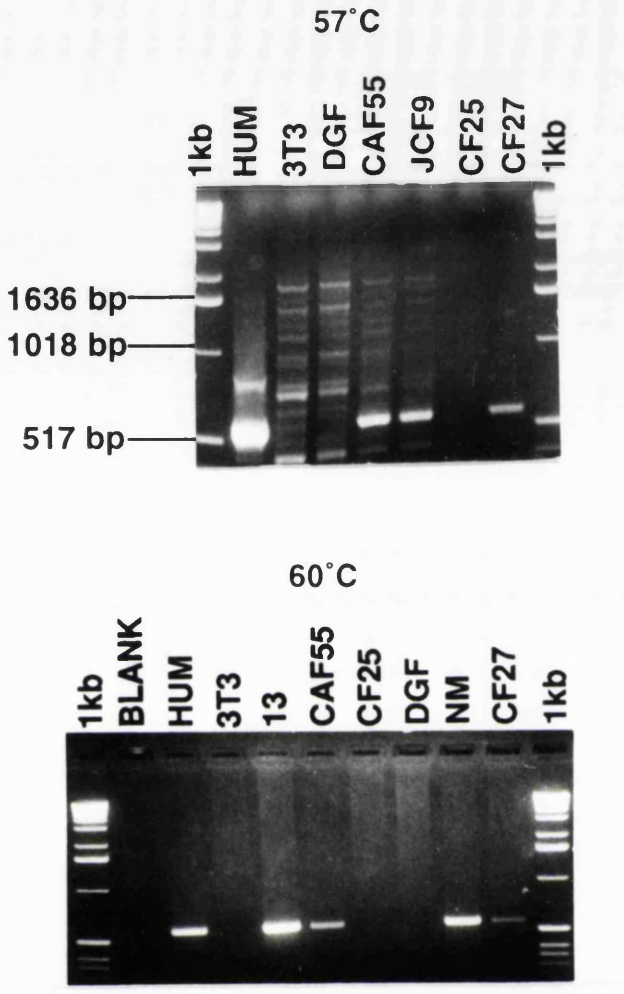


Figure 3.19. The effect of raising the annealing temperature on the specificity of a PCR reaction using oligonucleotide primers for 3' Alu 66. The upper photograph shows a dominant 550 bp band and the extra bands that result when an annealing temperature of 57°C is used. The extra bands are derived from mouse sequences and are seen in mouse 3T3 DNA and all the mouse/human hybrids (DGF, CAF55, JCF9). CF25 and CF27 are hybrids in which Chinese hamster cells were used and do not show any extra bands. The lower photograph shows the disappearance of the mouse-specific bands if the annealing temperature is increased to 60°C.

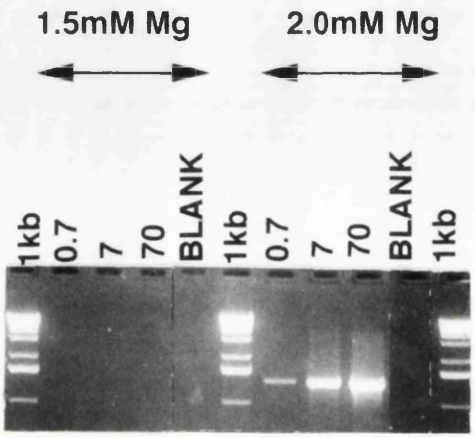


Figure 3.20. Example of the effect of magnesium concentration and the amount of template DNA on the sensitivity of a PCR reaction. Two PCR reactions were performed using the same oligonucleotide primers. In each reaction, the concentration of human DNA in the templates was 0.7 ng, 7 ng, and 70 ng. The magnesium concentrations used were: 1.5 mM (left) and 2.0 mM (right). No product is seen on the left whereas on the right, a PCR product is obtained in all 3 lanes indicating that these primers require the higher magnesium concentration in order to work. In the experiment shown on the right it can be seen that increasing the amount of DNA template increases the amount of PCR product obtained.

The amount of product obtained also depended on the amount of template DNA added. In general, the more template used, the greater the yield of PCR product. This is also illustrated in figure 3.20, where, at a magnesium concentration of 2.0 mM, only a very faint band is seen on electrophoresis of the product of a PCR reaction to which 0.7 ng of template DNA had been added. Increasing the amount of template DNA to 7 ng and then 70 ng resulted in a step-wise increase in the PCR product yield. Where possible, 100 ng of template was used, though PCR was capable of giving product visible on agarose gel electrophoresis with as little as 0.25 ng.

The duration of the annealing and denaturation steps depended on the length of the products being amplified. Allowing 60 sec per kb of product to be amplified seemed to be sufficient for full length amplification to occur.

Hence, the factors that affect the specificity of a PCR reaction are primarily the annealing temperature, and the magnesium concentration of the PCR reaction mix. The yield of PCR product can be optimised by adjusting the duration of the annealing and extension times and using sufficient template DNA. Clearly, designing oligonucleotide primers and optimising the conditions for their use in PCR reactions is a labour intensive process and could only be justified if a high percentage of the inter-Alu sequences yielded successful STS.

3.5.5 *STS designed for 3' Alu probes.*

It was possible to design successful STS for 5 of the six different 3' Alu probes. The one failure was 3' Alu 169, which was 340 bp long and had a human insert (in between the Alu primers) of only 280 bp. Long poly (dA) tails are present at both ends of the human sequence and, although primers were designed from the first readable sequence after these tracts, the resulting PCR product was only 50 bp long. When used in PCR reactions, 3' Alu 169 primers produced bands in human, mouse and the yeast, Saccharomyces cerevisiae. As 3' Alu 169 appeared to act as a low grade repeat on screening cosmid library filters (section 3.4 and table 3.2) it was not investigated further.

The first set of primers designed for 3' Alu 71 gave low signal intensity at the annealing temperature predicted by their Tm. When the annealing temperature was reduced in order to try to increase product yield, background amplification of mouse and yeast DNA occurred. A second set of primers were, therefore, made, both 24 mers rather than 21 mers. The new primer at the T3 end was identical to the original, but with the addition of three extra bases. The new primer at the T7 end was derived from a different part of the sequence of the probe, starting 13 bases internal to the end of the first. Figure 3.21 shows the strong bands produced using the second set of primers at

the annealing temperature predicted by their Tm, and with 2.0 mM magnesium in the pre-mix. It was much easier to derive useful STS for the other four 3' Alu probes. All four yielded specific PCR products at the annealing temperature predicted from their Tm with a magnesium concentration of 1.5 mM (figure 3.21).

Although it had been possible to derive useful STS for five of the six inter-Alu probes (3' Alu 90, 78, 71, 66 and 62) the work involved in doing this would only be worthwhile if screening by PCR had distinct advantages over screening by hybridisation. Before comparing the results of screening with the two techniques, the development of a more detailed mapping panel of somatic cell hybrids will be described.

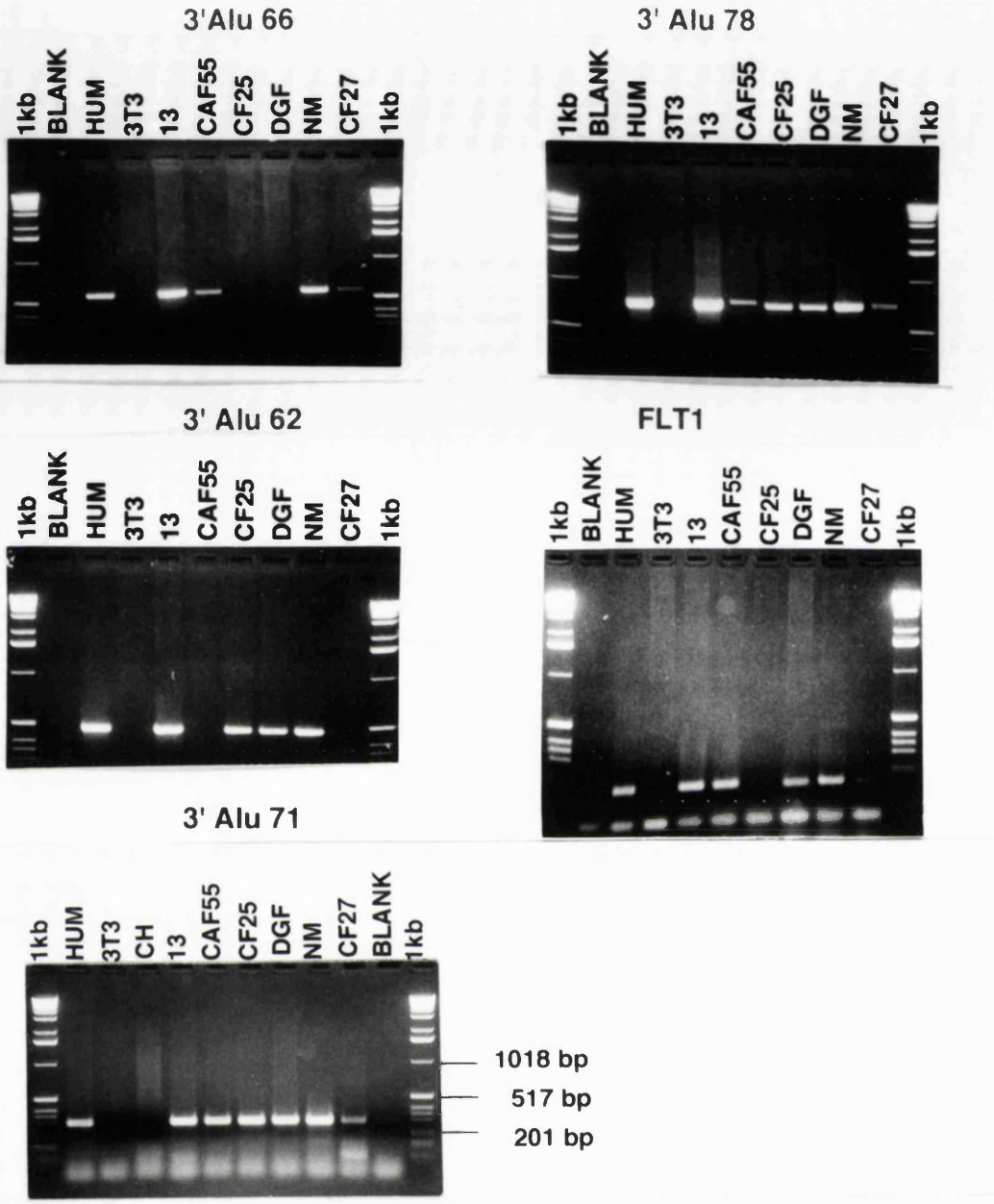


Figure 3.21. The results of electrophoresis of the PCR products obtained when STS for probes mapping to 13q12 were used to screen a panel of somatic cell hybrids. The hybrid panel is described in figure 3.22. The 550 bp band produced by 3' Alu 66 is not seen in DGF27C11 or CF25 but is seen in the other hybrids. An STS for the oncogene, FLT1, produces a 180 bp band in all the hybrids except CF25. STS for 3' Alu 78 and 71 both produce bands (600 bp and 300 bp in size respectively) in all of the hybrids. 3' Alu 62 produces a 500 bp band in CF25, DGF27C11 and NM. All the STS used show appropriate sized bands in genomic human DNA and a hybrid containing all of human chromosome 13 as its only human component (13) and no product is obtained from DNA extracted from mouse 3T3 cells or from Chinese hamster (CH) cells.

3.6 EXTENDING THE CHROMOSOME 13 SPECIFIC SOMATIC CELL HYBRID MAPPING PANEL

All of the unique 3' Alu probes had been localised above the proximal breakpoint in KBF11, in 13q12. However, the proximal KBF11 breakpoint and the one in DGF27C11 only divided 13q12 into two relatively large subsections. In order to map new probes as close as possible to the DG breakpoint it was essential to generate hybrids with breakpoints that closely flanked that of DG. Two patients, CA and JC, were identified with constitutional deletions of 13q12-13q14 which had predisposed them to retinoblastoma (Cowell et al, 1988). The proximal breakpoints of these deletions could subdivide 13q12 and allow further definition of the position of markers relative to the DG breakpoint. Hence, hybrids were created between cells from the lymphoblastoid cell lines from these two patients and mouse 3T3 cells. Individual hybrid clones were isolated and screened by PCR using oligonucleotide primers for the factor X gene (FACX), located in 13q34, and primers for a variety of exons of the retinoblastoma predisposition gene (RB1) which lie in 13q14 (Hogg et al, 1992).

One hybrid from patient CA, CAF5, showed a strong PCR product for FACX and a weak signal for RB1 indicating that the majority of the hybrid cells contained the deleted chromosome 13 but a few also carried the normal homologue. Single cell clones were isolated and two, CAF55 and CAF510 showed the presence of FACX but not RB1. One of the hybrid clones from patient JC, JCF9, showed the same profile as CAF55 without the need for sub-cloning. The presence of the deleted chromosome 13, but not the normal homologue, was confirmed by cytogenetic analysis of both hybrids (J. Cowell, pers. comm.).

Three other hybrids with breakpoints in 13q12 were obtained from other laboratories. CF25, and CF27 (Mohandas et al, 1979) were provided by Dr. D. Warburton who also generated and provided a third hybrid, NM. Figure 3.22 is an ideogram showing the regions of chromosome 13 that had been retained in these hybrids. Although cytogenetic analysis had suggested that these hybrids may be of use, the only way of deciding if the hybrids had divided 13q12 into subregions was to map them with all the STS available for the region, as discussed in the next section.

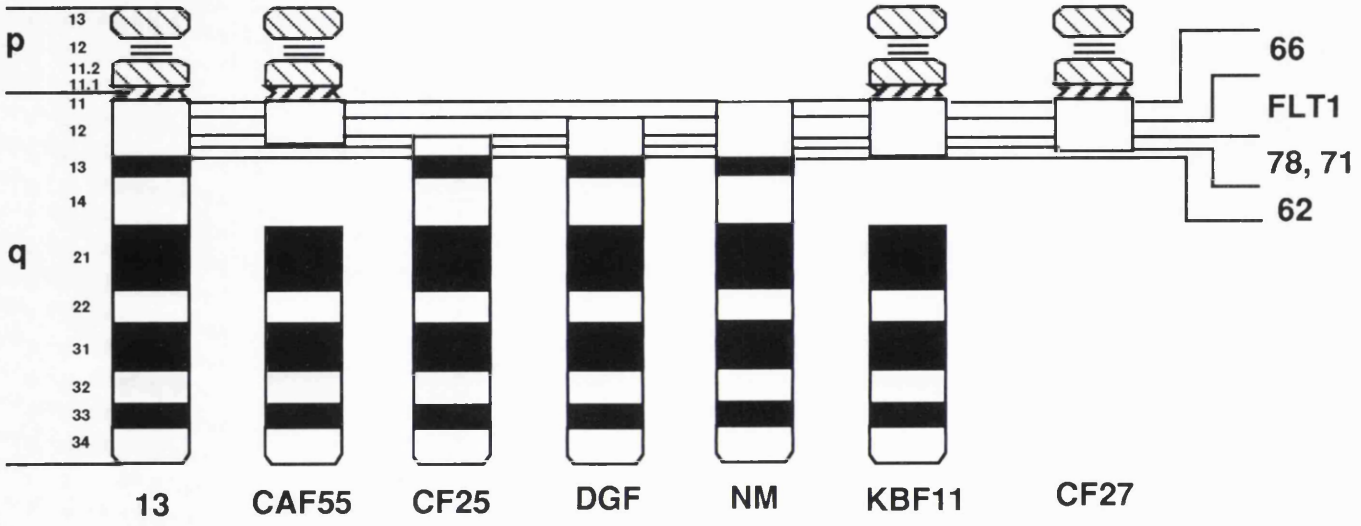


Figure 3.22. An ideogram of the extended chromosome 13 mapping panel which allows sub-division of the 13q12 region. The position of the STS for the four 3'Alu probes and the oncogene, FLT1, are shown.

3.7 MAPPING A PANEL OF SOMATIC CELL HYBRIDS WITH STS FROM 13q12.

Mapping with the somatic cell hybrid panel shown in figure 3.22 with STS for 3' Alu 66, 78, 71, 62 and the oncogene, FLT1, sub-divides 13q12 into 6 different regions. Photographs of the results of PCR screening of the hybrid mapping panel with these STS are shown in figure 3.21. It can be seen that all the STS produce single bands in amounts sufficient to be recognised following gel electrophoresis of the reaction products. The most difficult STS to use was that for FLT1, which required 2.0 mM magnesium to yield an interpretable amount of product and even then showed the most variation in the amount of product obtained from experiment to experiment.

The most centromeric of the breakpoints is that of NM which, as can be seen from figure 3.22, is in proximal 13q12. This breakpoint is proximal to that of DG and none of the 5 probes was absent from this hybrid. 3' Alu 66 is present in NM but absent in DGF27C11, confirming the FISH data placing 3'Alu 66 in proximal 13q12. 3'Alu 66 is, therefore, the only molecular marker from 13q12 which maps above the DG breakpoint. Immediately telomeric to the DG breakpoint is that of CF25. None of the 3' Alu probes were able to separate these two breakpoints. However, an STS for the oncogene FLT1 (Polymeropoulos et al, 1991) was present in DGF27C11 but absent in CF25 making it the closest marker below the DG breakpoint.

The breakpoints of CAF55 and JCF9 were indistinguishable by any of the probes analysed and in the interests of legibility, only hybrid CAF55 is shown on the mapping panels. The breakpoint of CF27 mapped to the same position as those of CAF55 and JCF9 with both 3' Alu 78 and 71 present, but 3' Alu 62 was absent in all three of these hybrids. All 5 probes were present in KBF11, making its proximal breakpoint the most telomeric of the hybrids on the panel. Data for KBF11 are not shown on the gel photographs in figure 3.21.

What of 3' Alu 90? It was present in all of the hybrids except NM which does not fit with the rest of the mapping data. FISH data had mapped it to the centromeres of many different human chromosomes and it was likely that its presence in the other hybrids resulted from its homology to the centromeres of other chromosomes and not from chromosome 13 sequences. 3' Alu 90 was, therefore, excluded from further analysis.

There were two major benefits from deriving STS for the 3' Alu probes. The first was the ease of screening the hybrids using PCR. Instead of having to grow the hybrids in bulk and make large amounts of genomic DNA from the cells, data could be obtained from just a few colonies of cells (section 2.4.6c). The reaction itself was also much quicker and could be completed within a few hours rather than the four or five days it would have taken to obtain a result from a hybrid using Southern blotting.

However, the main reason for developing STS for the 3' Alu probes was to be able to increase the accuracy of the results obtained with these probes, as PCR can recognise, and selectively amplify, a single copy target sequence in the presence of a large amount of background DNA. Did this work in practice? Certainly, the increased accuracy of PCR mapping did resolve the remaining confusion over 3' Alu 62 and 27. These two probes had the same restriction enzyme fingerprint and the same sequence, but, by hybridisation, mapped on different sides of the DG breakpoint (figure 3.15). However, when PCR was used, 3' Alu 62 was present in DGF27C11 which is different to the data obtained by hybridisation, and puts both probes on the same side of the DG breakpoint. The difference in hybridisation results must have been due to the presence of contamination or degradation in the probe made for 3' Alu 62.

More importantly, PCR mapping allowed more accurate localisation of probes relative to the DG breakpoint. The 3' Alu probes had been mapped relative to the DG breakpoint by Southern hybridisation and the absence of a recognisable band in DGF27C11 resulted in a given probe being mapped above the DG breakpoint. Table 3.3 shows the presence or absence of the 3' Alu probes in the hybrids using PCR and it is instructive to compare the results in hybrid DGF27C11 with those obtained by hybridisation (table 3.1). Three probes (3' Alu 78, 71 and 62) that were thought to be absent from DGF27C11 by Southern hybridisation (figure 3.15) were found to be present in this hybrid when screened using PCR (figure 3.21). Hence, even though the design and optimisation of oligonucleotide primers was very time consuming, the approach of PCR mapping was justified by the ease of screening using PCR and the increased accuracy of the results. Furthermore, by combining the results obtained using PCR and FISH, we were confident that FLT1 and 3' Alu 66 closely flanked the DG breakpoint. FISH data from St. Jude's Hospital, Memphis, estimated the physical distance between FLT1 and 3' Alu 66 to be around 1 Mb. It was unlikely that further hybrids would allow sub-division of this 1 Mb interval and, therefore, I decided to use FLT1 and 3' Alu 66 to screen YAC libraries in order to try and clone the DG breakpoint junction.

Table 3.3

PROBE	HYBRID					
	KBF11	CAF55	CF25	DGF27C11	NM	CF27
3' Alu 66	+	+	-	-	+	+
3' Alu 78	+	+	+	+	+	+
3' Alu 71	+	+	+	+	+	+
3' Alu 62	+	-	+	+	+	-
FLT1	+	+	-	+	+	+

A table showing the presence or absence of the various STS in the hybrids used in the mapping panel. The markers which most closely flanked the DG breakpoint were 3' Alu 66 above and FLT1 below. FISH data from Dr. D. Shapiro at St. Jude's hospital Memphis , put these two markers about 1 Mb apart, close enough to attempt to clone the breakpoint junction fragment using a walk in yeast artificial chromosomes (YACs).

3.8 YEAST ARTIFICIAL CHROMOSOME ANALYSIS

3.8.1 *Screening YAC libraries by hybridisation*

The two markers which most closely flanked the DG breakpoint were about 1 Mb apart. Attempting to clone a region of this size in cosmid vectors would have required the isolation and characterisation of many cosmid clones. However, the average size of the human DNA insert in YACs was over ten times that of the cosmids. Hence, a relatively small number of YAC clones should be sufficient to clone the region in between FLT1 and 3' Alu 66 and, therefore, the DG breakpoint.

The 4X YAC library available from the ICRF consists of 20,000 clones gridded out on a 20 cm x 20 cm hybridisation membrane. During the course of the study it was available in three different formats, each with different screening procedures. In the first generation of filters, the clones were spotted next to each other, 144 on the x axis and 144 on the y axis of the filter, with no gaps in between them. The screening protocol called for the library to be screened with ^{32}P labelled probe alone. The position of any given hybridisation signal was determined by overlaying the autoradiograph from the hybridised filter onto a 'control' autoradiograph (made from an identical filter which had been probed with pBR322 DNA). As the pYAC4 vector

contains the origin of replication and ampicillin resistance genes from pBR322, the position of every YAC clone on the filter was defined on the control autoradiograph. At the top of both the control autoradiograph and the YAC library filter were a series of identification dots. Lining these dots up allowed orientation of the autoradiograph from the probe of interest with the pBR322 probed control. An example of a typical screening experiment using this protocol is shown in figure 3.23a. In fact, accurate localisation of a potential positive clone proved to be very difficult and, of 4 potential positive clones identified by this method, none contained the sequence with which the filter had been probed.

In the second generation of filters, the YAC clones were gridded in arrays of 6x6, 24 arrays on each axis of the filter. Probing the filter with ^{35}S labelled yeast host DNA (section 2.7.1a) gave faint, but detectable, signal from every YAC. Probing the filter with a ^{32}P labelled probe produced hybridisation signal that stood out against the paler YAC background. A section of an autoradiograph from a second generation YAC filter probed with 3' Alu 66 is shown in figure 3.23b. and, by comparing this figure with 3.21a, it is clear that the clones are much more likely to be correctly localised. However, although the autoradiographs were much easier to read than before, difficulties in accurately localising the hybridisation signals remained. Indeed, of 32 positive clones identified with 5 different probes, only 13 contained the sequence of interest (table 3.4). In order to further improve the ease of localisation of the hybridisation signals, a third generation of filters have been made. In these filters the colonies are still spotted in arrays, but each square array has YACs spotted in duplicate so true positives can be differentiated from random 'spots'. The third generation filters were not available until the end of the thesis and so no data from them is presented.

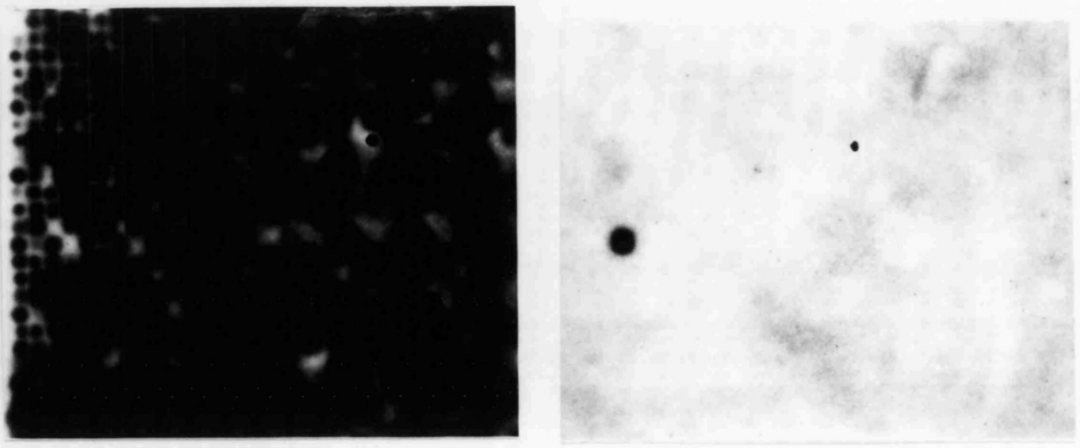


Figure 3.23.a. Screening the 'first generation' ICRF 4X YAC library by Southern hybridisation. The photograph on the left is a portion of the autoradiograph supplied with the library in which the filter has been probed with pBR322 to allow identification of the position of all the YAC clones. The photograph on the right shows a single potential positive clone on a portion of an autoradiograph after probing with 3' Alu 90.

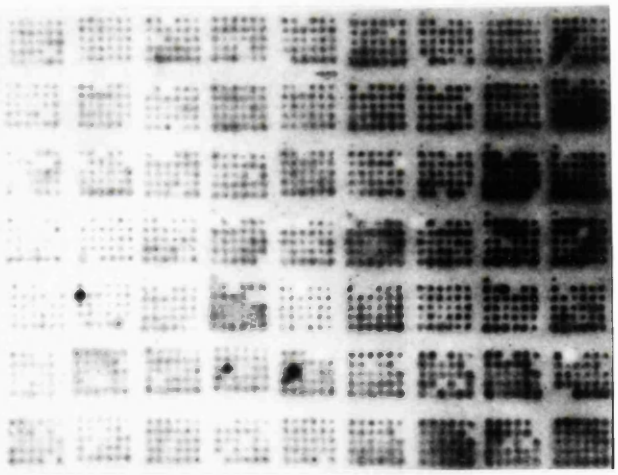


Figure 3.23.b. A photograph of the 'second generation' YAC library. Hybridisation with ^{35}S-labelled yeast host DNA gives a faint, background signal from all the YACs spotted on the filter allowing the 6x6 grid to be visualised. Three potential positive clones can be seen as dark dots on this portion of an autoradiograph of a filter after hybridisation with 3' Alu 66. Only the darkest of the 3 dots proved to be a true positive clone on further analysis.

Table 3.4

Probe	Screening method	Number of YAC clones identified on primary library screen	Number of YAC clones containing sequence of interest	Identification numbers of YACs
3' Alu 66	Hyb	8	1	A05136
	PCR	2	2	A05136 H11111
3' Alu 78	Hyb	7	5	A0483 E0480 H0884 H0984 A11111
3' Alu 90	Hyb	5	3	G0567 E0118 C1273
3' Alu 71	Hyb	2	0	
3' Alu 62	Hyb	6	4	G092 C12132 G02103 D06108
FLT1	PCR	1	1	G12067
cSKI	Hyb	4	0	
	PCR	1	1	E0851

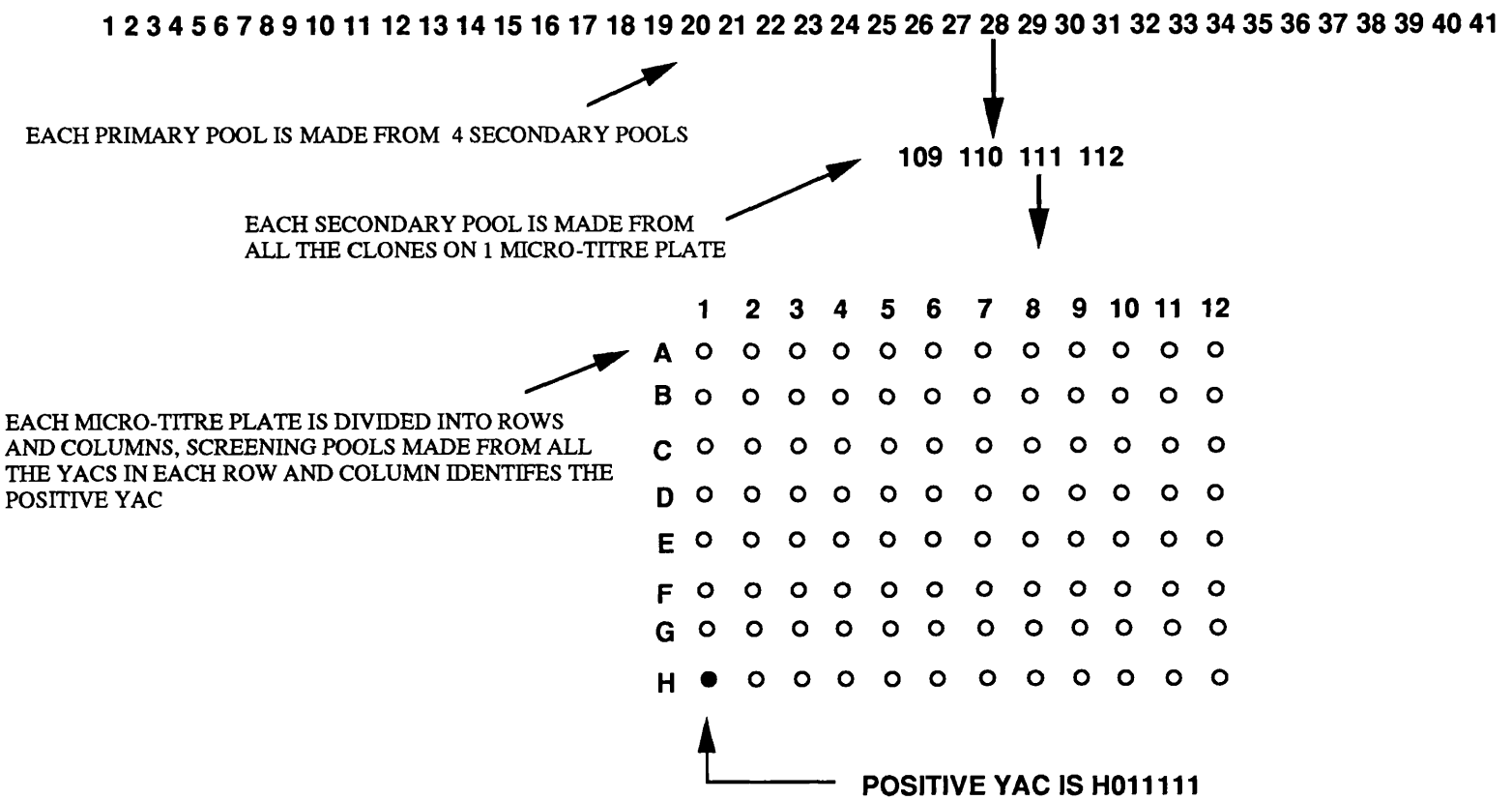
A comparison of the accuracy of screening YAC libraries by PCR and by hybridisation (Hyb). The number of YAC clones identified as being positive on initial screening of the library is compared with the number of these clones which actually contained the marker with which the library had been screened. When the second generation ICRF YAC library filters were screened by hybridisation, only 13 of the 32 YAC clones identified contained the probe of interest. However, all of the 3 clones identified by PCR were found to contain the sequence with which the library was screened, making PCR a much more accurate method of screening YAC libraries. The identification numbers of all the YACs that contained the relevant markers is also noted.

3.8.2. *Screening YAC libraries using PCR*

PCR is a more sensitive method of detecting the presence of a given sequence than hybridisation (section 3.7). Furthermore, some genetic loci, such as FLT1, are only available as STS defined by a pair of oligonucleotide primers. The ICRF 4X library has, therefore, recently been made available as a series of YAC pools which can be screened in a three stage procedure which is shown diagrammatically in figure 3.24. Screening the YAC library using PCR was a much more laborious procedure than Southern hybridisation but, when positive clones are identified, they almost always contain the sequence of interest. In this study, the 3 YACs that were identified with 2 different STS all contained the sequence of interest.

Difficulties did, however, occur due to the very low amounts of template DNA in the primary pools and potential interference of the PCR reaction by impurities in the YAC pools which were made from unpurified, lysed YAC cells. An example of a screening experiment is shown in figure 3.25a. Following gel electrophoresis of the PCR products obtained from each of the 41 primary YAC pools screened, none apparently contained the sequence of interest. However, when this gel was blotted and probed with a ^{32}P labelled PCR product which had been obtained using the same primers, hybridisation identified a band in lane 17. The concentration of DNA in the primary pools is between 30-100 ng/µl and, since each pool is made up of DNA from 400 different yeast clones, as little as 0.15 ng of each clone is present in the 2 µl of the pooled DNA used as the template for the PCR reaction. Furthermore, the majority of the DNA in the pool will be of yeast host origin. The average yeast clone in the ICRF library will contain 600 kb of human DNA and 15 Mb of yeast DNA. It would seem logical, therefore, to increase the amount of template DNA added to the reaction in order to circumvent the problem of poor PCR product yields. However, increasing the amount of template DNA used may lead to interference of the PCR reaction due to impurities in the DNA sample. An example of the 'poisoning' of a PCR reaction following the addition of excess template DNA is shown in figure 3.25.b, which shows the results of using 3' Alu 66 primers to screen a YAC pool known to contain a YAC identified by 3' Alu 66. The addition of 1 µl of template DNA resulted in far more PCR product than adding only 0.1 µl. However, when 3 µl of template DNA was added to a 50 µl PCR reaction, the impurities in the sample resulted in complete inhibition of PCR product formation. From figure 3.25.b it would appear that accurate screening of a PCR based YAC library should, therefore, use no more than 1 µl of each pool. In fact, 2 µl of template was shown to be the optimal amount (data not shown). Hybridisation of radiolabelled PCR product from the positive control to blots made from the gel electrophoresis of the primary pool screening should also be routinely performed.

Figure 3.21



A diagramatic representation of screening the ICRF 4X YAC library by PCR. On the primary screen, pool 28 was positive. This pool was made up of DNA from all the clones on 4 microtitre plates (109-112). A pool made from all the clones on plate 111 was positive. Screening pooled DNA from each row and column on plate 111 gave positives in row H, column 1. The positive YAC was H 01 111.

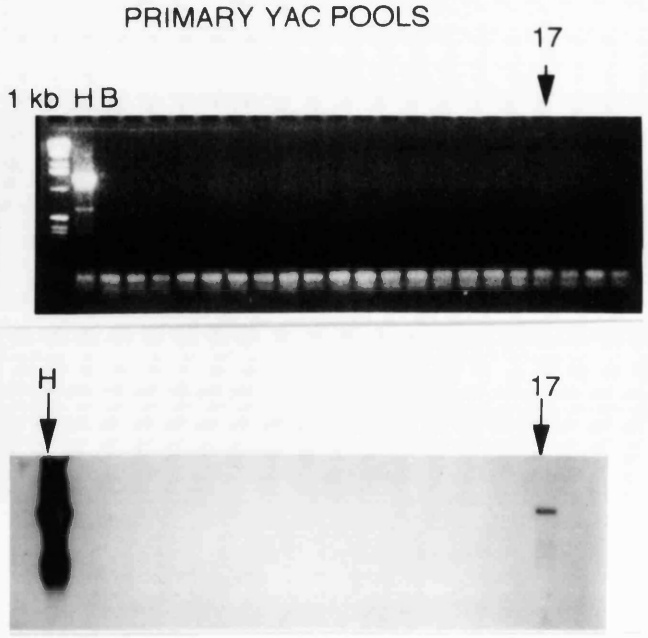


Figure 3.25.a. The upper photograph shows an ethidium bromide stained agarose gel containing the PCR products obtained after screening primary YAC pools with primers for the left hand end of YAC 5136. A strong band is seen in lane H, in which human DNA was used as the template for the PCR reaction. No signal is seen in lane B (the blank) and pool 17 (arrowed) also shows no visible product. When a Southern blot of this filter was probed with radiolabelled product obtained from the human control PCR reaction an appropriately sized band is seen in lane 17 (lower photo).

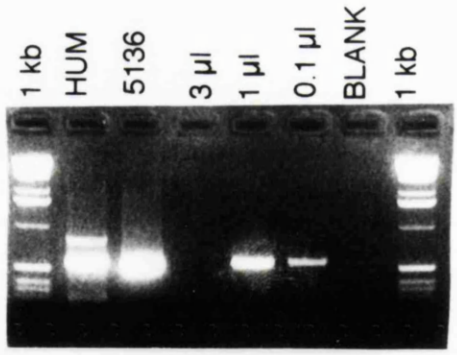


Figure 3.25.b. A photograph showing the inhibitory effects of contaminants in PCR pools on PCR amplification. The templates for the PCR reaction (using primers for 3' Alu 66) were 100 ng of genomic human DNA (HUM), 100 ng of YAC 5136 and different volumes of a YAC pool known to contain YAC 5136. Large amounts of PCR product were obtained from human and 5136. Using 1.0 μl of the YAC pool gave more PCR product than 0.1 μl. However, when 3 μl of the pool was used in the PCR reaction the contaminants in the template resulted in complete inhibition of PCR product formation.

3.8.3. *Analysis of potential positive YAC clones*

Potential positive clones arrive as stab cultures in YAC broth agar and are colony purified by plating out onto uracil deficient medium which should select for those colonies that carry YACs. Due to the difficulties in identifying the exact position of a clone on the gridded YAC library, a given 'positive' clone may not contain the sequence of interest, or only a proportion of the colonies may do so. Before the YACs are analysed further careful assessment of single colonies is necessary to prove the presence of the sequence of interest . Two different methods were used; hybridisation to YAC colonies in situ, and PCR directly from colonies.

In the first method, individual YAC colonies were gridded out on agar plates, and grown until they were approximately 3 mm in diameter. The colonies were then transferred to disks of hybridisation membrane and prepared for hybridisation as described in section 2.9.4. Probing the disks with the sequence that was initially used to identify the positive clones should give strong signal if the colonies are 'true' positives. Figure 3.26a. shows the results of the hybridisation of 3' Alu 62 to three potential positive YACs. Filter I, containing colonies from YAC 12132, is clearly strongly positive. Filter II contains colonies from YAC 092 and produces equivocal signal similar to that from the colonies from YAC 10135 on filter III. From these hybridisation results, it was not possible, therefore, to ascertain whether YACs 092 and 10135 did indeed contain 3'Alu 62.

In the second method, oligonucleotide primers for 3' Alu 62 were used to screen YAC colonies directly. YAC DNA template was added to the PCR reaction mix by touching the surface of the colony with a sterile pipette tip and then using this tip to stir the PCR. In figure 3.26b. it can be seen that strong, unequivocal signal was seen from 4 of the 6 YACs. It can be clearly seen that YAC 12132, which was found to contain 3' Alu 62 by hybridisation, does so by PCR. YAC 092, which gave equivocal results by hybridisation, also contains 3' Alu 62 whereas YAC 10135, which gave hybridisation results which were similar to those of YAC 092, does not. This type of PCR screening was much faster and far more accurate than the filter hybridisation procedure and was used whenever suitable oligonucleotide primers were available. If no primers were available for a molecular marker, YAC colonies identified by hybridisation were regarded as a 'potential' positive clone until the presence of the probe had been proven by Southern hybridisation of enzyme digested YAC clone DNA.

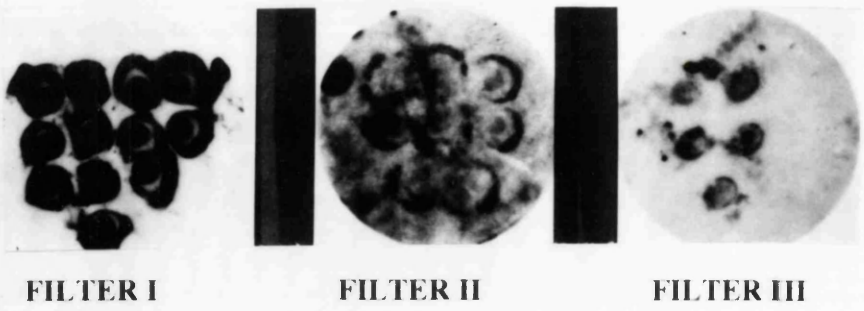


Figure 3.26.a. Individual colonies from three different YAC pools originally identified with 3' Alu 62 in the 4X library were grown on separate plates and transferred to filters. After denaturation and hybridisation with radiolabelled 3' Alu 62 strong signal is seen over all of the colonies on filter I but it is difficult to decide whether the signals on filters II and III represent specific hybridisation even though all were washed to the same stringency.

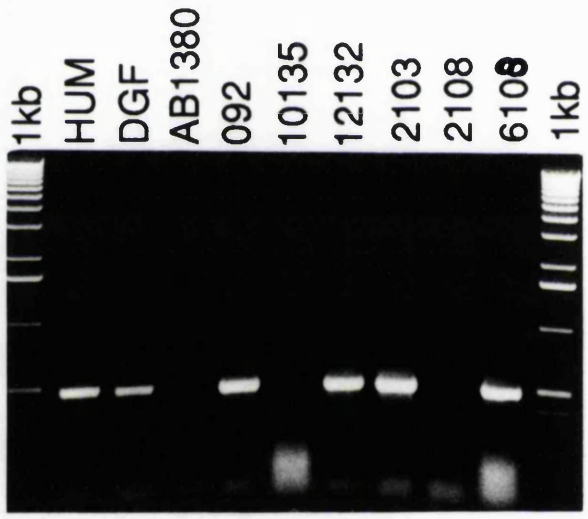


Figure 3.26.b. A photograph of the products of a PCR reaction using primers for 3' Alu 62. A DNA product is seen in human (HUM) and the hybrid DGF27C11 but not in the yeast host, AB1380. In all, 6 YACs were assessed for the presence of 3' Alu 62 using PCR by touching the surface of the YAC colonies with sterile pipette tips which were then used to stir the PCR premix. In 4 of the 6 YACs, a positive signal was obtained. YAC 092, which gave strong signal by hybridisation on filter I in figure 3.26.a, is positive using PCR. YAC 12132 (filter II) is also positive in spite of giving a much lower intensity of hybridisation signal. YAC 10135 (filter III), however, is negative by PCR. Clearly, PCR is a more accurate than hybridisation in determining the presence or absence of specific sequences in colonies containing YACs.

3.8.4 *Estimation of the size of YACs*

Knowing the size of a YAC is important if it is to be useful in physical mapping. Standard agarose gel electrophoresis is only accurate for the size estimation of DNA fragments of around 20 kb in length but, as the average size of YACs in the ICRF library is around 700 kb, other methods of size estimation are required. Pulsed field gel electrophoresis (PFGE) is able to separate fragments of over 1 Mb in length (Schwartz and Cantor, 1984). This technique involves switching the direction of the voltage applied to an agarose gel and running at a controlled temperature for a prolonged length of time.

The conditions used in PFGE have a profound effect on the size separation of the technique. Although the molecular basis of PFGE is not clearly understood, it is known that changes in the voltage and the time the electric field is applied in a particular direction (switch time) are important. Three different sets of conditions were used, all using a 1% agarose gel and 0.5% TBE buffer cooled to 10°C. The shortest run was at 173 volts for 22 hr with a switch time of 70 sec. These conditions provided good separation of fragments up to 700 kb in length. However, since the size of the YACs is not known at the outset, the conditions that were used most frequently were 150 volts for 48 hr with a switch time of 80 sec which allow fragments of up to 900 kb to be distinguished. For very large fragments is was necessary to change the conditions during the PFGE run, and the conditions needed to separate fragments of 1 Mb or more were: 200 volts with a switch time of 70 sec for 15 hr followed by 12 hr with a switch time of 120 sec. A comparison between the first and third sets of conditions noted above is shown in figure 3.27. It is clear that the first set of conditions (shown in the upper photograph) do not allow 5136, a 1.3 Mb YAC, to be distinguished from the band representing the limiting mobility. In the lower photograph, in which the third set of conditions was used, this large YAC is easily seen. Table 3.5 shows that using all three sets of PFGE conditions, it was possible to estimate the sizes of the YACs investigated in this study, even though they ranged in size from 225 kb to 1.3 Mb.

Once the PFGE has been completed the gel is blotted as described in section 2.7.4a. It is not sufficient, however, to simply probe the gel with the molecular marker used to isolate the YAC since this would miss the occasional finding of a yeast clone which contains two YACs. Figure 3.28 illustrates this point. Analysis of a PFGE gel by transillumination alone reveals the presence of a 500 kb band in YAC 6108. Probing the blot made from this gel with total human DNA identified two YACs in this sample, only one of which had been visible by transillumination. Whether the two bands represented the incorporation of two separate YACs into a single host cell or whether two different YAC clones had been inoculated into the same stab culture is uncertain.

Table 3.5

Probe with which YAC was identified	ID number of YAC	Size of YAC (kb)
3'Alu 66	A05136	1300
	H11111	650
3'Alu 78	0483	745
	A11111	850
	H0884	450
	H0984	450
	E0480	745
3'Alu 62	D06108-Larger YAC	500
	-Smaller YAC	225
FLT1	G12067	1200
cSKI	E0851	350

Details of the sizes, as assessed by PFGE, of YACs identified by key probes in the analysis of the DG breakpoint.

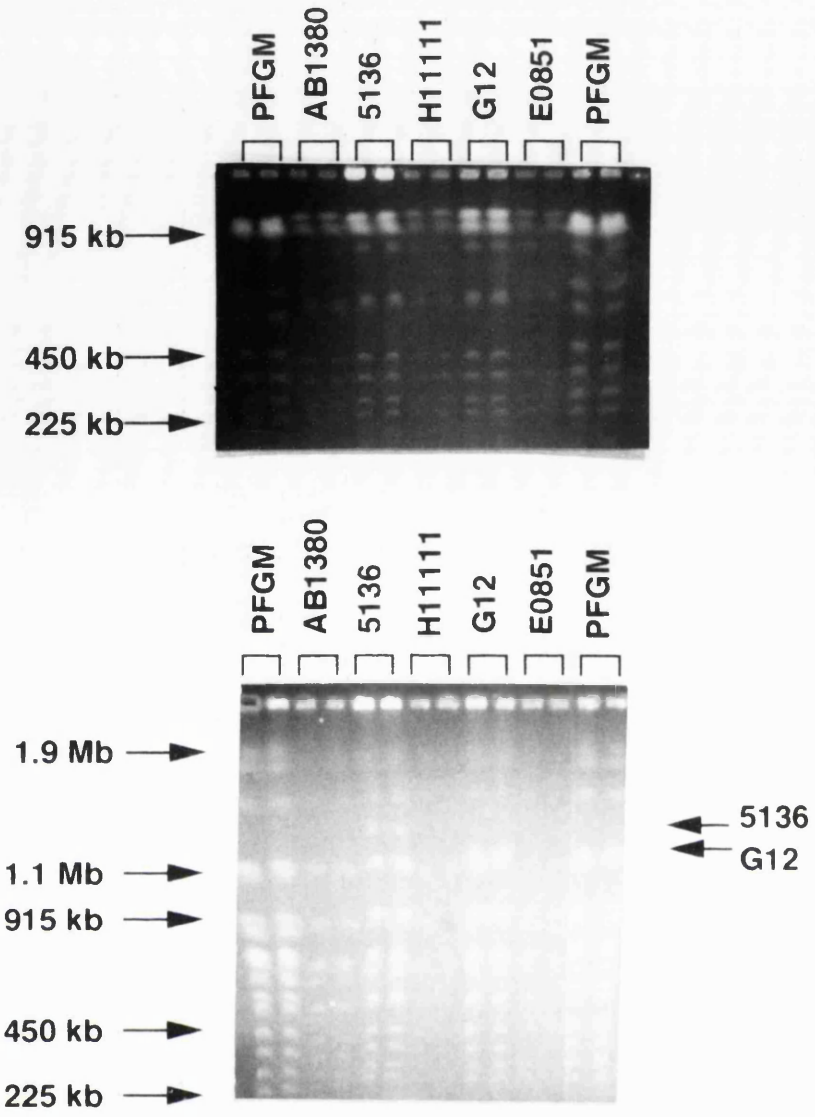


Figure 3.27. The effect of changing the PFGE conditions on the separation of large DNA fragments. The sizes of selected PFGE markers (PFGM) are shown on the left of each photograph. In the upper photograph the conditions used were 173 V, a switch time of 50 sec and a run time of 22 hr. Each YAC sample was loaded in 2 adjacent lanes as this allowed extra bands due to the presence of YACs to be differentiated from random 'spots' on the gel. In the lower photograph, plugs from the same batch were run at 200 V with a switch time of 70 sec for 15 hr then the switch time was changed to 120 sec for 12 hr. The separation of large DNA fragments is much better under the conditions used to produce the lower photograph where a 1.3 Mb YAC from sample 5136 and a 1.2 Mb YAC from sample G12 can clearly be seen (right hand arrows), and much larger bands on the PFGE markers are now visible (left hand arrows).

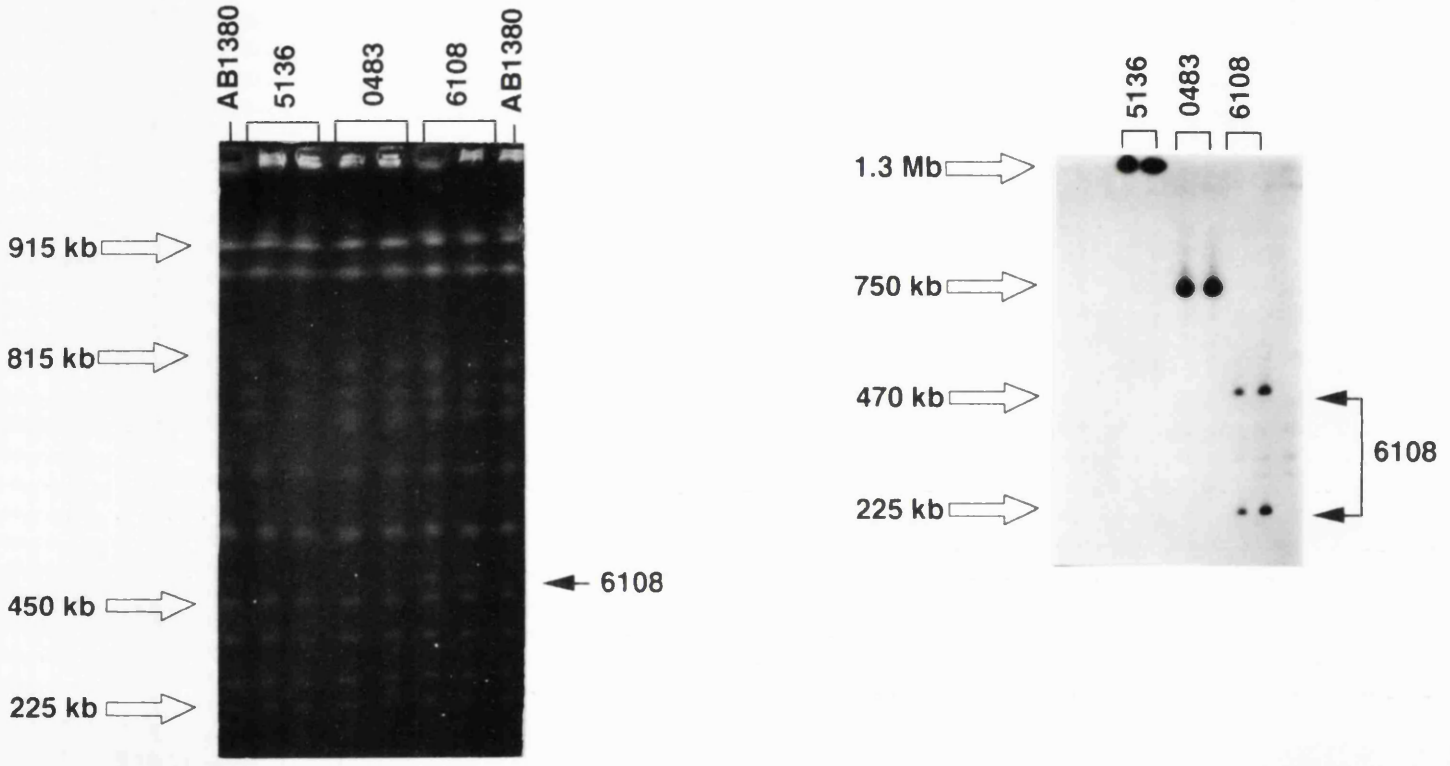


Figure 3.28. The importance of probing YACs separated by PFGE with total human DNA. The left hand photograph shows conventional transillumination of a PFGE gel . The sizes of selected normal chromosomes of AB1380 are given by the open arrows. Only one YAC is seen in lane 6108 (arrowed on the right of the photo). However, when the DNA on this gel was probed with total human DNA, (right hand photo) it is clear that sample 6108 contains two YACs, both of which are arrowed, whereas 5136 and 0483 contain one YAC each. The size of each of the YACs is given by the open arrows.

3.8.5. *Characterising the ends of YAC clones*

The decision to use YACs in a physical mapping strategy was based on the fact that they carry large human inserts. When a YAC is identified by a molecular marker which closely flanks the breakpoint the hope is that it will span the breakpoint junction. There are two ways in which this can be ascertained. The YAC can be used as a probe in FISH experiments against metaphase chromosomes from cell lines carrying the rearrangement. If hybridisation signal is seen on both derivative chromosomes, it is very likely that the YAC crosses the breakpoint. FISH analysis also has the advantage of readily identifying chimeric YACs in which signal will be seen on all the normal chromosomes from which the human insert has been derived. The disadvantages of this method are that if the YAC only overlaps the breakpoint by, say, 5 kb the amount of signal seen on one side of the breakpoint may be insufficient to be recognised. Secondly, if the YAC does not cross the breakpoint the FISH analysis will not provide reagents for YAC walking to isolate one that does. Techniques have been developed which allow the ends of the human insert in a given YAC to be isolated and these ends can then be used to probe somatic cell hybrids containing a single derivative chromosome (DGF27C11) or as the start of a YAC walk. The two methods that I used to capture the ends of the human inserts in YACs both used PCR. The first, vectorette-end rescue, (Riley et al, 1990) depended on the presence of a restriction enzyme site close to the end of the insert. The second method, Alu-vector PCR, would only be successful if an Alu repeat lay within approximately 3 kb of the YAC end.

The vectorette technique is described in section 2.7.5b. It involves digestion of the YAC with a restriction enzyme and the ligation of a vectorette (bubble) sequence to the resulting fragments. In order to generate relatively small DNA fragments, restriction enzymes such as Rsa 1 or Hinf 1 are used. If the enzyme used to digest the YAC DNA is changed, the sequence of the bubble has to be altered to ensure that the ends of the DNA fragments that result from the digestion of the YAC are compatible with the ends of the bubble. Primers specific to either end of the YAC vector are used in a PCR reaction in combination with primers which are identical (**not** homologous) to the bottom strand of the bubble. If an enzyme recognition site exists within 2 - 3 kb of the YAC cloning site, the YAC vector primer will initiate the amplification a single strand of DNA between the vector and the bubble. This single strand will act as a template to which the bubble primer can anneal, allowing production of a double stranded PCR product. In this way, only those fragments which have been primed by the YAC vector sequence will be amplified. Primers have been designed for the left hand end (HYAC-C) and the right hand end (1091) of the pYAC4 vector (Riley et al, 1990) and their positions relative to the cloning site are shown diagrammatically in figure 3.35a. Examples of the results of vectorette end rescue for a series of five YACs is shown in figure 3.29. DNA in agarose plugs was used in the experiment shown

(though I generally found it easier to use solution DNA prepared from the YACs). For each of the five YACs, figure 3.29a illustrates that the DNA in the plugs, which was not degraded, had been completely digested, in this case with Rsa 1. The digested fragments of YAC DNA were ligated to an Rsa 1-specific bubble and aliquots of the resulting product used as the DNA template for a PCR reaction using both HYAC-C and bubble primers. The resulting PCR product will contain a length of YAC vector sequence (232 bp for HYAC-C, 168 bp for 1091) which could be removed by digesting the end rescued fragments with Eco R1 (the pYAC4 vector cloning site). The resulting fragments could then be gel purified as shown in figure 3.29b. Two bands are seen for four of the five YACs; the constant 232 bp band represents vector sequence and the other band, which is of different size in each YAC, is the end of the human insert. Note that clone D06108, which was shown to contain two YACs in figure 3.28, has at least two bands other than the 232 bp vector sequence.

The sizes of the end rescue products derived from selected YACs is shown in table 3.6 and it can be seen that the fragments obtained from the right hand end of YACs (1091 primer) were generally smaller than those from the left hand end; for example the human specific insert in YAC 0483 was only 45 bp in length. The PCR products from the 1091 end generally gave a constant band, 250 bp long, which occasionally would overly the digested human specific fragment (e.g. in YAC 5136) making it impossible to generate a human specific fragment for hybridisation. If either of these situations arose, the YAC DNA was digested with Hinf 1 instead of Rsa 1, in order to try and generate a YAC end that would be useful as a probe. A bubble designed to anneal to Hinf 1 cut fragments was ligated to the digested DNA fragments and PCR performed using a primer specific for the Hinf 1bubble and a primer for either of the ends of the YAC vector. However, on every occasion, the fragment generated by Hinf 1 digestion was a similar size to the Rsa 1 product or even smaller.

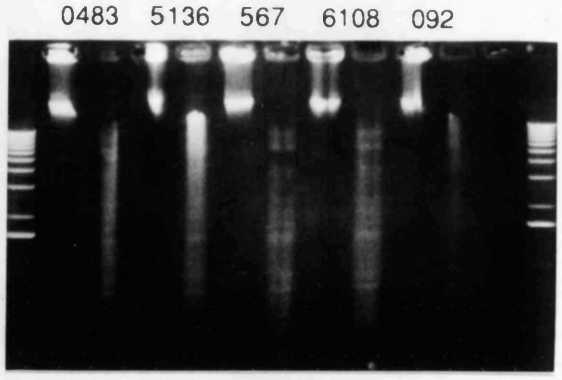


Figure 3.29.a. Preparation of YAC DNA samples for Riley vectorette end rescue. For each of the 5 YACs (identified by their library number) an agarose plug of YAC DNA has been cut into 3 equal pieces. 2 of the 3 pieces have been digested with Rsa 1 and the remaining piece was left undigested. When the undigested DNA from this plug is run alongside one digested with Rsa 1, it is clear that the original YAC plugs were not degraded and that digestion was complete in each case.

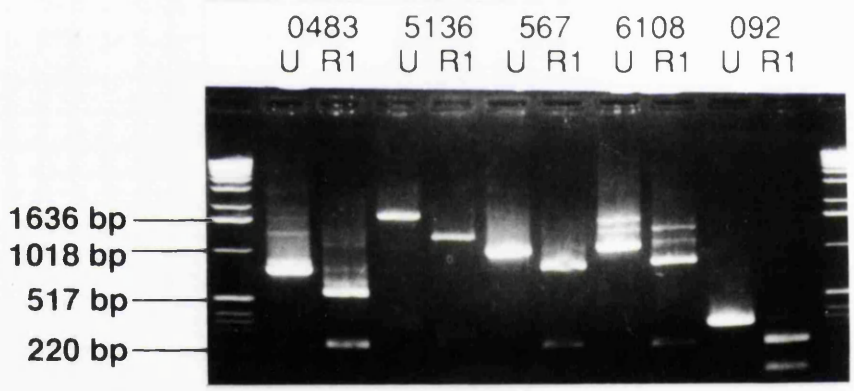


Figure 3.29.b. This photograph shows the results of PCR amplification of vectorette libraries made from the remaining Rsa 1 digested piece of the YAC plugs noted in figure 3.29.a. Primers for the HYAC-C end of the YAC vector were used with primers specific for the 'bubble' sequence to amplify the left hand end of each YAC. Aliquots of the PCR product were digested with Eco R1 to remove the YAC end sequences. The undigested (U) and digested (R1) samples were run side by side, although the size of the human insert varied from YAC to YAC, the vector fragment is constant at 232 bp. The undigested end rescue sample of YAC 6108 has at least two dominant bands related to the presence of two YACs in this clone (figure 3.28).

Table 3.6

Probe with which YAC was identified	YAC ID number	Size of human specific part of end clone (bp) derived following Rsa 1 digestion of YAC DNA	
		HYAC-C(left) end	1091(right) end
3'Alu 66	A05136 H11111	1400 900	280 550
3' Alu 78	A0483	550	45
3'Alu 90	G0567	700	500
3'Alu 62	G092	250	not obtained
FLT1	G12067	1400	220
cSKI	EO851	750	344

Details of the sizes of the human specific sections (i.e. after Eco R1 digestion) of the end rescue products of YACs identified by probes important in the mapping of the DG breakpoint. Of note is the fact that the left hand ends of the YACs are larger than the right, a finding which is difficult to explain.

The second method of generating end clones from YACs did not involve enzyme digestion or ligation of a vectorette but depended on the amplification of the human DNA sequence in between the first Alu repeat in the human insert of the YAC and a primer specific to one or other end of the YAC vector. However, if no Alu repeat was present within about 3 kb of the end of the human inset, no product was produced and so this method was much more dependent on the intrinsic structure of an individual YAC than that of vectorette PCR. Hence, Alu-vector PCR was not the method of choice for isolating YAC ends and was only used on one occasion in an attempt to derive useful probes from the ends of YACs G12067 and A05136 in which vectorette end rescue had failed to produce specific hybridisation probes. The results, shown in figure 3.30, indicate that both YACs contain Alu sequences as a variety of bands are amplified from each YAC using Alu primers alone. When the PCR reaction is performed with both Alu primer and primers to one or other end of the YAC vector, any 'extra' bands represent sequences in between the first Alu sequence in the YAC and the YAC vector primer. No extra bands are seen at either end of YAC A05136, but an extra 400 bp band is seen from the left hand end of YAC G12067.

End rescue often resulted in the production of DNA fragments which behaved as repetitive probes in hybridisation experiments. Indeed, of the 14 DNA fragments rescued by this method only one (the right hand end of YAC G0567) acted as a unique sequence. However, by sequencing these repetitive YAC end clones, unique STS could be generated on the three of the four occasions on which it was attempted. Indeed, as described in sections 3.8.6 and 3.8.7, useful markers for both ends of YAC G12067 and the left hand end of A05136 were successfully derived by this method. The aim of deriving markers from the end clones of YACs was to try to acquire contigs of YACs around the DG breakpoint in the expectation that one or more YACs would contain the region of the DG breakpoint. Hence, YACs above and below the DG breakpoint were analysed and the DNA sequences contained within them compared.

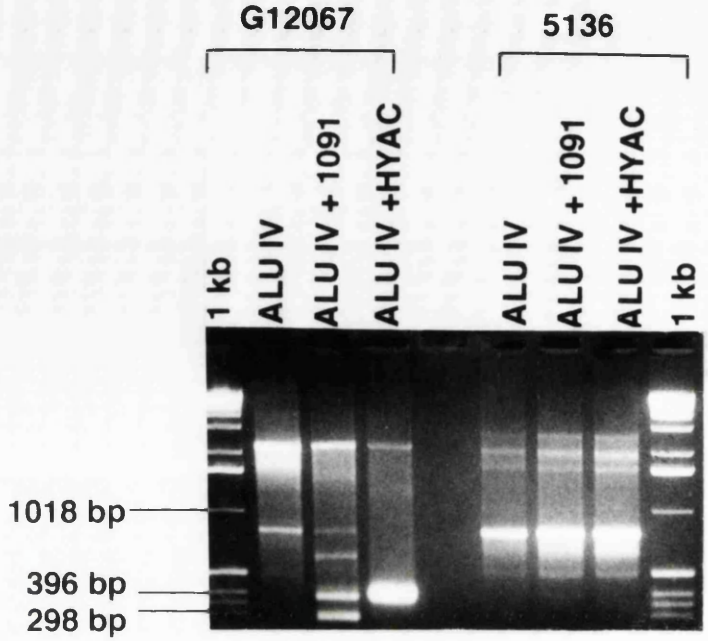


Figure 3.30. Another method of rescuing the human end of a YAC. Two YACs are used, G12067 and A05136. For each, 3 separate PCR reactions are performed and the products of each reaction are run side by side. In the first reaction, the primer used is Alu IV alone. In the second reaction, Alu IV and primers for the right hand end of the YAC vector (1091) are used and in the third PCR reaction Alu IV and primers for the left hand end of the YAC vector (HYAC-C) are employed. Amplification with Alu IV primers alone will amplify any 'inter-Alu' sequences in the same way as that described in figure 3.7. Any extra bands in the lanes with HYAC or 1091 primers represent the DNA fragment in between the vector primer and the first Alu repeat in the human insert of the YAC. No extra bands are seen in YAC 5136 when vector primers are used. In YAC G12067 the HYAC primer yields an extra 400 bp band and the 1091 primer yields 3 bands ranging in size from 700 bp to 290 bp which all probably result from the Alu primer annealing to multiple sites within a repetitive sequence at the end of the YAC.

3.8.6. *Analysis of YACs mapping below the DG breakpoint*

The isolation of YACs from 13q12 proceeded alongside the fine structure mapping of the 13q12 region (section 3.7). YACs were, therefore, initially identified using all of the 3' Alu probes that acted as unique sequences using hybridisation. As more accurate mapping data was obtained, those YACs which were obviously some distance from the DG breakpoint were excluded from further investigation. As shown in figure 3.22, the closest marker below the breakpoint was FLT1. However, before describing the characterisation of the sole YAC identified by FLT1, YACs identified by the 3' Alu probes that mapped below the DG breakpoint will be described.

6 potential positive YAC clones were identified with 3' Alu 62 of which 4 (G092, C12132, G02103 and D06108) were found to contain this marker by PCR (figure 3.26b). One clone, D06108, contained 2 YACs as shown in figure 3.28 and end rescue of this clone led to the amplification of two different sized PCR products (figure 3.29b.). The ends of another YAC, G092, which was also identified by 3' Alu 62, were rescued using the vectorette PCR technique. When these end clones were used as probes, both contained repetitive sequences and could not be used to screen the YAC library in order to generate YAC contigs below the DG breakpoint. 3' Alu 62 was found to be the marker which was furthest from the DG breakpoint and, hence, these YACs were not investigated further. Similarly, before it was realised that 3' Alu 90 mapped to the centromeres of many different chromosomes (table 3.2), YACs isolated with this marker were investigated. 3' Alu 90 identified five apparently positive clones. However, after screening the five clones by PCR only one, G0567, actually contained the STS for this Alu marker. Both ends of YAC G0567 were rescued and the right hand end behaved as a unique probe for Southern hybridisation, the only end clone to do so in this study. Although these YACs may prove useful in other mapping studies, the location of 3' Alu 90 and 62 relative to the DG breakpoint precluded their further use in the analysis of the DG translocation.

Until the STS for FLT1 became available, the closest distal markers to the DG breakpoint were 3' Alu 78 and 71. Although two potential positive YACs were identified with 3' Alu 71 further analysis revealed that neither contained the probe. Six potential positive clones were identified with 3' Alu 78 of which five were found to be true positives when analysed using both hybridisation and PCR. It was important to decide if any of these five YACs were identical in order not to waste time analysing them independently. Bearing in mind the success of comparisons of size and restriction enzyme fingerprints in differentiating between Alu clones (section 3.3.3a), I decided to compare the sizes of the five YACs and the band patterns they produced on Alu fingerprinting.

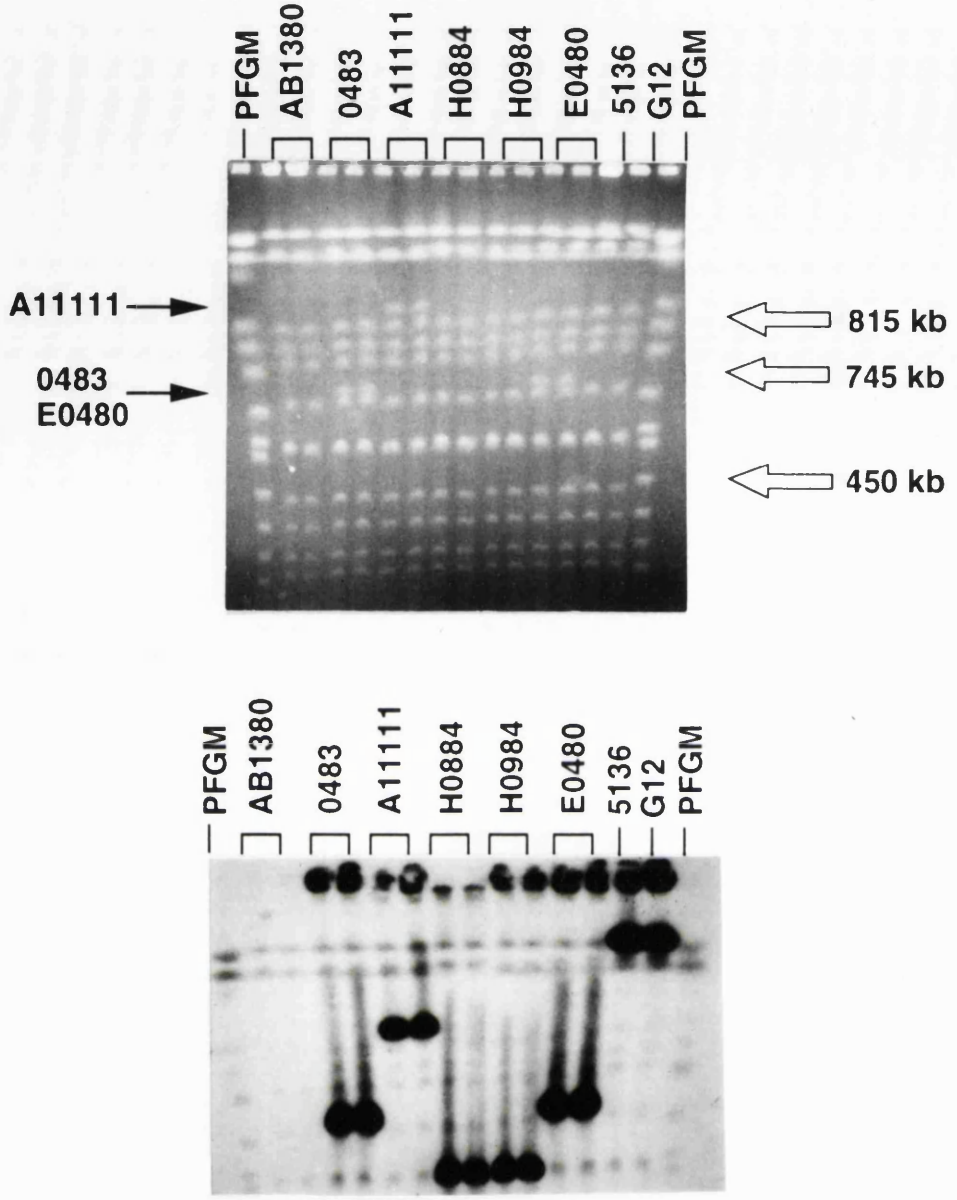


Figure 3.31. PFGE of YACs identified by 3' Alu 78. The upper photograph is of an ethidium bromide stained PFGE analysis of the 5 YACs (A0483, A11111, H0884, H0984, E0480) identified using the 3' Alu 78 probe. These YACs are run in duplicate. 2 other YACs are shown (A05136, G12067) as are the yeast host (AB1380) and yeast markers of known size (PFGM). The sizes of selected yeast markers are indicated by the open arrows. 2 YAC clones carry identical size YACs (A0483, E0480) and the only other visible YAC is found in A11111; all 3 visible YACs are arrowed. The lower photograph shows the results of probing the gel with total human DNA. H0884 and H0984 carry identical size YACs as do A0483 and E0480.

Samples of the five YACs, which had all been lysed in agarose plugs, were separated by PFGE and figure 3.31 shows that there were two pairs of YACs of identical size. A0483 and E0480 were both 745 kb long, and H0884 and H0984 were each 450 kb in length.

Two separate Alu fingerprints were generated for each YAC; one using the ALE 1 primer (derived from the 5' end of the Alu repeat), the other using the ALE 3 primer (derived from the 3' end of the Alu sequence). The use of these primers is described in Cole et al (1991) . ALE1 and ALE3 hybridise far less stringently than the Alu primers that I used to make the Alu PCR libraries. The fingerprints obtained are shown in figure 3.32 and it can be seen that YACs H0884 and H0984, which were the same size on PFGE, have identical fingerprints. Indeed, when the left hand ends of these two YACs were isolated and sequenced they were found to be the same. However, the fingerprint pattern of A0483 closely resembles that of H0884 even though A0483 is a much larger YAC. Conversely, the other pair of YACs which were the same size, A0483 and E0480, produced very different fingerprints. Furthermore, isolation of the left hand end of A0483 showed the fragment to be much smaller than either end of E0480 and to have a completely different sequence (P. Baird pers. comm.). It appears, yet again, that information from different sources needs to be obtained to make accurate assessments of whether clones are identical or not.

In order to determine whether YAC A0483 spanned the DG breakpoint, both ends of YAC A0483 were isolated by vectorette end rescue and used as probes to blots shown in figure 3.33. The left hand (HYAC-C) end, which was 600 bp long, hybridised to a 0.7 kb fragment from A0483 and a weaker 1 kb fragment in human DNA following digestion with Hind III. The human component of the right hand (1091) end of A0483 was only 45 bp long but still produced faint signal with human DNA and produced a strong signal at 22 kb on Hind III digested DNA from A0483. Both end clones from A0483 gave a multi-banded background when used to probe AB1380 DNA and a smear with 3T3 DNA which made their presence in DGF27C11 difficult to determine.

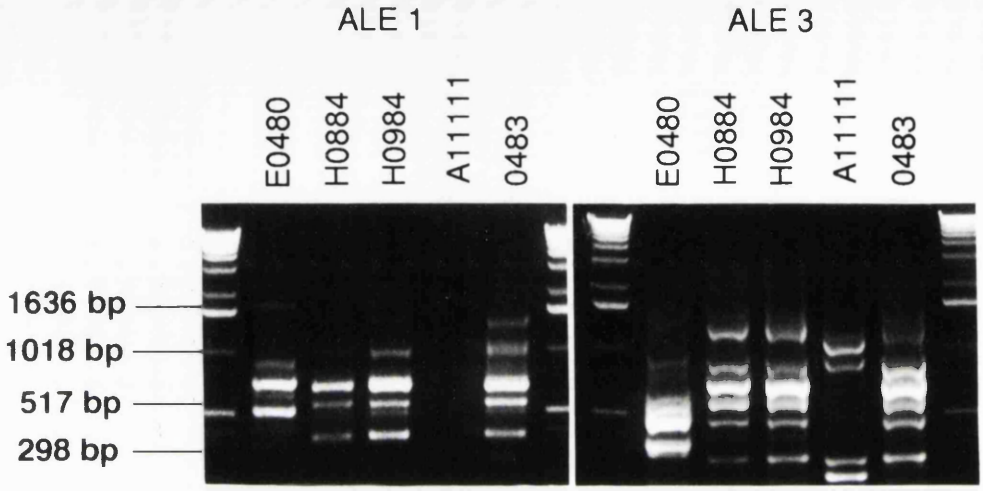


Figure 3.32. An ethidium bromide stained gel showing the results of PCR amplification of 5 YACs using two different Alu primers. DNA from all 5 YACs identified by 3' Alu 78 was amplified with two different Alu primers, ALE 1 and ALE 3. H0884, H0984 and A0483 give identical band patterns using both primers, whereas the fingerprint of E0480, identical in size to A0483, is very different.

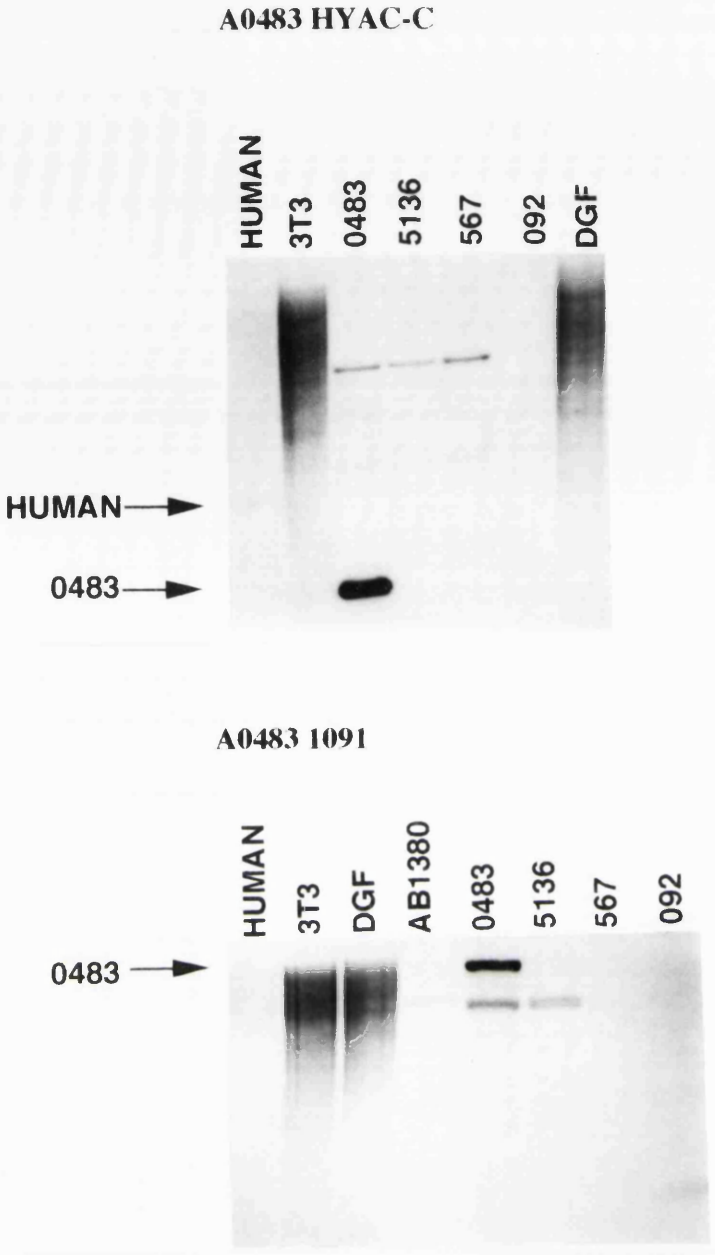


Figure 3.33. When the left (HYAC-C, upper photo) and right (1091, lower photo) end rescued clones from YAC A0483 are used as probes in Southern hybridisation experiments, both give a repetitive smear with mouse 3T3 DNA. This makes it impossible to determine their presence or absence in DGF27C11. The multi-banding pattern seen in the lanes containing the other YACs (A05136, 00567, G092) and the yeast host (AB1380) indicates either cross reactivity with yeast sequences or contaminating yeast DNA in the probe used. However, both YAC ends hybridised strongly to the YAC from which they were derived (arrowed) but the HYAC-C end also recognised a band in genomic human DNA.

The demonstration that the FLT1 oncogene mapped below the DG breakpoint and was closer to it than any of the 3' Alu probes, meant that a major effort was concentrated on isolating a YAC for this locus. The ICRF YAC library was screened using PCR primers for FLT1 and a single YAC, G12067 (G12), was isolated. Using PFGE, G12 was found to be 1.2 Mb long (figure 3.27) and as such was potentially big enough to span the estimated 1 Mb distance in between FLT1 and 3' Alu 66. However, when G12 was used as the DNA template in a PCR reaction using 3' Alu 66 primers, no product was found indicating that this YAC did not contain both flanking markers. It is possible that the locus for 3' Alu 66 is located some distance away from the DG breakpoint and that YAC G12 still spans it. If this were indeed the case, then one end of the human insert in G12 would map above the DG breakpoint (and be absent from DGF27C11) and the other would map below the DG breakpoint (and be present in DGF27C11). Both ends of G12 were, therefore, isolated using vectorette end-rescue. Before Eco R1 digestion, the 1091 end was 550 bp long, the HYAC-C end 1.6 kb in length (figure 3.34a). The human specific fragment was isolated from each of the end-clones by digesting them with Eco R1 and the resulting DNA fragments were used as probes to the blots shown in figure 3.34.b. Both ends were present in control human DNA, DGF27C11 and G12 but only the 1091 end hybridised to DNA from PGME, a hybrid containing normal chromosome 13 as its only human component (figure 3.34b). This strongly suggested that G12 was chimeric, the HYAC-C end being derived from a different chromosome. Neither probe acted as a fully unique sequence in hybridisation experiments and it could be that the stringent washing that was necessary to reduce background smearing had removed recognisable signal from the PGME DNA.

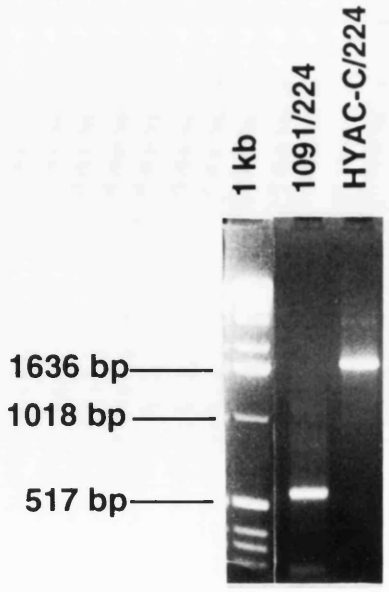


Figure 3.34.a. Results of the vectorette end rescue of YAC G12067.

Using primers to the 1091 end of the YAC and the Riley vectorette bubble (224) a 550 bp fragment was obtained, whereas when primers from the HYAC-C end are used a 1.6 kb product is seen. Neither product has been digested with Eco R1 so both still include YAC vector sequences.

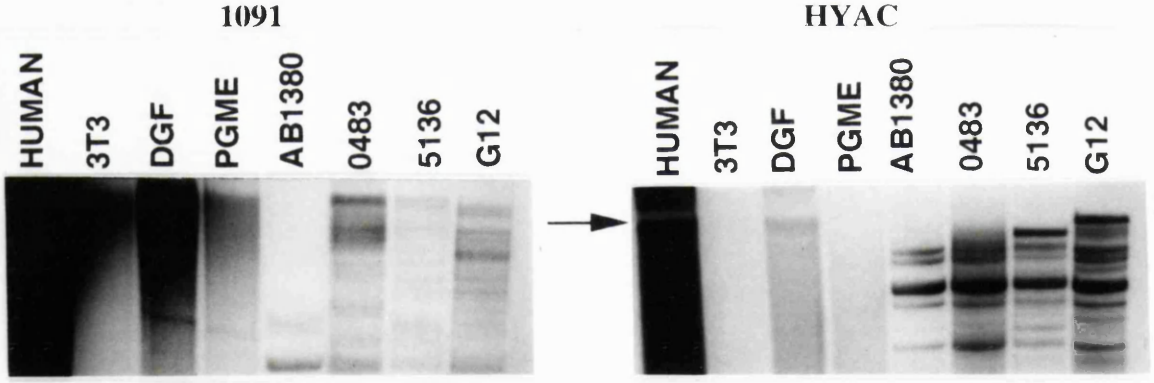


Figure 3.34.b. When the end rescued fragments from both ends of YAC G12 are used as probes for Southern hybridisation, the 1091 end gives a smear in total human and DGF27C11 DNA as well as in the PGME hybrid which contains all of normal chromosome 13. By contrast, although the HYAC-C end shows a 22 kb band in human, DGF27C11 and G12 (arrowed) there is no signal seen in PGME indicating that this probe is not derived from chromosome 13 and that G12 may well be chimeric.

Hybridisation results had been misleading in the past and, as it was essential to determine the map position of the ends of G12, I decided to sequence both ends of G12 and try to develop unique STS. Sequencing was performed as described in section 2.8. In order to allow separation of the two DNA strands, PCR amplification of the vectorette library was performed using non-biotinylated HYAC-C or 1091 and biotinylated 224 primers. Initially, 224 primers were used for the positive strand and HYAC-C or 1091 primers were used for the negative strand. However, the length of YAC vector sequence that lay in between the sequencing primer and the human insert led to little useful human sequence being obtained. To overcome this problem a new set of primers were designed, internal to HYAC-C and 1091, and the primers for the bubble were also moved closer to the cloning site. Thus, primer RA2 was used when sequencing the 1091 end, and primers 32194 and 32427 for the HYAC-C end and the bubble respectively. The sequence of the YAC vector and the vectorette bubble as well as the position of all the primers used for PCR and sequencing are shown in figure 3.35.

The first pair of primers designed from the sequence obtained from the HYAC-C end of G12 yielded a unique STS. The position of the primers relative to human sequence from the HYAC-C end-rescue fragment are shown diagrammatically in figure 3.36a. When these primers were used in PCR reactions a 1400 bp band was produced. As shown in figure 3.36b, this band was present in control human DNA and in DNA from G12 but not in any of the chromosome 13 hybrids, including PGME, confirming that the HYAC-C end of G12 is not derived from chromosome 13. In order to obtain further evidence that G12 was chimeric, G12 DNA was used as a probe for FISH to normal metaphase spreads. The results are shown in figure 3.38 and it is clear that signal is present on both chromosome 13 and chromosome 2. This result suggests that the HYAC-C end of G12 is derived from chromosome 2.

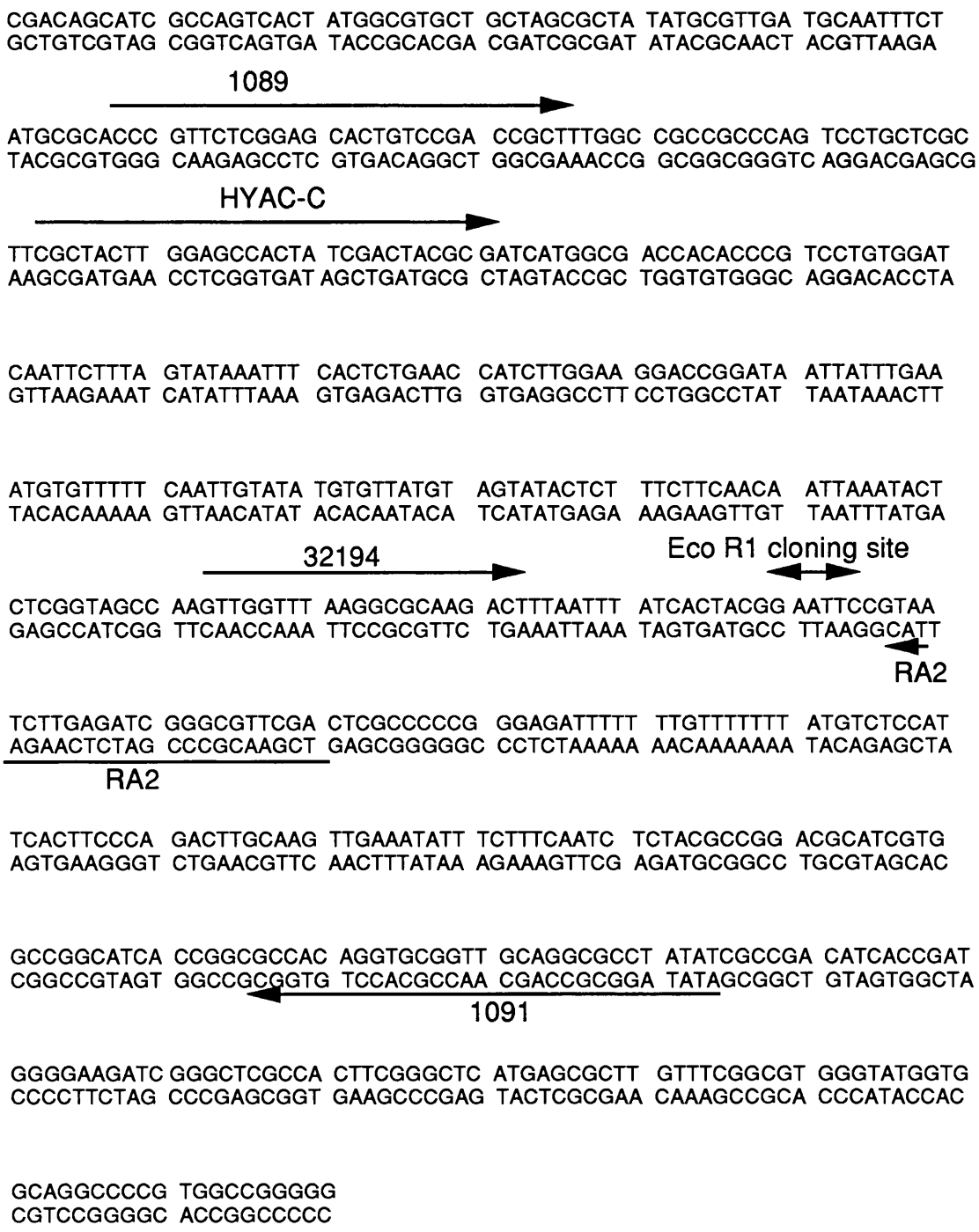


Figure 3.35.a. The DNA sequence of a portion of the pYAC4 vector showing the position of the Eco R1 cloning site and the oligonucleotide primers used in the PCR analysis of YAC end clones.

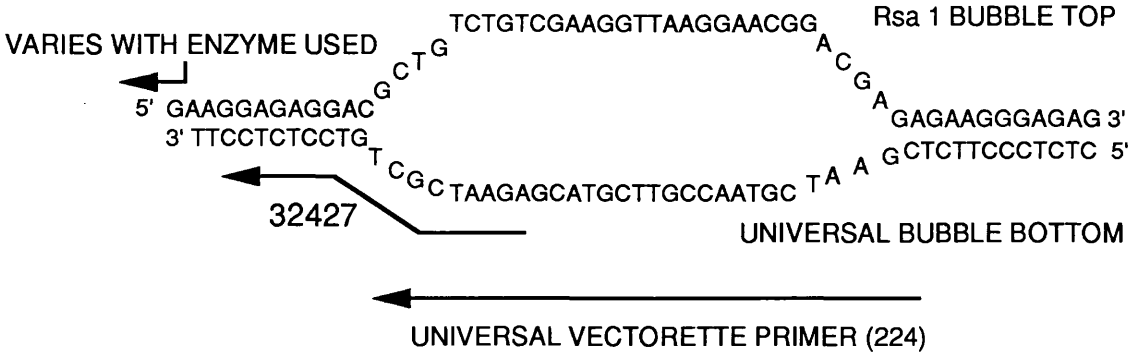
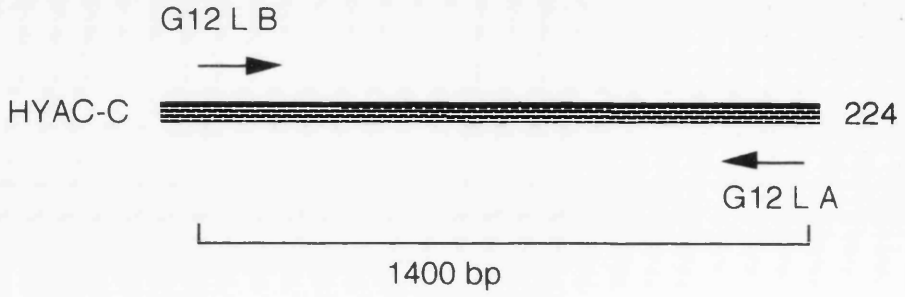


Figure 3.35.b. The DNA sequence of the Rsa 1 Riley vectorette bubble showing the position of the oligonucelotide primers used in the PCR analysis of YAC end clones.

a:



b:

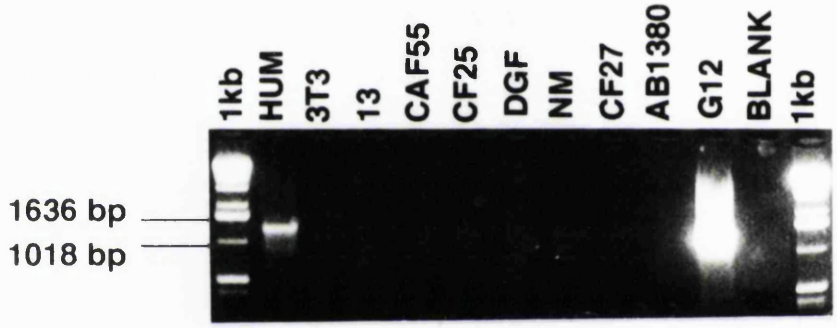
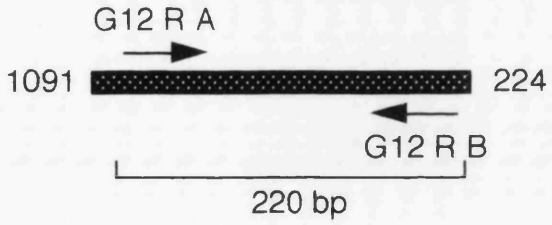


Figure 3.36. The position of the primers generated for the left hand end of YAC G12067 relative to the YAC vector sequence (HYAC-C) and the Riley vectorette bubble sequence (224) are shown in (a). When these primers were used in the PCR screening of a chromosome 13 specific somatic cell hybrid mapping panel (b) they would be expected to amplify a 1400 bp fragment. A 1400 bp PCR product is only seen in lanes containing human DNA (HUM) and DNA from the original YAC (G12). Although very faint bands are seen in CF25 and NM no product is seen with the mono-chromosome hybrid, PGME, implying that this end of the YAC is not located in chromosome 13.

Two different pairs of primers had to be designed from the sequence obtained from the 1091 end of YAC G12 before specific PCR amplification could be produced. The position of the primers relative to the shorter human specific end-rescue fragment are shown diagrammatically in figure 3.37a, and it can be seen that a 220 bp fragment would be produced. Figure 3.37b shows the results of using primers G12R A and G12R B in the PCR screening of a panel of chromosome 13-specific somatic cell hybrids. The presence of a 220 bp band in hybrid CF25, from which FLT1 is absent, suggests that the 1091 end of G12 is more telomeric than FLT1.

A schematic diagram of G12 is shown in figure 3.39 in which the relative positions of the CF25 breakpoint, FLT1, and the junction of the chromosome 13 and chromosome 2 portions of G12 are noted. Creating nested deletions of G12 or using the YAC as a probe against the ICRF chromosome 13 specific cosmid library would enable the isolation of the most centromeric sequence in G12, but these techniques are labour intensive. Section 3.9 describes additional chromosome 13 markers which map even closer to the telomeric end of the DG breakpoint than FLT1 and it is envisaged that they will be used in the future to isolate further YACs which will hopefully cross the DG breakpoint.

a:



b:

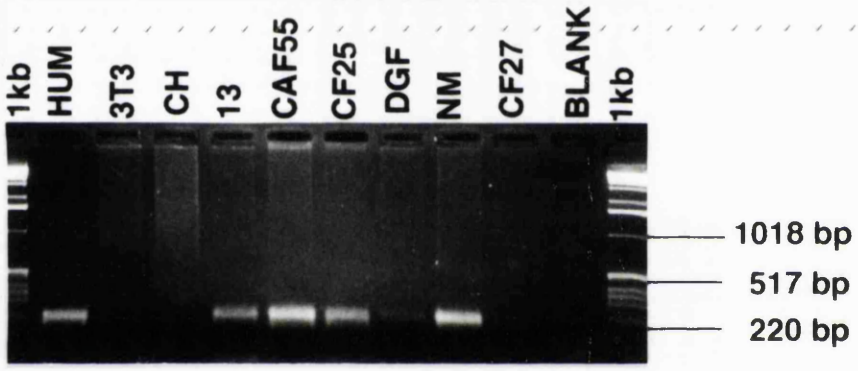


Figure 3.37. The position of the primers generated for the right hand end of YAC G12067 relative to the YAC vector sequence (1091) and the Riley vectorette bubble sequence (224) are shown in (a). When these primers were used in PCR screening they would be expected to amplify a 220 bp band. The correct sized band is seen in DNA from human (HUM), G12 and all 5 of the chromosome 13-specific somatic cell hybrids, though the bands in DGF27C11 and CF27 are faint. The presence of this DNA sequence in hybrid CF25 (from which FLT1 is absent) suggests that this end of YAC G12 is more telomeric than FLT1.

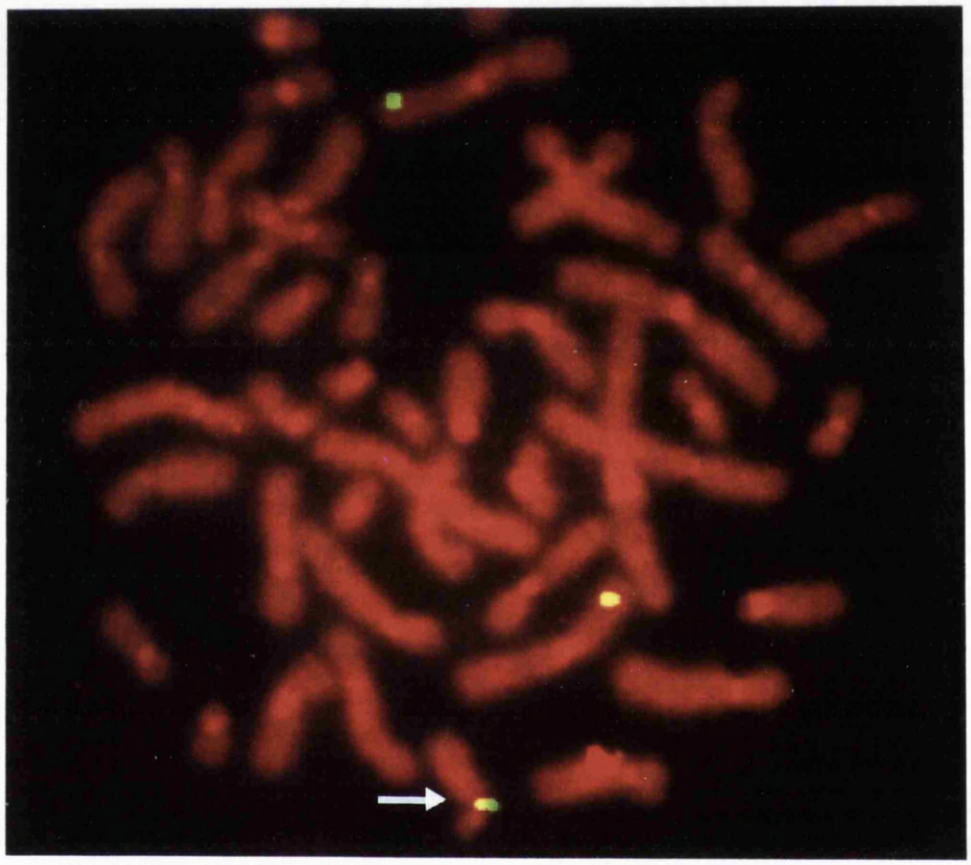


Figure 3.38. The results of using YAC G12067 as a probe for FISH against a normal metaphase spread. The chromosomes are shown in red and the hybridisation signal is seen as yellow/green dots. Three hybridisation signals are seen. The two upper signals are seen on the short arm of chromosome 2 and the lower signal (arrowed) is located on the proximal 13q. The other copy of chromosome 13 is not seen on this metaphase spread.

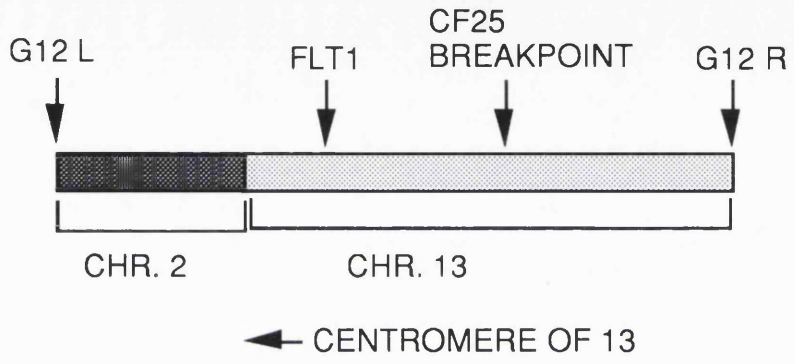


Figure 3.39. Schematic representation of YAC G12067 (not drawn to scale) showing the relative positions of the breakpoint in hybrid CF25, the STS for FLT1 and the end clones. The left hand end (G12L) of the human insert was on chromosome 2. Using two other STS it was possible to orientate this YAC on chromosome 13 relative to the breakpoint in CF25. The right hand end of the YAC (G12R) was shown to be present in CF25 (figure 3.37) but FLT1, the probe with which G12067 was identified, was absent from this hybrid. The most centromeric sequence of the chromosome 13 component of this YAC is in the middle of the YAC next to the chromosome 2 fragment rather than at the end of the YAC where it could be more easily rescued.

3.8.7. *Analysis of YACs above the DG breakpoint*

The only molecular marker from chromosome 13 that was absent from DGF27C11 was 3' Alu 66 and this probe was used to screen the ICRF YAC library by hybridisation. Five potential positive clones were identified of which only one, YAC A05136 (5136), was shown to contain 3' Alu 66 using both PCR and hybridisation. PFGE analysis of 5136 showed that it was 1.3 Mb long (figure 3.27) but, unlike G12, it did not appear to be chimeric by FISH and produced a discreet signal in 13q12 alone (figure 3.40). STS for FLT1, 3' Alu 78 and 71 were not present in YAC 5136, but this did not necessarily mean that this YAC did not span the breakpoint.

In order to ascertain if 5136 did indeed cross the DG breakpoint, the ends of 5136 were rescued by vectorette PCR. It can be seen from figure 3.41 that the 1091 end was 500 bp long but an extra band, 340 bp in length, was also seen on the agarose gel. Eco R1 digestion of this 500 bp product cleaved off the YAC vector sequence but the human specific insert then co-migrated with the contaminating 340 bp band. When the DNA corresponding to this band was isolated and hybridised to a mapping panel (figure 3.42) a faint band at 22 kb was seen in Hind III digested human DNA. In mouse 3T3 DNA a faint background smear was observed which was also seen in DGF27C11. Unexpectedly, a multi-band pattern was seen with AB1380 control DNA and, therefore, with all the YACs. In fact, the YAC that gave the strongest signal was not 5136, from which the end clone was derived, but A0483 a YAC that had been identified by 3' Alu 78. Initially it was thought that the primer pools had possibly been contaminated by A0483 DNA since experiments immediately preceding the end rescue of 5136 had used this YAC. YAC 5136 had been identified with 3' Alu 66 and YAC A0483 with 3' Alu 78, and these two probes were known to map a good distance apart, separated by the breakpoints in DGF27C11 and CF25. It seemed unlikely, therefore, that the 1091 end of 5136 had recognised homologous sequences in A0483. The 1091 end of A0483 gave a very similar hybridisation pattern to that of 5136 (figure 3.33) and, at the time, this was taken as evidence that made the contamination theory more likely. However, since the completion of the thesis the full sequence of the 1091 end of 5136 has been run against the EMBL database and it shows a 40% homology to α-satellite sequences, suggesting it may be anchored in the centromere.

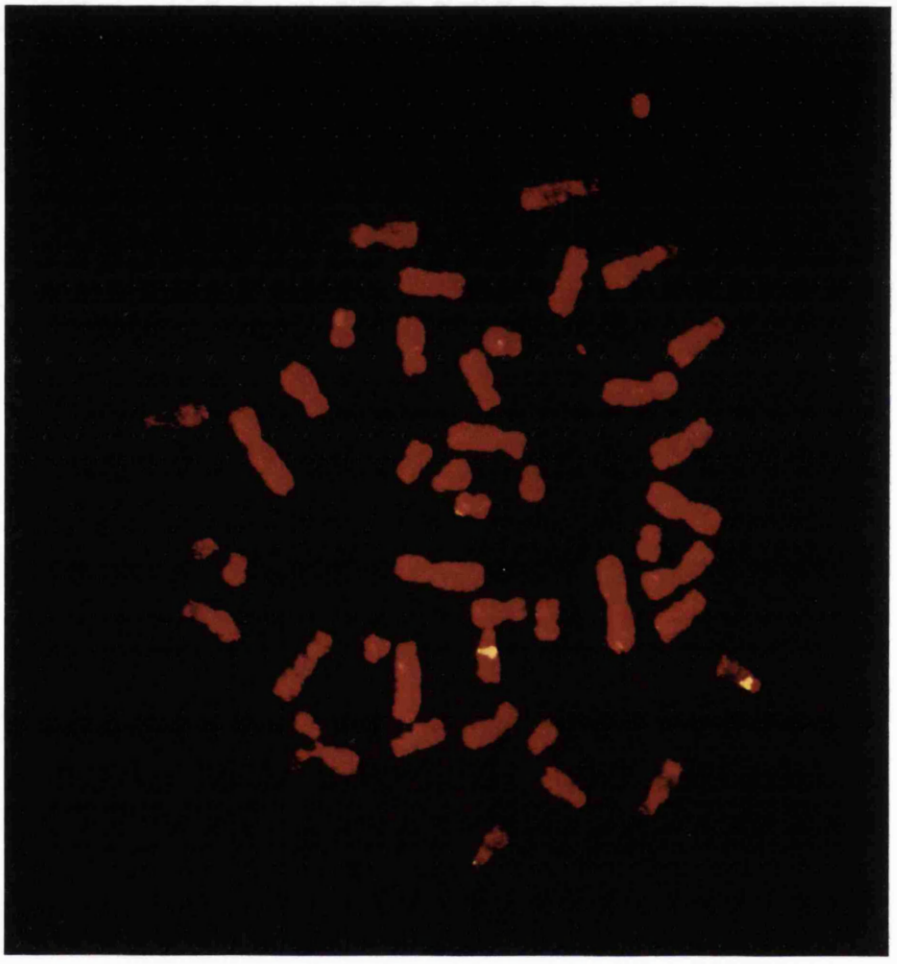


Figure 3.40. The results of using YAC A05136 as a probe for FISH against a normal metaphase spread. The chromosomes are shown in red and the hybridisation signals in yellow/green. Two hybridisation signals are seen, both on proximal 13q, with no cross-hybridisation seen on other chromosomes suggesting that this YAC is not chimeric.

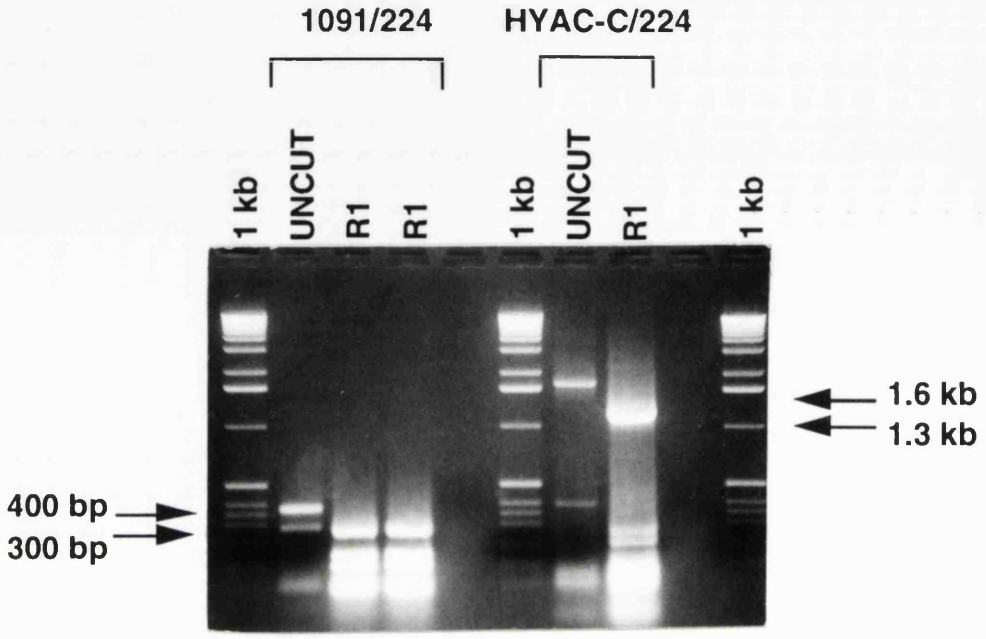


Figure 3.41. Results of the vectorette end rescue of YAC A05136. Using primers to the right hand end (1091) of the YAC vector and to the vectorette bubble (224) a 400 bp fragment is obtained but there is also a contaminating 300 bp band (both arrowed on the left). When the 400 bp end rescue product is digested with Eco R1 the resulting human fragment overlies the contaminating band. The HYAC-C end rescue product is 1.6 kb long and digestion with Eco R1 releases a 1.3 kb human fragment (both arrowed on the right).

In order to try and isolate a unique probe for the 1091 end of 5136 the vectorette end-rescue procedure was repeated using Hinf 1 to digest the YAC DNA. The resulting end rescue fragment was smaller than that obtained with Rsa 1 and, when it was used as a probe, it behaved as a repetitive sequence. A third attempt to generate a unique hybridisation probe from the 1091 end of 5136 involved using Alu-vector PCR as described in section 3.8.5. However, as shown in figure 3.30, no extra bands were generated when Alu and 1091 primers were used in combination compared to when Alu primers were used alone. This suggested that no Alu sequences existed within about 3 kb of the 1091 end of 5136 and meant that a useful probe for this end of the YAC could not be derived by this technique.

The HYAC-C end of 5136 was a 1.6 kb fragment and the human insert released after Eco R1 digestion was 1.3 kb (figure 3.41). When used as a probe, this end recognised 5136 strongly, but only gave a faint band in human DNA. Unfortunately, a smear of strong cross hybridisation was seen with mouse DNA (figure 3.42) making it impossible to determine the presence or absence of this probe in the human/mouse hybrid DGF27C11.

Both the 1091 and HYAC-C ends of 5136 were, therefore, sequenced in order to try to obtain unique STS. The ends were amplified using biotinylated bubble primer (224) and non-biotinylated primers for each of the YAC ends (1091 or HYAC-C) in two separate PCR reactions. The DNA strands were separated and the positive strand was captured on streptavidin coated magnetic beads as described in section 2.8. Each strand was then sequenced using the internal primers shown in figure 3.35. The HYAC-C end was sequenced in both directions; from the YAC vector (HYAC-C) and the bubble (224) ends. Three separate primers were developed and their position relative to each other is shown diagrammatically in figure 3.43.

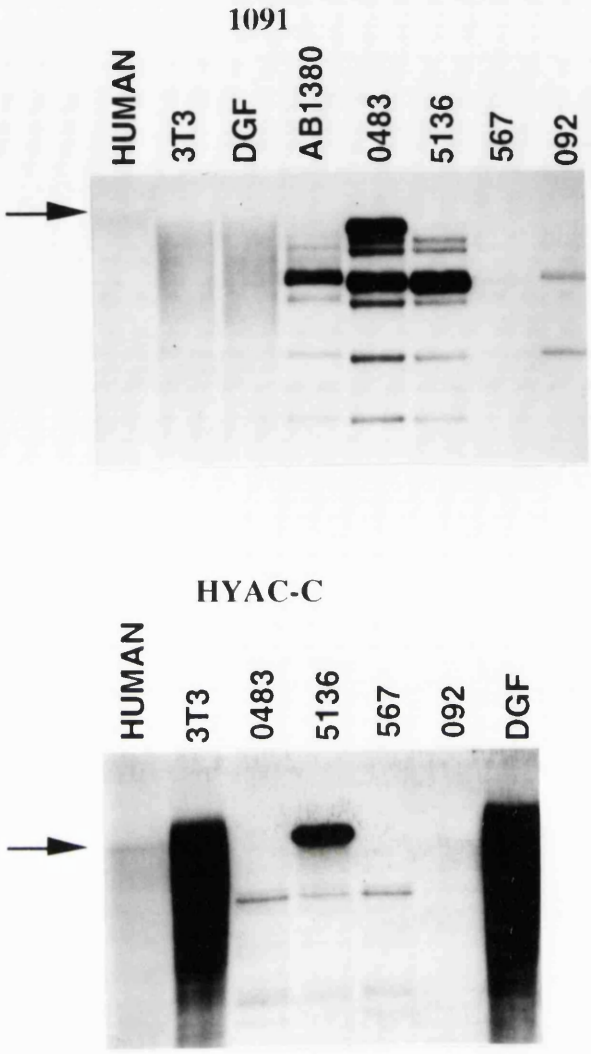


Figure 3.42. Results of using the human specific fragments of the left and right hand ends of YAC A05136 as probes in Southern hybridisation. The right hand(1091) end, shown in the upper photograph, gives a faint band with human DNA (arrowed) but only background smears in 3T3 and DGF27C11. There is a strong multi-band pattern in the yeast host (AB1380) DNA and all of the YACs, the strongest being in YAC A0483. This is possibly due to cross hybridisation of the contaminating band seen in figure 3.40 though a more likely explanation is that the probe is acting as a low grade repeat. The left hand (HYAC-C) end, shown in the lower photograph, gives a faint band with human DNA, a strong band in the YAC from which it was derived and such strong background in 3T3 that the presence or absence of this sequence cannot be determined in the hybrid DGF27C11.

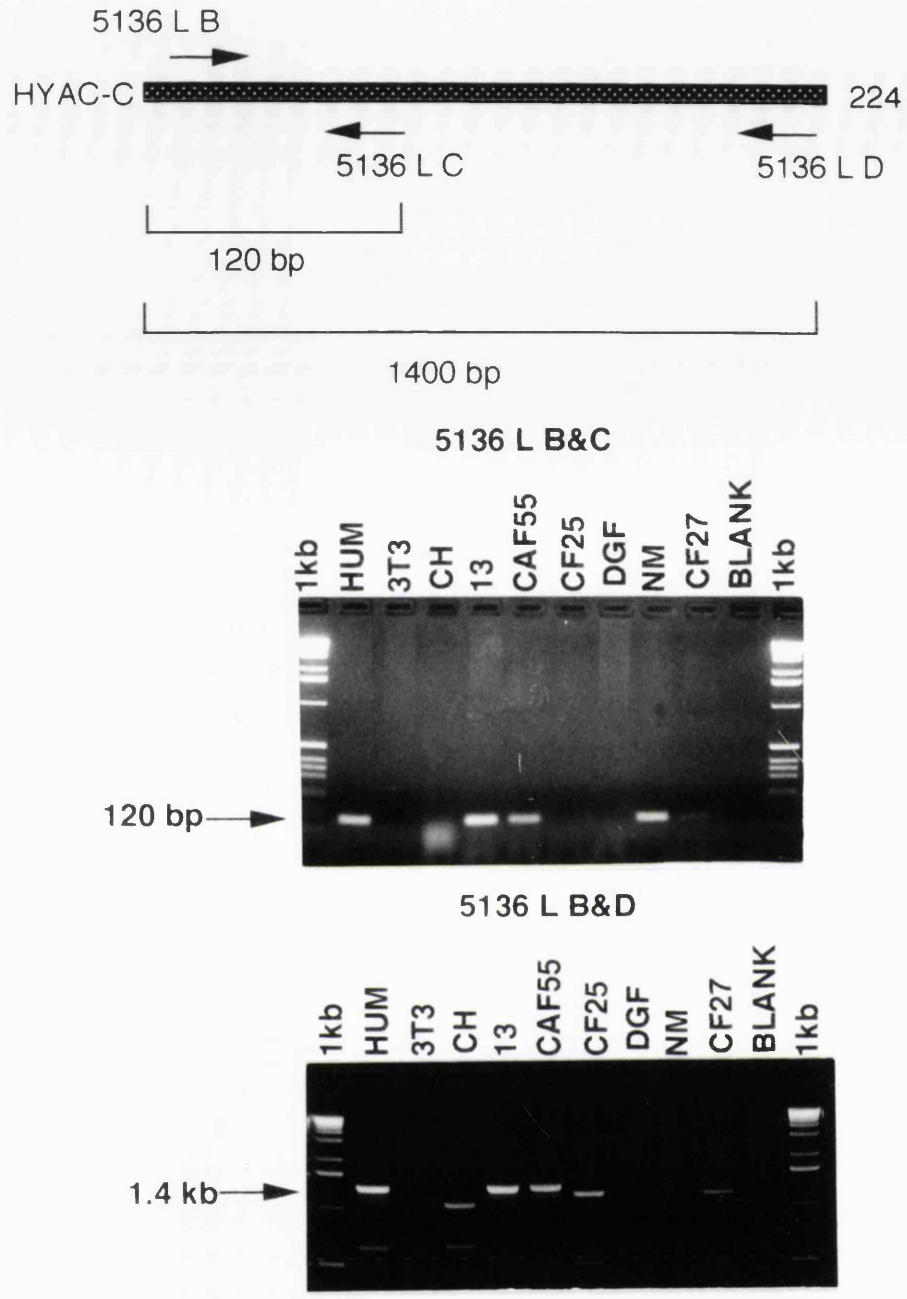


Figure 3.43. A diagram (above) of the left hand (HYAC-C) end of YAC A05136 showing the position of the primers used in PCR analysis. The two photographs show the results obtained when these primers were used to screen a chromosome 13-specific somatic cell hybrid mapping panel. Primers B and C generated a 120 bp product (arrowed, upper photo) whereas using primers B and D a 1.4 kb fragment was amplified (arrowed, lower photo). Primers B and D gave a background multi-band pattern in all lanes. Both sets of primers amplified products that mapped to the same position as 3' Alu 66: present in NM but absent in DGF27C11.

Primer 5136 LB (left, B) was derived from the end of the human fragment in 5136 that was closest to the cloning site of the pYAC4 vector. Primer 5136 L.C. (left, C) was derived from a sequence 120 bp into the human fragment and primer 5136 L.D. (left, D) from section of the end clone next to the bubble. Using primers 5136 L B and C a 120 bp product, corresponding to the end of the HYAC-C end-rescue product closest to the pYAC4 cloning site, was obtained using PCR. When mapped on a panel of chromosome 13 specific somatic cell hybrids the 120 bp band was found to be absent from DGF27C11 but present in NM as shown in figure 3.43. Hence, this STS for the HYAC-C end-rescue product did not cross the breakpoint and mapped to the same position on the panel as 3' Alu 66, the probe with which 5136 had been identified. As shown in the lower photograph in figure 3.43, a PCR product corresponding to the full length of the HYAC-C human insert of 5136 could be amplified using primers 5136 L B and D. This STS mapped to the same position as the smaller STS (amplified using primers 5136 LB and C) although the results were made more difficult to interpret by the presence of smaller, background bands. Neither of these STS were present in YAC G12 proving that the most centromeric sequence of G12 was not present in the left hand end of 5136 (data not shown). Although the STS for the HYAC-C end of 5136 had not crossed the DG breakpoint they did enable a YAC walk to be undertaken, although there was no way of knowing if the walk was proceeding towards or away from the breakpoint. Primers 5136 B and C were used to screen the ICRF YAC library and five potential positive clones were identified. These clones had not yet arrived at the end of my thesis, but will be analysed in the same way as the other YACs described in this thesis.

The 1091 (right hand) end of 5136 contained a human insert of only 114 bp, all of which could be sequenced from one direction. The full length sequence is shown in figure 3.44. This figure also notes the position of the 6 primers that were designed to try and generate a unique STS from this end-rescue fragment. No combination of these six primers led to the amplification of a unique PCR product. This was probably due to the fact that the 1091 end of 5136 was a repetitive fragment, as had been suggested by the hybridisation data.

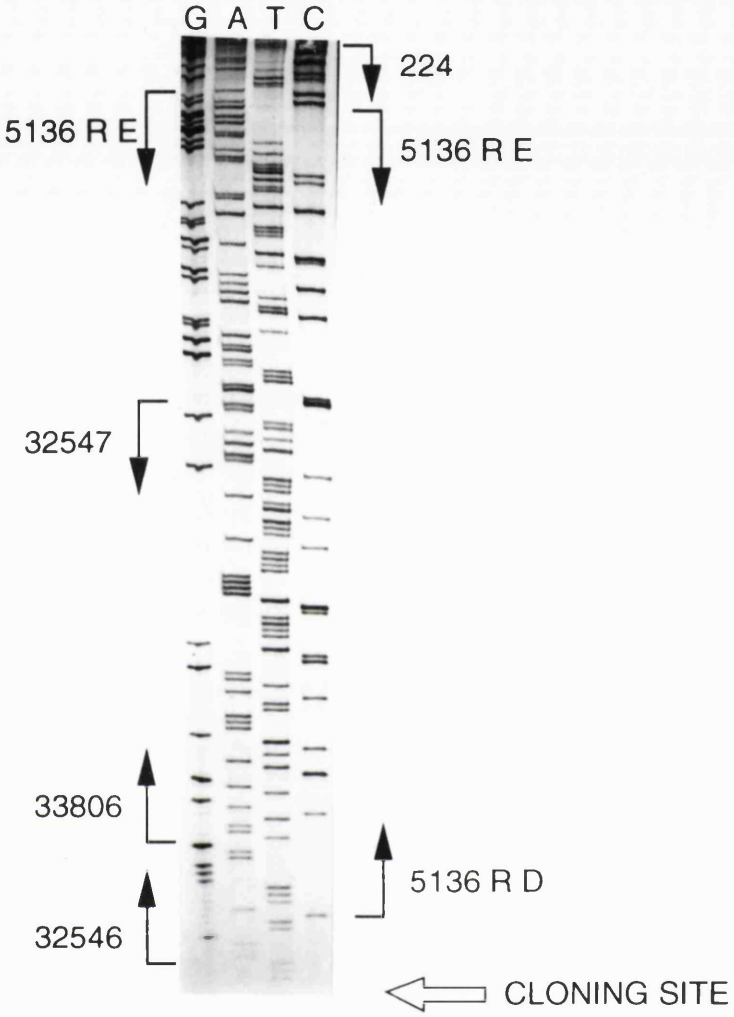


Figure 3.44. The full length DNA sequence of the human specific insert of the right hand (1091) end of YAC 5136. The positions of all the primers that were developed to try and obtain a specific STS are shown. No combination of primers produced a unique PCR product.

The fact that the right hand end of 5136 was a repetitive sequence made it impossible to determine whether the end of this YAC crossed the DG breakpoint, or if it was present in YAC G12. If the right hand end of 5136 overlapped with the proximal 'internal' fragment of YAC G12 then one or other YAC must have crossed the DG breakpoint. In an attempt to solve this, Alu fingerprints were obtained for both YACs using various Alu primers. If the fingerprints obtained from two YACs using PCR shared a number of bands it would imply that they shared a common DNA sequence. As seen in figure 3.32 (which shows the similarity between the fingerprints obtained from various YACs using two different Alu primers) the use of more than one Alu primer increases the certainty that two YACs really do contain identical sequences. Hence, fingerprints were obtained from the two YACs using primers 3' Alu, BK33, Alu IV alone and a combination of AluIV and BK33 primers. The banding patterns obtained (data not shown) were very different and there were no bands of identical size in both YACs which would warrant sub cloning and further analysis.

YAC 5136 had been identified by screening the ICRF YAC library by hybridisation. Screening the library using PCR identified a second YAC, H11111. PCR analysis showed that the left hand end of 5136 was not present in H11111 and this new YAC is either completely contained within 5136 or overlaps its repetitive right hand end. Either way, rescuing the ends of H11111 may allow YAC walking in the opposite direction to that made possible by the STS for the left hand end of 5136. A 900 bp human fragment was isolated from the right hand (1091) end of H11111 and a 550 bp human sequence from the left (HYAC-C). Both gave repetitive signals when used as probes for Southern hybridisation and their presence in DGF27C11 and YAC 5136 could not be assessed. Both are being sequenced in order to try and develop unique STS for the ends.

A different approach is also being taken using 5136, G12 and H11111. All these YACs are to be used as probes in FISH experiments using the lymphoblastoid cell line from DG . If any of these YACs crosses the breakpoint by 5 kb or more, signal should be seen on both derivative chromosomes as well as the normal chromosome 13.

3.9. ANALYSIS OF FURTHER MARKERS FROM 13q12

During the course of this study, four additional molecular markers from 13q12 have recently become available. None of these have, as yet, been used to screen the ICRF YAC library. Two probes (c12.2 and FLT3) were provided by Dr. D. Shapiro of St. Jude's Hospital, Memphis and the primer sequences of the other two (D13S115 and D13S120) have just been published (Bowcock et al, 1992). Figure 3.45 shows the results of screening the chromosome 13 specific mapping panel with STS for D13S115 and D13S120. D13S115, also known as MS34, is present in NM but absent in DGF27C11 and, therefore, maps to the same region as 3' Alu 66. D13S120, also called 1353, is present in all the hybrids on the chromosome 13 specific mapping panel and maps to the same position as 3' Alu 78 and 71. Neither is present in YACs 5136 or G12. The genetic distance between D13S115 and D13S120 is an estimated 15 cM which corresponds to a physical distance of about 15 Mb. It is unlikely, therefore, that D13S120 will be of use in mapping the DG breakpoint.

This has to be contrasted with data from the other probes that have been recently acquired. FLT3, another member of the FLT family of oncogenes, also maps to 13q12, and is thought to be 350 kb more centromeric than FLT1 ((Rosnet et al, 1993) and Dr. Shapiro pers. comm.). Both FLT1 and FLT3 were originally isolated from rodent cells and the mouse and human sequences are homologous for much of their lengths. From Dr. Shapiro we had access to the FLT3 cDNA (pTH3L) and a FLT3 cosmid (c172B2). The FLT3 cDNA was used as a probe against blots of human, 3T3 and DGF27C11 DNA. The signal strength was very low in the human lane and in the lanes from both mouse and DGF27C11 a stronger signal was seen, which was taken to be rodent in origin (data not shown). The sequence of the FLT3 cDNA has been published (Rosnet et al, 1991) and I developed primers from the 3' untranslated region of the sequence (which would be more likely to approximate the genomic DNA than sequence from the coding region) which recognised a 300 bp STS. This STS was present in mouse and hamster DNA and hence gave signal in all the rodent/human hybrids. However, it was absent in the yeast host, AB1380, and YACs 5136 and G12 (figure 3.46a). Screening the ICRF YAC library with this STS should allow a FLT3 YAC to be identified which will be closer to the breakpoint than YAC G12 (which was identified by FLT1). However, data from the second marker obtained from Dr. Shapiro makes it is unlikely that the oncogene FLT3 itself has been disrupted by the DG breakpoint. Using high definition FISH analysis, he identified a cosmid, c12.2, which mapped 450 kb centromeric to FLT3. When this cosmid was used as a probe (with Cot 1 competition) against a blot of Eco R1 digested DNA strong signal was seen both in human and DGF27C11. This blot is shown in figure 3.46b. In order to obtain a single copy probe

from c.12.2, the cosmid was digested with Eco R1, the products of the digest separated by gel electrophoresis and blotted. When the blot was probed with total human DNA, only one band that was visible on UV transillumination of the ethidium bromide stained gel did not hybridise to sheared human DNA. When DNA from this was used as a probe it acted as a low grade repeat but did produce signal from DGF27C11. It is likely, therefore, that c12.2 maps below the DG breakpoint and constitutes the closest telomeric marker. Figure 3.47 shows the order of molecular markers relative to each other and to the breakpoints of hybrids NM, DGF27C11 and CF25. Genetic and physical distances are noted where they are known. FISH data from St. Jude's puts c12.2 about 250 kb away from 3' Alu 66 making it very likely that YACs can be obtained which contain both flanking markers and, therefore, the breakpoint junction fragment from DG.

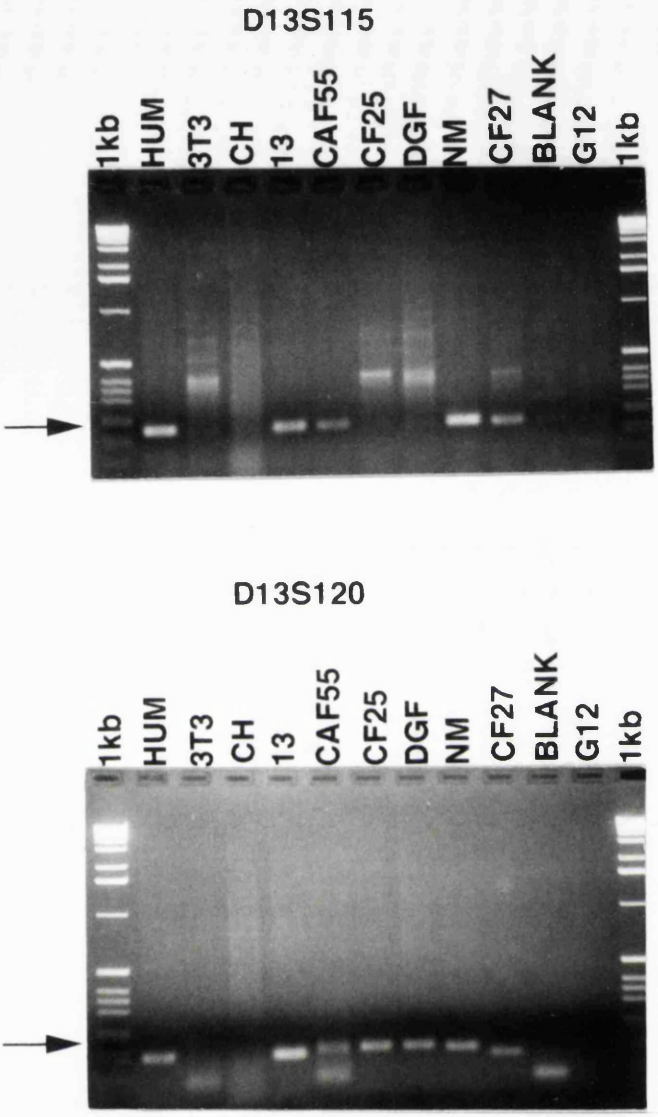


Figure 3.45. The results of screening a chromosome 13-specific somatic cell hybrid mapping panel and YAC G12067 with STS for D13S115 and D13S120 . D13S115 (upper photograph) is present in NM (arrowed) but absent in CF25 and DGF27C11 and so maps to the same region as 3' Alu 66. D13S120 (lower photograph) is present in all the hybrids (band arrowed) and , therefore, maps to same region as 3' Alu 78 and 71.

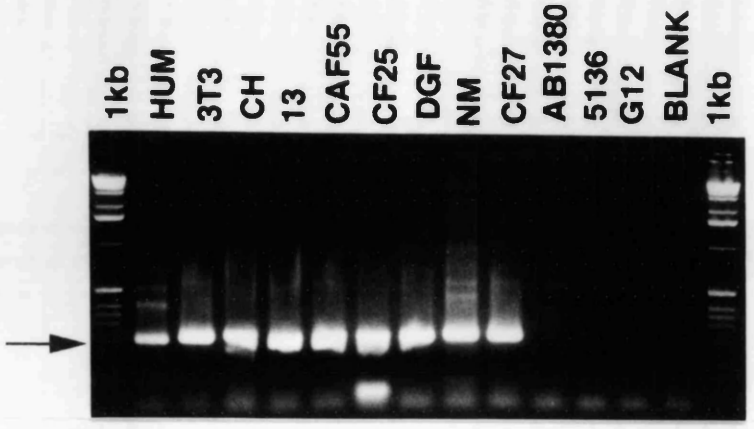


Figure 3.46.a. Screening a series of somatic cell hybrids and YACs with an STS for the 3' untranslated region of FLT3. A 237 bp band (arrowed) is present in mouse (3T3) and hamster (CH) DNA and, therefore, is seen in all the rodent/human hybrids. No amplification is seen in the yeast host (AB1380) or YACs A05136 and G12067.

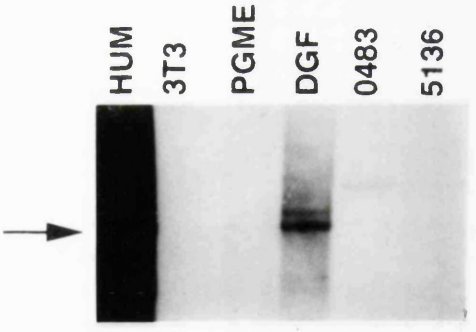


Figure 3.46.b. Results of using the cosmid, c12.2 (with Cot 1 DNA competition) as a probe for Southern hybridisation against a blot of Rsa 1 digested DNA The DNA samples were human (HUM), mouse (3T3), a somatic cell hybrid contaning all of chromosome 13 (PGME), DGF27C11 and two YACs, A0483 and A05136. Some repetitive smearing is seen and the human lane is 'burnt out' but an band 8 kb in size (arrowed) is seen in human, DGF27C11 and faintly in PGME, but not in mouse or the two YACs.

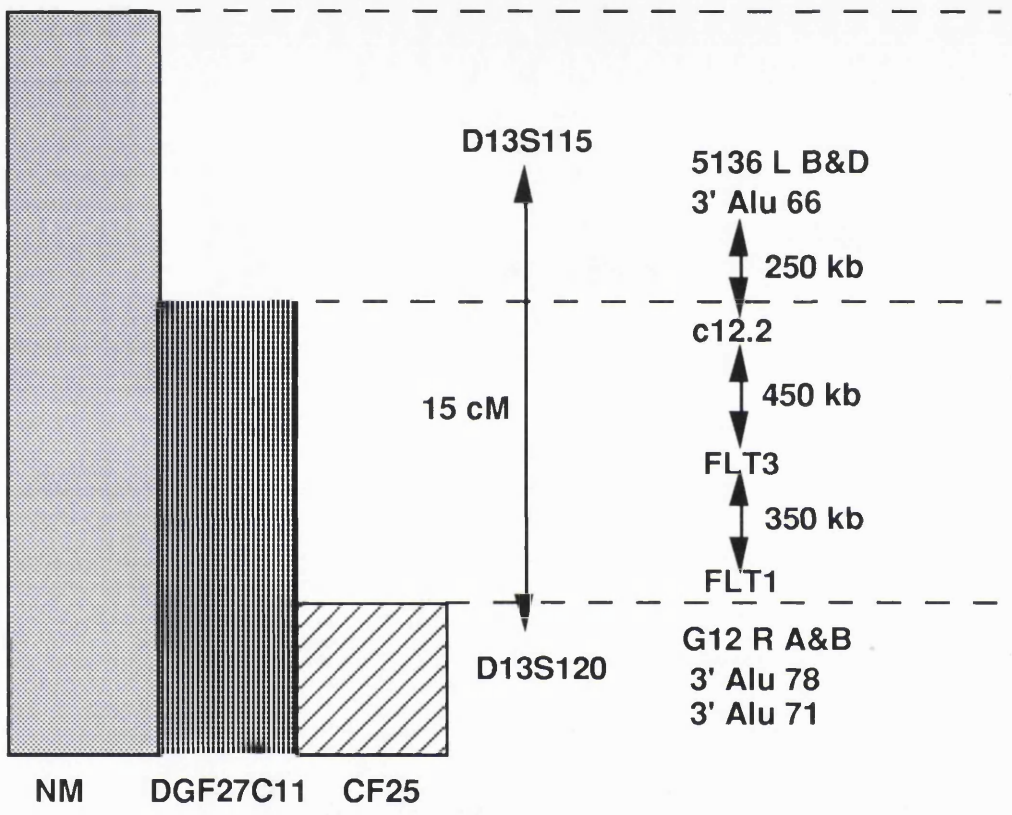


Figure 3.47. A diagram showing the relative position of the molecular markers that most closely flank the DG breakpoint and the breakpoints in hybrids NM, DGF27C11 and CF25. Genetic and physical distances are included where they are known. The map positions of the STS derived from the left hand end of YAC A05136 (5136 L B&D) and the right hand end of YAC G12067 (G12 R A&B) are also shown.

3.10 ANALYSIS OF CANDIDATE GENES AROUND THE DG BREAKPOINT

The only candidate genes that map near to the DG breakpoint on chromosome 13 are FLT1 and FLT3 which are discussed in sections 3.8 and 3.9. respectively. On chromosome 1 two potentially important genes map to 1q22; the gene for the neural associated stem cell leukaemia factor, NSCL-1 and the oncogene cSKI

NSCL-1 maps to 1q22. The sequence of the cDNA is published and there is much less homology to rodent sequences in both 3' and 5' untranslated regions than in the open reading frame (Lipkowitz et al, 1992). I attempted to derive unique STS from each untranslated region as if one was present in DGF27C11 and one was not, it would be conclusive evidence that the DG breakpoint interrupts this gene. The 5' untranslated region was very GC rich and the initial primers, which attempted to amplify a 400 bp fragment, gave smears due to secondary structure formation. Increasing the annealing temperature of the PCR reaction, adding 10% di-methyl sulphoxide (DMSO) or adding a commercially available inhibitor of secondary structure formation (Perfect Match) all failed to resolve the problem. The primers were remade to only amplify a 200 bp fragment and gave a unique STS. The opposite problem was encountered in the 3' untranslated region. In an attempt to develop an STS that extended right to the end of the recorded sequence, the initial primers included the first few bases of the long poly-A tail and no product was obtained. Moving the primers 200 bp away from the tail resulted in the amplification of unique product. Neither product was found in mouse making both STS useful for screening somatic cell hybrids but, as shown in figure 3.48, neither produced a product from DGF27C11 indicating that the whole of this gene is telomeric to the DG breakpoint.

The second candidate gene is the oncogene, cSKI (Sutrave and Hughes, 1991), which is present in DGF27C11 (figure 3.6). The probe is a full length cDNA and figure 3.49 shows the results of probing a Pst 1 digested blot with this sequence. The lymphoblastoid cell line from the patient shows no rearrangement and the pattern seen in DGF27C11 is a composite of mouse and human bands. Instances have been reported of abnormalities in genes which do not cause obvious changes in the patterns obtained using cDNA hybridisation. I decided, therefore, to identify YACs containing cSKI and ascertain of they crossed the DG breakpoint. Screening the ICRF YAC library with the cDNA probe for cSKI resulted in 4 potential positive clones. cDNA probes do give weaker signal than genomic probes and the positives were faint, so it is, perhaps, not surprising that further analysis revealed that none of these contained the cDNA sequence. The sequence of the chicken cSKI gene is available on the EMBL database (EMBL ref.: GDCSKI3) and I derived a 420 bp STS from the 3' untranslated region of this sequence. The STS gave identical size bands in human, mouse and hamster DNA so it was useless for screening somatic cell hybrids. However, as no signal was seen in

AB1380, the STS for cSKI could be used to screen YAC libraries by PCR. A single YAC, E0851, was obtained from the ICRF YAC library. It is 350 kb long and rescuing the ends by vectorette PCR produced a 344 bp human fragment from the 1091 end and a 750 bp human sequence from the HYAC-C end. When the 1091 end was used as to probe a blot containing DNA from human, mouse, DGF27C11, yeast host and YAC E0851, a 1.4 kb band was seen in the human lane, but not in the lanes containing DGF27C11 or mouse DNA. On the same blot, an 8 kb band was seen in YAC E0851 but not in the yeast host (data not shown). The relevance of this finding is unclear. It may be that YAC E0851 spans the DG breakpoint and that the SKI oncogene has been disrupted by the translocation. However, it is also possible that YAC E0851 is chimeric and the HYAC-C end is derived from a different chromosome, explaining its absence in DGF27C11. This needs to be investigated further by sequencing the HYAC-C end of E0851 and developing STS for it as described for YACs G12 and 5136. The HYAC-C end of E0851 behaves as a high grade repeat and there is little chance of generating a useful molecular marker from it, even if it is sequenced. Using E0851 as a probe for FISH against the lymphoblastoid cell line from DG will also be done as this YAC potentially crosses the DG breakpoint. It is envisaged that both the positional cloning approach and the assessment of the role of candidate genes will proceed in tandem in order to try and clone the DG breakpoint junction and ascertain if genes important in Nb development have been disrupted by this genetic change.

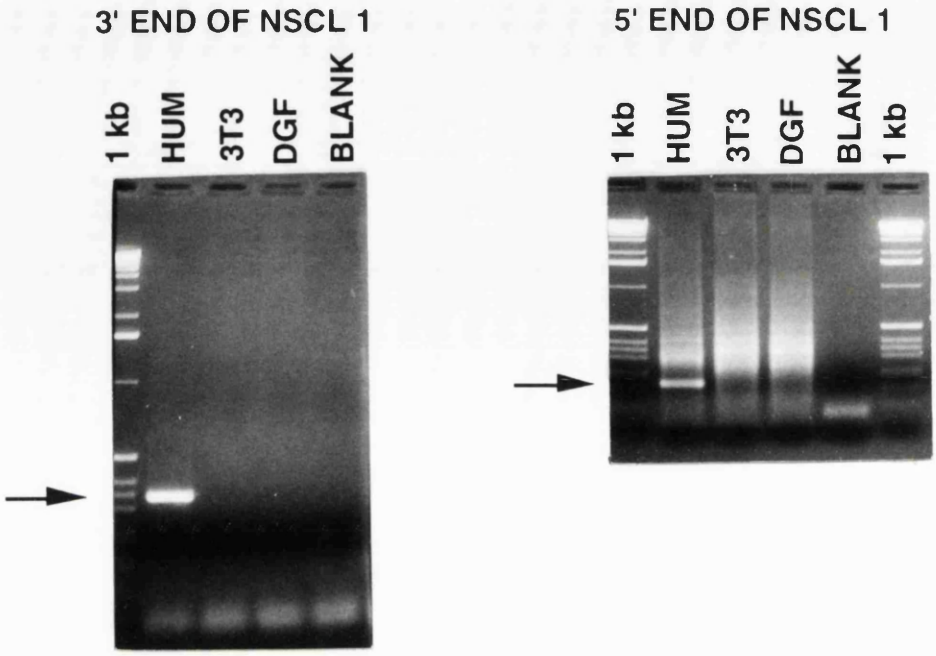


Figure 3.48. Results of using STS for the 3' (left hand photo) and 5' (right hand photo) untranslated regions of the NSCL1 gene to map DGF27C11. Single bands (arrowed) are seen in the lanes in which human genomic DNA (HUM) was used as the template for the PCR reaction. No signal is produced in DGF27C11 by either STS indicating that the whole of the NSCL1 gene is more telomeric on chromosome 1 then the DGF breakpoint. The background smearing seen when the primers for the 5' untranslated region were used is due to secondary structure formation in this very GC rich region.

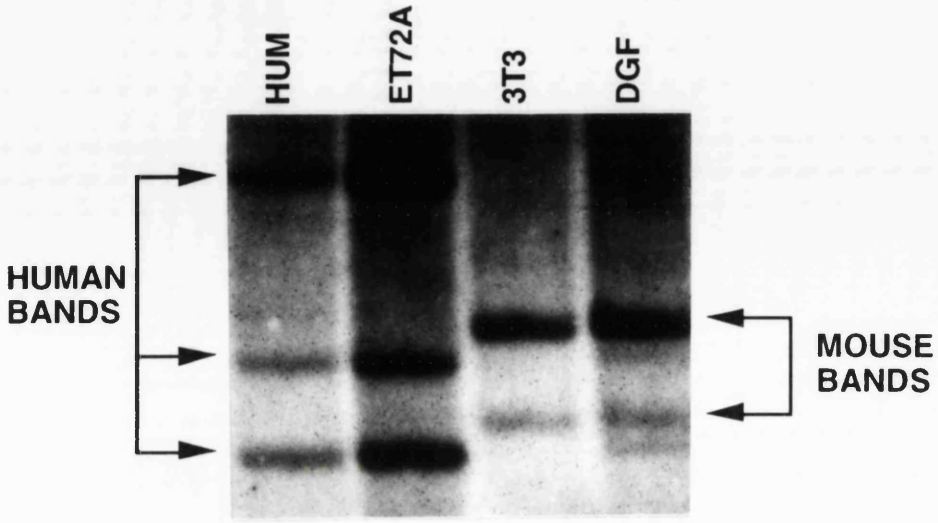


Figure 3.49. Results of using a cDNA probe for cSKI to probe a blot of Pst 1 digested DNA from human (HUM), the lymphoblastoid cell line from patient, DG (ET72A), mouse (3T3) and DGF27C11. Three bands (arrowed on left) are seen in human DNA and an identical pattern of bands is seen in ET72A. Two different sized bands (arrowed on right) are seen in mouse. All five bands are seen in DGF27C11, indicating that cSKI is present in this hybrid.

DISCUSSION

DISCUSSION

4.0 OVERVIEW

The aim of this thesis has been to isolate and characterise the breakpoint junction fragment of a constitutional t(1;13)(q22;q12) found in a patient with ganglioneuroblastoma. Two approaches were taken. The majority of the study concentrated on attempting to clone the breakpoint junction fragment in order to ascertain if a gene or genes important in neural development have been disrupted by this genetic change. This approach, called positional cloning, has been successful in the past in identifying a number of genes important in the development of malignancy (section 1.7). The second approach was to identify known, candidate genes in the vicinity of the DG breakpoint and investigate whether they have been disrupted by the translocation in this patient. Only those genes which code for proteins important in neural development or the control of cell growth would be investigated. The basis of both approaches is the assumption that the translocation observed in DG predisposed him to the development of ganglioneuroblastoma.

4.1 CONSTITUTIONAL CHROMOSOME ABNORMALITIES

In patients with genetic disease, analysis of constitutional chromosome abnormalities has proved to be a useful way of identifying the position of genes responsible for the phenotype (section 1.6.4). However, if the t(1;13)(q22;q12) was a normal variant or if incidence of constitutional chromosome changes in the normal population was high, detailed investigation of a single case of a translocation involving a chromosomal region not previously implicated in Nb development would be difficult to justify.

Seven population studies on the frequency of constitutional chromosome abnormalities have been performed (reviewed in Tawn and Earl, 1992). Five of the studies were performed in newborns, one in 7-8 year old children and one in normal adults. The frequency of constitutional abnormalities of the autosomes varied between 0.22% to 0.72%. The inv(9)(p11;q13) has been excluded from these figures as it is thought to be a polymorphic variant (Tawn and Earl, 1992). Translocations are the most frequently seen karyotypic change, comprising 85% of the abnormalities seen. Robertsonian translocations were identified in 0.09% (Hook and Hamerton, 1977) to 0.22% (Hansteen et al, 1982) of normal individuals although the consequences of these is unknown as the genetic material is unlikely to be rearranged. In the normal population the frequency of reciprocal translocations is between 0.09% (Hook and Hamerton, 1977) and 0.25% (Tawn and Earl, 1992). Although the figure of 0.09% comes from the largest study (56,952 individuals) its accuracy is suspect as the study

analysed unbanded karyotypes. Hook (1989) reviewed his previous data and concluded that although banding was not necessary to detect Robertsonian translocations, 51% of balanced reciprocal translocations would be missed if banding was not performed. Extra, structurally abnormal, chromosomes and autosomal pericentric inversions together account for less than 15% of the abnormalities of karyotype identified.

A small study of 329 adult patients with a variety of solid tumours (Richard et al, 1992) concluded that the frequency of karyotype abnormalities is increased in patients with malignancy. Four translocations were seen; two balanced reciprocal translocations (t(3;15)(q13.3;q21), t(3;6)(q13.3;q22)) and two Robertsonian translocations (both of which were t(13q;14q)). The incidence of abnormalities of constitutional karyotype was 1.22%, six fold higher than reported in normal individuals. However, these results should be interpreted with caution as the number of individuals studied was relatively small. In patients with haematological malignancies the frequency of congenital chromosome abnormalities was twice that of the normal population but this increase was not thought to be statistically significant (Alimena et al, 1985, Beneitez et al, 1987). Constitutional chromosome translocations are, therefore, relatively infrequent in both the general population and in patients with malignancy. The t(1;13)(q22;q12) has not been previously reported in normal individuals. Taken together, these facts support the notion that the rearrangement is causally related to the phenotype.

Very little data exists on the incidence of abnormalities of constitutional karyotype in patients with Nb. The latest of only two published studies three constitutional translocations were found in 44 patients, a frequency of 6.8% (Bown et al, 1993). Two of the three translocations were inherited from a normal parent and the third, a t(2;11)(p23;q22) appeared to be a de novo case. All three patients had stage 4 Nb. The earlier study (Moorhead and Evans, 1980) found two reciprocal, balanced translocations in 37 patients with Nb. One patient, with stage II Nb, had a t(4;7) which was present in unaffected relatives. The second patient presented with ganglioneuroblastoma and carried a t(11;16)(q23;q24); a genetic change that was also identified in the patient's unaffected mother. Patient DG, however, is the sole member of his nuclear family carrying the t(1;13)(q22;q12). The balanced, reciprocal translocation he carries either represents a previously unreported, random genetic event or a true, predisposing mutation. The decision to investigate the breakpoint so intensively is undeniably a 'high risk' strategy, but potentially offers a way of identifying a gene important in the development of 'low grade' Nb.

4.2 GENERATING FLANKING MOLECULAR MARKERS.

4.2.1 λ *phage analysis.*

The initial attempts at deriving molecular markers that closely flanked the DG breakpoint involved screening λ phage libraries. Using a commercially available chromosome 13 specific λ phage library (sections 2.9 and 3.3.1) the titre of phage obtained was similar to that quoted in the manufacturer's literature. However, the non-recombination rate was four times higher than that quoted by the supplier (68% as opposed to 14%). It is not clear why so many λ phage clones did not contain human DNA. One would imagine that, if the library had degraded or if it had been used incorrectly, the titre of phage obtained would have also decreased. It could be that the sample of 24 clones was sufficiently small for a disproportionate number of non-recombinants to have been selected by chance. With λ phage there is no counterpart of the X-gal colour selection system which allows identification of bacterial colonies containing plasmids into which insert DNA has successfully been cloned. As only six clones were investigated, sample size may also explain why all the clones from which human DNA was successfully isolated acted as repetitive sequences . The proportion of repetitive clones will vary according to the region under study but as the library was made from all of chromosome 13, small inter-regional variations should not have had an effect. If λ phage analysis had been the only way to derive chromosome 13 specific sequences, much more intensive efforts would have been put into optimising the use of this, or other, λ phage libraries. The major disadvantage of using these libraries was that the cloned sequences could have been located anywhere on chromosome 13. However, in early 1989 the technique of Alu-PCR was described (Nelson et al, 1989)and seemed to offer a rapid method of cloning molecular markers derived solely from the region of interest.

4.2.2 *The Alu repeat family*

Repetitive DNA sequences constitute 25-30% of the human genome. Some DNA repeats are clusters of relatively short sequences and are not interspersed with other sequence types. These clusters are generally located at the centromeres and telomeres of chromosomes and are thought to have a structural role, though their precise function is unknown. Other repetitive elements are interspersed throughout the genome. Minor families such as the 'mini-satellite' sequences (Jeffreys et al, 1985) or the variable number tandem repeats (VNTR) sequences (Nakamura et al, 1987) make up only a small proportion of the interspersed repetitive DNA. The two major families are the Alu and LINE sequences and it is these repetitive elements that have been so

valuable in the cloning of human DNA fragments from complex sources using the technique of Interspersed Repetitive Sequence PCR (IRS-PCR).

Alu is the commonest family of short interspersed repetitive sequences in the human genome. In humans each Alu element is a dimer, 300 bp in length, composed of two monomers separated by a poly(A) sequence. A poly (A) tract is also found at the 3' end of the Alu element and the regions flanking the Alu repeats are often AT rich (Jelinek and Schmid, 1982). The monomers are not identical. The left hand monomer is 130 bp long but the right hand monomer is 161 bp in length as it contains a 31 bp insertion. In rodents, Alu sequences exist as a 120 bp monomer and there is significant sequence divergence between the human and rodent Alu repeats (Britten et al, 1988). Alu sequences have broad homology with the 7 sl RNA, an abundant cytoplasmic RNA which functions as a component of the signal recognition particle that mediates the translocation of secretory particles across the endoplasmic reticulum (Ullu and Tschudi, 1984). It is likely that the genes coding for 7 sl RNA formed the first Alu repeat sequences. Alu is not a single sequence and, within humans, can be divided into distinct sub-families; Alu J and Alu S. The Alu S family itself can be further divided into three groups; a, b, and c (Jurka and Smith, 1988) and many smaller variations exist within each group. Other authors have identified the same sequence divergence but have given the sub-families different names (classes I-IV)(Britten et al, 1988). The Alu families retain strong homology to each other and a consensus sequence has been derived (Kariya et al, 1987). The average sequence homology between 125 different Alu family members was 84% (Jurka and Smith, 1988) but certain parts of the Alu element have been found to be highly conserved. The two most highly conserved regions are nucleotides 23-47 and 245-260 (Kariya et al, 1987).

There are $9x10^5$ copies of the Alu repeat in the human haploid genome which constitutes approximately 10% of the total DNA (Britten et al, 1988). If Alu elements were randomly distributed there would be an Alu repeat in every 4 kb of genomic DNA, but, in fact, Alu rich and Alu poor regions exist. In the Alu poor regions individual Alu elements are 10 kb or more apart, whereas in areas which are Alu rich the average spacing is 1 kb or less. Where repeats are less than 500 bp apart they are preferentially orientated in the same direction (Moyzis et al, 1989).

The biology of LINES differs substantially from that of Alu. The major human LINES family is L1 or Kpn 1 and has a consensus sequence of 6.4 kb. Although most LINES have an identical 3' end, many family members have deletions or rearrangements of the 5' section of the sequence. Although there are $1x10^4$ copies of the 5' end of the L1 sequence per haploid genome the 3' end is found five times more

frequently than this (Singer and Skowronski, 1985). When Alu and LINES sequences are used as probes for in-situ hybridisation, both are fifty times under-represented in centromeric and telomeric regions. However, they are inversely represented along the rest of the chromosomal structure with L1 sequences found in the Giemsa/Quinacrine positive bands and Alu sequences hybridising strongly to the Giemsa/Quinacrine negative or Reverse bands (Korenberg and Rykowski, 1988). This difference in chromosomal location is important to bear in mind when choosing which repetitive element to use as the basis of IRS-PCR cloning.

The DG breakpoint on chromosome 13 is in 13q12, a Giemsa-negative band and a region shown to be Alu rich using in-situ hybridisation . Therefore, I decided to use primers to one or other end of the Alu sequence to amplify fragments of genomic DNA from a suitable template using IRS-PCR.

4.2.3 *Choice of Alu primer*

The ideal Alu primer for these circumstances would amplify only sequences of human origin, but also recognise as broad a range of sub-families of the human Alu repeat as possible. Primers derived from the 3' end of the Alu sequence will amplify genomic DNA which is found between Alu elements whose 3' ends point towards each other. Conversely, primers to the 5' end of the Alu repeat will amplify an entirely different set of genomic fragments, those found between Alu elements whose 5' ends are orientated towards each other (fig. 3.7). Due to the limited proficiency of Taq DNA polymerase, the maximum size of fragment we could expect to amplify would be around 3 kb. It is unlikely, therefore, that this method will amplify any genomic sequences from Alu poor regions in which individual Alu elements can be up to 10 kb apart. Using a single Alu primer, the number of genomic sequences amplified from very Alu rich regions may also be low because the Alu elements are often orientated in the same direction. Clearly, Alu-PCR does not provide anything approaching complete representation of the genomic DNA in the template, but, if the primers and the DNA template are chosen with care, it does allow rapid amplification of a selection of sequences from the region of interest.

The consensus Alu sequence derived by Kariya et al (Kariya et al, 1987) and the relative position of the Alu primers available at the time I made the Alu-PCR libraries are shown in figure 4.1. The primers to the 3' end of the Alu repeat were TC65 (Nelson et al, 1989), Alu I and Alu IV (Cotter et al, 1990). Of these, Alu IV is the closest to the 3' end of the Alu repeat and TC65 is the furthest away. The proximity of a primer to the end of the Alu element is an important consideration as the inter-Alu fragments that are amplified by Alu-PCR will have Alu 'tails' at either end. If these tails are long, the resulting cloned sequences may act as repetitive probes even if the genomic inter-Alu fragments are unique. Alu I encompasses nucleotides 245-260 which are highly

conserved between Alu sub-families. However, even the extreme 3' end of the Alu sequence is relatively well conserved and I decided that on theoretical grounds Alu IV offered the best balance between encompassing the maximum number of Alu elements and generating single copy probes and used a variant of this primer, Alu 8559 which comprises Alu IV with a restriction enzyme recognition site at the 3' end.

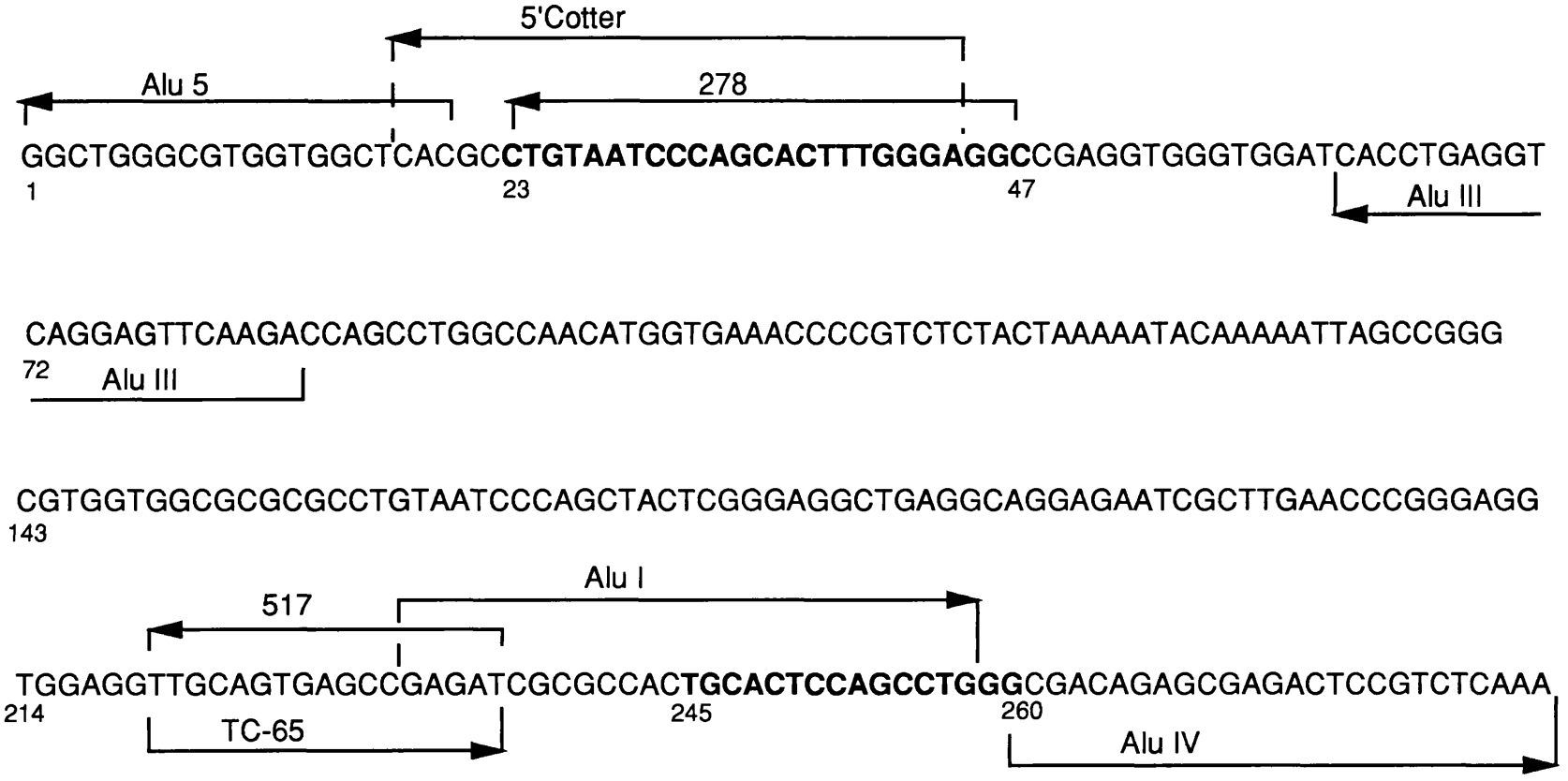


Figure 4.1. A diagram of the consensus Alu sequence (Kariya et al,'87) showing the position and orientation of the Alu primers discussed in section 4.2.3 (after Nelson,'91). The Alu sequence is shown 5' to 3' with selected nucleotides numbered. The two highly conserved regions are shown in bold type. The Alu primers are shown by arrows, which also denote the direction of extension in PCR reactions. The arrows that point to the left represent primers whose actual sequence is the reverse complement of the sequence shown.

For the 5' end of the Alu sequence the situation differed in that, although nucleotides 23-47 were highly conserved, nucleotides 1-20 were deleted in twelve of the fifty different Alu family members studied by Kariya et al (Kariya et al, 1987).. Primer 517 (Nelson et al, 1989) was actually closer to the 3' end of the Alu consensus than to the 5' end and would amplify 237 nucleotides of the Alu sequence at each end of the inter-Alu fragment. Any fragments cloned after amplification with the 517 primer would almost definitely act as repetitive sequences. Alu III (Cotter et al, 1990) is a 21 mer, 62 nucleotides from the 5' end of the Alu consensus sequence and an inter-Alu fragment generated by this primer would contain 82 nucleotides of the Alu sequence at each end and may act as repetitive probes. Alu 5 (Breukel et al, 1990), situated at the 5' end of the Alu repeat (nucleotides 1-20), would be deleted in 24% of Alu family members and is likely to amplify fewer inter-Alu fragments than the other primers. Primer 278 (Nelson et al, 1989) encompasses the highly conserved bases 23-47 but does not include the first 20 nucleotides of the Alu repeat. I used a variant of this primer, Alu 5684, extended in order to try and increase the species specificity (Cotter, pers. comm.).

4.2.4 *Choice of template DNA*

The ideal template for the Alu-PCR library would include 13q12 (the region of the DG breakpoint) but as little extra human chromosome material as possible. Rodent/human somatic cell hybrids containing deleted portions of chromosome 13 fulfilled these criteria. If the Alu primers used were human-specific, no rodent sequences would be amplified and, therefore, all the PCR products would come from the human chromosome segments contained in each hybrid. Three hybrids were considered; KSF39C6, RHF14 and ICD, all shown diagrammatically in figure 3.5. ICD is a hamster/human somatic cell hybrid containing 13pter-13q14 as its only human component. KSF39C6 and RHF14 also contained 13pter-13q14 and the breakpoints in 13q14 were more centromeric than that of ICD. However, less was known about the other human chromosome content in KSF39C6 and RHF14 contained and by using ICD any human specific sequences generated had to come from chromosome 13.

4.2.5 *Alu-PCR amplification- discussion of results*

An Alu-PCR reaction was performed using the 3' Alu primer (Alu 8559) and genomic DNA from ICD as the template. The PCR products were cloned into a plasmid vector, characterised and used as probes for Southern hybridisation (sections 2.6 and 3.3.2). The resulting clones were all human-specific, not even those which acted as repetitive sequences in human hybridised to genomic DNA from mouse 3T3 or Chinese hamster. The close proximity of the Alu 8559 primer to the 3' end of the Alu sequence obviated the need to routinely use competitive DNA 'blocking'. All the autoradiographs

shown in figure 3.15. were obtained without any competitive DNA being used and with only moderately stringent washing of the hybridisation filter. Although the human lanes of probes 3'Alu 62 and 71 show some background smearing which may have been reduced with the use of Cot 1 competition, no smearing is seen in the lanes from the hybrid DGF27C11. The advantage of not routinely using competition is not just based on considerations of time and cost but, more importantly, competitive DNA may decrease the signal strength in already weak hybridisations. Indeed, the major problem of using 3'Alu PCR products as probes in Southern hybridisation was the possibility of false negative results (section 3.7) and the routine use of competitive DNA may further exacerbate this problem.

Using the Alu 8559 primer, however, was not an unqualified success. The problem was that a small number of sequences dominated the library. Only six different sequences (3'Alu 169, 90, 66, 78, 71, and 62) were identified, despite screening 180 clones. Of these six sequences 3'Alu 90 was a centromeric sequence which hybridised to the centromeres of many different chromosomes, 3'Alu 169 was a low grade repeat and the remaining four sequences were derived from the centromeric third of 13q12. It is unlikely that no other clones had been amplified. Certainly, agarose gel electrophoresis of the PCR products derived from ICD using the Alu 8559 primer did show five dominant bands (figure 3.8) but against a background of smeared signal which was taken to represent other Alu-PCR products. Transferring the DNA from the gel shown in figure 3.8 to a hybridisation membrane and probing it with a labelled aliquot of the Alu-PCR product would have identified these extra bands if they existed. However, having analysed 180 separate clones from the library, would I have expected to identify other sequences if they were present? Looking critically at figure 3.8 it can be estimated that the five dominant bands are at least 30 times more intense than the background smear. If we assume that the background smear is composed of 50 separate bands then in order to characterise each of the background bands once I would have had to analyse 30 clones from each of the dominant bands. Therefore, the minimum number of clones I would have to analyse would be 5 x 30 + 50, i.e.: 200 clones. In practice, the number of clones to be analysed would be around four times that number if a complete characterisation of all the clones was required. When the 3' Alu-PCR library was made from ICD, 360 clones were gridded for analysis. I analysed the first 180 and colleagues from an affiliated laboratory (Dr. C. Mitchell) were to analyse the remainder. Of the 80 clones they analysed no further sequences that acted as single copy probes were identified. Why is it that these six sequences numerically dominated the library? Competition between templates is a well known phenomenon in PCR (Nelson, 1991) and it is possible that the six cloned sequences were flanked by Alu sequences that hybridised more strongly to the Alu 8559 primer or were more accessible in terms of secondary DNA structure. There is published data to which we can compare the results

obtained with this library. Bull et al (1993) developed an Alu-PCR library using ICD as the DNA template and A1S as the 3' Alu primer. A1S is very similar to Alu 8559 as shown in figure 4.1. However, the pattern of amplified sequences was different to that obtained using Alu 8559. Although four inter-Alu sequences were amplified from 13pter-13q12, four clones contained DNA from 13q14, one clone mapped to 13q13-14 and one sequence was found to be of hamster origin. Eight other clones were not assigned to a specific region of chromosome 13. It would be interesting to know whether the four sequences from 13pter-13q12 were identical to the ones obtained in this study but there was insufficient data in the manuscript to allow the sequences from the two libraries to be compared. It appears, therefore, that small changes in the sequence of the Alu primer can have profound effects on the amplification of specific inter-Alu clones. Direct comparisons between the results obtained with different Alu primers are rare, though primers A1, Alu IV, and TC65 are thought to yield comparable results (Nelson, 1991). It is not clear whether changing the primer would have allowed the amplification of a greater number of sequences from the region of interest.

60% of the clones derived from ICD using the Alu 8559 primer were repetitive sequences, a much higher percentage than that found by other groups. Cotter (1990) used the closely related Alu IV primer to amplify sequences from a somatic cell hybrid containing one homologue of chromosome 5 as the only human component and found only 15% of the clones to be repetitive. However, that figure was based on the percentage of clones which hybridised strongly to total human DNA and, using the same technique, only 18% (22 out of 122) of the clones I generated would have been classed as being repetitive. The remaining 42% of repetitive clones were only identified as not behaving as unique sequences when they were used as probes in Southern hybridisation. It may also be that the 13pter-13q14 region in ICD is rich in inter-Alu sequences that are repetitive, so that the repeats are amplified in disproportionately large numbers. Using the A1S Alu primer and ICD as the DNA template, Bull et al (1993) reported 18 different probes of which 6 (30%) were repetitive. However, 15 of the 59 colonies studied were found to contain the same repetitive clone and a second repeat was found on 9 occasions. Indeed, of the 59 colonies, 29 (49%) were repetitive sequences, a figure similar to that I obtained. Generating internal fragments from repetitive probes (section 3.3.3c) was not successful in generating new single copy sequences. In retrospect, having isolated only four single copy sequences from the 3' Alu-PCR library some further analysis of the repetitive clones may have been justified. I could have repeated the hybridisation using stringent competition, accepting that for some probes this would reduce the intensity of hybridisation of the unique portion of the sequence. Sequencing a few selected repetitive clones and deriving oligonucleotide primers for them may have allowed the identification of unique STS. However, by the

time the 3' Alu-PCR library had been fully characterised I was already analysing the 5' Alu-PCR library which, I hoped would yield more useful clones. Furthermore, I had already isolated a 1.3 Mb YAC with probe 3' Alu 66 and a 1.2 Mb YAC with FLT1. There was evidence that FLT1 and 3' Alu 66 were only 1 Mb apart (section 3.7) and at that stage it seemed that instituting a YAC walk would be a more useful way of characterising the region of the DG breakpoint than developing further probes from the 3' Alu-PCR library. In reality, YAC walking proved more difficult than expected (sections 3.8.6, 3.8.7 and 4.3.7) and the identification of further markers from the region of the DG breakpoint would have been very helpful.

Different problems were encountered with the library made from the PCR products of 5' Alu amplification of ICD DNA at 58°C (section 3.3.4). Ten different sequences were identified and all acted as repetitive probes, producing smears in both human and mouse genomic DNA. They were, therefore, useless for mapping DGF27C11, a mouse/human somatic cell hybrid. Close analysis of figure 3.8 does indeed suggest that the majority of the sequences derived from ICD are of rodent origin as the pattern of bands most closely resembles that seen in the lane derived from Chinese hamster DNA. The amplification of rodent sequences could be suppressed by raising the annealing temperature as shown in figure 3.9. At an annealing temperature of 72°C no signal was seen when rodent DNA was used as the DNA template and only 3 separate bands were seen on gel electrophoresis of the PCR product derived from ICD (section 3.3.4. and figure 3.14). Unfortunately, all three of these bands hybridised strongly to total human DNA.

The difficulty in using primers derived from the 5' end of the Alu element to generate unique sequences has been noted by several authors. Indeed, a recent review (Nelson, 1991) noted that none of the four primers derived from within the first 85 nucleotides of the 5' end of the consensus Alu sequence were human specific. Primers derived from the 5' end of the Alu consensus all seem to amplify rodent sequences, and the one I used also amplified a disproportionate number of repetitive sequences from the human DNA in ICD. There are reports of 5' Alu primers amplifying sequences that have proved to be useful in mapping experiments. Cole et al (1991) successfully generated STS from an irradiation-fusion hybrid containing 10-15 Mb of human DNA from the X chromosome using the ALE1 primer (derived from the conserved region of the 5' end of the Alu repeat) and the ALE34 primer (derived from the 3' end of the Alu sequence). They found that the number of sequences generated by ALE1 was less than that obtained with ALE34 (18 compared to 25) and that the sequences generated were more difficult to clone into a plasmid vector. One of the STS was hamster specific but the rest identified single copy, human specific sequences. Two of the Alu-PCR products were successfully used as probes in Southern hybridisation but it was not noted whether

or not these were derived using the ALE1 primer (Cole et al, 1991). By using ALE1 and ALE34 in combination a further 20 clones were generated but it is uncertain how many of these would have yielded sequences useful for mapping. However, Cole et al (1991) estimated that they had isolated a sufficient 'density' of Alu-PCR clones that acted as unique sequences to identify one Alu-PCR marker per 170 kb of genomic DNA. Indeed, using the products of Alu-PCR amplification with ALE1 and ALE34 alone and in combination they developed STS which were sufficiently closely spaced to allow the generation of a YAC contig without recourse to YAC end cloning (Cole et al, 1992a, 1992b).

There are now more choices of DNA template than at the time the library was made. However, rather than using a single template one could use a variation on the technique of differential hybridisation (Bernard et al, 1991). The breakpoint in hybrid CF25 is just below that of DG and the breakpoint in NM just centromeric (figure 3.22). The distance between the breakpoints of CF25 and NM encompasses the DG breakpoint and is less than a single chromosome band. Using Alu primers inter-Alu sequences could be amplified from the hybrid, PGME, which contains all of human chromosome 13 as its only human component. The resulting clones would then be gridded onto microtitre plates and the DNA transferred to hybridisation membranes. Using the same Alu primer inter-Alu fragments would then be amplified from CF25 and NM. An aliquot of the PCR product from each hybrid in turn would be radio-labelled and hybridised to the filter containing Alu-PCR clones from PGME. Those clones that hybridise to the PCR product of NM but not to that of CF25 are chromosome 13 specific (as they are found in PGME) and are derived from the portion of 13q12 containing the DG breakpoint. This method has the advantage of identifying clones from a small portion of a single chromosome with more certainty than PCR of microdissected chromosome regions or radiation hybrids in which DNA fragments from outside the region of interest are frequent contaminants. Obviously, the success of this method depends on the number of individual Alu-PCR clones that are generated. The major problem encountered using the Alu 8559 and 5684 primers was that there were insufficient clones generated to provide a reasonable density of markers from the region of interest. If I was to repeat the experiments I would use the ALE1 and ALE34 primers, both singly and in combination as described by Cole et al (1991). The density of Alu-PCR clones they obtained was far greater than I did. The somatic cell hybrid (ICD) from which I derived the Alu-PCR libraries is approximately 35 Mb long and the only Alu-rich region within it is 13q12 (10-12Mb long). Deriving five unique probes from the region gives a marker density of one per 7 Mb of genomic DNA, compared to one per 170 kb of DNA isolated by Cole. Although using the ALE1 and ALE34 primers would be expected to amplify more inter-Alu sequences the difference is probably too

large to be accounted for by the primers alone. Indeed, even using a single Alu primer (ALE1) Cole generated 25 clones from 10-15 Mb of DNA so it could be that there is an intrinsic difference between the regions under study. At the time I was analysing the Alu-PCR libraries derived from ICD other investigators were also using Alu-PCR to derive probes from the same region. The results obtained by Bull et al (1993) are discussed earlier in this section but neither he nor Chakravarti's group (pers. comm.) obtained anything approaching the density of clones generated by Cole et al (1991).

Is generating new probes from 13q12 now less important than at the outset of this thesis? As mentioned in section 1.7, high density genetic maps are being generated for all chromosomes using microsatellite markers (Cox Matise et al, 1994; Gyapay et al, 1994). The maps are of two kinds: framework maps which consist of highly informative loci the position of which relative to their 'neighbours' is known (the odds of the markers being positioned in a given location as opposed to anywhere else are greater than 1000:1) and comprehensive maps which consist of more markers the position of which is less accurately established. However, the recombination distances between markers located in the pericentromeric region are often large. This may indicate the relative paucity of microsatellite sequences in these regions or interference with the recombination mechanisms compared to the rest of the chromosome which would increase the recombination distance. In any event, even in the second generation Genethon human genetic linkage map (Gyapay et al, 1994) there are only 66 markers for 13q and only 4 of these are located in a region corresponding to that between CF25 and NM. The most comprehensive recent map of chromosome 13 (Washington et al, 1993) has 33 markers located in 13q14-13pter and of these 18 map to a region defined by the centromeric breakpoint in CF25 and 13pter (a region which is not further subdivided in the publication). Furthermore, none of the other recent maps of chromosome 13 divide the region between CF25 and NM any more accurately (Bowcock et al, 1993a; Bowcock et al, 1993b; Gerken et al,1993; Petrukhin et al, 1993). Hence, it appears that there are fewer available molecular markers from 13q12 than from some other regions of the genome and that even now it is difficult for investigators to use those markers that are available to directly isolate and map YACs from the region of interest.

4.2.6 *Cloning of Alu-PCR sequences*

Alu-PCR products amplified from ICD using Alu 8559 and Alu 5684 primers were blunt-end cloned into the Eco RV site of Bluescript SK plasmid and the ligated plasmids used to transform DH5α competent cells (section 2.6.3.b.). Blunt-end cloning is very inefficient. The transformation efficiency obtained with the Alu 8559 PCR product was $2x10^5$ colonies per μg of plasmid DNA compared to $5x10^8$ colonies per μg

with a control plasmid (section 3.3.2.c.). Whereas Cole et al (1991) found that the PCR products generated using primers to the 5' end of the Alu repeat were more difficult to clone, those generated from ICD by primer Alu 5684 gave a transformation efficiency of $3x10^8$, considerably higher than that of the Alu 8559 primers and almost equal to that of the control plasmid. This was originally thought to be a feature of the genomic DNA region under study rather than an intrinsic property of the Alu primer used. However, only 4% of the colonies resulting from the cloning of Alu 5684 amplified PCR products were white on the X-gal selection system and in half of these the plasmid did not contain a human DNA fragment. This corresponds to a 98% non-recombination rate; the majority of the colonies had been transformed with what was essentially control plasmid. Using the Bluescript SK blue/white screening system for the detection of non-recombinants was only partly successful in identifying bacterial colonies containing plasmids into which human DNA had been inserted. The selection depends on the role of the lacZα gene in the conversion of X-gal to a blue coloured product. The lacZα gene is situated at the cloning site and if a DNA insert is cloned into the linearised plasmid, the gene is disrupted and white bacterial colonies result, whereas if the linearised plasmid simply religates with no DNA insert, the gene remains functional. However, the process of linearisation and cloning can often introduce frame shift mutations in lacZα and even if the plasmid religates without a cloned DNA insert, the bacterial colonies resulting from transformation with the plasmid will be white.

Other methods of cloning the products of PCR amplification into plasmids have since been described, which attempt to minimise the problems of low cloning efficiency and a high frequency of non-recombination. One of the problems with the blunt end cloning of PCR products is that Taq DNA polymerase adds an extra deoxyadenosine (dA) onto the 3' end of the PCR product. Some authors designed PCR primers which had restriction enzyme sites at the ends. Both the PCR products and the plasmid into which they were to be ligated were digested with the appropriate enzyme thereby creating complementary sticky ends (Brooks-Wilson et al, 1990). However, the ends of the PCR products were often difficult to digest and the presence of restriction sites in the middle of the amplified PCR products also proved to be a problem. Another method involved using a dideoxynucleotide (ddT) to create a vector which had a single extra T residue on each arm. The ddT-tailed vector can then ligate to the PCR product through the formation of phosphodiester bonds between the 5'-P of the vector and the 3'-OH group of the overhanging A from the PCR product (Holton and Graham, 1990). Treating the vector with alkaline phosphatase appeared to further reduce the incidence of non-recombination. The use of alkaline phosphatase, restriction enzymes or T4 DNA ligase was avoided all together by a method which used PCR amplification of both inter-Alu fragments and the vector into which they were to be cloned. A 12 nucleotide

overhang was created in each and the resulting long sticky ends annealed without the need for DNA ligase (Aslanidis and de Jong, 1990). Formal comparisons between these methods have not been performed.

If the aim of generating the Alu-PCR library is simply to isolate a series of YACs from the region of interest then the cloning of inter-Alu fragment can be avoided all together by using the PCR product as a complex probe for hybridisation directly to a gridded YAC library (Monaco et al, 1991). However, the region of chromosome 13 contained in ICD is so large that attempting to link the large number of YACs obtained by this method would have been very laborious. If a small region of interest can be defined using differential hybridisation (section 4.2.5) then the hybridisation of inter-Alu sequences generated from that region directly to YAC filters may identify YACs more rapidly than using a number of individual probes.

4.2.7 *Characterisation of individual inter-Alu clones*

The aim of constructing the 3' and 5' Alu PCR libraries from ICD was to isolate unique sequences from 13q12 and it was, therefore, important to be able to exclude repetitive clones from further analysis and the identify which clones contained identical inserts. The number of repetitive clones found in the libraries is a feature of the template and the primers used (section 4.2.6). In order to compare the success of individual techniques, it is important to discuss the methods by which repetitive sequences can be identified. In the 3' Alu PCR library the colonies of transformed bacteria that contained plasmids with human DNA inserts were gridded out and transferred to hybridisation membranes. The membranes were then probed with sheared total human DNA. This method identified 18% of the colonies as being repetitive but a further 42% of the remainder acted as repeats on Southern hybridisation. Clearly, a more efficient method of identifying repeats was necessary. Those clones that did not hybridise to human DNA could be amplified using the Alu primer with which the library was made and size separated on agarose gels. The gel could then be blotted and probed with total human DNA a second time. Certainly, when used in the analysis of the 5' Alu library, this second round of screening was successful in identifying the repetitive sequences present (section 3.3.4). However, as mentioned in section 4.2.5, it would have been useful to select a few of these repetitive clones for further analysis. Although the isolation of restriction fragments from selected clones was not successful in identifying any unique probes it may have been that useful STS could have been obtained from some of the clones, increasing the number of useful molecular markers derived from the Alu-PCR libraries.

In the 3'Alu library the 30 cloned DNA fragments that acted as single copy probes on Southern hybridisation were made up of only 6 different sequences and I had,

therefore, performed many unnecessary hybridisation experiments. The optimal way of identifying probes that are identical was through comparison of their size, both undigested and following digestion with restriction enzymes (figures 3.11.b. 3.12). It is important not to 'over-interpret' small differences in the position of a band on a gel which may be due to the amount of DNA loaded or irregularities in the agarose gel. For example, 3'Alu 66 and 77 were thought to be two separate probes on the basis of a slight difference in the size of bands produced on enzyme digestion (figure 3.11.b & figure 3.12) whereas subsequent sequence analysis showed them to be identical. If there is any doubt, the clones should be digested with different restriction enzymes and the patterns of fragments obtained compared once more. As inter-Alu fragments are usually between 0.5 and 3 kb in length it is important to use frequently cutting restriction enzymes. Rsa1 and Taq1 were used in this study, other groups (Bull et al, 1993) used Sal1 and Rsa1. The comparison of clones on the basis of size gave much more reproducible results than 'self to self hybridisation'. Making a blot of size separated Alu-PCR products and probing the filter with one of the clones gave autoradiographs which showed evidence of cross-hybridisation, even to products which were different in size to the one used as a probe (figure 3.11). If two clones are identical they will produce the same size bands when they are used as probes in Southern hybridisation experiments and will obviously have the same sequence but these methods are far too laborious for routine comparison of cloned sequences.

4.2.8 *Use of Alu-PCR clones in mapping experiments*

I found Southern hybridisation to be an insensitive method of assessing the presence or absence of a given Alu-PCR clone in a panel of somatic cell hybrids. Indeed, three unique probes generated from the 3' Alu library (3' Alu 62, 71, 78) were thought to be absent from DGF27C11 by hybridisation but, using STS derived from these sequences, were found to be present in this hybrid by PCR (section 3.7). It is not surprising that hybridisation gave such a high number of false negative results. The Alu derived probes are genomic sequences, often less than 1 kb in length, and the labelled probe is, therefore, hybridising to a DNA fragment of identical length in the hybrid. The target DNA will constitute less than a millionth of the total DNA content of the hybrid and even if 20 μg of hybrid DNA per lane is loaded onto the gel, the amount of target DNA will be no more than 20 pg. PCR, however, is known to be able to successfully amplify DNA targets of this concentration. There have been reports of the use of inter-Alu sequences both as probes and as STS (Cole et al, 1991, Cotter et al, 1991) but no direct comparisons of the results obtained with the two methods were made.

Southern hybridisation is a useful and rapid way of screening cosmid and YAC libraries for the presence of a given Alu-PCR clone and can be accomplished in a fraction of the time taken by PCR based methods (sections 3.4. and 3.8.1.). However, screening the ICRF YAC library with 3' Alu 66 using Southern hybridisation only identified one positive clone whereas screening by PCR identified two (section 3.8.7.). PCR is again the more sensitive of the two techniques, but only when close attention is paid to the methods by which the PCR products are assessed. Visual identification of bands on an ethidium bromide stained gel on which the products of PCR screening of the YAC library have been size separated may miss positive clones. Indeed, the 'extra' YAC containing 3'Alu 66 was only identified when the gel was blotted and probed to further increase sensitivity (figure 3.25.a).

The development of STS is a time consuming process involving sequencing of both ends of the Alu-PCR probe, the selection of appropriate primers and the optimisation of PCR conditions for their use (section 3.5.). The long poly d(A) tracts found at the 3' ends of Alu repeat sequences complicate sequence analysis and decrease the choice of sequence from which primers can be designed. In spite of contributing to the difficulty of designing STS, the poly d(A) tails at the ends of the Alu repeats have been found to be a rich source of probes such as the VNTR polymorphisms that have been identified from the $(CA_4)n$ microsatellite sequence (Charlieu et al, 1992). The use of VNTR from 3' Alu tails has the similar practical disadvantages to that of inter-Alu sequences in that the polymorphisms may act as low grade repeats (Bernard and Wood, 1992) and may be difficult to identify in genomic DNA due to template competition (Charlieu et al, 1992).

Several attempts may be necessary to design primers for a given sequence which will amplify a unique STS. For example, the first set of primers designed for 3' Alu 71 amplified an identical sized product from human, mouse and AB1380 yeast genomic DNA (Michalski and Cowell, 1993). A second attempt, combined with close attention to optimising the PCR conditions, led to the development of a successful STS (figure 3.21.). Other authors report similar difficulties in designing STS from Alu-PCR sequences, deriving useful STS from around 70% of Alu-PCR clones (Cole et al, 1991), which is comparable to the figure of 83% that I obtained.

Mapping STS or hybridisation probes on a panel of somatic cell hybrids allows their position relative to the breakpoints of the various hybrids included in the panel to be determined. As a method of determining presence or absence of a sequence in a particular hybrid STS mapping is unrivalled but problems can still occur. 3' Alu 90 was thought to map below the DG breakpoint in 13q12 as it was present in hybrids CAF55, KBF11, CF25, CF27 and DGF27C11 (see figure 3.22. for hybrids and Michalski and Cowell, 1993). However, when hybrid NM became available (which includes all of

13q12) 3' Alu 90 was found to be absent. Hybridisation of 3' Alu 90 to the ICRF chromosome 13 specific cosmid library had identified several positive clones one of which was used as a probe for in-situ hybridisation to a normal chromosome spread. The cosmid identified by 3' Alu 90 was found to hybridise to the centromeres of many human chromosomes (section 3.4.). Hybrid DGF27C11 contains the centromere of chromosome 1 and CF25 the centromere of another human chromosome and it was these regions and not 13q12 that shared homology with 3' Alu 90 (Michalski and Cowell, 1994).

The use of cosmids as probes for in-situ hybridisation has many advantages in mapping studies. The position of the signal on the chromosome allows direct visualisation of the location of the probe of interest without the need to construct panels of somatic cell hybrids. As the technology used in FISH studies has improved it is now possible to estimate the genetic distance between markers. Indeed, Dr. David Shapiro has been able do so for markers in which we are interested (section 3.9). However, cosmid technology also has its problems. Cosmids isolated by 3'Alu 78 were not adequately visualised on several occasions and it is thought that there are regions of the genome where the secondary DNA structure inhibits adequate hybridisation of the labelled cosmid 'probe'. If cosmids are used to assess the presence or absence of a sequence in a hybrid they may yield false negative results. If only 5 kb of the human insert contained in a given cosmid is present in the hybrid that is being assessed the signal strength may be below the limits of detection.

The use Alu-PCR clones as probes for Southern hybridisation, as STS with PCR or to detect cosmids for use in FISH each has advantages and drawbacks. Basing conclusions as to the location of a given Alu-PCR clone on one method alone is risky but using all three methods in combination will give accurate data. It is a matter of debate as to whether the acquisition of more accurate mapping data fully justifies the time input involved in performing all three techniques.

4.3 YEAST ARTIFICIAL CHROMOSOME ANALYSIS

4.3.1 *Choice of YAC library*

Using Alu-PCR cloning and the molecular mapping techniques discussed in section 4.2.8, markers which closely flank the DG breakpoint had been isolated. The next step was to assemble a contig of cloned DNA sequences that contained markers on both sides of the breakpoint and hence the breakpoint junction fragment. Contig assembly is a laborious process as each cloned DNA fragment has to be characterised and its position relative to the other fragments determined. The longer the DNA fragments are, the fewer will be needed to span a given stretch of genomic DNA. YAC

technology allows the cloning of fragments of genomic DNA 1 MB or more in length, at least 25 times longer than had been achieved with cosmids. In order to span distances of 1 MB or more the screening and analysis of YAC libraries remains the method of choice.

At the time flanking markers for the DG breakpoint had been derived two separate libraries were available to us; the first had been created in Dr. Maynard Olson's laboratory (Burke et al, 1987) and the second by the ICRF (Lehrach and et al, 1990). The YACs in the Olson library averaged 300 kb in length and there was a 40-60% incidence of chimerism (Green and Olson, 1990). The ICRF library was chosen as it contained YACs which were over twice as long and the incidence of chimerism was only 20% (M. Ross pers comm). Both libraries were constructed from partial digestion of genomic human DNA with Eco R1. Libraries made from partial digests have the advantage of containing overlapping clones as the enzyme will have cut at different sites in different genomic DNA fragments. The presence of overlapping clones allows genomic walking within the library whereas in libraries constructed from complete digestion with rare cutting restriction enzymes contig assembly would not be possible. Complete digestion with rare cutting enzymes would also produce fragments that would be biased by the G+C distribution and CpG methylation patterns in the source DNA (Burke, 1990).

The Olson library contained 5 'genome equivalents' of DNA, whereas the ICRF library, which is made from a human cell line containing four X chromosomes, gives a two-fold genome coverage for the autosomes and a four fold coverage for the X chromosome. In retrospect I should have used both libraries in order to maximise the number of YACs identified with a given molecular marker. Only two YACs were identified with 3' Alu 66 and this delayed the development of a YAC contig above the DG breakpoint. The situation was even worse for FLT1, with the isolation of only a single, chimeric, YAC. Markers which did not closely flank the DG breakpoint identified larger numbers of YACs; five separate clones for 3' Alu 78, and four for 3' Alu 62 (section 3.8.4).

YAC libraries have since been made from individual human chromosomes (or chromosomal regions), using either flow sorted chromosomes or somatic cell hybrids containing the region of interest. With flow sorted chromosomes the amount of DNA available for cloning is low and therefore the majority of single chromosome region libraries have been constructed from somatic cell hybrids (Schlessinger et al, 1991). The human specific YACs can be differentiated from the larger number of YACs containing rodent inserts by hybridisation with an Alu probe. The advantages of using a monochromosome library are greater representation of the region of interest and a reduction in the rate of chimerism (probably due to the fact that there are fewer human

sequences with which homologous recombination can occur (Schlessinger et al, 1991). However, no chromosome 13-specific YAC library was available during the course of this study.

4.3.2 *Screening YAC libraries*

Initially the ICRF YAC library was only available as filters for Southern hybridisation. The advantage of screening the library by hybridisation for the presence of a sequence of interest was ease and speed as a single labelling and hybridisation was all that was required. However, even with the second generation of filters less than half of the YACs identified as potential positive clones actually contained the sequence of interest (section 3.8.1). The reason for the high error rate was the difficulty of accurately localising the position on the filter of the positive signal on the autoradiograph. The result was that over half the positive YACs that had been identified by hybridisation were not picked and this severely limited the generation of YAC contigs. Hybridisation is also not as sensitive as PCR. Indeed, only one positive clone (A05136) was identified when 3' Alu 66 was used as a probe for hybridisation but a second clone (H11111) was identified by PCR (section 3.8.7). At the primary pool stage the sensitivity of PCR alone, however, was not sufficient to allow the identification of the second YAC (H11111) on transillumination of an ethidium bromide stained gel. The gel had to be blotted and probed with 3' Alu 66 in order for the positive primary pool to be apparent. The reason for the low yield of specific PCR product was the small amounts of template DNA from the YAC of interest and the presence of contaminants in the template DNA. Only 60-200 ng of DNA was added to the PCR mix, but this represented 400 separate YAC clones. Therefore, only 0.15 ng of the yeast colony of interest was added to the PCR mix, of which around 7% was YAC DNA (the rest being DNA from the yeast host). It is interesting to note that in the PCR screening of the Olson library 100 ng of DNA from 1920 yeast colonies was added as the DNA template which works out as 0.05 ng from each yeast colony; much less that the 0.15 ng employed using the ICRF library (Green and Olson, 1990). Indeed, Green and Olson found they could still detect single copy sequences when 100 ng of DNA derived from 7680 YAC clones (0.013 ng of DNA from each yeast colony) was used. Although these amounts of template DNA appear minute they have to be compared with the amounts used in the amplification of single copy sequences from the much larger human genome. The genome of Saccharomyces cerevisiae is about 1.4 $x10^7$ bp, 200 times smaller than that of humans. The average YAC is 600 kb in length and so will only increase the size of the genome of the yeast clone to 1.46 x 10^7 bp. Hence, using 100 ng of DNA from 400 YAC clones in the ICRF library is equivalent to using 50 ng of human DNA as the template for a PCR reaction and 100 ng from the Green

and Olson library (1920 clones) equates to 10 ng of human DNA as a template for PCR. Both of these figures are well within the limits of sensitivity for the PCR reaction.

Other methods of library screening have been developed which allow hybridisation of pools of Alu-PCR products to library filters. Monaco et al (1991) reported a technique allowing uncloned Alu-PCR products from somatic cell hybrids to be used as probes on the second generation ICRF filters. The products of PCR amplification of a given hybrid were labelled by the random priming method and the repetitive elements competed out using a 2 day pre-association protocol. However, as seen with the Alu-PCR cloning of ICD, the molar ratios of the individual amplification products may differ widely and the weaker ones may fail to produce an identifiable signal. This problem may be compounded by the relative inefficiency of the random priming method when labelling complex mixtures of probes which will be used with stringent competition. Cole et al (1992b) developed a protocol which overcame these problems by first cloning the individual Alu-PCR products and then re-amplifying them using primers specific to the cloning vector. Equimolar amounts of the Alu-PCR products were pooled and then labelled by linear PCR amplification using a reaction mix containing ^{32}P dCTP instead of non-radioactive dCTP. If I did generate a series of region specific PCR clones by differential hybridisation (section 4.2.5) then I would use the method of Cole et al (1992b)to identify positive clones from the ICRF YAC library, although it still may not be as sensitive as PCR screening. Optimising the stringency of competition of repetitive sequences is vitally important as if too much is used the sensitivity will fall. Conversely, if too little is employed then the autoradiograph will be uninterpretable and the filter will probably be unusable for a prolonged period of time. The availability of replacement filters is not great and such an event can lead to prolonged delays in the identification of YACs of interest.

Once the yeast colonies that have been identified by either screening method have been picked it is important to colony purify them. The stab cultures may contain yeast colonies which have lost the YAC of interest or may even have been contaminated with other clones. Streaking out the stab culture on uracil deficient medium selects against colonies which have lost YACs. The selection is not, however, absolute and several yeast colonies that grew on uracil deficient medium did not contain YACs following analysis by PFGE and hybridisation to radiolabelled total human DNA (data not shown). Yeast clones which have been transformed with more than one YAC will only require the presence of a single YAC in order to grow on uracil deficient medium and may randomly lose the other(s). Out of the 14 different yeast clones characterised, one (D06108, section 3.8.6) contained two separate YACs in spite of colony purification, a figure which is in keeping with published data which suggests

that 10% of yeast clones contain more than one YAC (Schlessinger et al, 1991). Streaking out clones allows individual colonies to be screened for the presence of the probe with which they were identified and, again, PCR screening was found to be more accurate than hybridisation (section 3.8.3 and figure 3.23).

4.3.3 *Characterising individual YACs*

The two most important factors that decide the usefulness of a YAC in long range mapping are its size and the presence of chimerism. In order to prevent shearing and degradation of YACs, yeast clones grown in solution were embedded in agarose plugs before lysis (section 2.7.3.b). PFGE was performed to size-separate the host yeast chromosomes and the YAC. It was occasionally possible to visualise the YAC on UV transillumination of the ethidium bromide stained pulsed field gel but often the YACs were hidden behind yeast host chromosomes and were only seen when the gel was probed with human DNA (figure 3.31). Even if the YAC is readily visualised on transillumination it is important to probe a blot made from the gel with human DNA to exclude the presence of more than one YAC in the yeast clone (figure 3.28).

Changes in the conditions used for PFGE were needed for accurate assessment of the size of very large YACs (section 3.8.4. and figure 3.27). The YACs ranged in size from 225 kb to 1.3 Mb with an average size of 680 kb, which is comparable to the published average insert size for the ICRF library of 620 kb (Larin et al, 1991).

The identification of chimerism in YACs is important if they are to be used in the generation on contigs or in genomic walking. The length of the human insert DNA in YACs is such that chimerism is a much more frequent event in YAC cloning than in cosmids or phage. Chimerism can arise by two different mechanisms. The first occurs before transformation of the YAC construct into the yeast host and involves religation of DNA fragments from non-contiguous regions of the genome. Because all the fragments have been digested with the same restriction enzyme (Eco R1 in the ICRF and Olson libraries), they all have complementary sticky ends. The second mechanism occurs following transformation of two separate YACs into the yeast host. If the human DNA inserts share a region of homology it is possible that a recombination event could occur. Homologous recombination, often involving Alu elements, is common in yeast constructs and has been used in the manipulation of YACs for mapping or gene rearrangement (section 4.4.0). Few chimeric YACs have been studied in sufficient detail to ascertain which mechanism is the more important but a detailed investigation of a single chimeric YAC showed that homologous recombination involving an Alu element was responsible for the genetic rearrangement (Green et al, 1991).

The easiest way of detecting chimeric YACs is to use the YAC as a probe for FISH to spreads of normal metaphase chromosomes. If signal is seen on two different chromosomes then the YAC is either chimeric or contains sequences which have strong

homology to two different regions. YAC G12067 was shown to be chimeric by FISH with signal on both chromosome 13q and chromosome 2 (figure 3.38). If a YAC gives a single signal it is unlikely to be chimeric, though if the recombination event had occurred within 1 kb or so of the end of the YAC the signal from this short, chimeric end may be below the level of detection. FISH technology was not available in the laboratory until the last few months of the project but, given its speed and power, I would use it as the first method of assessing YACs for chimerism in the future. The two examples of FISH in the thesis (figures 3.38 and 3.40) were performed by labelling DNA made from lysed whole yeast clones. Methods exist for labelling the Alu-PCR products generated from YACs and are claimed to produce less background signal as no yeast DNA sequences are amplified (Baldini et al, 1992, Lengauer et al, 1993).

The most accurate method of determining the presence of chimerism is to rescue each end of the human insert of the YAC (section 4.3.5) and map it using a panel of somatic cell hybrids. The left hand end of YAC G12067 was shown to be absent from chromosome 13 by this method (figure 3.36). End cloning has the added advantage of determining which end is from the region of interest, allowing molecular markers from that end to be used in genomic walking strategies or the generation of contigs. Indeed, mapping the right hand end of YAC G12067 on a panel of somatic cell hybrids proved that it was further away from the DG breakpoint than the probe with which the YAC had been identified and the orientation of the human insert in the YAC could be determined (figure 3.36).

4.3.4 *Mapping an individual YAC relative to the DG breakpoint*

YACs that have been isolated using molecular markers that closely flank the DG breakpoint (3' Alu 66 above the breakpoint and FLT1 below) may contain human inserts that span the breakpoint junction. It was important to ascertain whether a given YAC contained markers that were known to map both proximal and distal to the DG breakpoint. However, YACs A05136 and H11111, which had been identified with 3' Alu 66, did not contain FLT1 and YAC G12067, which had been identified with FLT1, did not contain 3' Alu 66. It was still possible that one of these three YACs did cross the breakpoint as the flanking markers may map some distance away from the translocation junction. The position of individual YACs relative to translocation breakpoints can be assessed by two methods: mapping with sequences derived from the ends of the YAC and mapping by FISH.

4.3.5 *YAC end cloning*

The isolation and cloning of the ends of the human inserts in YACs can be accomplished by methods that rely on the end of the insert lying close to either a restriction enzyme site or an Alu repeat.

The first method that was reported involved digesting the YAC with a restriction enzyme. Many fragments would be produced but the 'left hand' end of the human insert could be rescued as it would be attached to the end of the YAC vector that contained plasmid sequences coding for the origin of replication (ori) and ampicillin resistance (amp). After digestion with an appropriate restriction enzyme, the fragments are religated under conditions that favour circularisation. Only the fragment of human insert DNA that is attached to the left hand end of the YAC vector forms a functional plasmid. This plasmid can be used to transform bacterial cells and the human specific DNA insert can then be recovered using the miniprep technique. However, the restriction enzyme used must be one which does not cut inside the 'plasmid' part of the YAC vector and hence Xho 1 was used, which cuts rather infrequently. The circularisation and cloning of large fragments was inefficient and yields were low. The other obvious limitation is that only one end is rescued (Nelson, 1990). Both ends can be rescued by inverse PCR but here again success depends on the presence of two restriction enzyme sites at either end of the YAC insert which are not found in the YAC vector (Silverman et al, 1991).

I chose to use the technique of vectorette end rescue (sections 2.7.5.b and 3.8.5) described by Riley et al (1990) as it recovers both ends of the YAC insert and allows the use of frequently cutting restriction enzymes. If a given restriction enzyme does not produce end clones of usable size the bubble can be modified so another enzyme can be employed. The two restriction enzymes I used were Rsa 1 and Hinf 1 (section 3.8.5), but the presence of a Hinf 1 site between the 1091 primer and the Eco R1 cloning site means that Hinf 1 digestion is only useful for deriving end clones from the left hand end of YACs. However, there were significant problems generating usable molecular markers using vectorette end rescue. The right hand ends of the YACs that were recovered by this technique were frequently smaller than the left hand ends and as there is no biological reason for such a size selection, chance alone can be the only explanation.

Of the 14 end clones generated, only one acted as a unique sequence on Southern hybridisation. The end clones that acted as repetitive sequences were sequenced and oligonucleotide primers designed in order to develop unique STS. This allowed the unequivocal localisation of both ends of YAC G12067 and one end of YAC A05136 but repeated attempts at deriving unique sequence from the short right hand end of A05136 did not yield a usable marker. One of the limitations of the concept of end cloning is if a given YAC end is repetitive then useful markers cannot be derived. The only choice is to generate random clones from the YAC, either by sub-cloning fragments obtained by restriction enzyme digestion or by amplifying internal sequences by Alu-PCR.

Vector-Alu PCR (sections 2.7.5.a. and 3.8.5) relies on the presence of an Alu repeat within 2 to 3 kb of the end of the YAC if PCR amplification of the end clone is to be successful and there are areas of the genome where Alu sequences are infrequent (section 4.2.2). Also, the orientation of the Alu repeat cannot be predicted so primers for both the 3' and 5' ends of the Alu repeat have to be used in separate experiments. Even if vector-Alu sequences are isolated, they will contain Alu sequences and are more likely to function as repetitive probes in hybridisation than end clones obtained by methods using enzyme digestion. I would only use this method if it was not possible to generate end clones of usable size with vectorette end rescue.

4.3.6 *Assessing the position of a YAC relative to a translocation breakpoint by FISH*

In patients such as DG, who carry reciprocal chromosome translocations, FISH can be used to rapidly assess the position of a YAC relative to the breakpoint. The method involves hybridising a YAC from the region of interest to a spread of metaphase chromosomes from the lymphoblastoid cell line (carrying both derivative chromosomes). If the YAC does indeed cross the breakpoint it may produce signal on both derivative chromosomes. However, if one end of the YAC only overlaps the breakpoint by a small distance (less than 5 kb) the intensity of the signal produced from this short fragment may be too low to be seen. Also, this technique does not produce any markers with which genomic walking can be attempted. Alu-PCR products derived from YACs have been used for mapping the position of YACs by FISH. However, when looking for what may be small regions of overlap it is preferable to use whole YAC DNA as the probe, as all of the human DNA insert is labelled. Using Alu-PCR products for FISH produces signal only from those regions in between Alu repeats sequences. If the end of a given YAC does not have any Alu repeats within it, the human DNA from that region will not be amplified and the resulting 'probe' will lack a potentially important region of the YAC insert.

4.3.7 *Generating YAC contigs*

The development of contigs of overlapping YAC clones allows the assembly of contiguous lengths of genomic DNA much greater than could be contained within a single YAC. Indeed, the entire Saccharomyces Pompe genome has been spanned by YACs (reviewed in Hoheisel and Lehrach, 1993). However, the isolation and characterisation of YAC contigs is difficult due to the frequent occurrence of chimerism in YAC clones. Perhaps the simplest method of generating a contig involves screening primary pools of YACs by PCR using STS from the region of interest (Cole et al, 1992a). If a primary pool is identified by two STS either each STS has identified different YACs within the primary pool or each STS has identified the same YAC. Using libraries with a 4 times representation of the genome, if the two STS really are

related, a number of primary pools should be identified as containing both molecular markers. The chances of a single primary pool containing both STS by chance alone (i.e. on different YACs in the same pool) is 25% but the probability of this occurring in two pools is only 4.5%, and only 0.5% if three primary pools containing both markers are identified. If the number of shared primary pools for any pair of STS is sufficiently high, a contig can be generated without the need to identify and screen separate YAC clones. This method is also more tolerant of chimerism than other techniques which depend on YAC end rescue and fingerprinting. Once the contig has been generated, only those YACs which contain more than one STS need to be isolated, allowing effort to be concentrated on YACs which extend the coverage of genomic DNA in the region of interest. The first drawback of this technique is that it relies on having a high density of STS in the region of interest. Cole et al (1992) used seven STS to generate a contig from Xq26. Although the STS isolated from 13q12 were developed using the same methods reported by Cole et al (1992) only 4 were found. The second requirement is the use of a YAC library with multiple genome equivalents of human DNA and the ICRF library has only two equivalents of chromosome 13 which decreases the power of the method.

Another widely used method of generating YAC contigs relies on generating fingerprints of the human DNA insert in YAC clones that have been isolated using a given probe or with probes that are known to be closely linked. Two main methods are used, both of which rely on the use of repetitive DNA sequences. In the first, the pattern of bands amplified from different YACs by Alu-PCR is compared. Bands of equal size are taken to be identical and to indicate that the two YACs overlap. Bands which are present in one YAC but not another represent regions of human DNA which are not shared by the two YACs. The dissimilar bands can be excised from gels on which the products of PCR amplification have been separated and then be reamplified for use as probes in genomic walking experiments.

The second method involves digestion of solution DNA made from the YAC clones with a frequently cutting restriction enzyme and probing blots made from the electrophoretic separation of the digest with highly repetitive Alu sequences, or moderately repetitive probes (occurring up to 2000 times in the genome) such as pTR5 (Schlessinger et al, 1991). Again, bands of identical size are assumed to be identical and bands that are present in one YAC but not another are thought to be derived from human DNA found in one YAC but not the other. The disadvantage of this second method is that a given band cannot easily be isolated for use as a probe. The use of either method for generating probes for genomic walking will be compromised by the presence of chimerism in either YAC as bands of dissimilar size will then result from co-cloned rather than contiguous DNA. A more fundamental problem is the assumption

that bands which are identical in size represent the same DNA sequence. This may be a statistically valid assumption if the bands are probed with or produced by sequences such as pTR5 or the L1 repeats that occur around 2000 times per haploid genome but is unjustifiable if the much more commonly occurring Alu sequences are used. However, if, after Alu-PCR amplification, YACs share **several** bands of identical size, this does provide more conclusive evidence of homology in a section of the human genomic DNA that the YACs contain. For example, figure 3.32 shows the amplification pattern obtained from five different YACs all of which have been identified by a single probe (3' Alu 78). YACs H0884, H0984 and A0483 all share four identical bands when amplified using ALE 1 primers. A further YAC (E0480) shared three bands with the others whereas no bands were amplified from the last YAC, A11111. The implication was that H0884, H0984 and A0483 contained human inserts that were very similar to each other. The insert in YAC E0480 shared less homology with the first three YACs and that of A11111 had no identifiable overlap with the others by Alu-PCR fingerprinting. I found restriction enzyme fingerprinting to be a useful way of deciding if Alu-PCR clones of identical size were indeed identical (section 3.3.3b and figure 3.11.b). Hence, a restriction enzyme digest of all five PCR products was performed but the number of fragments derived was so large that many co-migrated, making it impossible to obtain an accurate fingerprint from each band.

Stronger evidence of homology came from the Alu-PCR amplification of the same YACs using an Alu primer which was orientated in the opposite direction (ALE3, derived from the 5' end of the Alu repeat). Again, YACs H0884, H0984 and A0483 share an identical band pattern, this time of six bands. YAC E0480 shares three bands with the above YACs and A11111 (which must have some homology to the others as it was identified with the same probe) shares two bands of identical size. Indeed, H0884 and H0984 were eventually found to contain identical human inserts (section 3.8.6). However, YAC A0483, which was indistinguishable by Alu-PCR fingerprinting, was found to be 200 kb larger than H0884 (figure 3.31). It could be that Alu repeats with homology to the ALE 1 and ALE3 primers were absent from the extra 200 kb genomic fragment and this underscores the point that considerable lengths of genomic DNA can be 'missed' by fingerprinting.

The only certain way of ascertaining the relative order of YACs in a contig is to determine the presence or absence of known molecular markers in each YAC in the contig. The markers can either be previously isolated probes or YAC end clones, derived by any of the methods detailed in section 4.3.5, but it is important to map all of these sequences to the region of interest. The relative position of the two YACs identified by 3' Alu 66 has not, as yet, been completely determined. However, the absence of the STS for the right hand end of YAC A05136 in YAC H11111 suggests that the latter YAC is either completely contained within the former or overlaps the left

hand end of A05136 (section 3.8.7). The difficulty determining the relative position of these two YACs was due to the fact that all four end clones acted as repetitive probes on Southern hybridisation and a unique STS could not be developed from the right hand end of YAC A05136. If the end clones of YAC H11111 yield useful STS it will be possible to ascertain the relative orientation of AO5136 and H11111.

4.3.8 *Future analysis to define a YAC contig crossing the DG breakpoint*

FISH analysis suggests that 3' Alu 66 and cosmid c12.2 may be as close as 250 kb apart; a distance that could easily be spanned by a single YAC (figure 3.47). Two YACs (A05136 and H11111) have already been isolated using 3' Alu 66 and the orientation of the left hand end of YAC 5136 relative to the DG breakpoint and to YAC H11111 has been established. The right hand end of YAC A05136 is repetitive and, therefore unsuitable for further mapping studies. STS need to be developed from both ends of YAC H11111 and their position relative to the DG breakpoint and YAC A05136 determined, producing a contig of two YACs. The size of this YAC contig can be extended by further analysis of the five primary YAC pools that were identified by PCR screening of the ICRF YAC library using STS for the left hand end of YAC A05136 (section 3.8.7). The resulting YAC contig above the DG breakpoint would then consist of seven YACs. I would determine the size of each YAC by PFGE and the presence of chimerism by FISH. YACs which were not chimeric by FISH would be used as probes for FISH against metaphase chromosomes of the lymphoblastoid cell line from DG in order to ascertain if any crossed the DG breakpoint. In order to orientate the YACs relative to each other I would isolate the ends of all the YACs and map the end clones on the chromosome 13 hybrid panel (to check for chimerism), and on blots made from digested DNA from all seven YACs in the contig. It may be that this contig would contain one or more YACs that spanned the DG breakpoint and would provide a high density of probes from the region.

Cosmid c12.2 maps below the DG breakpoint and should be used to isolate YACs. The cosmid could be hybridised directly to the ICRF YAC library with stringent DNA competition to block repetitive DNA sequences. However, incomplete blocking of repetitive elements will render the filter unusable and too stringent competition will decrease the sensitivity of the screening. It should be possible to isolate unique molecular markers from c12.2. Eco R1 digestion of c12.2 did not yield a single copy probe (section 3.9) but digestion with another restriction enzyme may allow a unique probe to be identified. STS could be generated from c12.2 by Alu-PCR or by sequencing both ends of the human insert using primers derived from the cosmid vector. The STS could be mapped on the chromosome 13 specific somatic cell hybrid panel and used to screen YAC libraries (both MRC and ICRF) by PCR.

Once a contig that contains one or more YACs that span the DG breakpoint has been assembled it is important to confirm that the genomic DNA contained within the contig is not rearranged as homologous recombination can lead to the loss of sections of a contiguous DNA fragment. Certain regions are difficult to clone in YACs or, once cloned, are intrinsically unstable and prone to frequent recombination events. Telomeric sequences and tandemly repeated DNA sequences seem to pose a particular problem (Neil et al, 1990). Such deletions can be recognised by comparing pulsed field gel map is obtained from the YACs in the contig with that made from genomic DNA. Differences in the methylation patterns of genomic DNA and DNA contained in YACs can cause confusion but if the maps obtained with a number of rare cutting restriction enzymes are compared it should be possible to confirm the fidelity of the contig (Boldog et al, 1993, Schlessinger et al, 1991).

4.4 CLONING GENES FROM THE REGION OF INTEREST

4.4.1 *Identification of genes in the human insert of YACs*

The aim of isolating YACs which span the breakpoint junction fragment was to ascertain if a gene important in neural development has been damaged by the translocation. Initially, it would be important to establish that the breakpoint junction fragment contained a gene of any sort and then to determine the potential function of that gene.

One approach would be to generate a high definition map of the human DNA contained within the YAC spanning the DG breakpoint and narrow down the region of the translocation junction to a size amenable to DNA sequencing. Comparison of the sequence of the breakpoint junction with that of known genes in the EMBL data-base would allow assessment of its homology to known genes.

Various methods of mapping YACs have been reported. The approach originally suggested by Burke et. al. (1987) involved partial digestion of YAC DNA, creating fragments of different lengths which could then be separated by PFGE. Hybridising blots made from these pulsed field gels with probes specific for the right and left ends of the YAC vector allowed orientation of restriction enzyme sites within the YAC. Another method of deriving smaller sized fragments from the human insert of the YAC depends on homologous recombination. A construct consisting of a telomeric sequence, a selectible marker and an Alu repeat element is introduced into the YAC of interest. The Alu sequence in the construct then recombines with Alu repeats in the human insert (Pavan et al, 1990). This event results in the loss of all YAC sequences distal to the site of integration. The transformed yeast clones are then grown on media

which select for the presence of the fragmentation vector (HIS3 rather than TRP1). The resulting transformants will contain human inserts of various lengths depending on which Alu repeat in the YAC has been the target for the recombination event.

Sub-cloning smaller fragments generated by either method into λ phage or cosmids and using them for genomic walks will eventually lead to the identification of a small DNA fragment which represents the region of the breakpoint junction. However, formal mapping approaches are very labour intensive and slow. Methods have been developed which allow the isolation of genes from whole YACs on the basis of the gene's structure or its expression.

4.4.2 *Screening cDNA libraries*

In order to produce a functional protein genes are first transcribed into mRNA. The mRNA that is being expressed in a given tissue can be reverse transcribed into cDNA using the enzyme reverse transcriptase and many tissue specific cDNA libraries are now available. Screening a cDNA library with a YAC spanning the DG breakpoint could potentially identify transcribed sequences from the region of interest.

The first technique for screening cDNA libraries with YACs involved isolating YAC DNA and hybridising it to filters made from the gridded clones of a cDNA library. However, the need for stringent competition of repetitive sequences and the large size of the YAC probes led to poor signal intensity and only approximately 10% of the cDNA clones that were homologous to sequences in the human DNA were identified (Elvin et al, 1990). A second approach aimed to reduce the complexity of the probe to be hybridised to the cDNA library by sub-cloning the YAC into λ phage. The position of the λ clones relative to the breakpoint was ascertained by hybridising them to blots made from PFGE separated digests of YAC DNA. λ clones that mapped close to the breakpoint were then used to screen cDNA libraries. This approach was successful in cloning the gene for Menke's disease (Mercer et al, 1993) and the t(3;8) translocation breakpoint from a patient with hereditary renal carcinoma (Boldog et al, 1993) A further strength of the method is the generation of a high resolution map of the region of interest. However, mapping a large number of λ phage clones is time consuming and both groups commented on the weak hybridisation signals obtained when some of the phage clones were used to screen the cDNA library.

The ideal method of screening a cDNA library would selectively amplify those clones which have homology to the YAC of interest, thereby increasing the sensitivity of the procedure without the need for laborious sub-cloning. Several techniques have been developed which use PCR amplification of cDNA clones homologous to the YAC of interest; the major difference between them being the method of selecting the cDNA clones to be amplified. DNA from the YAC of interest is digested with a restriction

enzyme and transferred to a hybridisation membrane. The membrane is hybridised with cDNA clones which have not been radiolabelled. Either the membrane or the cDNA clones have been previously 'quenched' with DNA in order to block all the repetitive elements. cDNA clones that are not homologous to the YAC DNA are eluted by washing whereas those that are specific to the region of interest remain bound to the membrane. The specific cDNA clones are amplified using PCR either directly from the membrane(Parimoo et al, 1991) or following elution from the membrane (Lovett et al, 1991). These methods, though quicker than sub-cloning, were subject to technical difficulties, particularly in the handling of the membranes and the stringency of the washes needed to remove non-specific DNA clones. Recently, biotinylated fragments of the YAC of interest have been used to 'capture' specific cDNA clones by immobilising them on streptavidin coated magnetic beads. The cDNA clones can then be eluted from the beads and amplified with specific PCR primers. The initial method required the biotinylation of purified YAC DNA (Morgan et al, 1992). However, a more recent technique produces biotinylated human specific sequences from the YAC clone of interest without the need for DNA purification. Human specific fragments from the YAC of interest, are amplified by Alu-PCR with incorporation of biotin in the same reaction by using biotinylated nucleotides in the PCR mix (Forster and Rabbitts, 1993). The resulting biotinylated fragments are hybridised to random cDNA clones which have been made with priming sites for PCR on both ends. The biotinylated fragments are recovered on streptavidin coated magnetic beads. The captured cDNA clones from the region of interest are then eluted from the beads and amplified with specific oligonucleotide primers. Although it is possible that the region of the breakpoint junction may not be amplified by Alu-PCR, this method does have the advantages of speed and sensitivity.

However, no matter how elegant the procedure for screening cDNA libraries becomes, there are serious problems with this approach. Firstly, it assumes that the cDNA library has been made from tissue in which the gene is expressed. In attempting to clone genes coding for HLA-A loci involved in immuno-recognition cDNA libraries made from B lymphocytes or spleen were used as it was known that these tissues were involved in producing immunoglobulin. Hence, genes important in immuno-recognition must be present and constitutively expressed in libraries made from lymphocytes or spleen and indeed this was the case (Parimoo et al, 1991). However, the function of the putative Nb tumour suppressor gene is unknown. One could guess that it should be expressed in foetal adrenal tissue but it may encode a transacting differentiation factor produced by the liver or the brain. Which tissue specific cDNA library do we screen? Even if the correct tissue is chosen, it may be that the Nb tumour suppressor gene is only expressed at high levels for a brief window in foetal development and is only

transcribed at very low levels at other times. Again, how do we know which age of foetus to derive the cDNA library from? Within a given tissue only around 10% of the 10^5 genes in the human genome are expressed and the level of expression varies from 1 to 200,000 copies per cell (Patanjali et al, 1990). Screening a cDNA library derived from the correct tissue at the correct stage of differentiation may still result in the gene of interest being missed due to the vast excess of clones from more frequently transcribed genes. Attempts have been made to 'normalise' the numbers of cDNA clones within a library so as to increase the chances of cloning 'rare' transcripts (Patanjali et al, 1990). However, even in the normalised cDNA libraries, a given cDNA sequence may be identified only once for every 10^5 clones screened. All libraries still contain repetitive sequences and the majority of the clones identified are found to be 'orphans'; unrelated in structure to any known gene (Grausz and Auffray, 1993). Even more concerning was the report that over half the cDNA clones in the Genethon library were not of human origin but may have been cloned from yeast and bacterial contaminants (Savakis and Doelz, 1993). Clearly, any clone identified by cDNA library screening must be mapped back to genomic DNA (in our case to the DG breakpoint) to confirm that it comes from the region of interest. Studies of its functional relevance would then need to be performed (section 4.4.3).

Screening cDNA libraries is undeniably a powerful method of identifying transcribed genes from the region of interest. Once a single YAC spanning the DG breakpoint has been identified I would use the biotin-streptavidin cDNA capture method (Forster and Rabbitts, 1993) to isolated cDNA clones homologous to the human insert of the YAC. However, because of the significant problems that exist with cDNA cloning I would also attempt to clone genes located in the YAC insert using methods that rely on the structure of genes rather than their expression.

4.4.3 *Cloning genes on the basis of their structure*

Two groups of techniques have been described which utilise the structure of genes to clone sequences of interest, exon trapping and the cloning of CpG islands.

4.4.3.a *Exon trapping and amplification*

These methods depend on the intron-exon structure of the gene. Most genes are composed of transcribed sequences (exons) separated by non-transcribed DNA (introns) the function of which is unclear. When mRNA is transcribed from genomic DNA, exons are joined together, and correct splicing is dependant on the presence of splice sites at the 5' and 3' ends of the exons. The splice sites are not specific to a given gene, in fact all 5' and 3' splice sites are thought to be generically compatible. Techniques have been devised which allow capture of the 3' or both 3' and 5' splice sites from exons contained within random genomic DNA sequences.

Initial reports detailed the technique of 'exon trapping' by which the 3' ends of exons could be cloned (Duyk et al, 1990). Genomic DNA was digested with a restriction enzyme (BstY1) and cloned into a vector (pETV-SD) which contained an antibiotic resistance gene and an exon trap cassette. This cassette consisted of the first exon of the human β globin gene including its splice donor followed by an intron sequence and the α-β-galactosidase gene. Distal to the α-β-galactosidase gene was the cloning site for the digested genomic DNA. Following cloning of human DNA, the constructs were transfected into a retroviral packaging cell line. During the retroviral life cycle, genomic sequences of non-viral origin are correctly spliced and may be recovered as cDNA. Any 3' splice acceptor sites within the cloned human DNA would be spliced onto the 5' splice donor site of the human β globin gene, resulting in loss of the intervening intron sequences of the β globin gene and the α-β galactosidase gene. The opportunity for correct splicing was increased by using virions from the first retroviral packaging line to transfect a second packaging line. The product was then amplified by infecting COS cells where the presence of the SV40 promoter ensured amplification of the cDNA product. The episomal cDNA construct was then used to transform bacteria which were grown on media containing antibiotic and X-gal selection. White colonies contained plasmids in which the splice donor in the construct had joined up with a splice acceptor in the cloned human DNA resulting in loss of the α-β–galactosidase gene. Exon sequences could then be recovered by standard plasmid lysis methods or by PCR. Although this method was elegant in design the multiple transfections resulted in it being very laborious.

Exon amplification, reported by Buckler et al (1991) selects exons on the basis of the presence of both 3' and 5' splice sites and only involves a single transfection step. DNA from the region of interest is digested and cloned into a plasmid construct consisting of functional 5' and 3' splice sites of the HIV 1 tat gene, separated by an intron which contains the cloning site. The HIV 1 tat gene is itself flanked by β globin gene sequences. The construct is transfected into COS cells where its SV40 origin of replication ensures amplification of the construct and the SV40 promoter facilitates high levels of transcription. Again, processing of the transcript will result in any exon in the cloned human DNA being spliced to the HIV tat splice sites. Two or three days after transfection, cytoplasmic RNA is isolated from the COS cells and the exon sequences are rescued by RNA based PCR analysis using primers to the β globin gene sequences that flank the HIV 1 tat gene splice sites. This method was used successfully in the isolation of the human glycerol kinase gene from Xp21 (Walker et al, 1993).

However, both methods have several important drawbacks. The splice sites of some genes may not be compatible with the sites in the constructs. A few human genes may not have any introns and would not be able to be cloned. More importantly, are the false positives that result from the cloning of intron sequences that resemble splice sites (cryptic splice sites). It is estimated that cryptic splice sites may form as many as 20% of the positive clones identified (Buckler et al, 1991). It would be important to check that the sequences identified do indeed represent transcribed sequences by hybridising them to Northern blots of RNA from different species (section 4.4.3) and using them to identify unique cDNA clones. The other disadvantage with the exon capture methods is that it would be very time consuming to screen the number of clones that would be obtained from a large YAC. In spite of all the difficulties, exon capture is a very direct way of identifying transcribed sequences from the region and, in conjunction with cDNA cloning, has the best track record at identifying important genes.

4.4.3.b *Cloning of CpG islands*

Although human DNA is highly methylated, 1% of the genome is composed of stably unmethylated sequences (Bird, 1992). These unmethylated sequences exist as discreet CpG rich islands, usually one to two kb in length. It is now clear that most, if not all, CpG islands are associated with the 5' ends of genes (Bird, 1992). Restriction enzymes, such as Not 1, Mlu 1, BssHII, Sac II, that recognise CpG islands cut very infrequently and, therefore, create long DNA fragments after digestion. Long range maps of large DNA fragments can be made using PFGE to separate the products obtained from a given DNA fragment following digestion with one of these restriction enzymes. Although these pulsed field maps are very useful in determining the gross structure of a given DNA fragment, they do not allow easy sub-cloning of the DNA sequences in the region of the enzyme cutting sites which may represent the 5' ends of important genes. Hence, techniques have been developed which allow cloning of the DNA sequences adjacent to rare-cutting restriction enzyme cutting sites (Zabarovsky et al, 1993). The method involves digestion of the DNA of interest with a frequently cutting restriction enzyme (such as Bam H1) and circularisation of the resulting products. The circularised digests are then cut with Not 1 and the resulting linear digestion products, which comprise a Not 1 recognition site and the DNA flanking it, cloned into a plasmid vector. The plasmids are then used to transform bacteria and the resulting colonies comprise a library of not 1 'linking clones' from the region of interest. These clones can then be mapped back to the region of interest and those clones that correspond to the region of the genetic abnormality (the translocation breakpoint in the case of DG) can either be sequenced and used to develop specific oligonucleotide primers for the region, or used as probes to cDNA libraries. The success of the method obviously depends on the methylation pattern and purity of the DNA of interest. In

order to avoid contamination by yeast sequences, which are frequently unmethylated, either the YAC would have to be separated from the other yeast chromosomes, or the resulting library would have to be screened for the presence of human specific sequences which would not be easy as the size of the DNA fragments is small. Any genes which are methylated in the tissue from which the library was developed would be relatively under-represented. The major problem with the method is that not all genes have CpG islands at their 5' ends and, hence, important genes may be 'missed' by this technique. Even though the method has problems, it is considerably quicker than any of the exon capture techniques of identifying genes on the basis of their structure and is worth considering alongside cDNA library screening.

4.4.4 *Assessing the relevance of a cloned gene in the disease process*

Once a transcribed sequence from the region of the DG breakpoint has been identified its relevance in the evolution of the Nb phenotype needs to be established. It is relatively easy to establish that a given sequence is a gene. Important genes are conserved in evolution and should, therefore, be conserved among species which can be established by screening a 'zoo-blot' of RNA from various species (Monaco et al, 1986). As mentioned in section 4. 4.2, genes will have CpG rich areas at their 5' ends. Analysis of the sequence should identify the presence of open reading frames (Gray et al, 1982) enhancers (Weber et al, 1984) or promoter regions (Allen et al, 1988). The real difficulty comes in proving that a given gene is responsible for the Nb phenotype when we have no clue as to the pathophysiology of the disease. If the gene sequence showed close homology to transcription factors or known oncogenes one could infer that the gene was important in Nb development but that would not constitute proof of causality. When the human glycerol kinase gene was identified (Walker et al, 1993) there was the opportunity to compare the sequence to the prokaryotic equivalent, but there are no phylogenetically equivalent genes for the Nb tumour suppressor gene.

In the case of a gene potentially involved in tumour development one would expect the gene to be mutated in Nb tumours. Unfortunately, no tumour tissue from patient DG was deep frozen, but the presence of mutations of the putative Nb suppressor gene could be looked for in other tumours. The majority of aggressive Nb tumours are initially treated with chemotherapy and the tumour tissue that is resected after chemotherapy is largely necrotic so care would have to be taken to ensure that mutations were looked for in viable tissue and not in areas composed of necrosis or stroma. However, in Rb the mutations that have led to tumour development can be too subtle to be determined by Southern blotting. Sophisticated techniques such as single strand conformational polymorphism (SSCP) analysis, that can detect single base pair changes, may be necessary to detect mutations (Hogg et al, 1992).

Similarly, one may expect that a gene involved in suppressing tumour development may not be expressed in tumours but that is not the case in Rb where a transcript is seen in 60% of tumours (Friend et al, 1986) and in the tumours associated with mutation of p53 the abnormal transcript acts in a dominant negative fashion (section 1.3.5). In Wilms tumour (WT) it is now clear that more than one gene is involved in the disease phenotype and the majority of sporadic WT do not show abnormalities in the WT1 gene (section 1.3.4.a). It is likely that more than one gene is important in Nb development (section 1.5.5) and this will further complicate the assessment of whether the putative gene at the DG breakpoint is causally involved in the genesis of Nb. One method of acquiring convincing evidence of the role of the gene in Nb development would be to do the sort of elegant in-situ hybridisation studies of gene expression in tumour and foetal tissue that were performed by Pritchard Jones et al (Pritchard-Jones and Fleming, 1991) for WT as described in section 1.3.4.a. The introduction of the putative gene into Nb cells in tissue culture should reverse the malignant phenotype. This could be accomplished by micro cell mediated chromosome transfer (Stanbridge, 1990) which allows the transfer of whole chromosomes or chromosome fragments into cultured cells. Better still would be the transfer of just the gene of interest into cultured cells by the use of a YAC vector (Pavan et al, 1990). There are two problems with this approach. Firstly, only high grade Nb tumours grow in tissue culture and hence it would be impossible to test the relevance of any gene we have identified from DG, a patient with low grade Nb, in cell cultures derived from low grade tumours. It could be that high grade and low grade Nb are different diseases that share a common histological phenotype, in a way analogous to the L1 lymphocyte of acute lymphoblastic leukaemia being a phenotype that includes many genetically different disorders. If this was the case, then the gene isolated from DG may have no effect on high grade Nb tumours in culture even though it is central to the development of low grade Nb. Secondly, even if the cultured cells did start to behave in a less malignant fashion this would not prove that the loss of the gene that had been transferred to the cultured cells was vital in tumorigenesis. The transfer of p53 into cell lines from many different malignancies results in a slowing of cell growth and less malignant behaviour (Stanbridge, 1990) even though p53 abnormalities are only secondary events associated with disease progression rather than being responsible for tumour development. Coming full circle, the identification of patients with Nb who carry constitutional abnormalities of the putative Nb suppressor gene would be strong evidence for its role in Nb tumour development.

4.5 CANDIDATE GENES AND THE DG BREAKPOINT

The majority of the work in this thesis was directed towards isolating and cloning the breakpoint junction fragment of the constitutional t(1;13)(q22;q12) translocation found in patient DG. However, three candidate genes which map in the region of the translocation breakpoint were analysed in order to ascertain whether they had been disrupted by this genetic change. FLT3 maps to 13q12, but is at least 450 kb away from the DG breakpoint by FISH (figure 3.47) and is , therefore, very unlikely to have been damaged by the DG translocation. However, NSCL1 and SKI map to 1q22 and warranted further investigation.

The genetic change in DG was a translocation, not a subtle point deletion, and we had a somatic cell hybrid (DGF27C11) containing just the derivative chromosome 1. Perhaps the most convincing way of determining whether a candidate gene had been disrupted by the translocation was used for NSCL1. It was possible to design human specific STS for both the extreme 3' and 5' ends of the published gene sequence and prove by PCR analysis that both were absent from DGF27C11 (figure 3.48 and section 3.10). If the gene had been damaged by the translocation the STS for one end would have been present in DGF27C11 and the STS for the other end would have been absent.

If the oncogene, cSKI, had been disrupted by the genetic change then hybridisation of a cDNA probe for that gene to DGF27C11 should produce a band of different size to that found in normal human DNA. If the cDNA probe was hybridised to the lymphoblastoid cell line derived from DG (line ET72A) then two bands would be seen, one of normal size and one which migrated differently. This was not found to be the case (figure 3.49 and section 3.10). However, it could be that the extreme 5' or 3' end of the SKI gene had not been included in the cSKI probe and, therefore, that the SKI gene had been disrupted without any changes in the size of bands produced on hybridisation being apparent. Indeed, when a cDNA probe for the RET locus was hybridised to blots of DNA from 50 patients with MEN 2A no altered DNA fragments were seen, but PCR analysis of the RET gene detected mutations in 20 of 23 families with the disease (Mulligan et al, 1993). The sequence of the chicken SKI gene was available from the EMBL database and I intended to design STS from both 3' and 5' ends of the SKI sequence in a similar way described for NSCL1. However, the 3' end of the SKI gene was not species specific and produced an identical sized PCR product in both human and mouse making it unsuitable for screening DGF27C11, a mouse/human hybrid. Re-designing the primers did not allow identification of a PCR product which was specific for human DNA. If the SKI gene had been disrupted by the DG translocation a YAC identified by this gene would span the breakpoint. Using the STS

for the 3' end of SKI, a single YAC (E0851) was identified from the ICRF YAC library. Both ends of this YAC have been isolated but were found to act as repetitive probes on Southern hybridisation (section 3.10). Deciding whether YAC E0851 does span the breakpoint will depend on generating STS from the end clones or using the YAC as a probe for FISH against metaphase chromosomes from DGF27C11.

Each report of the cloning of a gene which maps to 13q12 or 1q22 raises the question of whether time and resources should be directed to assessing the position of that gene relative to the DG translocation breakpoint. The involvement of a given gene in the development of other malignancies would suggest that assessment of its role in the DG translocation may be profitable. One such gene is PBX1 (formerly called PRL) which maps to 1q23. A sub-set of childhood B cell leukaemias is characterised by the presence of a t(1,19) in the leukaemic clone. The breakpoint on chromosome 19 involved the E2A gene, which codes for a helix-loop-helix protein functioning as an enhancer for the transcription factors E12 and E47 (Nourse et al, 1990). The region of 1q23 at the breakpoint junction was a novel gene, PBX1, which had some homology to the homeobox genes important in development. Initially it was thought that the breakpoints in 1q23 were tightly clustered in a defined region (Nourse et al, 1990) but subsequent analysis of a larger number of leukaemic cells and cell lines revealed that the breakpoints occurred in a 50 kb region on 1q23. Furthermore, none of the breakpoints seemed to involve any of the exons of the PBX1 gene (Mellentin et al, 1990). In the t(9;22), Philadelphia, translocation the breakpoints in the ABL gene are also widely scattered, so there is a precedent for this finding. However, the fact that such a large and ill-defined region of 1q23 is potentially involved makes it difficult to assess whether the gene has been disrupted by the DG translocation. In the absence of information as to the extent of the breakpoint cluster region within PBX1, the best approach would be to identify YACs containing the PBX1 gene and see if they cross the DG breakpoint by FISH.

Other solid tumours also show genetic rearrangements in 1q21. A t(1;19) translocation has been seen in meningiomas (Prempree et al, 1993) and a der(16)t(1;16)(q21;q13) translocation has been identified in 2 of 20 Ewing's tumours (Douglass et al, 1990) which are more frequently characterised by the t(11;22). It may be that different genes have been disrupted in each of these translocations but it is also possible that abnormalities of a single gene located in 1q21 have occurred.

4.6 CONCLUDING REMARKS

The pace of technological advance during the period of this project has been rapid. At the outset of the study, Alu-PCR had only just been reported and YAC technology was in its infancy and not available in the host laboratory. Similarly, very

few probes and hybrids were available for mapping 1q22 and 13q12. The analysis of the DG translocation breakpoint has benefited from the use of many new methods and reagents, but has also encountered some of the problems associated with utilising recently described techniques. The decision as to how many new methods should be tried was difficult. The time constraints implicit on any PhD project dictate that one has to be selective and focused, hopefully without being blinkered. However, one of the major lessons has been that corroboration of data using different techniques is essential. The project has essentially revolved around the attempt to use positional cloning to isolate the DG translocation breakpoint. At the outset (section 1.7) it was obvious that the decision to approach the breakpoint from the chromosome 13 or the chromosome 1 'end' depended on the availability of molecular markers for the region. Approaching the breakpoint from chromosome 13 was justified in the light of the fact that the development of somatic cell hybrids for the region allowed the division of 13q12 (which is 10-12 Mb long) into six separate regions. However, the major problem with the project was the inability to derive sufficient molecular markers to allow a YAC contig to be readily constructed or a long range map of the genomic region to be easily devised. There still remains a relative paucity of probes from 13q12; nowhere near the one marker per 150 kb that would allow the construction of YAC contigs without the necessity for YAC walking and whether this represents difficulties intrinsic to the region or the under-use of available cloning techniques is unclear. Because few diseases appear to map to 13q12, the region has certainly attracted less attention than other parts of the genome. During the course of the PhD, I did attempt to map markers from 13q12 as they became available but of the few markers that exist for the region the majority were only described at the conclusion of my experimental work. In retrospect I should have spent more time attempting to derive markers for the region which would have increased my chances of constructing both a YAC contig and a long range physical map, accepting the possibility of being left with an incomplete map of 13q12 as the sole result of my work. In my analysis of YACs it may have been more fruitful to attempt to derive further markers from within the YACs which could then be used to identify further YAC clones rather than concentrating so much effort on trying to isolate end clones. However, I have managed to define the region of the DG breakpoint and generate flanking markers which may be as little as 250 kb apart. These, together with the YACs that have been isolated on either side of the DG breakpoint will serve as useful reagents in the future cloning of the breakpoint junction fragment. and may lead to the identification of a gene important in the development of Nb.

APPENDIX 1

APPENDIX 1.1 ALU PRIMERS

The sequences of the Alu primers used are given below. The size and number of PCR products varied with the template DNA used. For the construction of an Alu-PCR library the following PCR conditions were used: 94°C for 45 sec, 57°C for 30 sec, 72°C for 2 min, 30 cycles with a magnesium concentration of 1.5 mM. For 'Alu fingerprinting' the reaction conditions were 94°C 30 sec, 55°C 30 sec 72°C 1 min, for 30 cycles with a magnesium concentration of 1.5 mM.

NAME OF OLIGO	SEQUENCE OF OLIGO (5' TO 3')
ALE 1	GCCTCCCAAAGTGCTGGGATTACAG
ALE 3	CCACTGCACTCCAGCCTGGG
BK 33	CTGGGATTGCAGGCGTGAGCC
Alu 8559 (Alu IV)	CAGAATTCGCGACAGAGCGAGACTCCGTCTT
Alu 5684	CGGGATCCCAAAGTGCTGGGATTACAGGCGTG

APPENDIX 1.2 PLASMID PROMOTERS

The sequences of the Bluescript SK T3 and T7 promoters are given below. The PCR conditions shown will amplify inserts up to 2 kb in size. If the primers are being used to amplify inserts of known size the extension time should be reduced appropriately (1 min/ kb insert).

NAME OF OLIGO	SEQUENCE OF OLIGO (5' TO 3')	PCR CONDITIONS
BLUESCRIPT T3	ATTAACCCTCACTAAAG	94°C 45 sec 48°C 45 sec 72°C 2 min 30 cycles, 1.5mM Mg.
BLUESCRIPT T7	AATACGACTCACTATAG	

APPENDIX 1.3 STS FOR 3'ALU PROBES

It was possible to derive STS from five of the 3' Alu probes. 3' Alu 90 maps to the centromeres of many human chromosomes, the other four STS map to 13q12 (figures 3.21 and 3.22).

NAME OF OLIGO	SEQUENCE OF OLIGO (5' TO 3')	SIZE OF PCR PRODUCT IN bp	PCR CONDITIONS
3' ALU 66	1: CAGAAGGTGTGAGATGAGGACAC 2: TAGTTACACTGGAAGCCGGACCT	550	94°C 30 sec 60°C 30 sec 72°C 30 sec 30 cycles, 1.5 mm Mg.
3' ALU 71	1: GTCATACTTTGGCCAAATCTCGAT 2: ACTTGGTTTGTCACATTCTTGGTG	400	94°C 45 sec 55°C 45 sec 72°C 45 sec 35 cycles, 2.0 mM Mg.
3' ALU 78	1: AGTTTATAGGAGGCCGTAGTTTTGG 2: CCAAACTTGAAGGGGATGAGAATA	600	94°C 30 sec 58°C 30 sec 72°C 30 sec 30 cycles, 1.5mM Mg.
3'ALU 62	1: CCATTGGACAATACATGAACAGTT 2: TTCCAAGTTCATAGGCCTTCTGAA	510	AS FOR 3' ALU 78
3' ALU 90	1: CATAGAAAGTGCTAGGATTTACAAGGATG 2: TGCTTGATCACATTTACAGATGCTTTTCC	400	AS FOR 3' ALU 78

APPENDIX 1.4 CHROMOSOME 13 STS

The following STS were obtained for known markers on chromosome 13. Primers for exon 8 of RB were designed by A. Hogg and those for FLT1 by Polymeropoulos (Polymeropoulos et al, 1991). The primers for Factor VII /intron are given in Abbott and Povey (1991). The sequences of the primers for 1353 (D13S120) and MS34 (D13S115) were published in Bowcock et. al (1992). I designed the primers for FLT3 from the 3' untranslated region of the gene, the sequence of which was published in Rosnet et al (Rosnet et al, 1991) bases 3186-3207 and 3399-3423 inclusive.The sequences of the primers, the reaction conditions and the sizes of the expected PCR products are shown below.

NAME OF OLIGO	SEQUENCE OF OLIGO (5' TO 3')	SIZE OF PCR PRODUCT IN bp.	PCR CONDITIONS
EXON 8 RB	1: GACCTAAGTTATAGTTAGAATACTTC 2: CATGCTCATAACAAAAGAAGTAAA	316	94°C 30 sec 55°C 30sec 72°C 30 sec 35 cycles, 1.5 mM Mg.
FACTOR VII INTRON	1: CTTGTCCTTTGGATCAGTCCACGGA 2: TAATCCTAGTGGGACAGGGACTGGT	376	AS FOR EXON 8 RB
FLT1	1: TTTGGCCGACAGTGGTGTAA 2: AGGACCAAACCATGTCTGTC	220	94°C 45 sec 55°C 45 sec 72°C 45 sec 35 cycles, 2.0 mMMg.
MS34 (D13S115)	1: TGTAAGGAGAGAGAGATTTCGACA 2: TCTTAGCTGCTGGTGGTGG	160	94°C 1 min 60°C 1 min 72°C 1 min 35 cylcles, 1.5 mM Mg.
1353 (D13S120)	1: CAGACACCACAACACACATT 2: ATGACCTAGAAATGATACTGGC	200	AS FOR MS34
FLT3	1: ACTTTTCTCTAAGATGCTGTCTG 2: AATCATATTTAGCTTTCCCTTAGCA	237	94°C 45 sec 52°C 45 sec 72°C 45 sec 30 cycles, 1.5 mM Mg.

APPENDIX 1.5 CHROMOSOME 1 STS

The following STS were obtained for known markers on chromosome 1. NGFB3 and SPECT were published by Abbott and Povey (1991). I designed those for cSKI from the 3' untranslated region of the chicken cSKI sequence (bases 441-465, 936-960 inclusive of GDCSKI30 available in the EMBL database. Primers to both the 3' and 5' untranslated regions of NSCL1 were designed from the sequence given in Lipkowitz et al (1992). The 5' primers were bases 1-25 and 146-166 the 3' primer bases 2226-2249 and 22441-22464.

NAME OF OLIGO	SEQUENCE OF OLIGO (5' TO 3')	SIZE OF PCR PRODUCT IN bp.	PCR CONDITIONS
SPECT	1: TAGAGAAGAAAGGCTGCCCAGTAGG 2: CCTGGAACCAATGCCCCATGGATAA	500	94°C 45 sec 64°C 45 sec 72°C 2 min 30 cycles, 1.5 mM Mg
NGFB3	1: GGAAGCATCCTGTGTGCTGATGCTG 2: CTAACTTCTCCAAGGGAGGAGCCAG	400	AS FOR SPECT
cSKI	1: GCCGCAGATTCTCAACTCGGTGCTG 2: CTCCTCCTTGCCCGTGTAATCCTGG	500	94°C 30 sec 55°C 30 sec 72°C 30 sec 40 cycles, 1.5 mM Mg
NSCL1 5'	1: AGGTGTGTGTGTCTGAGGGTGTGTG 2: GTGACCAGCTTTCTAGACTCTC	250	AS FOR cSKI
NSCL1 3'	1: GTTGTGCACTTTGGGGTAGAGTGA 2: GGCCTGCTCTTGGGAGTATCAGGA	400	94°C 30 sec 60°C 30 sec 72°C 30 sec 30 cycles, 1.5 mM Mg.

APPENDIX 1.6.1 VECTORETTE BUBBLES

The sequences of the Rsa 1 and Hinf 1 bubble tops and the universal bubble bottom are given below. Equimolar amounts of one bubble top and the universal bubble bottom are heated to 65°C for 15 min and allowed to cool to room temperature over a 2 hour period. YAC DNA was cut with Rsa 1 or Hinf 1 and the residual enzyme activity destroyed by heat inactivation .100 ng of the annealed bubble (Rsa 1 or Hinf 1 depending on which enzyme was used) was added to YAC DNA and ligated overnight as described in section 2.7.5.b.

BUBBLE NAME	BUBBLE SEQUENCE (5' TO 3')
Rsa 1 BUBBLE TOP	GAAGGAGAGGACGCTGTCTGTCGAAGGTAAGGAACGGACGAGAGAAGGGAGAG
Hinf 1 BUBBLE TOP	ACAGTTGAAGGAGAGGACGCTGTCTGTCGAAGGTAAGGAACGGACGAG
UNIVERSAL BUBBLE BOTTOM	CTCTCCCTTCTCGAATCGTAACCGTTCGTACGAGAATCGCTGTCCTCTCCTTC

APPENDIX 1.6.2 VECTORETTE PCR

Using the primers given in this table it was possible to rescue both ends of YACs to which the bubbles noted in Appendix 1.6 had been ligated. Primer HYAC-C is derived from the left hand end of the pYAC4 vector, primer 1091 from the right hand end and primer 224 from the YAC bubble. The position of all the primers relative to the vector is shown in figure 3.35. In order to rescue the left hand end primers HYAC-C and 224 were used, for the right hand end 1091 and 224. The initial PCR conditions were 94°C 1min, 67°C 2 min, 72°C 2 min for 35 cycles using a magnesium concentration of 2.0 mM. If background bands appeared the magnesium concentration was reduced to 1.5 mM. For products of 1 kb or under the times of the denaturing, annealing and extension steps was reduced to 45 secs in subsequent reactions.

NAME OF PRIMER	SEQUENCE OF PRIMER (5' TO 3')
HYAC-C	GCTACTTGGAGCCACTATCGACTACGCGAT
1091	CGATATAGGCGCCAGCAACCGCACCTGTGG
224	CGAATCGTAACCGTTCGTACGAGAATCGCT

APPENDIX 1.7 PRIMERS USED FOR SEQUENCING YAC END CLONES

The following primers were used to sequence the YAC end clones generated by vectorette PCR as described in sections 2.7.5.b. . For the amplification of the DNA template to be sequenced primers for HYAC-C and 224 were used to rescue the left hand YAC end clone, primers 1091 and 224 for the right hand end clone. The 224 primers noted in appendix 1.6 were biotinylated to allow separation of DNA strands prior to sequencing as described in section 2.8.1. The sequencing reactions were performed using the method in section 2.8.2-3. The positon of the primers relative to the YAC vector sequence is shown in figure 3.35. Primer RA2 was used to sequence the 1091 end clone, primer 32194 the HYAC-C end clone and primer 32427 to sequence the 'negative' strand of both end clones.

NAME OF SEQUENCING PRIMER	SEQUENCE OF PRIMER (5' TO 3')
RA2	TCGAACGCCCGATCTCAAGATTA
32194	GTTGGTTTAAGGCGCAAGAC
32427	GAATCGCTGTCCTCTCC

APPENDIX 1.8 YAC END CLONE PRIMERS

Primers were designed for each of the end clones of YACS A05136 and G12067. Unique STS were derived for both ends of G12067 and the HYAC-C end of AO5136. Two STS were derived for the HYAC-C end of YAC AO5136, a small product resulting from using primers 5136 Left B and C and a larger product from the use of primers 5136 Left B and D. The sequences, reaction conditions and the sizes of the PCR products expected are given in the table below.

NAME OF OLIGO	SEQUENCE OF OLIGO (5' TO 3')	SIZE OF PCR PROCUCT IN bp.	PCR CONDITIONS
G12 RIGHT	1: GGCTTATTTCTGTTTAGATCC 2: GGAACTACTAATTACCCTAAG	220	94°C 45 sec 56°C 45 sec 72°C 45 sec 30 cycles, 1.5 mM Mg
G12 LEFT	1: GCAGTTTTCTGCATTTCAGAATCAGG 2: CTTTAATCTTGGATGACTGCTATCCC	1300	94°C 1 min 54°C 2 min 72°C 2 min 35 cycles, 1.5 mM Mg
5136 LEFT-(LBC)	B: GGGTTTGAGAACCTATTTATCTC C: GGATCCATGACAGAATTATTGTA	113	94°C 30 sec 56°C 30 sec 72°C 30 sec 30 cycles, 1.5 mM Mg
5136 LEFT-(LBD)	B: GGGTTTGAGAACCTATTTATCTC D: CGCGTCAGTAGCAGTA	1400	94°C 40 sec 52°C 40 sec 72°C 40 sec 30 cycles, 2.0 mM Mg

APPENDIX 2

APPENDIX 2.1 PROBES FROM 13q

Appendices 2.1, 2.2 and 2.3 show details of the previously reported probes that were used to characterise the DGF hybrids and the mapping panels. Each probe is identified by the names of the locus and the probe. The size of the probe is given next followed by the enzyme required to cut it out of the plasmid vector. The size of the bands obtained when probes are used in hybridisation is given in kb. If a probe recognises more than one band the sizes are noted directly whereas if the probe is polymorphic and recognises different sized bands in different individuals, the sizes are given prefixed by: 1 or 2. The stringency of washing necessary to minimise background without losing signal strength is noted. Finally, the map position of the probes and references for the probes are shown.

PROBES FROM CHROMOSOME 13q

Locus	Probe	Size of probe (kb)	Cloning site	Size of band on southern blot (kb)	Enzyme used to digest DNA on blot	Washing conditions	Location	Reference
D13S1	p7F12	4.2	Hind III	4.2	Hind III	0.5xSSC, 30°C	13q12	Cavenee et al,'84
D13S6	pHU10	2.3	Hind III	2.3	Hind III	0.5xSSC, 30°C	13q12	Dryja et al, '84
D13S93	WC71	1.4	Hind III	22.0	Eco R1	2xSSC, 65°C	13q12	P. Bull pers comm
D13S100	KC21.1	4.4	Hind III	22.0	Eco R1	2xSSC, 65°C	13q12	P. Bull pers comm
RB1	RB 3.8	3.8	Hind III	6.8	Eco R1	2xSSC, 50°C	13q14	Lalande et al,'84
D13S2	p9D11	2.3	Hind III	22.0	Eco R1	1xSSC, 30°C	13q21	Cavenee et al,'84
D13S3	p9A7	1.0	Hind III	22.0	Eco R1	2xSSC, 65°C	13q32	Cavenee et al,'84

APPENDIX 2.2 PROBES FROM CHROMOSOME 1p

Locus	Probe	Size of probe (kb)	Cloning site	Size of band on southern blot (kb)	Enzyme used to digest DNA on blot	Washing conditions	Location	Reference
D1Z2	p1-79	0.9	Eco R1/Hind III	multiple small bands	Eco R1	2xSSC 65°C	1pter	Borg et al, '92
D1S80	pMCT 118	3.1	Pst 1	VNTR	Eco R1	2xSSC, 65°C	1p36	O'Connell et al, 1989
D1S96	p1-45	4.5	Eco R1	2.0	Rsa 1	2xSSC 65°C	1p36 1.2-2	Weith et al, 1989
FUCA1	pAF3	0.9	Eco R1	0.9	Eco R1	2xSSC, 65°C	1p34	Middleton-Price,'90
FGR	pFGR	2.0	Eco R1/Bam H1	9.0	Eco R1	2xSSC, 65°C	1p34	Middleton Price,'90
MYCL	pLmyc10	1.8	Eco R1/Hind III	6.6	Eco R1	2xSSC, 65°C	1p32	Nau et al,'85
D1S17	p3-18	1.3	Eco R1	1.3	Eco R1	2xSSC, 65°C	1p31	Dracopoli et al, '89
D1S20	p5-21	1.0	Eco R1	1.0	Eco R1	2xSSC, 65°C	1p22	Dracopoli et al '89
D1S12	p2-05	3.5	Eco R1	3.5	Eco R1	2xSSC, 65°C	1p22	Dracopoli et al, '89
D1S9	p1-04	2.1	Eco R1	2.1	Eco R1	2xSSC, 65°C	1p22	Dracopoli et al,'89
AMY1	pEB8	0.55	Eco R1/Hind III	12 &8	Pst 1	2xSSC, 65°C	1p21	M.Meisler, pers comm.
D1S13	p5-22	4.4	Eco R1	4.4	Eco R1	2xSSC, 65°C	1p13	Dracopoli et al, '89
D1S14	p6-02	5.5	Eco R1	5.5	Eco R1	2xSSC, 65°C	1p13	Dracopoli et al,'89
NRAS	pMRC3	6.0	Eco R1	4.3 & 1.7	Taq 1	2xSSC, 65°C	1p13	Middleton Price,'90

APPENDIX 2.3 PROBES FROM CHROMOSOME 1q

Locus	Probe	Size of probe (kb)	Cloning site	Size of band on southern blot (kb)	Enzyme used to digest DNA on blot	Washing conditions	Location	Reference
SKI	cSKI	1.1	Eco R1	3 bands	Pst 1	2xSSC, 65°C	1q22	Nomura et al, '89
				20 & 9	Eco R1			
MUC	pMUC10	1.8	Eco R1	12	Eco R1	0.2xSSC, 65°C	1q22	Gendler et al, '87
AT3	pA62	1.6	Pst 1	1: 10 2: 5.5 & 5	Pst 1	2xSSC, 65°C	1q21-23	Middleton Price,'90
APOA2	pB3	1.3	Pst 1	20	Eco R1	1xSSC, 65°C	1q22	Middleton Price et al,'88
APCS	pSAp11	0.9	Pst 1	11.5 or 7.1	Msp 1	2xSSC, 65°C	1q22	Floyd-Smith et al,'86
REN	pHRnES1.9	0.7	Eco R1/SST 1	4.3	Eco R1	2xSSC, 65°C	1q41	O'Connell et al, 1989
				8.7	Hind III			

REFERENCES

REFERENCES

Abbott C and Povey S (1991) Development of human chromosome-specific PCR primers for characterization of somatic cell hybrids. Genomics 9:73-77.

Abramson DH, Ellsworth RM, Kitchin DF , Tung G (1984) Second nonocular tumors in retinoblastoma survivors, are they radiation-induced? Ophthalmology 91:1351-1355.

Adams JM (1985) Oncogene inactivation by fusion of chromosomes in leukaemia. Nature 315:542-543.

Alimena G, Billstrom R, Casalone R, Gallo E, Mitelman F , Pasquali F (1985) Cytogenetic pattern in leukaemic cells of patients with constitutional chromosome abnormalities. Cancer Genetics Cytogenetics 16:207-218.

Allen ND, Cran DG, Barton SC, Hettle S, Reik W , Surani MA (1988) Transgenes as probes for active chromosomal domains in mouse development. Nature 333:852-5.

Ames BN, Durston WE, Yamasaki E , Lee FD (1973) Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. Proc.Natl.Acad.Sci.USA 70:2281-2285.

Angrist M, Kauffman E, Slaugenhaupt SA, Matise TC, Puffenberger EG, Washington SS, Lipson A, Cass DT, Reyna T, Weeks DE , et al (1993) A gene for Hirschsprung disease (megacolon) in the pericentromeric region of human chromosome 10. Nature Genetics 4: 351-356.

Aslanidis C and de Jong PJ (1990) Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 18:6069-6074.

Baird PN, Santos A, Groves N, Jadresic L , Cowell JK (1992) Constitutional mutations in the WT1 gene in patients with Denys-Drash syndrome. Hum.Mol.Genet. 1:301-305.

Baker DL, Reddy UR, Pleasure D, Thorpe CL, Evans AE, Cohen PS , Ross AH (1989) Analysis of nerve growth factor receptor expression in human neuroblastoma and neuroepithelioma cell lines. Cancer Res 49: 4142-6.

Baker SJ, Markowitz S, Fearon Er, Willson JKV , Vogelstein B (1990) Suppression of human colorectal carcinoma cell growth by wild-type p53. Science 249: 912-915.

Baldini A, Ross M, Nizetic D, Vatcheva R, Lindsay EA, Herach H , Siniscalco M (1992) Chromosomal assignment of human YAC clones by fluorescence in situ hybridisation: use of single-yeast-colony PCR and multiple labelling. Genomics 14:181-184.

Ballabio A (1993) The rise and fall of positional cloning? Nature Genetics 3: 277-279

Basu TN, Gutmann DH, Fletcher JA, Glover TW, Collins FS , Downward J (1992) Aberrant regulation of ras proteins in malignant tumour cells from type 1 neurofibromatosis patients. Nature 356:713-715.

Beckwith JB, Kiviat NB , Bonadio JF (1990) Nephrogenic rests, nephroblastomatosis, and the pathogenesis of Wilms' tumor. Pediatr Pathol 10:1-36.

Beckwith JB and Perrin EV (1963) In situ neuroblastomas: a contribution to the natural history of neural crest tumors. Am.J.Pathol. 43:1089-1104.

Benedict WF, Banerjee A, Mark C , Murphree AL (1983) Nonrandom chromosomal changes in untreated retinoblastomas. Cancer Genet.Cytogenet. 311-333.

Beneitez J, Valcarcel E, Ramos C, Ayuso C , Sanchez Cascos A (1987) Frequency of constitutional chromosome alterations in patients with hematological malignancies. Cancer Genetics Cytogenetics. 24:345-354.

Berberich S, Hyde-DeRuyscher N, Espenshade P, Cole M (1992) max encodes a sequence-specific dNA-binding protein and is not regulated by serum growth factors. Oncogene 7:775-779.

Bernard LE, Brooks-Wilson AR , Wood S (1991) Isolation of DNA fragments from a human chromosomal subregion by Alu PCR differential hybridisation. Genomics 9:241-246.

Bernard LE and Wood S (1992) Human chromosome 5 sequence primer amplifies Alu polymorphisms on chromosome 2 and 17. Genome 36:302-309.

Bernstein ML, Leclerc JM, Bunin G, Brisson L, Robison L, Shuster J, Byrne T, Gregory D, Hill G, Dougherty G, et al (1992) A population-based study of neuroblastoma incidence, survival, and mortality in North America. J Clin Oncol 10:323-329.

Bernstein LR, Ferris DK, Colburn NH, Sobel ME (1994) A family of mitogen-activated protein kinase-related proteins interacts in vivo with activator protein-1 transcription factor. J. Biol. Chem 269 9401-4.

Biegel JA, Meek RS, Parmiter AH, Conard K , Emanuel BS (1991) Chromosomal translocation t(1;13)(p36;q14) in a case of rhabdomyosarcoma. Genes Chromosom Cancer 3:483-4.

Biegel JA, White PS, Marshall HN, Fujimori M, Zackai EH, Scher CD, Brodeur GM , Emanuel BS (1993) Constitutional 1p36 deletion in a child with neuroblastoma. Am.J.Hum.Genet. 52:176-182.

Billaud M, Isselbacher KJ , Bernards R (1993) A dominant-negative mutant of Max that inhibits sequence-specific DNA binding by Myc proteins. Proc Natl Acad Sci USA 90: 2739-2743.

Bird A (1992) The essentials of DNA methylation. Cell 70:5-8.

Bishop JM (1987) The molecular genetics of cancer. Science 235:305-311.

Blair DG, Oskarsson M, Wood TG, McClements WL, Fischinger PJ , Vande Woude GG (1981) Activation of the transforming potential of a normal cell sequence: a molecular model for oncogenesis. Science 212:941-943.

Blenis J (1993) Signal transduction via the MAP kinases: Proceed at your own RSK. Proc. Nat Acad Sci. USA 90: 5889-5892.

Boguski M.S. (1994) Bioinformatics Current Opinion in Genetics and Development 4:383-388.

Boland RP (1974) The neurocristopathies. A unifying concept of disease arising in neural crest maldevelopment. Human Pathol. 5:409-429.

Boldog FL, Gemmill RM, Wilke CM, Glover TW, Nilsson AS, Chandrasekharappa SC, Brown RS, Li FP , Drabkin HA (1993) Positional cloning of the hereditary renal carcinoma 3;8 chromosome translocation breakpoint. Proc.Natl.Acad.Sci.USA 90:8509-8513.

Bonaiti-Pellie C, Chompret A, Tournade M-F, Hochez J, Moutou C, Zucher J-M, Steschenko D, Brunat-Mentigny M, Roche H, Tron P, Frappaz D, Munzer M, Bachelot C, Dusol F, Sommelet-Olive D , Lemerle J (1992) Genetics and Epidemiology of Wilms' tumor: the French Wilms' tumor study. Med.Pediatr.Oncol. 20:284-291.

Borg A, Zhang Q-X, Alm P, Olsson H , Sellberg G (1992) The retinoblastoma gene in breast cancer - allele loss is not correlated with loss of gene protein expression. Cancer Research 52:2991-2994.

Bourhis J, de Veille F, Wilson GD, Hartmann O, Terrier LMJ, Boccon GL, McNally NJ, Lemerle J, Riou G , Benard J (1991) Combined analysis of DNA ploidy index and N-myc genomic content in neuroblastoma. Cancer Res 51:33-36.

Bowcock A, Osborne-Lawrence S, Barnes R, Chakravarti A, Washington S , Dunn C (1993a) Microsatellite polymorphism linkage map of human chromosome 13q. Genomics 15:376-386.

Bowcock A, Gerken S C, Barnes R, Shiang R, Jabs E W, Warren A C, Antionarakis s, Retief A E, Vergnaud G, Leppert M, Lalouel J-M, White R L, Cavalli-Sforza L L. (1993b) The CEPH consortium linkage map of human chromosome 13 Genomics 16: 486-496.

Bown NP, Pearson AD , Reid MM (1993) High incidence of constitutional balanced translocations in neuroblastoma [letter]. Cancer Genet Cytogenet 69:166-7.

Breukel C, Wijnen J, Tops C, van der Klift H, Dauwerse H , Meera Khan P (1990) Vector-Alu PCR: a rapid step in mapping cosmids and YACs. Nucleic Acids Res. 18:3097.

Britten RJ, Baron WF, Stout DB , Davidson EH (1988) Sources and evolution of human Alu repeated sequences. Proc.Natl.Acad.Sci.USA 85:4770-4774.

Brodeur GM (1989) Clinical significance of genetic rearrangements in human neuroblastomas. Clin Chem 35/7(B) B38-B42.

Brodeur GM, Azar C, Brother M, Hiemstra J, Kaufman B, Marshall H, Moley J, Nakagawara A, Saylors R, Scavarda N , et al (1992) Neuroblastoma. Effect of genetic factors on prognosis and treatment. Cancer 70:1685-1696.

Brodeur GM, Green AA, Hayes RA, Williams KJ, Williams DL , Tsiatis AP (1981) Cytogenetic features of neuroblastomas and cell lines. Cancer Res. 41:4678-4686.

Brodeur GM, Hayes FA, Green AA, Casper JT, Wasson J, Wallach S , Seeger RC (1987) Consistent N-myc copy number in simultaneous or consecutive neuroblastoma samples from sixty individual patients. Cancer Res 47:4248-53.

Brodeur GM and Seeger RC (1986) Gene amplification in human neuroblastomas: basic mechanisms and clinical implications. Cancer Genet Cytogenet 19:101-11.

Brodeur GM, Seeger RC, Barrett A, Berthold F, Castleberry RP, D'Angio G, De Bernardi B, Evans Ae, Favrot M, Freeman AI, Haase G, Hartmann O, Hayes FA, Helson L, Kemshead J, Lampert F, Ninane J, Ohkawa H, Philip T, Pinkerton CR, Pritchard J, Sawada T, Siegel S, Smith EI, Tsuchida Y , Voute PA (1988) International criteria for diagnosis, staging, and response to treatment of patients with neuroblastoma. J.Clin.Oncol. 6 :1874-1881.

Brodeur GM, Sekon GS , Goldstein MN (1977) Chromosome aberrations in human neuroblastomas. Cancer 40:2256-2263.

Brooks-Wilson AR, Goodfellow PN, Povey S, Nevanlinna HA, de Jong PJ , Goodfellow PJ (1990) Rapid cloning and characterisation of new chromosome 10 DNA markers by Alu element-mediated PCR. Genomics 7:614-620.

Buckler AJ, Chang DD, GRaw SL, Brook D, Haber DA, Sharp PA , Housman DE (1991) Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. Proc.Natl.Acad.Sci USA 88:4005-4009.

Bull PC, Barwell JA, Hannah HT-L, Pautler SE, Higgins MJ, Lalande M , Cox DW (1993) Isolation of new probes in the region of the Wilson disease locus, 13q14.2 -> q14.3. Cytogenet. Cell Genet. 64:12-17.

Burke DT (1990) YAC cloning: Options and problems. GATA 7:94-99.

Burke DT, Carle GF , Olson MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236:806-812.

Call KM, Glaser T, Ito CY, Buckler AJ, Pelletier J, Haber DA, Rose EA, Kral A, Yeger H, Lewis WH, Jones C , Housman DE (1990) Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumour locus. Cell 60:509-520.

Caron H, van Sluis P, van Hoeve M, J. dK, Bras J, Slater R, Mannens M, Voute PA, Westerveld A , Versteeg R (1993) Allelic loss of chromosome 1p36 in neuroblastoma is of preferential maternal origin and correlates with N-myc amplification. Nature Genetics 4:187-190.

Cavenee WK, Leach R, Mohandas T, Pearson P , White R (1984) Isolation and regional localisation of DNA segments revealing polymorphic loci for human chromosome 13. Science 228:501-503.

Charlieu JP, Laurent A-M, Carter DA, Bellis M , Roizes G (1992) 3'Alu PCR: a simple and rapid method to isolate human polymorphic markers. Nucl.Acids Res. 20:1333-1337.

Chatten J and Voorhess M (1967) Familial neuroblastoma. New England Journal Medicine 277:1230-1236.

Chen P-L, Chen Y, Bookstein R , Lee W-H (1990) Genetic mechanisms of tumor suppression by the human p53 gene. Science 250:1576-1580.

Cheng JM, Hiemstra JL, Schneider SS, Naumova A, Cheung NK, Cohn SL, Diller L, Sapienza C , Brodeur GM (1993) Preferential amplification of the paternal allele of the N-myc gene in human neuroblastomas. Nature Genetics 4: 191-194.

Chitayat D, Friedman JM , Dimmick JE (1990) Neuroblastoma in a child with Wiedemann-Beckwith syndrome. Am.J.Med.Genet. 35:433-436.

Christiansen H and Lampert F (1988) Tumour karyotype discriminates between good and bad prognostic outcome in neuroblastoma. Br.J.Cancer 57:121-126.

Christiansen H, Schestag J, Christiansen NM, Grzeschik K-H , Lampert F (1992) Clinical impact of chromosome 1 aberrations in neuroblastoma: a metaphase and interphase cytogenetic study. Genes, Chroms. Cancer 5:141-149.

Clarke AR, Purdie CA, Harrison DJ, Morris RG, Bird CC, Hooper ML , Wylie AH (1993) Thymocyte apoptosis induced by p53-dependent and independent pathways. Nature 362:849-852.

Clausen N, Andersson P , Tommerup N (1989) Familial occurrence of neuroblastoma, von Recklinhausen's neurofibromatosis, Hirschprung's Agangliosis and Jaw-winking syndrome. Acta.Paed.Scand. 78:736-741.

Cohen D, Chumakov I, Weissenbach J. (1993) A first generation physical map of the human genome. Nature 366: 698-701.

Cohn SL, Salwen H, Quasney MW, Ikegaki N, Cowan JM, Herst CV, Kennett RH, Rosen ST, DiGiuseppe JA , Brodeur GM (1990) Prolonged N-myc protein half-life in a neuroblastoma cell line lacking N-myc amplification. Oncogene 5:1821-1827.

Cole CG, Goodfellow PN, Bobrow M , Bentley DR (1991) Generation of novel sequence tagged sites (STSs) from discrete chromosomal regions using Alu-PCR. Genomics 10: 816-826.

Cole CG, Dunham I, Coffey AJ, Ross MT, Meier-Ewert S, Bobrow M , Bentley DR (1992a) A Random STS strategy for construction of YAC contigs spanning defined chromosomal regions. Genomics 14:256-262.

Cole CG, Patel K, Shipley J, Sheer D, Bobrow M, Bentley DR , Dunham I (1992b) Identification of region-specific yeast artifical chromosomes using pools of Alu element-mediated polymerase chain reaction probes labeled via linear amplification. Genomics 14: 931-938.

Cole MD (1986) The myc oncogene: its role in transformation and differentiation. Ann.Rev.Genet. 20: 361-384.

Collins FS (1992) Positional cloning: let's not call it reverse anymore [news]. Nature Genetics 1: 3-6.

Collins S and Groudine M (1982) Amplification of endogenous myc related DNA sequences in a human myeloid leukaemia cell line. Nature 298: 679-681.

Comings DE (1973) A general theory of carcinogenesis. Proc.Natl.Acad.Sci.USA 70: 3324-3328.

Cook S & McCormick F. (1994) Ras blooms on sterile ground Nature 369: 361-362.

Cotter FE, Das S, Douek E, Carter NP , Young BD (1991) The generation of DNA probes to chromosome 11p23 by Alu PCR on small numbers of flow-sorted 22q-derivative chromosomes. Genomics 9: 473-480.

Cotter FE, Hampton GM, Nasipuri S, Bodmer WF , Young BD (1990) Rapid isolation of human chromosome-specific DNA probes from a somatic cell hybrid. Genomics 7: 257-263.

Cowell JK (1982) Double minutes and homogeneously staining regions: gene amplification in mammalian cells. Ann.Rev.Genet. 16: 21-59.

Cowell JK. Manipulation of somatic cell hybrids for the analysis of the human genome. In: Rooney DE, Czepulkowski BH, ed. Human Cytogenetics a Practical Approach 2nd Edition. II ed. Oxford: IRL Press, 1992: 235-292. vol 9).

Cowell JK and Mitchell CD (1989) A somatic cell hybrid mapping panel for the regional assignment of human chromosome 13 specific DNA sequences. Cytogenet.Cell Genet. 52: 1-6.

Cowell JK, Rutland P, Hungerford J , Jay M (1988) Deletion of chromosome region 13q14 is transmissible and does not always predispose to retinoblastoma. Hum.Genet. 80: 43-45.

Cowell JK, Rutland P, Jay M , Hungerford J (1986) Deletions of the esterase-D locus from a survey of 200 retinoblastoma patients. Hum. Genet. 72: 164-167.

Cox Matise T, Perlin M , Chakravarti A (1994) Automated construction of genetic linkage maps using an expert system (multiMap):application to 1257 human microsatellite markers. Nature Genetics 6: 384-390.

Crespo P, Xu N, Simonds WF, Gutkind JS (1994) Ras-dependent activation of MAP kinase pathway mediated by G-protein βγ subunits. Nature 369: 418-420.

Croce CM and Nowell PC (1985) Molecular basis of human B cell neoplasia. Blood 65: 1-7.

Crouch DH, Fisher F, Clark W, Jayaraman PS, Goding CR , Gillespie DA (1993) Gene-regulatory properties of Myc helix-loop-helix/leucine zipper mutants: Max-dependent DNA binding and transcriptional activation in yeast correlates with transforming capacity. Oncogene 8: 1849-1855.

D'Angio GJ, Evans AE , Koop CE (1971) Special pattern of widespread neuroblastoma with a favourable prognosis. Lancet 1: 1046-1049.

Dalla-Favera R, Wong-Staal F , Gallo RC (1982) Onc gene amplification in promyelocytic cell line HL60 and primary leukaemic cells in the same patient. Nature 299: 61-63.

DeCaprio JA, Ludlow JW, Figge J, Shew J-Y, Huang C-M, Lee W-H, Marsilio E, Paucha E , Livingston DM (1988) SV40 large tumour antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. Cell 54: 275-283.

Djabali M, Selleri L, Parry P, Bower M, Young BD , Evans GA (1992) A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. Nature Genetics 2: 113-8.

Doolittle RF, Hunkapiller MW, Hood LE, Devare SG, Robbins KC, Aaronson SA , Antoniades HN (1983) Simian sarcoma virus onc gene, v-sis, is derived from the gene or genes encoding a platelet-derived growth factor. Science 221: 275-277.

Douglass EC, Rowe ST, Valentine M, Meyer WH , Thompson EI (1990) A second nonrandom translocation, der(16)t(1;16)(q21;q13), in Ewing sarcoma and peripheral neuroectodermal tumor. Cytogenetics and Cell Genetics 53: 87-90.

Downward J, Yarden Y, Mayes E, Scrace G, Totty N, Stockwell P, Ullrich A, Schlessinger J , Waterfield MD (1984) Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. Nature 307: 521-527.

Dracopoli NC, Harnett P, Bale SJ, Stanger BZ, Tucker MA, Housman DE , Kefford RF (1989) Loss of alleles from the distal short arm of chromosome 1 occurs late in melanoma tumor progression. Proc Natl Acad Sci U S A 86: 4614-8.

Draper GJ, Sanders BM , Kingston JE (1986) Second primary neoplasms in patients with retinoblastoma. Br.J.Cancer 53: 661-671.

Dryja TP, Rapaport JM, Weichselbaum R , Brusn GAP (1984) Chromosome 13 restriction fragment length polymorphism. Hum.Genet. 65: 320-324.

Dryja TP, Rapaport JM, Joyce JM, Petersen RA (1986) Molecular detection of deletions involving band q14 of chromosome 13 in retinoblastomas. Proc. Nat. Acad. Sci. USA 83: 7391-7394.

Dryja TP, McGee TL, Reichel E, Hahn LB, Cowley GS, Yandell DW, Sandberg MA, Berson EL (1990) Nature 343: 364-366.

Duyk GM, Kim S, Myers RM , Cox DR (1990) Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. Proc.Natl.Acad.Sci.USA 87: 8995-8999.

Dyson N, Howley PM, Munger K , Harlow E (1989) The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. Science 243: 934-936.

Egan SE & Weinberg RA (1993) The pathway to signal achievement. Nature 365: 781-783.

Ellmeier W, Aguzzi A, Kleiner E, Kurzbauer R , Weith A (1992) Mutually exclusive expression of a helix-loop-helix gene and N-myc in human neuroblastomas and in normal development. Embo J 11: 2563-2571.

Elvin P, Slynn G, Black D, Graham A, Butler R, Riley J, Anand R , Markham AF (1990) Isolation of cDNA clones using yeast artificial chromosome probes. Nucl.Acids Res. 18: 3913-3917.

Emery LG, Shields M, Shah NR , Garbes A (1983) Neuroblastoma associated with Beckwith-Wiedemann syndrome. Cancer 52: 176-9.

Eva A, Robbins KC, Andersen PR, Srinivasan A, Tronick SR, Reddy EP, Ellmore NW, Galen AT, Lautenberger JA, Pap as TS, Westin EH, Wong-Staal F, Gallo RC , Aaronson SA (1982) Cellular genes analogous to retroviral onc genes are transcribed in human tumour cells. Nature 295: 116-119.

Evan GI, Wyllie AH, Gilbert CS, Littlewood TD, Land H, Brooks M, Waters CM, Penn LZ , Hancock DC (1992) Induction of apoptosis in fibroblasts by c-myc protein. Cell 69: 119-128.

Evans AE, Chatten J, D'Angio GJ, Gerson JM, Robinson J , Schnaufer L (1980) A review of 17 IV-S neuroblastoma patients at the Children's Hospital of Philadelphia. Cancer 45: 833-839.

Evans AE, D'Angio GJ, Propert K, Anderson J , Hann HW (1987) Prognostic factors in neuroblastoma. Cancer 59: 1853-9.

Evans AE, D'Angio GJ , Randolph J (1971) A proposed staging for children with neuroblastoma. Children's cancer study group. Cancer 27: 374-378.

Evans R (1988) The steroid and thyroid hormone receptor superfamily. Science 240: 889-895.

Fearon ER and Vogelstein B (1990) A genetic model for colorectal tumorigenesis. Cell 61: 759-767.

Feinberg AP (1993) Genomic imprinting and gene inactivation in cancer. Nature Genetics 4: 110-113.

Feinberg AP and Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal.Biochem. 132: 6-13.

Feingold M, Gheradi GJ , Simons C (1971) Familial neuroblastoma and trisomy 13. Amer.J.Dis.Child. 121: 451.

Finklestein JZ, Klemperer MR, Evans A, Bernstein I, Leikin S, McCreadie S, Grosfeld J, Hittle R, Weiner J, Sather H , Hammond D (1979) Multiagent chemotherapy for children with metastatic neuroblastoma: a report from Childrens Cancer Study Group. Med Pediatr Oncol 6: 179-88.

Fischer EH, Charbonneau H, Tonks NK (1991) Protein tyrosine phosphatases: a diverse family of intracelluar and transmembrane enzymes. Science 253: 401-406.

Floyd-Smith G, Whitehead AS, Colton HR , Francke U (1986) The human c-reactive protein gene (CPR) and serum amyloid P component gene (APCS) are located on the proximal long arm of chromosome 1. Immunogenetics 24: 171-176.

Fong C-T, Dracopoli NC, White PS, Merrill PT, Griffiths RC, Housman DE , Brodeur GM (1989) Loss of heterozygosity for the short arm of chromosome 1 in human neuroblastomas: correlation with N-myc amplification. Proc.Natl.Acad.Sci. 86: 3753-3757.

Fong CT, White PS, Peterson K, Sapienza C, Cavenee WK, Kern SE, Vogelstein B, Cantor AB, Look AT , Brodeur GM (1992) Loss of heterozygosity for chromosomes 1 or 14 defines subsets of advanced neuroblastomas. Cancer Res 52: 1780-5.

Forster A and Rabbitts TH (1993) A method for identifying genes within yeast artificial chromosomes: application to isolation of MLL fusion cDNAs from acute leukaemia translocations. Oncogene 8: 3157-3160.

Fountain JW, Wallace MR, Bruce MA, Seizinger BR, Menon AG, Gusella JF, Michels VV, Schmidt MA, Dewald GW , Collins FS (1989) Physical mapping of a translocation breakpoint in neurofibromatosis. Science 244: 1085-1087.

Francke U and Kung F (1976) Sporadic bilateral retinoblastoma and 13q-chromosomal deletion. Med.Paediatr.Oncol. 2: 379-385.

Franke F, Rudolph B, Christiansen H, Harbott J , Lampert F (1986) Tumour karyotype may be important in the prognosis of human neuroblastoma. J Cancer Res.Clin.Oncol. 111: 266-272.

Friend SH, Bernards R, Rogelj S, Weinberg RA, Rapaport JM, Albert DM , Dryja TP (1986) A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323: 643-646.

Friend SH, Horowitz JM, Gerber MR, Wang X-F, Bogenmann E, Li FP , Weinberg RA (1987) Deletions of a DNA sequence in retinoblastomas and mesenchymal tumours: organization of the sequence and its encoded protein. Proc.Natl.Acad.Sci.USA 84: 9059-9063.

Gazit A, Igarashi H, Chiu IM, Srinivasan A, Yaniv A, Tronick SR, Robbins KC , Aaronson S (1984) Expression of the normal human sis/PDGF-2 coding sequence induces cellular transformation. Cell 39: 89-97.

Gendler SJ, Burchill JM, Duhig T, Lamport D, White R, Parker M , Taylor-Papadimitriou J (1987) Cloning of partial cDNA encoding differentiation and tumor-associated mucin glycoproteins expressed by human mammary epithelium. Proc.Natl.Acad.Sci.USA 84: 6060-6064.

Gerken S, Leppert M, O'Connell P, Cavenee W, James D, Ballard L, Stauffer D, Elsner T, Plaetke R, Lalouel JM, White R (1993) A genetic linkage map with 29 loci spanning human chromosome 13q. Genomics 16: 515-519.

Gessler M, Poustka A, Cavenee W, Neve RL, Orkin SH , Bruns GAP (1990) Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping. Nature 343: 774-778.

Gibbs JB, Sigal IS, Poe M , Scolnick EM (1984) Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. Proc.Natl.Acad.Sci.USA 81: 5704-5708.

Gilbert F, Feder M, Balaban G, Lurie DK, Podolsky R, Rinaldt V, Vinikoor N , Weisband J (1984) Human neuroblastomas and abnormalities of chromosome 1 and 17. Cancer Res 44: 5444-5449.

Godbout R, Dryja TP, Squire J, Gallie BL , Phillips RA (1983) Somatic inactivation of genes on chromosome 13 is a common event in retinoblastoma. Nature 304: 451-453.

Grausz JD and Auffray C (1993) Strategies in cDNA programs. Genomics 17: 530-532.

Gray MR, Colot HV, Guarente L , Rosbash M (1982) Open reading frame cloning: identification, cloning, and expression of open reading frame DNA. Proc Natl Acad Sci U S A 79: 6598-602.

Green ED and Olson MV (1990) Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. Science 250: 94-98.

Green ED and Olson MV (1990) Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. Proc.Natl.Acad.Sci.USA 87: 1213-1217.

Green ED, Riethman HC, Dutchik JE , Olson MV (1991) Detection and characterisation of chimeric yeast artificial chromosome clones. Genomics 11: 658-669.

Grundy P, Koufos A, Morgan K, Li FP, Meadows AT , Cavenee WK (1988) Familial predisposition to Wilms' tumour does not map to the short arm of chromosome 11. Nature 336: 375-376.

Grundy RG, Pritchard J, Baraitser M, Risdon A , Robards M (1992) Perlman and Wiedemann-Beckwith syndromes: two distinct conditions associated with Wilms' tumour. Eur J Pediatr 151: 895-898.

Gu W, Schneider JW, Condorelli G, Kaushal S, Mahdavi V , Nadal-Ginard B (1993) Interaction of myogenic factors and the retinoblastoma protein mediates muscle cell commitment and differentiation. Cell 72: 309-324.

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, Bernardi G, Lathrop M, Weissenbach J (1994) The 1993-94 Genethon human genetic linkage map. Nature Genetics 7: 246-251.

Haber DA, Buckler AJ, Glaser T, Call KM, Pelletier J, Sohn RL, Douglass EC , Housman DE (1990) An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumour. Cell 61: 1257-1269.

Haccard O, Sarcevic B, Lewellyn A, Hartley R, Roy L, Izumi T, Erikson E, Maller J. (1993) Induction of metaphase arrest in cleaving Xenopus embryos by MAP kinase. Science 262: 1262-1265.

Hanawa Y, Sawada T , Tsunoda A (1990) Decrease in childhood neuroblastoma death in Japan. Med Pediatr Oncol 18: 472-475.

Hansteen I-L, Varslot K, Steen-Johnsen J , Langard S (1982) Cytogenetic screening of a new-born population. Clinical Genet. 21: 309-314.

Harbour JW, Lai S-L, Whang-Peng J, Gazdar AF, Minna JD , Kaye FJ (1988) Abnormalities in structure and expression of the human retinoblastoma gene in SCLC. Science 241: 353-357.

Harris H, Miller OJ, Klein G, Worst P , Tachibana T (1969) Suppression of malignancy by cell fusion. Nature 223: 363-368.

Hastie ND (1992) Dominant neative mutations in the Wilms tumour (WT1) gene cause Denys-Drash syndrome - proof that a tumour-suppressor gene plays a crucial role in normal genitourinary development. Hum.Mol.Genet. 1: 293-295.

Hawkins MM, Draper GJ , Kingston JE (1987) Incidence of second primary tumours among cancer survivors. Br.J.Cancer 56: 339-347.

Hecht F, Hecht BK, Northup JC, Trachtenberg N, Wood TS , J.D. C (1982) Genetic analysis of neuroblastoma-long range studies. Cancer Genetics and Cytogenetics. 7: 227-230.

Hempstead BL, Martin-Zanca D, Kaplan DR, Parada LF , Chao MV (1991) High-affinity NGF binding requires coexpression of the trk proto-oncogene and the low affinity NGF receptor. Nature 350: 678-683.

Hermans A, Heiserkamp N, von Linden M, van Baal S, Meijer D, van der Plas D, Wiedemann L.M, Groffen J, Bootsme D. (1987) Unique fusion of bcr and c-abl genes in Philadelphia chromosome positive acute lymphoblastic leaukaemia. Cell 51 33-40.

Hiemstra JL, Schneider SS , Brodeur GM (1994) High-resolution mapping of the N-myc amplicon core domain in neuroblastomas. Advances Neuroblastoma Res. 4: 51-57.

Higgins MJ, Hansen MF, Cavenee WK , Lalande M (1989) Molecular detection of chromosomal translocations that disrupt the putative retinoblastoma susceptibility locus. Mol.Cell Biol. 9: 1-5.

Hinds P, Finlay C , Levine AJ (1989) Mutation is required to activate the p53 gene for cooperation with the ras oncogene and transformation. J.Virol. 63: 739-746.

Hogg A, Onadim Z, Baird PN , Cowell JK (1992) Detection of heterozygous mutations in the RB1 gene in retinoblastoma patients using single- strand conformation

polymorphism analysis and polymerase chain reaction sequencing. Oncogene 7: 1445-1451.

Hoheisel JD and Lehrach H (1993) Use of reference libraries and hybridisation fingerprinting for relational genome analysis. FEBS 325: 118-122.

Holland J, Coffey AJ, Gianelli F, Bentley D (1993) Vertical integration of cosmid and YAC resources for interval mapping on the X-chromosome. Genomics 15: 297-304.

Hollstein M, Sidransky D, Vogelstein B , Harris CC (1991) p53 mutations in human cancers. Science 253: 49-53.

Holton TA and Graham MW (1990) A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. Nucleic Acids Res. 19: 1156.

Hook EB and Hamerton JL. Frequency of chromosome abnormalities detected in consecutive newborn studies - differences between studies - results by sex and by severity of phenotypic involvement.New York: Academic Press, 1977: 63-79.

Hook EB, Healy NP , Willey AM (1989) How much difference does chromosome banding make? Adjustments in prevalence and mutation rates of structural cytogenetic abnormalities. Ann. Hum. Genet. 53: 237-242.

Hopkinson DA, Mestriner MA, Cortner J , Harris H (1973) Esterase-D: a new human polymorphism. Ann.Hum.Genet. 37: 119-137.

Hunt JD and Tereba A (1990) Molecular evaluation of abnormalities of the short arm of chromosome 1 in neuroblastoma. Genes Chroms.Cancer 2: 137-146.

Ikeda I, Ishizaka Y, Tahira T, Suzuki T, Onda M, Sugimura T , Nagao M (1990) Specific expression of the ret proto-oncogene in human neuroblastoma cell lines. Oncogene 5: 1291-1296.

Innis MA. PCR Protocols: a guide to methods and applications.London: Academic Press, 1990:54-60.

Jeffreys AJ, Wilson V , Thein SL (1985) Individual-specific 'fingerprints' of human DNA. Nature 316: 76-9.

Jelinek WR and Schmid CW (1982) Repetitive sequences in eukaryotic DNA and their expression. Ann.Rev.Biochem. 51: 813-844.

Jensen NA, Rodriguez ML, Garvey JS, Miller CA , Hood L (1993) Transgenic mouse model for neurocristopathy: Schwannomas and facial bone tumors. Proc Natl Acad Sci U S A 90: 3192-3196.

Joshi VV, Cantor AB, Brodeur GM, Look AT, Shuster JJ, Altshuler G, Larkin EW, Holbrook CT, Silverman JF, Norris HT, Hayes FA, Smith EI , Castleberry RP (1993) Correlation between morphologic and other prognostic markers of neuroblastoma. Cancer 71: 3173-3181.

Jurka J and Smith T (1988) A fundamental division in the Alu family of repeated sequences. Proceedings of National Academy of Science USA 85: 4775-4778.

Kafatos FC, Mitsialis SA, Spoerel N, Mariani B, Lingappa JR , Delidakis C (1985) Studies on the developmentally regulated expression and amplification of insect chorion genes. Cold Spring Harb Symp Quant Biol 50: 537-547.

Kaplan DR, Martin-Zanca D , Parada L (1991) Tyrosine phosphorylation and tyrosine kinase activity of the trk proto-oncogene product induced by NGF. Nature 350: 158-160.

Kariya Y, Kato K, Hayashizaki Y, Himeno S, Tarui S , Matsubara K (1987) Revision of consensus sequence of human Alu repeats - a review. Gene 53: 1-10.

Kerem B-S, Rommens JM, Buchanan JA, Markiewicz D, Cox TK, Chakravarti A, Buchwald M Tsui L-C (1989) Identification of the cystic fibrosis gene: genetic analysis. Science 245: 1073-1080.

Kinnier Wilson LM and Draper GJ (1974) Neuroblastoma, its natural history and prognosis: a study of 487 cases. Br.Med.J. 3: 301-307.

Klein G (1991) Comparative action of myc and bcl-2 in B-cell malignancy. Cancer Cells 3: 141-143.

Klein R, Silos Santiago I, Smeyne RJ, Lira SA, Brambilla R, Bryant S, Zhang L, Snider WD , Barbacid M (1994) Disruption of the neurotrophin-3 receptor gene trkC

eliminates la muscle afferents and results in abnormal movements. Nature 368: 249-251.

Klein R, Smeyne RJ, Wurst W, Long LK, Auerbach BA, Joyner AL , Barbacid M (1993) Targeted disruption of the trkB neurotrophin receptor gene results in nervous system lesions and neonatal death. Cell 75: 113-122.

Knudson AG (1977) Genetics and etiology of human cancer. Adv.Hum.Genet. 8: 1-66.

Knudson AG (1971) Mutation and cancer: statistical study of retinoblastoma. Proc.Natl.Acad.Sci.USA 68: 820-823.

Knudson AG and Strong LC (1972) Mutation and cancer: a model for Wilms' tumour of the kidney. J.Natl.Cancer Inst. 48: 313-324.

Knudson AG and Strong LC (1972) Mutation and cancer: neuroblastoma and pheochromocytoma. Am.J.Hum.Genet. 24: 514-532.

Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, Kunkel LM (1987) Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organisation of the DMD gene in normal and affected individuals. Cell 50: 509-517.

Korenberg JR and Rykowski MC (1988) Human genome organisation: Alu, Lines, and the molecular structure of metaphase chromosome bands. Cell 53: 391-400.

Koufos A, Grundy P, Morgan K, Aleck KA, Hadro R, Lampkin BC, Kalbakji A , Cavenee WK (1989) Familial Wiedemann-Beckwith syndrome and a second Wilms' tumour locus both map to 11p15.5. Am.J.Hum.Genet. 44: 711-719.

Kunkel LM, Monaco AP, Middlesworth W, Ochs HD, Latt SA (1985) Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. Proc. Nat. Acad. Sci. USA 82: 4778-4782.

Kushner BH, Gilbert F , Helson L (1986) Familial neuroblastoma: Case reports, literature review and etiologic considerations. Cancer 57: 1887-1893.

Kwok SCM, Ledley FD, DiLella AG, Robson KJH, Woo SL (1985) Nucleotide sequence of a full-length complementary DNA clone and amino acid sequence of human phenylalinine hydroxylase. Biochemistry 24: 556-561.

Lalande M, Dryja TP, Schreck RR, Shipley J, Flint A, Latt SA (1984) Isolation of human chromosome 13-specific DNA sequences cloned from flow sorted chromosomes and potentially linked to the retinoblastoma locus. Cancer Genet.Cytogenet. 13: 283-295.

Lamb J, Harris PC, Lindenbaum RH, Reeders ST, Wilkie AOM, Buckle VJ, Barton NJ, Weatherall DJ (1989) Detection of breakpoints in submicroscopic chromosomal translocation, illustrating an important mechanism for genetic disease. Lancet. 819-824.

Lamballe F, Klein R , Barbacid M (1991) trkC, a new member of the trk family of tyrosine protein kinases, is a receptor for neurotrophin-3. Cell 66: 967-979.

Land H, Parada LF , Weinberg RA (1983) Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature 300: 596-602.

Lane DP and Crawford LV (1979) T-antigen is bound to a host protein in SV40 transformed cells. Nature 278: 261-263.

Larin Z, Monaco AP , Lehrach H (1991) Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. Proc.Natl.Acad.Sci.USA 88: 4123-4127.

Laureys G, Speleman G, Opdenakker G, Benoit Y , Leroy J (1990) Constitutional translocation t(1;17)(p36;q12-21) in a patient with neuroblastoma. Genes Chroms.Cancer 2: 252-254.

Lee EY-HP and Lee WH (1986) Molecular cloning of the human esterase D gene, a genetic marker of retinoblastoma. Proc.Natl.Acad.Sci. 83: 6337-6341.

Lehrach H and et al. Genetic and physical mapping.Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1990:39-81.

Leister I, Weith A, Bruderlein S, Cziepluch C, Kangwanpong D, Schlag P , Schwab M (1990) Human colorectal cancer: high frequency of deletions at chromosome 1p35. Cancer Res 50: 7232-7235.

Lengauer C, Speicher MR, Popp S, Jauch A, Taniwaki M, Nagaraja R, Riethman HC, Donis KH, D'Urso M, Schlessinger D , et al (1993) Chromosomal bar codes produced by multicolor fluorescence in situ hybridization with multiple YAC clones and whole chromosome painting probes. Hum Mol Genet 2: 505-12.

Leone A, Seeger RC, Hong CM, Hu YY, Arboleda MJ, Brodeur GM, Stram D, Slamon DJ , Steeg PS (1993) Evidence for nm23 RNA overexpression, DNA amplification and mutation in aggressive childhood neuroblastomas. Oncogene 8: 855-65.

Levine AJ, Momand J , Finlay CA (1991) The p53 tumour suppressor gene. Nature 351: 453-456.

Li FP and Fraumeni JF (1969) Soft-tissue sarcomas, breast cancer and other neoplasms: a familial syndrome? Ann.Intern.Med. 71: 747-752.

Lipkowitz S, Gobel V, Varterasian ML, Nakahara K, Tchorz K , Kirsch IR (1992) A comparative structural characterization of the human NSCL-1 and NSCL-2 genes. Two basic helix-loop-helix genes expressed in the developing nervous system. J Biol Chem 267: 21065-71.

Look AT, Hayes FA, Shuster JJ, Douglass EC, Castleberry RP, Bowman LC, Smith E , Brodeur GM (1991) Clinical relevance of tumor cell ploidy and N-myc gene amplification in childhood neuroblastoma: a Pediatric Oncology Group study. J Clin Oncol 9: 581-591.

Lovett M, Kere J , Hinton LM (1991) Direct selection: a method for the isolation of cDNAs encoded by large genomic regions. Proc.Natl.Acad.Sci.USA 88: 9628-9632.

Lowe SW, Schmitt EM, Smith SW, Osborne BA , Jacks T (1993) p53 is required for radiation-induced apoptosis in mouse thymoctyes. Nature 362: 847-849.

Lowy DR, Rands E , Scolnick EM (1978) Helper independent transformation by unintegrated Harvey sarcoma virus DNA. J.Virol. 26: 291-298.

Ma A, Moroy T, Collum R, Weintraug H, Alt FW , Blackwell TK (1993) DNA binding by N- and L-myc proteins. Oncogene 8: 1093-1098.

Malkin D, Li FP, Strong LC, Fraumeni Jr JF, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA , Friend SH (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250: 1233-1238.

Mannens M, Devilee P, Bliek J, Manjes I, De Kraker J, Heyting C, Slater RM , Westerveld A (1990) Loss of heterozygosity in Wilms' tumours, studied for six putative tumour suppressor regions, is limited to chromosome 11. Cancer Res. 50: 3279-3283.

Martin GS (1970) Rous sarcoma virus: a function required for the maintenance of the transformed state. Nature 227: 1021-1023.

Martinsson T, Weith A, Cziepluch C , Schwab M (1989) Chromosome 1 deletions in human neuroblastomas: generation and fine mapping of microclones from the distal 1p region. Genes Chroms. Cancer 1: 67-78.

McGee TL, Yandell DW , Dryja TP (1989) Structure and partial genomic sequence of the human retinoblastoma susceptibility gene. Gene 80: 119-128.

Mellentin JD, Nourse J, Hunger SP, Smith SD , Cleary ML (1990) Molecular analysis of the t(1;19) breakpoint cluster region in pre-B cell acute lymphoblastic leukemias. Genes Chromosom Cancer 2: 239-47.

Mercer JFB, Livingston J, Hall B, Payntr JA, Begy C, Chandrasekharappa S, Lockhart P, Grimes A, Bhave M, Siemieniak D , Glover TW (1993) Isolation of a partial candidate gene for Menkes disease by positional cloning. Nature Genetics 3: 20-25.

Metcalf D (1985) The granulocyte-macrophage colony-stimulating factors. Science 229: 16-22.

Michalski AJ, Cotter, FE, Cowell JK. (1992) Isolation of chromosome-specific DNAfrom an Alu polymerase chain reaction library to define the breakpoint in a patient with a constitutional translocation t(1;13)(q22;q12) and ganglioneuroblastoma. Oncogene 7 1595-1602.

Michalski AJ and Cowell JK (1993) Assignment of 4 sequence tagged sites to 3 sub-regions of 13q12 using a somatic cell hybrid mapping panel. Genomics 18: 141-143.

Michalski AJ and Cowell KK. Constructing a physical map around a constitutional t(1;13)(q22;q12) breakpoint in a patient with ganglioneuroblastoma.New York: Wiley-Liss Inc., 1994:79-85.

Michna BA, McWilliams NB, Krummel TM, Hartenberg MA , Salzberg AM (1988) Multifocal ganglioneuroblastoma coexistent with total colonic aganglionosis. J. Ped. Surg 23: 57-59.

Middleton-Price HR, Van Den Berghe JA, Scott J, Knott TJ , Malcolm S (1988) Regional chromosomal localisation of APOA2 to 1q21-1q23. Hum.Genet. 79: 283-285.

Miozzo M, Pierotti MA, Sozzi G, Radice P, Bongarzone I, Spurr NK , Della Porta G (1990) Human TRK proto-oncogene maps to chromosome 1q32-q41. Oncogene 5: 1411-1414.

Mitchell CD and Cowell JK (1989) Predisposition to retinoblastoma due to a translocation within the 4.7R locus. Oncogene 4: 253-257.

Mitchell CD, Ventris JA, Warr TJ , Cowell JK (1991) Molecular definition in a somatic cell hybrid of a specific 2:13 translocation breakpoint in childhood rhabdomyosarcoma. Oncogene 6: 89-92.

Mohandas T, Shapiro LJ, Sparkes RS , Sparkes MC (1979) Regional assignment of the steoid sulfatase-X-linked ichthyosis locus: Implications for a noninactivated region on the short arm of the human X chromosome. Proc.Natl.Acad.Sci.USA 76: 5779-5783.

Moley JF, Brother MB, Fong CT, White PS, Baylin SB, Nelkin B, Wells SA , Brodeur GM (1992) Consistent association of 1p loss of heterozygosity with pheochromocytomas from patients with multiple endocrine neoplasia type 2 syndromes. Cancer Res 52: 770-4.

Monaco AP, Lam VMs, Zehetner G, Lennon GG, Douglas C, Nizetic D, Goodfellow PN , Lehrach H (1991) Mapping irradiation hybrids to cosmid and yeast artificial chromosome libraries by direct hybridisation of Alu-PCR products. Nucleic Acids Res. 19: 3315-3318.

Monaco AP, Neve RL, Colletti FC, Bertelson CJ, Kurnit DM , Kunkel LM (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature 323: 646-50.

Monaco AP, Bertelson CJ, Middlesworth W, Colletti CA, Aldridge J, Fischbeck KH, Bartlett R, Pericak-Vance MA, Roses AD, Kunkel LM (1985) Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA segment. Nature 316: 842-845.

Moorhead PS and Evans AE. Chromosomal findings in patients with neuroblastoma.New York: Raven Press, 1980:109-118.

Morgan JG, Dolganov GM, Robbins SE, Hinton LM , Lovett M (1992) The selective isolation of novel cDNAs encoded by the regions surrounding the human interleukin 4 and 5 genes. Nucl.Acids Res. 20: 5173-5179.

Morton NE (1991) Parameters of the human genome. Proc. Nat. Acad. Sci. USA 88: 7474-7476.

Moyzis RK, Torney DC, Meyne J, Buckingham JM, Wu JR, Burks C, Sirotkin CM , Goad WB (1989) The distribution of interspersed repetitive DNA sequences in the human genome. Genomics 4: 273-289.

Mulligan LM, Kwok JB, Healey CS, Elsdon MJ, Eng C, Gardner E, Love DR, Mole SE, Moore JK, Papi L , et al (1993) Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. Nature 363: 458-60.

Mullis KB and Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. Methods in Enzymol. 155: 335-350.

Nagano H, Kano Y, Kobuchi S , Kajitani T (1980) A case of partial 2p trisomy with neuroblastoma. Japanese J. Human Genetics 25: 39-45.

Nakagawara A, Azar CG, Scavarda NJ , Brodeur GM (1994) Expression and function of TRK-B and BDNF in human neuroblastomas. Mol Cell Biol 14: 759-67.

Nakagawara A, Sasazuki T, Akiyama H, Kawakami K, Kuwano A, Yokoyama T , Kume K (1990) N-myc oncogene and stage IV-S neuroblastoma. Preliminary observations on ten cases. Cancer 65: 1960-1967.

Nakamura Y, Leppert M, O'Connell P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E , et al (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235: 1616-22.

Nau MM, Brooks BJ, Battey J, Sausville E, Gaxdar AF, Kirsch IR, McBride OW, Bertness V, Hollis GF , Minna JD (1985) L-myc, a new myc-related gene amplified and expressed in human small cell lung cancer. Nature 318: 69-73.

Negroni, Scarpa S, Romeo A, Ferrari S, Modesi A , Raschella G (1994) Decrease of proliferation rate and induction of differentiation by a mycN antisense DNA oligmoer in a human neuroblastoma cell line. Cell Growth Differ. 2: 511-518.

Neil DL, Villasante A, Fisher RB, VEntrie D, Cox B , Tyler-Smith C (1990) Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. Nucl. Acids. Res. 18: 1421-1428.

Nelson DL (1990) Current methods for YAC clone characterisation. GATA 7: 100-106.

Nelson DL (1991) Interspersed repetitive sequence polymerase chain reaction (IRS PCR) for generation of human DNA fragments complex sources. Companion to methods in Enzymol. 2: 60-74.

Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989) Alu polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. Proc.Natl.Acad.Sci.USA 86: 6686-6690.

Nevins JR (1992) E2F: a link between the Rb tumor suppressor protein and viral oncoproteins. Science 258: 424-429.

Nomura N, Sasamoto S, Ishii S, Date T, Matski M , Ishizaki R (1989) Isolation of human cDNA clones of ski and the ski-related gene, sno. Nucl.Acids Res. 17: 5489-5500.

Nourse J, Mellentin JD, Galili N, Wilkinson J, Stanbridge E, Smith SD , Cleary ML (1990) Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. Cell 60: 535-45.

Nowell P and Hungerford DA (1960) A minute human chromosome in human granulocytic leukaemia. Science 132: 1497.

O'Connell P, Lathrop GM, Nakamura Y, Leppert ML, Ardinger RH, Murray JL, Lalaouel JM , White R (1989) Twenty-eight loci form a continous linkage map of markers for human chromosome 1. Genomics 4: 12-20.

Onadim Z and Cowell JK (1991) Application of PCR amplification from paraffin embedded tissue sections to linkage analysis in familial retinoblastoma. J.Med.Genet. 28: 312-316.

Pal N, Wadey RB, Buckle B, Yeomans E, Pritchard J , Cowell JK (1990) Preferential loss of maternal alleles in sporadic Wilms' tumour. Oncogene 5: 1665-1668.

Parada LF, Tabin CF, Shih C , Weinberg RA (1982) Human EJ bladder carcinoma oncogene is a homologue of Harvey sarcoma virus ras gene. Nature 297: 474-478.

Parimoo S, Patanjali SR, Shukla H , Chaplin DD , Weissman SM (1991) cDNA selection: efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. Proc.Natl.Acad.Sci USA 88: 9623-9627.

Patanjali SR, Parimoo S , Weissman SM (1990) Construction of a uniform-abundance (normalised) cDNA library. Proc.Natl.Acad.Sci.USA 88: 1943-1947.

Pavan WJ, Hieter P , Reeves RH (1990) Generation of deletion derivatives by targeted transformation of human-derived yeast artificial chromosomes. Proc Natl Acad Sci USA 87: 1300-4.

Pavan WJ, Hieter P , Reeves RH (1990) Modification and transfer into an embryonal carcinoma cell line of a 360-kilobase human-derived yeast artificial chromosome. Mol Cell Biol 10: 4163-9.

Pelicci P-G, Knowles DM, Magrath I , Dalla-Favara R (1986) Chromosomal breakpoints and structural alterations of c-myc locus differ in endemic and sporadic forms of Burkitt lymphoma. Proc.Natl.Acad.Sci.USA 83: 2984-2988.

Pelletier J, Bruening W, Kashtan CE, Mauer SM, Manivel JC, Striegel JE, Houghtin DC, Junien C, Habib R, Fouser L, Fine RN, Silverman BL, Haber DA , Housman D

(1991) Germline mutations in the Wilms' tumour suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. Cell 67: 437-447.

Penn LJ, Laufer EM , Land H (1990) C-MYC: evidence for multiple regulatory functions. Semin Cancer Biol 1: 69-80.

Persico MG, Viglietto G, Martini G, Toniolo D, Paonessa G, Moscatelli C, Dono R, Vulliamy T, Luzzatto L, D'Urso M (1986) Isolation of huamn glucose-6-phosphate dehydrogenase (G6PD) cDNA clones; primary structure of the protein and unusual 5' non-coding region. Nucleic Acids Res. 14: 2511-2522.

Peter M, Michon J, Vielh P, Neuenschwander S, Nakamura Y, Sonsino E, Zucker JM, Vergnaud G, Thomas G , Delattre O (1992) PCR assay for chromosome 1p deletion in small neuroblastoma samples. Int J Cancer 52: 544-8.

Petrukhin KE, Speer MC, Cayanis E, de Fatima Bonaldo M, Tantravahi U, Soares MB, Fischer SG, Warburton D, Gilliam TC, Ott J (1993) A microsatellite genetic linkage map of human chromosome 13. Genomics 15: 76-85.

Poddighe PJ, Ramaekers FC, Smeets AW, Vooijs GP , Hopman AH (1992) Structural chromosome 1 aberrations in transitional cell carcinoma of the bladder: interphase cytogenetics combining a centromeric, telomeric, and library DNA probe. Cancer Res 52: 4929-34.

Polymeropoulos MH, Rath DS, Xiao H , Merril CR (1991) Dinucleotide repeat polymorphism at the human fms-related tyrosine kinase gene (FLT1). Nucleic Acids Res 19: 2803.

Prempree T, Amornmarn R, Faillace WJ, Arce CA , Nguyen TQ (1993) 1;19 translocation in human meningioma. Cancer 71: 2306-11.

Prins J, De Vries EGE , Mulder NH (1993) The myc family of oncogenes and their presence and importance in small cell lung carcinomas and other tumour types. Anticancer Research 13: 1373-1386.

Pritchard J, Barnes JM, Germond SM , Wallendszus KR (1989) Gender and survival in neuroblastoma. Lancet 1: 8633.

Pritchard-Jones K and Fleming S (1991) Cell types expressing the Wilms' tumour gene (WT1) in Wilms' tumours: implications for tumour histogenesis. Oncogene 6: 2211-2220.

Pritchard-Jones K, Fleming S, Davidson D, Bickmore W, Porteous D, Gosden C, Bard J, Call K, Housman DE, Van Heyningen V , Hastie N (1990) The candidate Wilms' tumour gene is involved in genitourinary development. Nature 346: 194-197.

Radice P, Sozzi G, Miozzo M, De BV, Cariani T, Bongarzone I, Spurr NK, Pierotti MA, Della PG (1991) The human tropomyosin gene involved in the generation of the TRK oncogene maps to chromosome 1q31. Oncogene 6: 2145-2148.

Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE , Feinberg AP (1993) Relaxation of imprinted genes in human cancer. Nature 362: 747-9.

Rauscher FJ, Morris JF, Tournay OE, Cook DM , Curran T (1990) Binding of the Wilms' tumour locus zinc finger protein to the EGR-1 consensus sequence. Science 250: 1259-1262.

Ray PN, Belfall B, Duff C, Logan C, Kean V, Thompson MW, Sylvester JE, Gorski JL, Schmickel RD, Worton RG. (1985) Cloning of the breakpoint of an X;21 translocation associataed with Duchenne muscular dystrophy. Nature 318: 672-675.

Reddy EP, Reynolds RK, Santos E , Barbacid M (1982) A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature 300: 6447-6458.

Riccardi VM, Sujansky E, Smith AC , Francke U (1978) Chromosome imbalance in the aniridia-Wilms' tumour association: 11p interstitial deletion. Pediatrics 61: 604-610.

Richard F, Muleris M, Couturier J, Gerbault-Seureau M, Lombard M , Dutrillaux B (1992) Constitutional balanced translocations in patients with solid tumours. Cancer Genet.Cytogenet. 61: 50-52.

Riley J, Butler R, Oligivie D, Finniear R, Jenner D, Powell S, Anand R, Smith JC , Markham AF (1990) A novel rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic acid Res. 18: 2887-2890.

Riordan JR, Rommens JM, Kerem B-S, Alon N, Rozamahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou J-L, Drumm ML, Iannuzzi MC, Collins FS, Tsui L-C (1989) Identifiaction of the cystic fibrosis gene: cloning and characterisation of complementary DNA. Science 245: 1066-1073.

Ritke MK, Shah R, Valentine M, Douglass EC , Tereba A (1989) Molecular analysis of chromosome 1 abnormalities in neuroblastoma. Cytogenet Cell Genet 50: 84-90.

Robertson CM, Tyrrell JC , Pritchard J (1991) Familial neural crest tumours. Eur J Pediatr 150 :789-792.

Romani M, De AA, Alhadeff B, Purrello M, Gluzman Y , Siniscalco M (1990) Preferential integration of the Ad5/SV40 hybrid virus at the highly recombinogenic human chromosomal site 1p36. Gene 95: 231-241.

Rommens JM, Ianuzzi MC, Kerem B-S, Drumm ML, Melmer G, Dean M, Rozmahel R, Cole JL, Kennedy D, Hidaka N, Zsiga M, Buchwald M, Riordan JR, Tsui L-C, Collins FS (1989) Identifiaction of the cystic fibrosis gene: Chromosome walking and jumping. Science 245: 1059-1065.

Rosnet O, Marchetto S, deLapeyriere O , Birnbaum D (1991) Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1R family. Oncogene 6: 1641-50.

Rosnet O, Stephenson D, Mattei. M-G, Marchetto S, Shibuya M, Chapman VM , Birnbaum D (1993) Close physical linkage of the FLT1 and FLT3 genes on chromosome 13 in man and chromosome 5 in mouse. Oncogene. 8: 173-179.

Roussel MF, Downing JR, Rettenmier CW , Sherr CJ (1988) A point mutation in the extracellular domain of the human CSF-1 receptor (c-fms proto-oncogene product) activates its transforming potential. Cell 55: 979-988.

Rozakis-Adcock M, McGlade J, Mbamalu G, Pelicci G, Daly R, Li W, Batzer A, Thomas S, Brugge J, Pelicci PG, Schlessinger J, Pawson T. (1992) Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicaated in activation of the Ras paathway by tyrosine kinases. Nature 360 689-692.

Ruderman JV (1993) MAP kinase and the activation of quiescent cells. Current opinion in Cell Biol. 5: 207-213.

Rudolph B, Harbott J , Lampert F (1988) Fragile sites and neuroblastoma: fragile site at 1p13.1 and other points on lymphocyte chromosomes from patients and family members. Cancer Genet. Cytogenet. 31: 83-94.

Saiki R K., Gelfand DH, Stoffel S, Scharf SJ, Highuchi R, Horn GT, Mullis KB , Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA , Arnheim N (1985) Enzymatic amplification of ß-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anaemia. Science 230: 1350-1354.

Sanger F, Nicklen S , Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc.Natl.Acad.Sci.USA 74: 5463-5467.

Santos E, Tronick SR, Aaronson SA, Pulciani S , Barbacid M (1982) T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes. Nature 298: 343-347.

Savakis C and Doelz R (1993) Contamination of cDNA sequences in databases [letter]. Science 259: 1677-8.

Sawada T, Sugimoto T, Tanaka T, Kawakatsu H, Ishii T, Matsumura T , Horii Y (1987) Number and cure rate of neuroblastoma cases detected by the mass screening program in Japan: future aspects. Med Pediatr Oncol 15: 14-17.

Sawada T, Todo S, Fujita K, Iino S, Imashuku S , Kusunoki T (1982) Mass screening of neuroblastoma in infancy. Am J Dis Child 136: 710-712.

Scheffer H, Van der Lelie D, Aanstoot GH, Goor N, Nienhaus AJ, Vand der Hout AH, Pearson PL , Buys CH (1986) A somatic cell hybrid mapping panel for the regional assignment of human chromosome 13 specific DNA sequences. Cytogenet.Cell Genet. 52: 971-973.

Scherly D, Nouspikel T, Corlet J, Ulca C, Bairoch A , Clarkson SG (1993) Complementation of the DNA repair defect in xeroderma pigmentosum group G cells by a human cDNA related to yeast RAD2. Nature 363: 182-185.

Schlessinger D, Little RD, Freije D, Abidi F, Zucchi I, Porta G, Pilia G, Nagaraja R, Johnson SK, Yoon J-Y, Srivastava A, Kere J, Palmieri G, Ciccodicola A, Montanaro V, Romano G, Casamassimi A , D'Urso M (1991) Yeast artificial chromosome-based genome mapping:some lessons from Xq24-q28. Genomics 11: 783-793.

Schneider SS, Hiemstra JL, Zehnbauer BA, Taillon-Miller P, Le Paslier D, Vogelstein B , Brodeur GM (1992) Isolation and structural analysis of a 1.2 megabase N-myc amplicon from a human neuroblastoma. Mol. Cell Biol. 12: 5563-5570.

Schroeder WT, Chao L-Y, Dao DT, Strong LC, Pathak S, Riccardi M, Lewis WK , Saunders GF (1987) Nonrandom loss of maternal chromosome 11 alleles in Wilms'tumours. Am.J.Hum.Genet. 40: 413-420.

Schwab M, Alitalo K, Klempnauer KH, Varmus HE, Bishop JM, Gilbert F, Brodeur GM, Goldstein M , Trent J (1983) Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour. Nature 305: 245-248.

Schwab M and Amler LC (1990) Amplification of cellular oncogenes: a predictor of clinical outcome in human cancer. Genes Chromosom Cancer 1: 181-193.

Schwab M, Ellison J, Busch M, Rosenau W, Varmus HE , Bishop JM (1984) Enhanced expression of the human gene N-myc consequent to amplification of DNA may contribute to malignant progression of neuroblastoma. Proc Natl Acad Sci USA 81: 4940-4.

Schwab M, Varmus HE , Bishop JM (1985) The human N-myc gene contributes to tumorigenic conversation of mammalian cells in culture. Nature 316: 160-162.

Schwartz DC and Cantor CR (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37: 67-75.

Sealey PG, Whittaker PA , Southern EM (1985) Removal of repeated sequences from hybridisation probes. Nucleic Acids Research. 13: 1905-1922.

Seeger RC, Brodeur GM, Salter H, Dalton A, Siegel SE, Wong KY , Hammond D (1985) Association of multiple copies of the Nmyc oncogene with rapid progression of neuroblastomas. N.Eng.J.Med. 313: 1111-1116.

Sherr CJ, Rettenmier CW, Sacca R, Roussel MF, Look AT , Stanley ER (1985) The c-fms proto-oncogene product is related to the receptor for the mononuclear phagocyte growth factor, CSF-1. Cell 41: 665-676.

Shih C and Weinberg RA (1982) Isolation of a transforming sequence from a human bladder carcinoma cell line. Cell 29: 161-169.

Shimada H, Chatten J, Newton WJ, Sachs N, Hamoudi AB, Chiba T, Marsden HB , Misugi K (1984) Histopathologic prognostic factors in neuroblastic tumors: definition of subtypes of ganglioneuroblastoma and an age-linked classification of neuroblastomas. J Natl Cancer Inst 73: 405-16.

Silber JH, Evans AE , Fridman M (1991) Models to predict outcome from childhood neuroblastoma: the role of serum ferritin and tumor histology. Cancer Res. 51: 1426-1433.

Silverman GA, Jockel JI, domer PH, Mohn RM, Taillon-Miller P , Korsmeyer SJ (1991) Yeast artificial chromosome cloning of a two-megabase-size contig within chromosomal band 18q21 establishes physical linkage between BCL2 and plasminogen activator inhibitor type-2. Genomics 9: 2190228.

Simon D, Knowles BB , Weith A (1991) Abnormalities of chromosome 1 and loss of heterozygosity on 1p in primary hepatomas. Oncogene 6: 765-70.

Singer MF and Skowronski J (1985) Making sense out of LINES: long interspersed repeated sequences in mammalian genomes. TIBS 10: 119-122.

Sirinelli D, Silberman B, Baudon JJ, Sinnassamy P, Gruner M , Montagne JP (1989) Beckwith-Wiedemann syndrome and neural crest tumors. A report of two cases. Pediatr. Radiol. 19: 242-5.

Slater RM and Mannens MMAM (1992) Cytogenetics and molecular genetics of Wilms' tumor of childhood. Cancer Genet.Cytogenet. 61: 111-121.

Smeyne RJ, Klein R, Schnapp A, Long LK, Bryant S, Lewin A, Lira SA , Barbacid M (1994) Severe sensory and sympathetic neuropathies in mice carrying a disrupted Trk/NGF receptor gene. Nature 368: 246-249.

Sotelo AC, Gonzalez CF , Fowler JW (1980) Complete and incomplete forms of Beckwith-Wiedemann syndrome: their oncogenic potential. J Pediatr 96: 47-50.

Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal Molecular Biology 98: 503-517.

Sparkes RS, Sparkes MC, Wilson MG, Towner JW, Benedict W, Murphree AL , Yunis JJ (1980) Regional assignment of genes for human esterase-D and retinoblastoma to chromosome band 13q14. Science 208: 1042-1044.

Spencer CA and Groudine M (1991) Control of c-myc regulation in normal and neoplastic cells. Advances Cancer Res. 56 : 1-48.

Squinto SP, Stitt TN, Aldrich TH, Davis S, Bianco SM, Radziejewski C, Glass DJ, Masiakowski P, Furth ME, Valenzuela DM , et al (1991) trkB encodes a functional receptor for brain-derived neurotrophic factor and neurotrophin-3 but not nerve growth factor. Cell 65: 885-93.

Squire J, Gallie BL , Phillips RA (1985) A detailed analysis of chromosomal changes in heritable and non-heritable retinoblastoma. Hum.Genet. 70: 291-301.

Srivatsan ES, Ying KL , Seeger RC (1993) Deletion of chromosome 11 and of 14q sequences in neuroblastoma. Genes Chromosom Cancer 7: 32-37.

Stanbridge EJ (1990) Human tumor suppressor genes. Ann.Rev.Genet. 24: 615-657.

Stock C, Ambros IM, Mann G, Gadner H, Amann G , Ambros PF (1993) Detection of 1p36 deletions in paraffin sections of neuroblastoma tissues. Genes Chromosom Cancer 6: 1-9.

Sukumar S, Notario V, Martin-Zanca D , Barbacid M (1983) Induction of mammary carcinomas in rats by nitroso-methylurea involves malignant activation of H-ras -1 locus by single point mutations. Nature 306: 658-661.

Sutrave P and Hughes SH (1991) The ski oncogene. Oncogene 6: 353-356.

Suzuki T, Bogenmann E, Shimada H, Stram D , Seeger RC (1993) Lack of high-affinity nerve growth factor receptors in aggressive neuroblastomas. J Natl Cancer Inst 85: 377-84.

Suzuki T, Yokata J, Mugishima H, Okabe I, Ookuni M, Sugimura T , Tereda M (1989) Frequent loss of heterozygosity on chromosome 14 in neuroblastoma. Cancer Res. 49: 1095-1098.

Tabin CJ, Bradley SM, Barymann CI, Weinberg RA, Papageorge AG, Scdnick EM, Dhar R, Lowy DR , Chang EH (1982) Mechanisms of activation of a human oncogene. Nature 300: 143-149.

Tabor S and Richardson CC (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc.Natl.Acad.Sci.USA 84: 4767-4771.

Takayama H, Suzuki T, Mugishima H, Fujisawa T, Ookuni M, Schwab M, Gehring M, Nakamura Y, Sugimura T, Terada M , al. e (1992) Deletion mapping of chromosomes 14q and 1p in human neuroblastoma. Oncogene 7: 1185-1189.

Tanaka K, Yanoshita R, Konishi M, Oshimura M, Maeda Y, Mori T , Miyaki M (1993) Suppression of tumorigenicity in human colon carcinoma cells by introduction of normal chromosome 1p36 region. Oncogene 8: 2253-2266.

Tanaka T, Slamon DJ, Shimada H, Shimoda H, Fujisawa T, Ida N , Seeger RC (1991) A significant association of Ha-ras p21 in neuroblastoma cells with patient prognosis. A retrospective study of 103 cases. Cancer 68:1296-1302.

Tawn EJ and Earl R (1992) The frequencies of constitutional chromosome abnormalities in an apparently normal adult population. Mutat Res 283: 69-73.

Taylor CP, McGuckin AG, Bown NP, Reid MM, Malcolm AJ, Pearson AD , Sheer D (1994) Rapid detection of prognostic genetic factors in neuroblastoma using fluorescence in situ hybridisation on tumour imprints and bone marrow smears. United Kingdom Children's Cancer Study Group. Br J Cancer 69: 445-451.

The I, Murthy AE, Hanningan GE, Jacoby LB, Menon AG, Gusella JF , Bernards A (1993) Neurofibromatosis type 1 gene mutations in neuroblastoma. Nature Genetics. 3: 62-66.

Thiele CJ, Reynolds CP , Israel MA (1985) Decreased expression of N-myc precedes retinoic acid-induced morphological differentiation of human neuroblastoma. Nature 313: 404-406.

Ullu E and Tschudi C (1984) Alu sequences are processed 7SL RNA genes. Nature 312: 171-172.

van Corven EJ, Hordijk PL, Medema RH, Bos JL, Moolenaar WH (1993). Pertussis toxin-sensitive activation of p21ras by G protein-coupled receptor agonists in fibroblasts. Proc. Nat. Acad. Sci. USA 90: 1257-1261.

Van Heyningen V, Bobrow M, Bodmer WF, Gardiner SE, Povey S , Hopkinson DA (1975) Chromosome assignment of some human enzyme loci. Mitochondrial malate dehydrogenase to 7, mannosphosphate isomerase and pyruvate kinase to 15, and probably, esterase-d to 13. Ann.Hum.Genet. 38: 295-299.

Varesco L, Thomas HJW, Cotrell S, Murday V, Fennell SJ, Williams S, Searle S, Sheer D, Bodmer WF, Frischauf A-M , Solomon E (1989) CpG island clones from a deletion encompassing the gene for adenomatous polyposi coli. Proc.Natl.Acad.Sci.USA 86: 10118-10122.

Varmus HE (1984) The molecular genetics of cellular oncogenes. Ann.Rev.Genet. 18: 553-612.

Vetrie D, Vorechovsky I, Sideras P, Holland J, Davies A, Flinter F, Hammarstrom L, Kinnon C, Levinsky R, Bobrow M, Smith CIE, Bentley DR (1993) The gene involved in X-linked agammaglobulinaemia is a member of the *src* family of protein-tyrosine kinases. Nature 361: 226-233.

Vogel F (1979) Genetics of retinoblastoma. Hum. Genet. 52: 1-54.

Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AMM , Bos JL (1988) Genetic alterations during colorectal-tumour development. N.Eng.J.Med. 319: 525-532.

Vogelstein B, and Kinzler KW (1992) p53 function and dysfunction. Cell 70 523-526.

Wada RK, Seeger RC, Brodeur GM, Einhorn PA, Rayner SA, Tomayki MM , Reynolds CP (1993) Human neuroblastoma cell lines that express N-myc without gene amplification. Cancer 72: 3346-3354.

Wadey RB, Pal NP, Buckle B, Yeomans E, Pritchard J, Cowell JK (1990) Loss of heterozygosity in Wilms' tumour involves two distinct regions of chromosome 11. Oncogene 5: 901-907.

Walker AP, Muscatelli F , Monaco AP (1993) Isolation of the human Xp21 glycerol kinase gene by positional cloning. Hum.Mol.Genet. 2: 107-114.

Warburton D, Yu M-T, Tantravahi U, Lee C, Cayanis E, Russo J, Fischer SG (1993) Regional localisation of 32 Not1 and Hind III fragments from a human chromosome 13 library by a somatic cell hybrid panel and *in situ* hybridisation. Genomics 16: 355-360.

Washington SS, Bowcock AM, Gerken S, Matsunami N, Lesh D, Osborne-Lawrence SL, Cowell J, Ledbetter DH, White RL, Chakravarti A (1993) A somatic cell hybrid map of human chromosome 13. Genomics 18: 486-495.

Waterfield MD, Scrace GT, Whittle, Stroobant P, Johnsson A, Wasteson A, Westermark B, Heldin C-H, Huang JS , Deuel TF (1983) Platelet-derived growth factor is structurally related to the putative transforming protein p28-sis of simian sarcoma virus. Nature 304: 35-39.

Weber F, de VJ , Schaffner W (1984) An SV40 "enhancer trap" incorporates exogenous enhancers or generates enhancers from its own sequences. Cell 36: 983-92.

Weinberg RA (1985) The action of oncogenes in the cytoplasm and nucleus. Science 230: 770-776.

Weiss R, Teich N, Varmus H , Coffin J. RNA tumour viruses.New York: Cold Spring Harbor Laboratory Press, 1985

Weith A, Martinsson T, Cziepluch C, Bruderlein S, Amler LC, Berthold F , Schwab M (1989) Neuroblastoma consensus deletion maps to 1p36:1-2. Genes Chroms.Cancer 1: 159-166.

White PS, Kaufman BA, Marshall HN , Brodeur GM (1993) Use of the single-strand conformation polymorphism technique to detect loss of heterozygosity in neuroblastoma. Genes Chromosom Cancer 7: 102-108.

Whyte P, Buchkovitch KJ, Horowitz JM, Friend SH, Raybuck M, Weinberg RA , Harlow E (1988) Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature 334: 124-129.

Wiedemann HR (1983) Tumours and hemihypertrophy associated with Wiedemann-Beckwith syndrome. Eur.J.Pediatr. 141: 129.

Wiggs J, Nordenskjeld M, Yandell D, Rapaport J, Grondin V, Janson M, Werelius B, Petersen R, Craft A, Riedel K, Lieberfarb R, Walton D, Wilson W , Dryja TP (1988) Prediction of the risk of hereditary retinoblastoma using DNA polymorphisms within the retinoblastoma gene. New Eng.J.Med. 318: 151-157.

Woods WG, Tuchman M, Bernstein ML, Leclerc JM, Brisson L, Look T, Brodeur GM, Shimada H, Hann HL, Robison LL , et al (1992) Screening for neuroblastoma in North America. 2-year results from the Quebec Project. Am J Pediatr Hematol Oncol 14: 312-319.

Xu G, O'Connell P, Viskochil D, Cawthon R, Robertson M, Culver M, Dunn D, Stevens J, Gesteland R, White R , Weiss R (1990) The neurofibromatosis type 1 gene encodes a protein related to GAP. Cell 62: 599-608.

Yokota J, Wada M, Shimosato Y, Tereda M , Sugimura T (1987) Loss of heterozygosity on chromosome 3, 13 and 17 in small-cell carcinoma and on chromosome 3 in adenocarcinoma of the lung. Proc.Natl.Acad.Sci.USA. 84: 9252-9256.

Young JL, Gloeckler Ries L, Silverberg E, Horm JW , Miller RW (1986) Cancer incidence, survival, and mortality for children younger than age 15 years. Cancer 58: 598-602.

Yunis JJ and Ramsay N (1978) Retinoblastoma and subband deletion of chromosome 13. Am.J.Dis.Child. 132: 161-163.

Zabarovsky ER, Kashuba VI, Pokrovskaya ES, Zabarovsky VI, Wang J-Y, Berglund P, Boldog F, Stanbridge EJ, Sumegi J, Klein G , Winberg G (1993) Alu-PCR approach to isolating Not1 -linking clones frm the 3p14-p21 region frequently deleted in renal cell carcinoma. Genomics 16: 713-719.

Zelent A (1994) Translocation of the RARa locus to the PML or PLZF gene in acute promyelocytic leukaemia. Br.J.Haematol. 86: 451-460.

Zhu X, Dunn JM, Goddard AD, Squire JA, Becker A, Phillips RA , Gallie BL (1992) Mechanisms of loss of heterozygosity in retinoblastoma. Cytogenet.Cell Genet. 59: 248-252.

PUBLICATIONS

Michalski AJ, Cotter, FE, Cowell JK. (1992) Isolation of chromosome-specific DNA from an Alu polymerase chain reaction library to define the breakpoint in a patient with a constitutional translocation t(1;13)(q22;q12) and ganglioneuroblastoma. Oncogene 7 1595-1602.

Michalski AJ and Cowell JK (1993) Assignment of 4 sequence tagged sites to 3 sub-regions of 13q12 using a somatic cell hybrid mapping panel. Genomics 18 :141-143.

Michalski AJ and Cowell KK. Constructing a physical map around a constitutional t(1;13)(q22;q12) breakpoint in a patient with ganglioneuroblastoma.New York: Wiley-Liss Inc., 1994:79-85.

Oncogene (1992), 7, 1595–1602

Isolation of chromosome-specific DNA sequences from an Alu polymerase chain reaction library to define the breakpoint in a patient with a constitutional translocation t(1;13) (q22;q12) and ganglioneuroblastoma

Antony J. Michalski, Finbarr E. Cotter & John K. Cowell

ICRF Molecular Oncology Group, Institute of Child Health, 30 Guilford Street, London, WC1N 1EH, UK

We describe the cytogenetic and molecular characterization of a t(1;13)(q22;q12) constitutional rearrangement occurring in a patient with a relatively benign form of neuroblastoma, called ganglioneuroblastoma. Somatic cell hybrids were generated between mouse 3T3 cells and a lymphoblastoid cell line from this patient, D.G. One isolated subclone, DGF27C11, contained the derivative chromosome, 1pter–q22::13q12–qter, but no other material from either chromosome 1 or 13. Using available DNA probes the 13 breakpoint was assigned proximal to all reported markers. In order to generate flanking markers to define this translocation further, an Alu polymerase chain reaction library was constructed from a somatic cell hybrid containing only the proximal, 13pter–13q14, region of chromosome 13. Seven unique sequences have been isolated from the library, three of which lie below and four of which lie above the 13q12 breakpoint. More precise mapping of the distal markers was achieved using a panel of somatic cell hybrids with overlapping deletions of chromosome 13. The paucity of probes in the 1q22 region has made a precise assignment of this breakpoint difficult, however it has been shown to lie distal to c-*SKI* and proximal to *APOA2*. This refined characterization of the breakpoint is a prerequisite for its cloning, which may yield genes important in the pathogenesis of ganglioneuroblastoma.

Introduction

Neuroblastoma (Nb) is the most common solid tumour in childhood, accounting for 8% of all malignancy. In rare families the Nb tumour phenotype segregates as an autosomal dominant trait. This led Knudson & Strong (1972) to propose that tumorigenesis was a consequence of a loss of function of a critical gene in embryonic tumour cells in a way similar to that proposed for retinoblastoma (Knudson, 1971). Ganglioneuroblastoma, although clinically and histologically less aggressive than Nb, is considered to be part of the same disease spectrum. Because Nb families are rare (Kushner *et al.*, 1986), it has not been possible to assign the Nb predisposition gene to a particular chromosome region using conventional linkage studies. Analyses of tumour cells, however, have demonstrated frequent involvement of chromosome 1 in structural rearrangements. Thus, 70% of stage III and IV tumours and 85% of cell lines have cytogenetically detectable abnormalities involving the short arm (1p) of chromosome 1 (Brodeur *et al.*, 1977; Gilbert *et al.*, 1984). Although consistent, these 1p abnormalities often occur against a background of other structural rearrangements such that it is difficult to determine which are causal and which are related to tumour progression. Loss of heterozygosity (LOH), for particular chromosome regions, has been used to provide formal proof of Knudson's two-hit hypothesis of oncogenesis in many different tumours. In Nb tumours which carry 1p deletions LOH has been demonstrated and, in 90% of cases (Weith *et al.*, 1989), involves the 1p36–pter region, although other groups find a much lower frequency of abnormalities in this region (Fong *et al.*, 1989). However, LOH has also been demonstrated in Nb for markers on chromosome 14 (Suzuki *et al.*, 1989). Whilst suggestive, the evidence does not provide conclusive evidence for the involvement of genes in 1p in tumorigenesis.

Analysis of constitutional chromosome abnormalities in individuals who develop particular tumour types has usually indicated the site of the predisposition gene. This has been the case for retinoblastoma (Yunis & Ramsay, 1978), Wilms' tumour (Riccardi *et al.*, 1978), neurofibromatosis (Fountain *et al.*, 1989; O'Connell *et al.*, 1989) and familial adenomatous polyposis (Varesco *et al.*, 1989). These structural abnormalities are predominantly deletions, but where translocations are identified the breakpoints usually interrupt the predisposition gene (Mitchell & Cowell, 1989). Few cases of Nb with constitutional chromosome abnormalities have been reported in the literature (Moorhead & Evans, 1980; Nagano *et al.*, 1980; Sanger *et al.*, 1984; Laureys *et al.*, 1990). The characterization of a t(1;13)(q22;q12) rearrangement occurring in a patient with ganglioneuroblastoma is described.

Results

Case report

Patient D.G. presented at 5 years of age with a 2-year history of gradually enlarging, left-sided neck mass which a magnetic resonance imaging (MRI) scan showed to be separate from the cervical spine but deep to the carotid vessels. The diagnosis of a cervical ganglioneuroblastoma was made. There were no other abnormalities on examination, his urinary catecholamine metabolite levels were normal and bone marrow aspirates and trephines were free of malignant cells. These features suggest that the tumour was relatively

Correspondence: J.K. Cowell
Received 2 June 1991; accepted in revised form 11 March 1992

benign. A well-encapsulated ganglioneuroblastoma was removed surgically and he received no chemotherapy. Cytogenetic analysis showed a constitutional t(1;13)(q22;q12) translocation; cytogenetic studies of his parents and two siblings are normal. Five years later he is well, with no cancer recurrence.

Analysis of somatic cell hybrids

A lymphoblastoid cell line was established from D.G. and the presence of the t(1;13) chromosome rearrangement was confirmed. Somatic cell hybrids were generated between these cells and mouse 3T3 cells. Individual hybrid colonies were investigated cytogenetically and in one, DGF27, most cells (Figure 1) apparently contained the derivative chromosome 1 (1pter–q21;13q12–qter). Single-cell clones were generated from this parental hybrid stock, producing a subclone, DGF27C11, which contained the der(1) chromosome.

Cytogenetic analysis of this hybrid failed to identify normal copies of either chromosome 1 or 13 or the derivative 13.

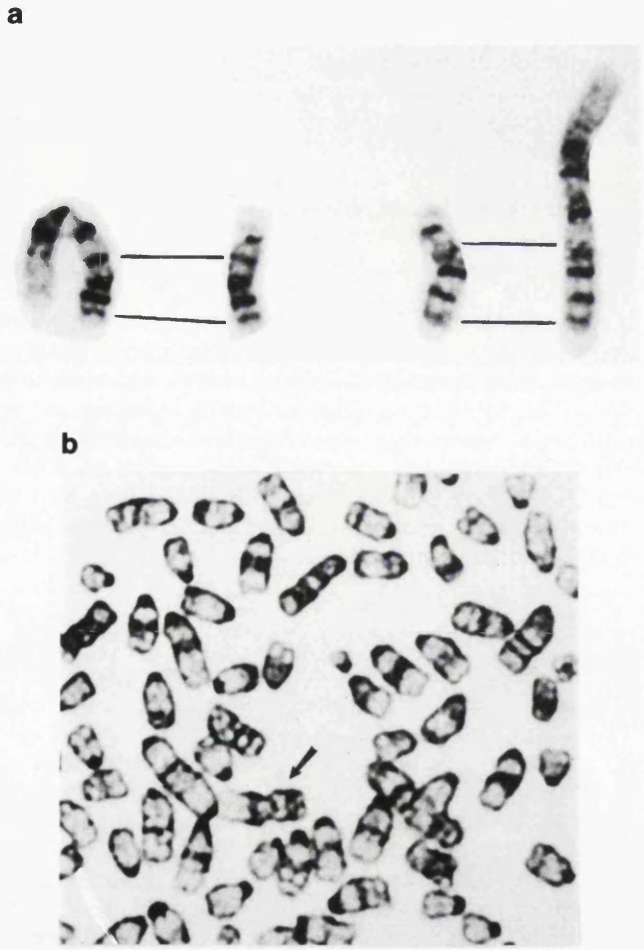


Figure 1 Partial trypsin–Giesma banded metaphase chromosomes from patient D.G. (**a**). The derivative (1;13) chromosome in hybrid DGF27C11 (**b**) is indicated by an arrow

Molecular analysis of DGF27C11

As cryptic rearrangements between human and rodent chromosomes can occur in somatic cell hybrids, hybrid DGF27C11 was analysed further using molecular probes from chromosomes 1 and 13.

An idiogram of normal copies of these chromosomes is shown in Figure 2 to illustrate the relative positions of the DNA probes used, and further information about these probes is given in Table 1. The presence of the retinoblastoma susceptibility gene, *RB1*, was demonstrated using polymerase chain reaction (PCR) to amplify exon 13. Under the conditions used for PCR, only the human sequence is amplified in a mouse background (P. Baird, personal communication). Amplification of exon 13 provides a rapid means of monitoring loss of the der(1) chromosome in this hybrid. The presence of *RB1* was further confirmed using conventional Southern blotting and probing procedures. There are few DNA probes with unequivocal assignments to the 13q12–13 region, but both D13S1 (Dryja *et al.*, 1984) and D13S6 (Cavenee *et al.*, 1984) were present in DGF27C11. Although located in the q12–q13 region of chromosome 13, their exact location is still unknown (Mitchell *et al.*, 1991). It is likely that the chromosome 13 breakpoint lies in the most proximal part of band q12, but with existing molecular probes we cannot yet exclude the presence of the pter–q12 region of chromosome 13 in the derivative chromosome.

Several probes were used to establish the position of the breakpoint on chromosome 1. The locus for the oncogene *SKI*, located on the long arm of chromosome 1, was present in DGF27C11, but more distal loci, *MUC*, *APCS* *APOA*$_2$ and REN, were not. As expected, short arm sequences for *AMY* and *MYCL* were present confirming the exclusive presence of der(1) in hybrid DGF27C11 (Figure 3).

The der(1) chromosome from D.G. has allowed us to establish the relative position of existing 1q probes more precisely than previously reported (showing *SKI* to be more centromeric than *APOA*$_2$, *APCS* and *MUC*).

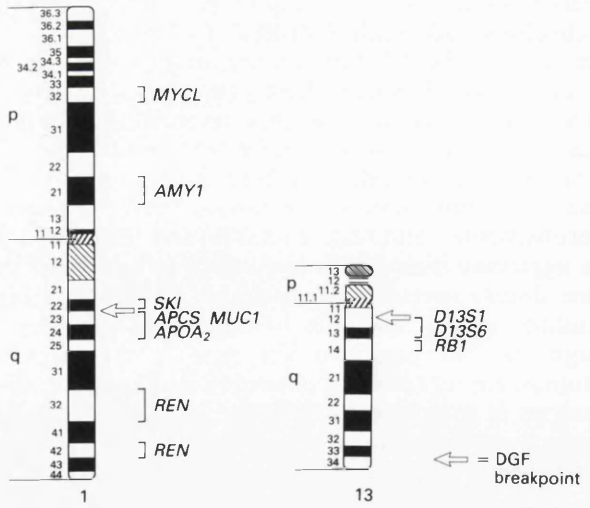


Figure 2 Idiogram of normal copies of chromosomes 1 and 13 showing the consensus map location of the probes used in the characterization of hybrid DGF27C11. The sites of the breakpoints involved in the reciprocal translocation are indicated by open arrows

Table 1 Description of the probes used to characterize DGF27C11

Locus	*Probe*	*Size (kb)*	*Enzyme*	*Reference*
Chromosome 1				
MYCL	pLmyc10	1.8	SmaI/EcoRI	Nau *et al.* (1985)
AMY1	pEB8	0.55	EcoRI/HindIII	M. Meisler (personal communication)
SKI	cski 1	1.1	EcoRI	Nomura *et al.* (1989)
APCS	pSAp11	0.92	PstI	Floyd-Smith *et al.* (1986)
APOA$_2$	pB3	1.3	PstI	Middleton-Price *et al.* (1988)
MUC	pmuc10	1.8	EcoRI	Gendler *et al.* (1987)
REN	pren	0.73	PvuII	T.J. Knott (personal communication)
Chromosome 13				
D13S1	p7F12	4.2	HindIII	Cavenee *et al.* (1984)
D13S6	pHU10	2.3	HindIII	Dryja *et al.* (1984)

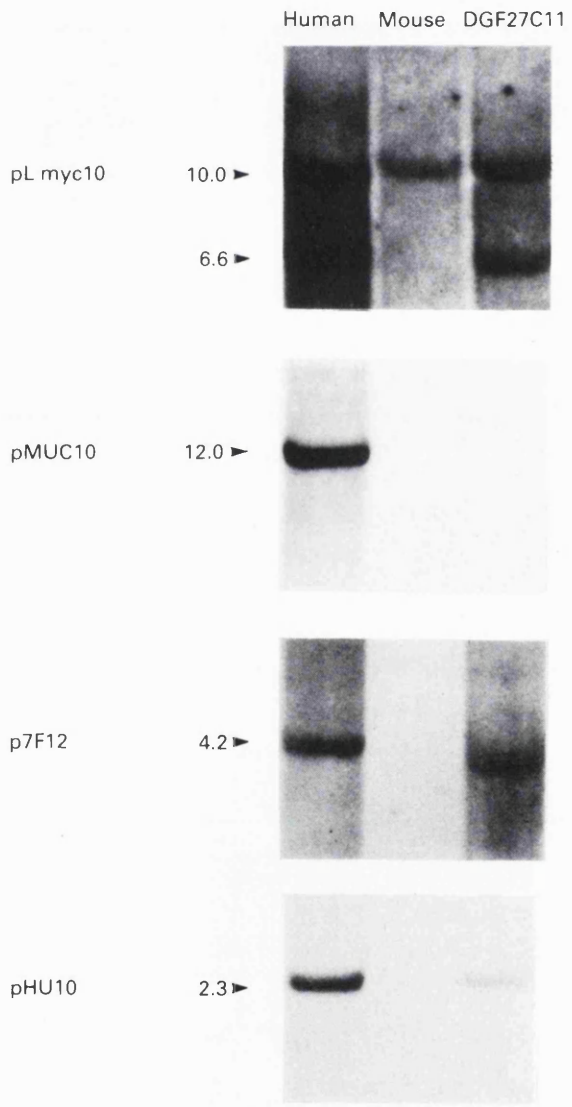


Figure 3 Southern blot analysis of DNA from hybrid DGF27C11. The hybridization profiles for 1p (Lmyc), 1q (pmuc) and proximal 13q (p7F12, pHU10) markers are shown, confirming the exclusive presence of the der(1) chromosome in this hybrid. In all cases the size of the bands following digestion with either EcoRI (Lmyc, pmuc) or HindIII (p7F12, pHU10) is indicated in kilobases. Note the cross-hybridization of the human L-*myc* probe with mouse DNA

Screening and characterization of Alu PCR-generated library

In order to characterize the position of the 13q12 breakpoint further, an Alu PCR library was made from a somatic cell hybrid (ICD) containing the 13pter–q14 region as the only human component (Scheffer *et al.*, 1986). Molecular characterization of ICD showed that the retinoblastoma predisposition gene is present, placing the breakpoint in the distal q14 or proximal q21 region.

In the first instance DNA from ICD was amplified using a single primer from the 3′ consensus region of the human Alu repetitive element as described by Cotter *et al.* (1990). This procedure selectively amplifies regions in the genome which contain Alu elements in opposite orientations. The limited proficiency of the Taq polymerase enzyme means that sequences amplified in these experiments are usually less than 2 kb long. A typical profile of DNA following Alu PCR amplification can be seen in Figure 4. Although a background smear can be seen there are typically 14–20 dominant bands which are more or less characteristic for the chromosome used. Under the conditions used no amplification was seen from mouse or hamster DNA.

As blunt-end cloning is very inefficient, only 2×10^5 colonies per μg were recovered. Using the X-gal selection system approximately 200 white colonies were gridded for further analysis. Hybridization to the colonies *in situ* with total human DNA identified 22 as containing repetitive sequences. The presence of a 30-bp Alu sequence at either end of the clones did not seem to result in positive hybridization under these circumstances. DNA from the remaining clones was amplified directly from the bacteria as outlined in the Materials and methods and resolved on 2% agarose gels. Fifty-four colonies did not yield inserts. Standard minipreps of plasmid DNA were made from 12 of these 54 samples and enzyme digestion confirmed the absence of cloned DNA.

In all, 57 colonies carrying inserts were recovered and used as probes against the mapping panel of somatic cell hybrids described in Figure 5. Thirty probes were demonstrated to be fully unique sequences acting as single-copy probes on the autoradiographs (Figure 6). Many of the inserts were similar in size and, therefore, potentially the same sequence (Figure 7). Double digests, using TaqI and RsaI in combination, provided 'fingerprints' of the inserts, enabling identification of identical sequences.

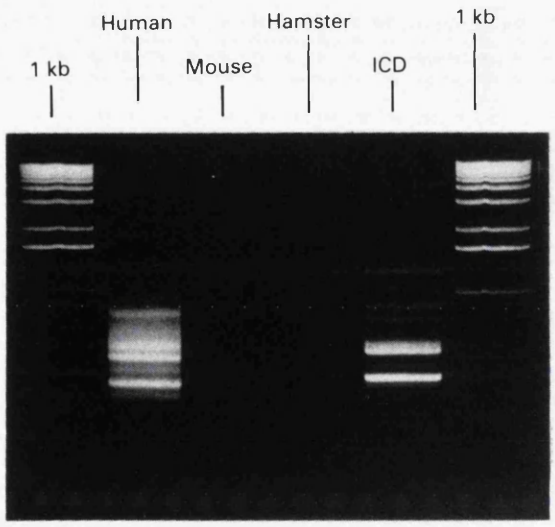


Figure 4 Typical profile following 3′ Alu PCR of human, mouse 3T3, Chinese hamster and ICD DNA

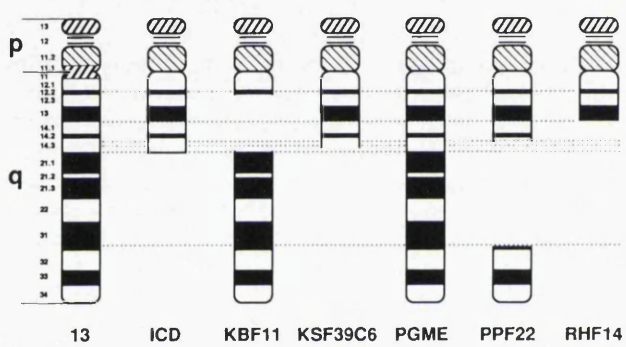


Figure 5 Idiogram of the somatic cell hybrid mapping panel used to map the position of the Alu PCR products by Southern hybridization. The evidence for the exact localization of the breakpoint in KBF11 is based on cytogenetic analysis (Cowell & Mitchell, 1989)

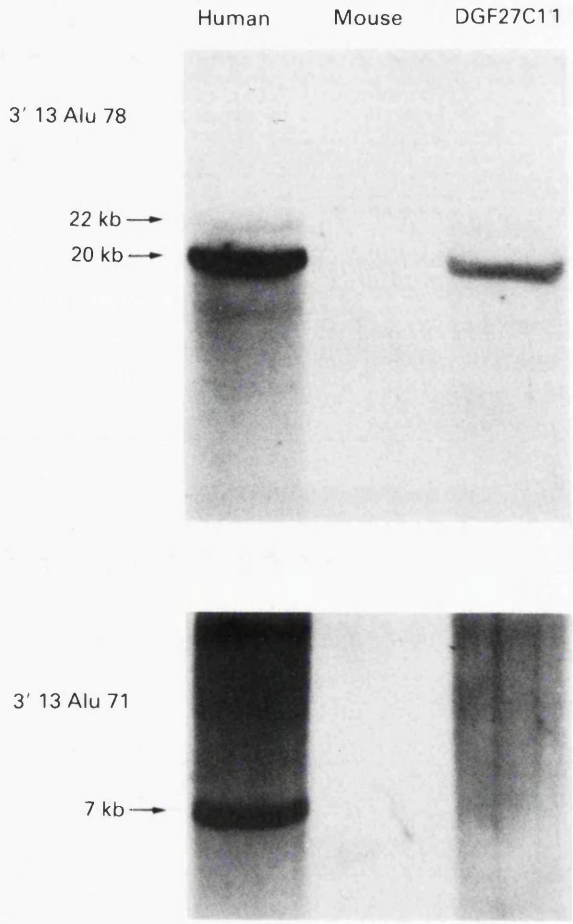


Figure 6 Pattern on autoradiography of Southern blots of EcoRI-digested human, mouse and DGF27C11DNA, probed with Alu PCR probes above and below the DGF breakpoint

The chromosomal localization of these seven probes (Table 2) was determined using the panel of hybrids shown in Figure 5. Four of the probes were absent from DGF27C11 and present in KBF11, localizing the D.G. breakpoint above that in KBF11, and confirming our original suggestion that the KBF11 breakpoint lies in the distal part of 13q12 (Cowell & Mitchell, 1989). However, it was not possible to say whether these probes are located on the short arm of chromosome 13. The remaining three probes were present in DGF27C11 and KBF11, placing them in the 13q12 region (Figure 8).

Since the 3′ Alu probes obviously represent a highly selected group of sequences, we tried to extend the representation of the region of interest by using a 5′ Alu primer. However, the 5′ primer amplified sequences from both mouse and Chinese hamster templates. The pattern of bands produced by this procedure was different in all three species, though some bands were

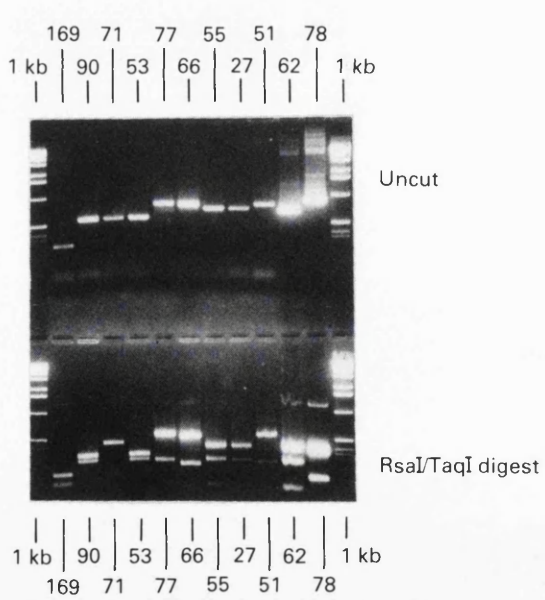


Figure 7 Fingerprint obtained by digesting unique 3′ 13 Alu PCR probes with TaqI and RsaI. 3′ 13 Alu 90 and 53 give identical patterns, as do 3′ 13 Alu 66 and 51.

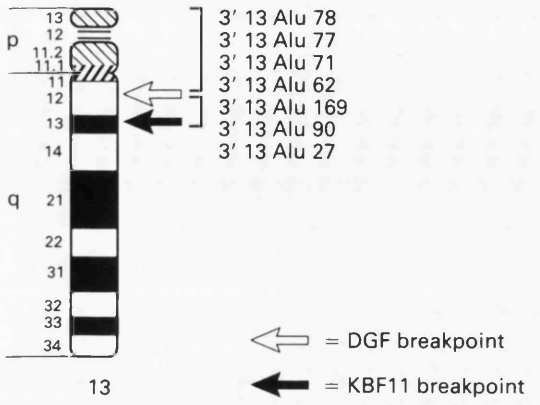


Figure 8 Idiogram showing the location of the eight unique Alu PCR-generated probes relative to the breakpoints in DGF27 (open arrow) and KBF11 (solid arrow)

Table 2 Characteristics of the seven unique 3′ 13 Alu PCR probes

3′ 13 Alu number	*Probe size (kb)*	*Size**	*Presence in DGF27C11*
169	0.34	22	Yes
90	0.55	20	Yes
78	0.90	22 and 20 doublet	No
77	0.85	4	No
71	0.55	7	No
62	0.75	5	No
27	0.75	7	Yes

*Size of band on autoradiograph following EcoRI digestion of total human DNA

shared. Progressively increasing the annealing temperature suppressed the amplification of rodent sequences but also decreased the number of human bands produced. At 72°C no rodent sequences were amplified but only three bands were produced from ICD, all of which unfortunately yielded repetitive sequences.

Discussion

There are now several examples of patients with specific genetic abnormalities and constitutional chromosome rearrangements in whom the breakpoint(s) interrupts a gene critical in development of the genetic condition. In retinoblastoma for example, the predisposition gene is known to lie in chromosome region 13q14 (Sparkes *et al.*, 1983). Shortly after the cloning of the candidate gene by Friend *et al.* (1986), we were able to show that a constitutional, predisposing translocation breakpoint lies within the *RB1* gene (Mitchell & Cowell, 1989). Several similar translocations have also been described by Higgins *et al.* (1989). Constitutional translocations have also been shown to interrupt critical genes on the X chromosome in Duchenne muscular dystrophy (Ray *et al.*, 1985) and on chromosome 17 in patients with neurofibromatosis type 1 (Viskochil *et al.*, 1990). Although the gene for aniridia has not yet been cloned, breakpoints have been described in 11p13, a region known to contain the aniridia locus (Moore *et al.*, 1986). So consistent is the finding that translocation breakpoints involve regions critical in determining a patient's phenotype that they warrant investigation. However, several of the translocations described previously in patients with Nb (Moorhead & Evans, 1980; Nagano *et al.*, 1980; Sanger *et al.*, 1984) are familial and not all translocation carriers develop Nb. This makes it unlikely that these rearrangements are critical in tumorigenesis. Therefore, our patient, D.G., represents one of only two cases with a potentially predisposing constitutional translocation in a patient with a neural crest tumour. In the other report (Laureys *et al.*, 1990), a t(1;17) translocation was described with the breakpoint in the region of 1p36. This breakpoint lies within the site frequently showing structural chromosome abnormalities as well as LOH (Martinsson *et al.*, 1989; Hunt & Tereba, 1990).

It is difficult to determine whether structural abnormalities in Nb tumours and cell lines are causes of tumorigenesis or consequences of it. Deletions of 1p are found almost exclusively in advanced-stage, aggressive tumours and are of poor prognostic significance (Christiansen & Lampert, 1988). It is possible, therefore, that disruption of genes in 1p35 is associated with disease progression rather than initiation. Indeed, several other tumour types show loss of 1p markers as they progress (Leister *et al.*, 1990). The fact that patient D.G. had a 'low-grade' type of tumour may indicate that a gene involved in predisposition to this type of Nb lies in the region 1q22. It is probable, therefore, that several different genes are involved in Nb development. In other tumours, too, mutations at more than one locus seem to be necessary for tumorigenesis. The development of metastatic colon carcinoma, for example, depends on the loss of at least three tumour-suppressor genes as well as the activation of the *ras* oncogene (Vogelstein *et al.*, 1988). In children's tumours, retinoblastoma proved to be explicable on the basis of loss of function of a single tumour-suppressor gene, but other tumours are more complicated. In Wilms' tumour, for example, although predisposition is related to deletion of genes from 11p13, LOH studies have shown that 11p15 may be the site of a second predisposition gene (Mannens *et al.*, 1988; Wadey *et al.*, 1990). In family linkage studies Wilms' tumour predisposition is not linked to markers on chromosome 11 (Grundy *et al.*, 1988; Huff *et al.*, 1988), implying that a third gene is associated with tumorigenesis. The identification of a constitutional translocation in a patient with ganglioneuroblastoma, therefore, provides an opportunity for isolating genes at the breakpoint which are important in tumorigenesis. To this end, we have generated unique markers flanking the breakpoint using the technique of Alu PCR.

The Alu family is the major group of short interspersed repeat DNA (SINES) in humans. It exists as a 280-bp dimer for which a computer-generated consensus sequence has recently been described (Kariya *et al.*, 1987). It is not entirely species specific, existing as a 170-bp monomer in rodent genomes. However, there is significant sequence divergence between humans and rodents (Jelinek & Schmid, 1982) and primers designed to maximize this divergence have been used to specifically amplify human sequences (Cotter *et al.*, 1990). The specificity of the 3′ Alu primers for human DNA has been noted previously, but our experience with the 5′ Alu primer yielded no single-copy human-

specific sequences. The pattern of the bands produced was species specific, with some shared bands. Increasing the stringency of the annealing step by progressively increasing the temperature led to a decrease in the number of rodent sequences amplified, but human sequences were also lost. Alu PCR methodology has been used to generate human-specific DNA sequences from somatic cell hybrids containing small amounts of human DNA in a large rodent background (Nelson *et al.*, 1989; Brooks-Wilson *et al.*, 1990; Cotter *et al.*, 1990; Cole *et al.*, 1991; Guzzetta *et al.*, 1991). Our results differ from those previously published in that we found a much higher percentage of repetitive sequences, 52% vs 15% (Cotter *et al.*, 1990) or 39% (Brooks-Wilson *et al.*, 1990). This may be specific to the region under study. The proportion of non-recombinants (white colonies containing no insert) was appreciable (34%), but there are few published data to compare this with. Damage to the β-galactosidase locus in the course of EcoRV digestion to linearize the plasmid, gel purification or excision and purification would explain this.

The 30 unique sequences isolated were a mixture of only seven different probes. The reason these sequences became so numerically dominant is unclear. We know that libraries generated by an Alu PCR method are by no means representative of the genomic DNA used as a template.

Although there are 9×10^5 copies of the Alu repeat in the haploid human genome, giving an average inter-Alu distance of 4 kb, they are not evenly distributed (Moyzis *et al.*, 1989). There seem to be Alu-poor regions with more than 10 kb in between each Alu repeat, far greater than any of our PCR products, and Alu-rich regions with an average spacing of 1 kb. However, repeats less than 500 bp apart are preferentially orientated in the same direction and hence would not be amplified using a single Alu primer. Indeed, Alu PCR only amplifies one sequence per 500–1000 kb of template DNA (Nelson *et al.*, 1989; Cotter *et al.*, 1991). Within the human Alu family, some members will have less homology to our primers than others and hence will not be efficiently amplified. Preferential ligation of some inter-Alu sequences may also occur. All of the above would explain the numerical dominance of our seven unique sequences.

The acquisition of flanking markers is a vital first step in the isolation and characterization of the t(1;13) breakpoint. Alu PCR methodology has enabled us to generate seven unique sequences, and we intend to use these to screen a yeast artificial chromosome (YAC) library either by Southern hybridization or by sequencing the probes, designing oligonucleotide primers and using PCR. Methodology for walking and jumping using PCR exists, enabling us to characterize this translocation more quickly than by using standard recombinant techniques (Green & Olson, 1990).

Materials and methods

Cells and cell culture

A lymphoblastoid cell line from the patient, D.G., bearing the t(1;13)(q22;q12) translocation, was established using Epstein–Barr virus (EBV), and maintained in RPMI-1640 medium supplemented with 10% fetal calf serum (Gibco/BRL) and 2 mM glutamine.

The tk^- 3T3 mouse cells were grown in monolayer in Dulbecco's modified Eagle medium (DMEM) and transferred using trypsin–EDTA. Prior to fusion, 3T3 cells were grown in 5-bromodeoxyuridine for 7 days to eliminate potential revertants from the tk^- phenotype.

Construction and analysis of hybrids

A total of 5×10^6 3T3 cells were harvested and mixed with 10^6 cells from the D.G. lymphoblastoid cell line. The mixed population of cells was washed in three changes of phosphate-buffered saline (PBS) to remove serum. Fusion was carried out in suspension as described by Cowell (1986) using 50% polyethylene glycol (mol. wt. 1000). Hybrid cells were cultured in HAT medium for 4–6 weeks when individual colonies were isolated by ring cloning (Cowell, 1986). When necessary, single cells were isolated under an inverted microscope and transferred to 96-well dishes (Linbro). Selected colonies were then grown in bulk for DNA and chromosome analysis.

Air-dried chromosome preparations and trypsin–Giemsa banding were carried out as described by Cowell (1980).

Molecular analysis of somatic cell hybrid DNA was carried out using standard Southern blotting and oligonucleotide labelling as described by Cowell & Mitchell (1989).

Generating Alu PCR products

Amplification was performed by PCR using Promega *Thermus aquaticus* (Taq) DNA polymerase and an automated Techne II thermal cycler. The reaction mixture was prepared for a final volume of 50 μl containing 200 μM of each deoxyribonucleotide triphosphate (dATP, dCTP, dTTP, dGTP, Pharmacia), 5 μl of 1:10 Taq polymerase buffer [500 mM potassium chloride, 100 mM Tris–HCl, 15 mM magnesium chloride, 0.1% gelatin (w/v), 1% Triton X-100] and oligonucleotide primers to 2 μM. A 100-ng aliquot of genomic DNA was used per reaction and 50 μl of mineral oil added to the tube to prevent evaporation. An initial denaturing step of 96°C for 15 min was performed. The sample was cooled to 60°C and 1 μl of the following was added: 2 μl of Taq DNA polymerase, 1 μl of 10 Taq buffer, 7 μl distilled water. Amplification was performed with 30 cycles following the sequence: 57°C for 30 s (annealing) 72°C for 2 min (extension) and 96°C for 45 s (denaturing). The final 72°C extension was increased to 10 min to ensure maximum synthesis of double-stranded DNA.

Aliquots of 5 μl of the reaction mixture were analysed by electrophoresis on a 2% agarose gel (TAE buffer, staining with ethidium bromide).

Alu primers used

The Alu primers used were as described by Cotter *et al.* (1990):

3′-CAGAATTCGCGACAGAGCGAGACTCCGTCTC
5′-CGGGATCCCAAAGTGCTGGGATTACAGGCGTG

Cloning of Alu PCR products from hybrid cell line ICD

Alu PCR products were purified on Sephadex G.50 columns, ethanol precipitated and resuspended in 10 μl of TE pH 8.0. Aliquots of 100 ng (1 μl) were ligated into the EcoRV site of the Bluescript SK vector (Stratagene) at a variety of vector to insert molar ratios (1:1 to 0.125:1), of which 0.125:1 was the most successful. Ligations were performed at 14°C for 16 h in a final volume of 20 μl with polyethylene glycol 8000 (Sigma) at 5% final concentration, with 200 units of T4 DNA ligase (New England Biolabs).

Each ligation was diluted 1:5 with TE and 5 μl of the diluted reaction mix was used to transform 100 μl of maximum efficiency DH5α-competent cells (BRL) according to

the manufacturer's instructions. Transformed cells were plated out onto Luria broth agar plates containing ampicillin (100 $\mu g\ ml^{-1}$) and X-gal (50 $\mu g\ ml^{-1}$). Colonies with disruption of the β-galactosidase gene (white colonies) were gridded onto ampicillin plates and analysed further.

Amplification and purification of Alu PCR plasmid inserts directly from colonies

Cells from each colony of interest were transferred with a sterile toothpick to the PCR mixture described above and were amplified using either the original 3′ or the 5′ Alu primers. A 5-μl aliquot of the reaction mixture was run on a 2% agarose gel to confirm size and concentration of the PCR product. About 1 μg of product was then run out on a 2% agarose gel, excised and the DNA purified (Geneclean, BIO 101) according to the manufacturer's instructions.

Hybridization of probes to somatic cell hybrid mapping panel

Samples of 20 μg of normal human, mouse 3T3, Chinese hamster and somatic cell hybrid DNAs were digested with EcoRI or HindIII, and separated by electrophoresis for 16 h at 30 mA in an 0.8% agarose gel. The DNA was transferred to Hybond N membranes (Amersham) by the method of Southern, denatured and neutralized according to manufacturer's instructions. Alu PCR probes were purified as described above and about 70 ng oligolabelled by the random priming method. Filters were hybridized in 4 × SSC, 10 × Denhardt's, 0.1% SDS and boiled, sonicated salmon sperm 50 $\mu g\ ml^{-1}$ at 65°C in a Hybaid oven. Washing was 2 × SSC, 0.1% SDS, at 65°C for 30 min.

Acknowledgements

We thank Dr J. Pritchard for his critical reading of the manuscript and Dr R. Shannon for bringing this patients to our attention and collecting the blood sample. We are grateful to Professor C.H.M.C. Buys for making the ICD hybrid available and to Dr I. Goldsmith, ICRF, for the synthesis of the oligonucleotide primers. A.J.M. is a clinical research fellow funded by a grant from the Medical Research Council of Great Britain. This work is supported, in part, by a grant from the Neuroblastoma Society.

References

Brodeur, G.M., Sekon, G.S. & Goldstein, M.N. (1977). *Cancer*, **40**, 2256–2263.

Brooks-Wilson, A.R., Goodfellow, P.N., Povey, S., Nevanlinna, H.A., De Jong, P.J. & Goodfellow, P.J. (1990). *Genomics*, **7**, 614–620.

Cavenee, W., Leach, R., Mohandas, T., Pearson, P. & White, R. (1984). *Am. J. Hum. Genet.*, **36**, 10–24.

Christiansen, H. & Lampert, F. (1988). *Br. J. Cancer*, **57**, 121–126.

Cole, C.G., Goodfellow, P.N., Bobrow, M. & Bentley, D.R. (1991). *Genomics*, **10**, 816–826.

Cotter, F.E., Hampton, G.M., Nasipuri, S., Bodmer, W.F. & Young, B.D. (1990). *Genomics*, **7**, 257–263.

Cotter, F.E., Das, S., Douek, E., Carter, N.P. & Young, B.D. (1991). *Genomics*, **9**, 473–480.

Cowell, J.K. (1980). *J. Natl. Cancer Inst.*, **65**, 955–961.

Cowell, J.K. (1986). In: *Human Cytogenetics, a Practical Approach*. Rooney, D.E. & Czepulkowski, B.H. (eds.). Oxford: IRL Press, pp. 202–218.

Cowell, J.K. & Mitchell, C.D. (1989). *Cytogenet. Cell Genet.*, **52**, 1–6.

Dryja, T.P., Rapaport, J.M., Weichselbaum, R. & Burns, G.A.P. (1984). *Hum. Genet.*, **65**, 320–324.

Floyd-Smith, G., Whitehead, A.S., Colten, H.R. & Franke, U. (1986). *Immunogenetics*, **24**, 171–176.

Fong, C.-T., Dracopoli, N.C., White, P.S., Merrill, P.T., Griffith, R.C., Housman, D.E. & Brodeur, G.M. (1989). *Proc. Natl. Acad. Sci. USA*, **86**, 3753–3757.

Fountain, J.W., Wallace, M.R., Bruce, M.A., Siezinger, B.R., Menon, A.G., Gusella, J.G., Michels, V.V., Schmidt, M.A., Dewald, G.W. & Collins, F.S. (1989). *Science*, **244**, 1085–1087.

Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M. & Dryja, T.P. (1986). *Nature*, **323**, 643–646.

Gendler, S.J., Burchill, J.M., Duhig, T., Lamport, D., White, R., Parker, M. & Taylor-Papadimitriou, J. (1987). *Proc. Natl. Acad. Sci. USA*, **84**, 6060–6064.

Gilbert, F., Feder, M., Balaban, G., Lurie, D.K., Podolsky, R., Rinaldt, V., Vinikoor, N. & Weisband, J. (1984). *Cancer Res.*, **44**, 5444–5449.

Green, E.D. & Olson, M.V. (1990). *Science*, **250**, 94–98.

Grundy, P., Koufos, A., Morgan, K., Li, F.P., Meadows, A.T. & Cavenee, W.K. (1988). *Nature*, **336**, 375–376.

Guzzetta, V., Montes de Oca-Luna, R., Lupski, J.R. & Patel, P.I. (1991). *Genomics*, **9**, 31–36.

Higgins, M.J., Hansen, M.F., Cavenee, W.K. & Lalande, M. (1989). *Mol. Cell Biol.*, **9**, 1–5.

Huff, V., Compton, D.A., Chao, L.Y., Strong, L.C., Geiser, C.F. & Saunders, G.F. (1988). *Nature*, **336**, 377–378.

Hunt, J.D. & Tereba, A. (1990). *Genes, Chrom. Cancer*, **2**, 137–146.

Jelinek, W.R. & Schmid, C.W. (1982). *Annu. Rev. Biochem.*, **51**, 813–844.

Kariya, Y., Kato, K., Hayashizaki, Y., Himeno, S., Taui, S. & Matsubara, K. (1987). *Gene*, **53**, 1–10.

Knudson, A.G. (1971). *Proc. Natl. Acad. Sci. USA*, **68**, 820–823.

Knudson, A.G. & Strong, L.C. (1972). *Am. J. Hum. Genet.*, **24**, 514–532.

Kushner, B.H., Gilbert, F. & Helson, L. (1986). *Cancer*, **57**, 1887–1893.

Laureys, G., Speleman, F., Opdenakker, G., Benoit, Y. & Leroy, J. (1990). *Genes Chrom. Cancer*, **2**, 252–254.

Leister, I., Weith, A., Bruderlein, S., Cziepluch, C., Kangwanpong, D., Schlag, P. & Schwab, M. (1990). **50**, 7232–7235.

Mannens, M., Slater, R.M., Heytig, C., Bliek, J., De Kraker, J., Coad, N., De Pagter-Holthuizen, P. & Pearson, P.L. (1988). *Hum. Genet.*, **81**, 41–48.

Martinsson, T., Weith, A., Cziepulch, C. & Schwab, M. (1989). *Genes Chrom. Cancer*, **1**, 67–78.

Middleton-Price, H.R., Van Den Berghe, J.A., Scott, J., Knott, T.J., & Malcolm, S. (1988). *Hum. Genet.*, **79**, 283–285.

Mitchell, C.D. & Cowell, J.C. (1989). *Oncogene*, **4**, 253–257.

Mitchell, C.D., Ventris, J.A., Warr, T.J. & Cowell, J.K. (1991). *Oncogene*, **6**, 89–92.

Moore, J.W., Hyman, S., Antonarakis, S.E., Mules, E.H. & Thomas, G.H. (1986). *Hum. Genet.*, **72**, 297–302.

Moorhead, P.S. & Evans, A.E. (1980). In: *Advances in Neuroblastoma Research*. Evans, A.E. (ed.). Raven Press: New York, pp. 109–118.

Moyzis, R.K., Torney, D.C., Meyne, J., Buckingham, J.M., Wu, J.R., Burks, C., Sirotkin, C.M. & Goad, W.B. (1989). *Genomics*, **4**, 273–289.

Nagano, H., Kano, Y., Kobuchi, S. & Kajitani, T. (1980). *Jpn. J. Hum. Genet.*, **25**, 39–45.

Nau, M.M., Brooks, B.J., Battey, J., Sausville, E., Gazdar, A.F., Kirsch, I.R., McBride, O.W., Bertness, V., Hollis, G.F. & Minna, J.D. (1985). *Nature*, **318**, 69–73.

Nelson, D.L., Ledbetter, S.A., Corbo, L., Victoria, M.F., Ramirez-Solis, R., Webster, T.D., Ledbetter, D.H. & Caskey, C.T. (1989). *Proc. Natl. Acad. Sci. USA*, **86,** 6686–6690.

Nomura, N., Sasamoto, S., Ishii, S., Date, T., Matsui, M. & Ishizaki, R. (1989). *Nucleic Acids Res.*, **17,** 5489–5500.

O'Connell, P., Leach, L., Cawthorn, R.M., Culver, M., Stevens, J., Viskochil, D., Fournier, R.E.K., Rick, D.C., Ledbetter, D.H. & White, R. (1989). *Science*, **244,** 1087–1088.

Ray, P.N., Belfall, B., Duff, C., Logal, C., Kean, V., Thompson, M.W., Sylvester, J.E., Gorski, J.L., Schmickel, R.D. & Worton, R.G. (1985). *Nature*, **318,** 672–675.

Riccardi, V.M., Sujansky, E., Smith, A.C. & Francke, U. (1978). *Pediatrics*, **61,** 604–610.

Sanger, W.G., Howe, J., Fordyce, R. & Purtilo, D.T. (1984). *Cancer Genet. Cytogenet.*, **11,** 153–159.

Scheffer, H., Van der Lelie, D., Aanstoot, G.H., Goor, N., Nienhaus, N.J., Van der Hout, A.H., Pearson, P.L. & Buys, C.H. (1986). *Cytogenet. Cell Genet.*, **52,** 971–973.

Sparkes, R.S., Murphree, A.L., Lingua, R.W., Sparkes, M.C., Field, L.L., Funderbuck, S. & Benedict, W.F. (1983). *Science*, **217,** 971–973.

Suzuki, T., Yokata, J., Mugishima, H., Okabe, I., Ookuni, M., Sugimura, T. & Terada, M. (1989). *Cancer Res.*, **49,** 1095–1098.

Varesco, L., Thomas, H.J.W., Cottrell, S., Murday, V., Fennell, S.J., Williams, S., Searle, S., Sheer, D., Bodmer, W.F., Frischauf, A.M. & Solomon, E. (1989). *Proc. Natl. Acad. Sci. USA*, **86,** 10118–10122.

Viskochil, D., Buchberg, A.M., Xu, G., Cawthorne, R.M., Stevens, J., Wolff, R.K., Culver, K., Carey, J.C., Copeland, N.G., Jenkins, N.A., White, R. & O'Connell, P. (1990). *Cell*, **62,** 187–192.

Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M.M. & Bos, J.L. (1988). *New Engl. J. Med.*, **319,** 525–532.

Wadey, R.B., Pal, N.P., Buckle, B., Yeomans, E., Pritchard, J. & Cowell, J.C. (1990). *Oncogene*, **5,** 901–907.

Weith, A., Martinsson, T., Cziepluch, C., Bruderlein, S., Amler, L.C., Berthold, F. & Schwab, M. (1989). *Genes, Chrom. Cancer*, **1,** 159–166.

Yunis, J.J. & Ramsay, N. (1978). *Am. J. Dis. Child.*, **132,** 161–163.

Erratum

In the paper *Isolation of chromosome-specific DNA sequences from an Alu polymerase chain reaction library to define the breakpoint in a patient with a constitutional translocation t(1;13)(q22;q12) and ganglioneuroblastoma* by Michalski *et al.*, *Oncogene* (1992) **7:** 1595–1602, Figure 6 was mislabelled. The correct version is reproduced below.

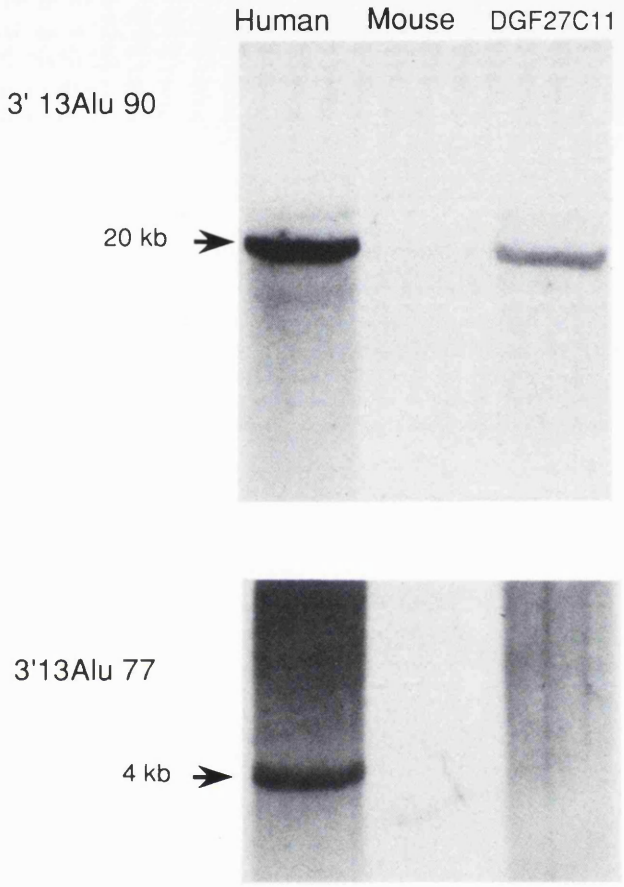


Figure 6 Pattern on autoradiography of Southern blots of EcoRI-digested human, mouse and DGF27C11DNA, probed with Alu PCR probes above and below the DGF breakpoint

SHORT COMMUNICATION

Assignment of Four Sequence-Tagged Sites to Three Subregions of 13q12 Using a Somatic Cell Hybrid Mapping Panel

A. J. MICHALSKI AND J. K. COWELL[1]

ICRF Oncology Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, United Kingdom

Received March 30, 1993; revised June 30, 1993

To define the position of a 13q12 breakpoint from a patient with ganglioneuroblastoma, a series of somatic cell hybrids carrying human chromosome translocations with breakpoints in the proximal part of chromosome 13 has been compiled. Sequence-tagged sites (STS) have been generated from a series of *Alu*-PCR probes previously shown to be in the 13q12 region. Together with an STS for the oncogene FLT1, these have been used to define the relative positions of the translocation breakpoints in the hybrids. In this way, four markers have been ordered in three subregions of 13q12 and reference breakpoints established. The refined physical map of 13q12 provides a series of reference markers with known locations and will be invaluable in the further characterization of breakpoints in this region.

A major goal in the construction of a physical map of a single human chromosome is the establishment of the linear order of DNA probes within a clearly defined subregion and one approach is to use the sequence-tagged sites (STS) generated for many probes and map them using panels of somatic cell hybrids. This approach establishes the linear order of the breakpoints in the somatic cell hybrid panel, which then defines subsections of the chromosome into which other probes can be placed. We recently described a series of *Alu*-PCR probes isolated from a somatic cell hybrid containing the proxmial half of chromosome 13 as the only human component (8). Although all of these probes mapped to the 13q12 region, they could only be assigned positions proximal or distal to a breakpoint in 13q12 derived from a patient (D.G.) with ganglioneuroblastoma and a constitutional t(1;13)(q22;q12) translocation. To subdivide the 13q12 region further, we have extended our panel of somatic cell hybrids and characterized the position of the breakpoints in these hybrids using STS from the *Alu*-PCR probes.

In a previous report (8) we localized a series of *Alu*-PCR-generated DNA clones within 13q12 using two hybrids, KBF11 and DGF2711 (a somatic cell hybrid derived from patient D.G.), which defined only two relatively large subsections of 13q12. In an attempt to establish a linear order for these probes, two new hybrids were created between mouse 3T3 cells and lymphocytes from two patients, C.A. and J.C., shown to have 13q12–q14 deletions (6) which predisposed them to retinoblastoma. Individual hybrids were isolated as described by Cowell (5) and screened by PCR using oligonucleotide primers from the factor VII gene (F7), which is located in 13q34 (1). Those hybrids yielding an appropriately sized PCR product were then analyzed using oligonucleotides designed to detect a variety of exons of the retinoblastoma predisposition gene (RB1), which lies in 13q14 (7). From patient C.A., one hybrid, CAF5, showed a strong PCR product for F7 and a very weak product for RB1 exon sequences. We interpreted this to mean that the majority of cells in the parental hybrid population contained the deletion chromosome 13 but that a few also contained the normal homologue. Single cell clones were isolated by picking individual cells (5) from CAF5, and two, CAF55 and CAF510, showed the presence of F7 but not RB1.

A single hybrid, JCF9, from the second patient, J.C., gave the same profile as CAF55 using PCR, without the need for subcloning. Cytogenetic analysis of CAF55 and JCF9 confirmed the presence of the deleted chromosome 13 in both cases and the absence of the normal homologue (data not shown). Hybrids CAF55 and JCF9 formed part of a panel of hybrids with breakpoints in 13q12, which included two other hybrids not isolated by us, CF25 and CF27 (9). An idiogram of this hybrid mapping panel is shown in Fig. 1.

Our previous attempts to sublocalize the *Alu*-PCR clones (8) involved Southern blotting, and some of the results were equivocal because of the low intensity of hybridization signal seen in some of the hybrids. We decided, therefore, to generate STS primers for the seven unique sequences described by Michalski *et al.* (8). Sequencing was carried out by direct PCR analysis using T3 or T7 primers to amplify the inserts which were cloned in the Bluescript vector. All seven clones showed a long poly(dA)$_n$ tract at both ends adjacent to the *Alu* repeats. The A-rich regions were usually so long that reading the sequence distal to them was difficult. This difficulty was compounded by the frequent band compression seen on the sequencing gel at the distal end of the poly(dA)$_n$ tract. As a result, very little unique se-

[1] To whom correspondence should be addressed. Telephone: (071) 242-9789, ext. 2194.

GENOMICS 18, 141–143 (1993)
0888-7543/93 $5.00

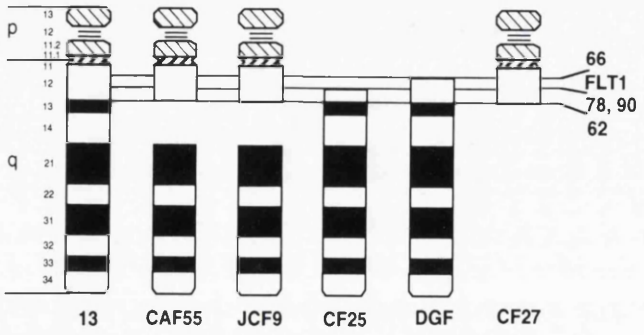


FIG. 1. An idiogram showing the location of the breakpoints in 13q12 in the somatic cell hybrid mapping panel and the relative position of the five STS.

quence was obtained from the clones, which limited the choice of primers that could be designed for the development of STS. The long poly(dA)$_n$ tracts meant that adequate sequence data could not be obtained from 17% of our probes, which is comparable to the 20% found by Cole *et al.* (3). This is perhaps not surprising since the clones reported here were all generated with the 3′ *Alu* primer described by Cotter *et al.* (4), which is derived from the same part of the *Alu* sequence as the ALE 1 primer used by Cole *et al.* (3), who also commented on the difficulty in obtaining good sequence data.

Despite the sequencing difficulties detailed above, oligonucleotide primers were designed for all seven clones (Table 1). Three of these STS, however, could not be used for mapping. Thus, although probe 3′13 *Alu* 169 gave a human-specific band on Southern hybridization, it is only a 250-bp inter-*Alu* segment with long poly(dA)$_n$ tails at both ends. Oligonucleotide primers were designed from the first readable sequence after these tracts, but the resulting product, which was present in DNA from human, mouse, and *Saccharomyces cerevisiae,* was only 50 bp long. Similarly, it was impossible to derive useful oligonucleotide primers for 3′13*Alu* 77, despite obtaining strong unique signals on hybridization, because of long poly(dA)$_n$ tracts and dense band compression on sequencing gels. 3′13*Alu* 71 gave weak signals on Southern hybridization and oligonucleotide primers designed from sequencing this probe also gave a PCR product in human, mouse, and yeast cells. Thus, only four of the *Alu*-PCR products were available for STS mapping using the somatic cell hybrid panel. Following PCR amplification, 3′13*Alu* 66 gave a multibanded background with human and mouse DNA as well as the single dominant product of the expected size. This background could be suppressed by increasing the annealing temperature to 62°C, but, in reducing the product yield as a consequence, the signal from some of the hybrids became very weak. The other three STS gave single-copy products. Therefore, once sequenced, 67% of the probes gave rise to usable human-specific single-copy STS, which, again, is comparable to the figure of 70% published by Cole *et al.* (3).

Using the four unique STS, it was possible to establish the linear order of the 13q12 breakpoints in the hybrid mapping panel. 3′13*Alu* 62 was absent from CAF55, JCF9, and CF27, making it the most distal of the four probes. 3′13*Alu* 78 was present in all the hybrids, including DGF27C11, which is at variance with our previous report (see below). 3′13*Alu* 90 mapped to the same region as 3′13*Alu* 78. 3′13*Alu* 66 was absent from DGF27C11 and CF25 but present in the other hybrids. Although the mapping panel could not exclude the possibility that this sequence mapped to the short arm of chromosome 13, its position immediately below the centromere on 13q was confirmed by *in situ* hybridization using a cosmid corresponding to this probe (data not shown).

Using these four probes alone, it was not possible to distinguish between the breakpoints in DGF27C11 and CF25. The FLT1 oncogene, however, has also been mapped to 13q12 (10) and, using oligonucleotide primers for this gene, it was possible to demonstrate its presence in DGF27C11 but not CF25, allowing us to determine the relative positions of these two translocation break-

TABLE 1

Summary of the Sequence Data of the Oligonucleotide Primers Designed for *Alu*-PCR Probes and the Size of the Product for Each STS

Name of STS	Sequence of oligonucleotide primer	Size of PCR product (bp)
T3	ATTAACCCTCACTAAAG	
T7	AATACGACTCACTATAG	
3′13*Alu*66	1: CAGAAGGTGTGAGATGAGGACAC	550
	2: TAGTTACACTGGAAGCCGGACCT	
3′13*Alu*90	1: CATAGAAAGTGCTAGGATTTACAAGGATG	400
	2: TGCTTGATCACATTTACAGATGCTTTTCC	
3′13*Alu*78	1: AGTTTATAGGAGGCCGTAGTTTTGG	600
	2: CCAAACTTGGAAGGGGATGAGAATA	
3′13*Alu*62	1: CCATTGGACAATACATGAACAGTT	510
	2: TTCCAAGTTCATAGGCCTTCTGAA	
FLT1	1: AGGACCAAACCATGTCTGTC	220
	2: TTTGGCCGACAGTGGTGTAA	

Note. All primers are shown 5′ → 3′ exactly as synthesized.

FIG. 2. An example of the STS mapping of the somatic cell hybrid panel. 3'13*Alu* 66 generates a 550-bp product using DNA from human and three of the hybrids (CAF55, JCF9, and CF27). There is a multiband pattern in 3T3 and all the mouse-based hybrids (DGF27C11, CAF55, and JCF9).

points. Although it was not the primary aim of our experiments, we were also able to show that D13S59, located in 13q14.3 (2), was absent from CAF55 and JCF9, putting the distal breakpoints of these deletions below this locus.

Screening hybrids by PCR using oligonucleotide primers for the 3'*Alu*-generated probes was much more rapid and used a fraction of the DNA needed for Southern blotting. The main advantage over Southern blotting, however, was accuracy, as PCR has the ability to recognize and selectively amplify a single-copy target sequence in the presence of large amounts of background DNA. Indeed, this increased sensitivity allowed us to reorder two of the probes relative to the DGF breakpoint. 3'13*Alu* 78 and 3'13*Alu* 62 were originally thought to be absent from DGF27C11 by Southern hybridization (8), but were found to be present using PCR and therefore below the DGF breakpoint.

We have thus defined a series of chromosomal breakpoints that subdivides the 13q12 band using a set of STS specific for that region. The hybrid mapping panel used in this analysis will allow the localization of other markers from this region in an attempt to saturate each subregion with probes that can be used to generate a contiguous physical map of the whole region. In particular, we have been able to define the position of a 13q12 breakpoint present in constitutional cells from a patient with ganglioneuroblastoma and a t(1;13)(q22;2) translocation. This resource will be invaluable in attempting to isolate genes disrupted by the rearrangement to assess their role in tumorigenesis.

ACKNOWLEDGMENTS

A.J.M. is supported by a grant from the Medical Research Council of Great Britain. We gratefully acknowledge the support of the UK Neuroblastoma Society for this project. The CF27 and CF25 hybrids were isolated by Dr. Mohandas and kindly provided by Dr. D. Warburton. The FLT1 oligonucleotide sequence was sent to us by Dr. D. Shapiro. We thank Sheila Giles for help with the manuscript preparation.

REFERENCES

1. Abbott, C., and Povey, S. (1991). Development of human chromosome-specific PCR primers for characterization of somatic cell hybrids. *Genomics* **9:** 73–77.
2. Bowcock, A. M., Farrer, L. A., Herbert, J. M., Bale, A. E., and Cacalli-Sforza, L. (1991). A contiguous linkage map of chromosome 13q with 39 distinct loci separated on average by 5.1 cM. *Genomics* **11:** 517–529.
3. Cole, C. G., Goodfellow, P. N., Bobrow, M., and Bentley, D. R. (1991). Generation of novel sequence-tagged sites (STSs) from discrete chromosomal regions using *Alu*-PCR. *Genomics* **10:** 816–826.
4. Cotter, F. E., Hampton, G. M., Nasipuri, S., Bodmer, W. F., and Young, B. D. (1990). Rapid isolation of human chromosome-specific DNA probes from a somatic cell hybrid. *Genomics* **7:** 257–263.
5. Cowell, J. K. (1992). Manipulation of somatic cell hybrids for the analysis of the human genome. *In* "Human Cytogenetics: A Practical Approach" (D. E. Rooney and B. H. Czepulkowski, Eds.) 2nd ed., Vol. 9, pp. 235–292, IRL Press, Oxford.
6. Cowell, J. K., Hungerford, J., Rutland, P., and Jay, M. (1989). Genetic and cytogenetic analysis of patients showing reduced esterase-D levels and mental retardation from a survey of 500 individuals with retinoblastoma. *Ophthal. Ped. Genet.* **110:** 117–127.
7. Hogg, A., Onadim, Z., Baird, P. N., and Cowell, J. K. (1992). Detection of heterozygous mutations in the RB1 gene in retinoblastoma patients using single strand conformation polymorphism (SSCP) analysis and polymerase chain reaction sequencing. *Oncogene* **7:** 1444–1451.
8. Michalski, A. J., Cotter, F. E., and Cowell, J. K. (1992). Isolation of chromosome-specific DNA from an Alu polymerase chain reaction library to define the breakpoint in a patient with a constitutional translocation t(1;13)(q22;2). *Oncogene* **7:** 1595–1602.
9. Mohandas, T., Shapiro, L. J., Sparkes, R. S., and Sparkes, M. C. (1979). Regional assignment of the steroid sulfatase-X-linked ichthyosis locus: Implications for a noninactivated region on the short arm of the human X chromosome. *Proc. Natl. Acad. Sci. USA* **76:** 5779–5783.
10. Rosnet, O., Stephenson, D., Mattei, M.-G., Marchetto, S., Masabumi, S., Chapman, V. M., and Birnbaum, D. (1993). Close physical linkage of the FLT1 and FLT3 genes on chromosome 13 in man and chromosome 5 in mouse. *Oncogene* **8:** 173–179.

Advances in Neuroblastoma Research 4, pages 79–85

CONSTRUCTING A PHYSICAL MAP AROUND A CONSTITUTIONAL t(1;13)(q22;q12) BREAKPOINT IN A PATIENT WITH A GANGLIONEUROBLASTOMA.

Antony J. Michalski MRCP, and John K. Cowell PhD.

Department of Haematology & Oncology,
Institute of Child Health,
30, Guilford Street,
London WC1N 1EH.

INTRODUCTION

Neuroblastoma (Nb) is the most common solid tumour of childhood, accounting for 8% of all paediatric malignancy and ganglioneuroblastoma, although clinically and histologically less aggressive, is considered part of the same disease spectrum. Amplification of the oncogene, Nmyc, identifies a subset of Nb with particularly poor prognosis, but as the majority of cases lack this finding, other genetic changes must exist (Seeger et al., 1985). In rare families, the Nb tumour phenotype segregates as an autosomal dominant trait. This led Knudson & Strong (1972) to propose that tumorigenesis was a consequence of a loss of function of a critical gene in embryonic development in a similar way to that proposed for retinoblastoma. As Nb families are rare (Kushner et al., 1986) it has not been possible to assign the Nb predisposition gene to a particular chromosome region using conventional linkage studies.

Analysis of tumour karyotypes and loss of heterozygosity (LOH) studies have suggested that the loss of genetic material from the short arm of chromosome 1 (1p36) may be important in tumour development (Brodeur et al., 1981), but these 1p abnormalities occur against a background of other structural rearrangements so that it is difficult to determine which are causal and which are related to tumour progression.

However, analysis of constitutional chromosome abnormalities in individuals who develop particular tumour types has usually indicated the site of the predisposition gene. This has been the case for retinoblastoma (Yunis and Ramsay, 1978) Wilms' tumour (Riccardi et al., 1978) and familial polyposis coli (Varesco et al., 1989). These structural abnormalities are usually deletions, but where translocations are identified the breakpoint usually interrupts a gene involved in the phenotype. Only two constitutional chromosome translocations have been described in patients with Nb (Michalski et al., 1992, Laureys et al., 1990). Laureys et al (1990) reported a t(1;17)(p36;q12-21), and its location in 1p36 correlates well with the tumour karyotype and LOH studies noted above. However, like Nmyc amplification, 1p abnormalities are found mainly in high stage tumours and may either be related to tumour progression, or denote a more

aggressive subset of Nb (Christiansen et al., 1992). There is no reason to believe that only a single tumour suppressor gene will be important in Nb. In Wilms' tumour, for example, although the WT1 gene in 11p13 has recently been cloned (Call et al., 1990, Gessler et al., 1990), LOH studies have shown that 11p15 may be the site of a second predisposition gene (Mannens et al., 1988, Wadey et al., 1990). Linkage studies in the very rare instances of familial Wilms' tumours suggest that predisposition is not linked to markers on chromosome 11 implying that yet a third gene is associated with tumorigenesis. It is important, therefore, to carefully investigate genetic changes potentially predisposing to Nb formation.

A patient with a ganglioneuroblastoma and a constitutional t(1;13)(q22;q12) translocation has been described (Michalski et al., 1992) and our aim was to isolate and characterise the translocation breakpoint fragment to see whether a gene critical to neural development has been disrupted by this genetic change.

METHODS, RESULTS AND DISCUSSION

Case Report: Patient DG presented at five years of age with a two year history of a gradually enlarging left-sided neck mass. He had no other clinical problems and his parents and two sisters were well. There was no family history of malignant disease. Examination showed a 10 cm x 7 cm neck mass but no other abnormalities. Magnetic resonance imaging (MRI) scan showed the mass to be separate from the cervical spine and deep to the carotid vessels. Biopsy showed the mass to be a cervical ganglioneuroblastoma but his urinary catecholamine metabolite levels were normal and bone marrow aspirate and trephines were free of malignant cells, suggesting that the tumour was relatively benign. A well encapsulated ganglioneuroblastoma was removed surgically and he received no further treatment. The tumour was not sent for karyotype studies but cytogenetic analysis of his lymphocytes showed a constitutional t(1;13)(q22;q12) translocation. Cytogenetic studies of his parents and two sibs are normal. Seven years later he is well with no tumour recurrence.

Somatic cell Hybrids: In order to map the breakpoint junction fragment, separation of the derivative (translocated) chromosomes from the normal copies of chromosomes 1 and 13 was essential. A lymphoblastoid cell line was established from DG and somatic cell hybrids constructed between DG lymphoblasts and mouse 3T3 cells as described in Cowell, (1992). Individual hybrid colonies were investigated cytogenetically and in one, DGF27, most cells contained the derivative chromosome 1 (1pter-q21;13q12-qter). Single cell stocks were generated from this parental hybrid producing a sub-clone, DGF27C11, which contained the der 1 but no other chromosome 1 or 13 material. The exclusive presence of the der 1 chromosome was confirmed by Southern blotting and the polymerase chain reaction (PCR) using markers on the short and long arms of chromosome 1 and 13q12-14.

Developing flanking markers by Alu-PCR: To isolate the breakpoint junction fragment from DGF27C11 its position relative to closely flanking molecular markers must be known. None of the molecular markers available at the time mapped near to the breakpoint so unique probes had to be isolated. As our laboratory has extensive experience with chromosome 13, we chose to approach the translocation from the

chromosome 13 end. There are many methods of developing novel unique probes, but Alu-PCR cloning ensures that the majority of the clones generated come from the region of interest.

Alu is the commonest human short interspersed repetitive element with 900, 000 copies per haploid genome. It is a 130 base pair (bp) monomer in rodents and a 300bp dimer in humans with significant sequence divergence between the two species (Jelinek and Schmid, 1982) Alu repeats are not evenly distributed within the genome; Alu rich and Alu poor regions exist (Moyzis et al., 1989). Using oligonucleotide primers for the 3' or 5' ends of the Alu repeat sequence PCR will allow amplification of 'inter-Alu' sequences in the DNA used as a template in the PCR reaction. These 'inter-Alu' sequences can be purified, cloned into the Bluescript plasmid vector , used to transform bacteria and characterised (Cotter et al., 1990, Nelson et al., 1989). We chose to use primers to the 3' end of the Alu repeat (as only human DNA was amplified) and used a somatic cell hybrid containing 13pter-q14 as its only human component as the DNA template. The only Alu rich region in the hybrid was 13q12; the region of the DGF breakpoint. Using the X-gal selection system, 200 white bacterial colonies were gridded for further analysis. 34% were non-recombinants (white colonies containing no insert) and 52% contained repetitive sequences; a much higher percentage than in other published studies (15% in Cotter et al, 1990, and 39% in Brooks-Wilson et al, 1990).This may be specific to the region of interest. In all, 30 probes were demonstrated to be fully unique sequences, acting as single copy probes when mapped by Southern hybridisation to a panel of somatic cell hybrids containing various deletions of chromosome 13 (Cowell and Mitchell, 1989). These 30 probes were a mixture of only 7 different unique sequences and the reason for their numerical dominance is unclear. These 7 probes, however, were found to flank the DGF breakpoint more closely than any other available marker (Michalski et al., 1992).

Confirming position of flanking markers by FISH and STS mapping: Two major problems persisted, however. Firstly, the somatic cell hybrid mapping panel did not allow us to exclude the possibility that probes proximal to the DGF breakpoint (3'13 Alu66 and 3'13 Alu77) mapped to 13p (where they would be too far away from DGF to be useful). Secondly, in circumstances when the molar ratio of target DNA is low Southern blotting may give false negative results (the absence of a band on the blot may not mean the absence of the probe sequence in the hybrid).

The position of the probes proximal to the DGF breakpoint was confirmed by isolating cosmids using the 3'13 Alu unique probes. These cosmids were nick translated, hybridised to normal metaphase spreads and viewed under a confocal microscope (fluorescent in-situ hybridisation (FISH)). Cosmids identified by 3'13 Alu66 and 3'13 Alu77 gave a strong hybridisation signal just below the centromere on 13q12. However, cosmids identified by 3'13 Alu90 (mapping just below the DGF breakpoint on our somatic cell hybrid panel) hybridised to the centromeres of several different human chromosomes and was, therefore, useless for further mapping of the DGF breakpoint (D. Shapiro, pers comm).

The accuracy of the mapping of the 3'13 Alu probes relative to the DGF breakpoint was improved using PCR. The probes were sequenced and oligonucleotide primers designed which recognised operationally unique DNA sequences termed sequence tagged sites (STS) (Olson et al., 1989). PCR would amplify the DNA in between the

probe-specific primers resulting in the presence of a band following gel electrophoresis of the PCR product (Cole et al., 1991). Sequencing was carried out by direct PCR analysis using oligonucleotide primers for the T3 and T7 promoter regions of the Bluescript vector into which the inserts had been cloned. Designing oligonucleotide primers for the 3'13 Alu generated probes was difficult due to long "A rich" regions at either end of the sequence adjacent to the Alu repeats. These poly (dA) tracts meant that adequate sequence data could not be obtained from 17% of our probes, which is comparable to the 20% found by Cole et al, (1991). For those probes that were successfully sequenced, oligonucleotide primers were designed as described by Hogg et al, (1992). Where it was possible to design oligonucleotide primers for the 3'13 Alu probes, PCR analysis was a much more accurate indicator of the presence of the probe in DGF27C11 than Southern hybridisation. Indeed, two probes thought to map above the DGF breakpoint by Southern hybridisation were reassigned as being below the breakpoint by this method. Table 1 shows the oligonucleotide sequences and reaction conditions of the STSs described above.

Table 1:
Summary of the sequence data of the oligonucleotide primers designed for Alu-PCR probes and the size of the product for each STS.

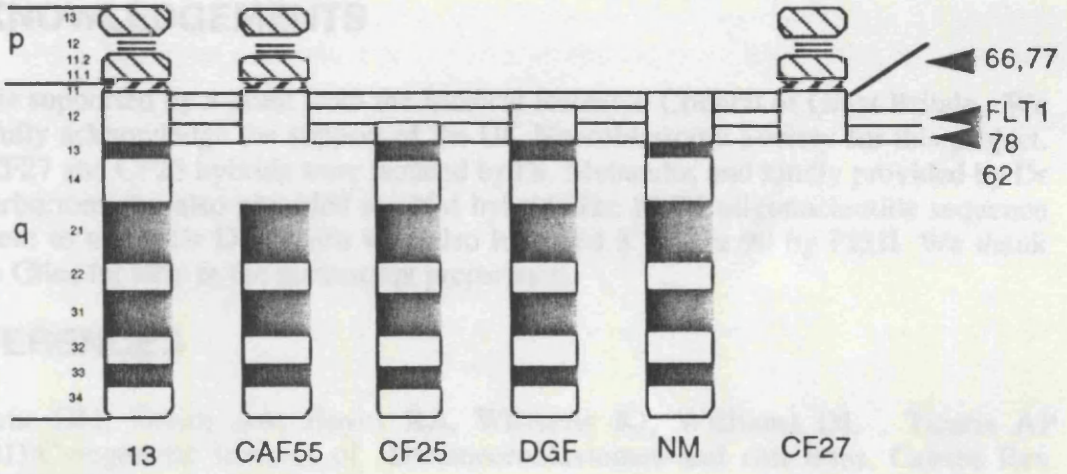
Name of STS	Sequence of oligonucleotide primers	Size of PCR product (bp)
T3	ATTAACCCTCACTAAAG	
T7	AATACGACTCACTATAG	
3'13 Alu 66	1: CAGAAGGTGTGAGATGAGGACAC 2:TAGTTACACTGGAAGCCGGACCT	550
3'13 Alu 78	1:AGTTTATAGGAGGCCGTAGTTTTGG 2:CCAAACTTGGAAGGGGATGAGAATA	600
3'13 Alu 62	1:CCATTGGACAATACATGAACAGTT 2:TTCCAAGTTCATAGGCCTTCTGAA	510
FLT1	1:AGGACCAAACCATGTCTGTC 2:TTTGGCCGACAGTGGTGTAA	220

All primers are shown 5' -> 3' exactly as synthesised. 100 ng of hybrid DNA was added to 50 pmol of 3' and 5' oligonucleotide primers for the STS. Amplification was performed by PCR using the technique noted in Michalski et al (1992). The reaction conditions for the Alu PCR STSs were 30 cycles of: 94°C for 30 secs (denaturing), 58°C for 30 secs (annealing), and 72°C for 30 secs (extension) at a final Mg^{2+} of 1.5mM. For the FLT1 STS the reaction conditions were 35 cycles of: 94°C for 45 secs (denaturing), 55°C for 45 secs (annealing), and 72°C for 45 secs (extension) at a final Mg^{2+} of 2mM. 5 ml aliquots of the reaction mixture were analysed by electrophoresis on a 2% agarose gel (TAE buffer, staining with ethidium bromide) and visualised on a UV transilluminator.

The development of a more detailed somatic cell hybrid mapping panel, dividing 13q12 into 3 sub regions, allowed the Alu-PCR probes to be ordered relative to each other. The FLT1 oncogene has recently been mapped to 13q12 (Rosnet et al 1993) and oligonucleotide primers defining a unique STS were designed allowing FLT1 to be

included in the panel of probes (Michalski and Cowell, 1993). The recent addition of hybrid, NM, to the panel supports the data obtained by FISH and confirms the position of 3'13 Alu 66 and 77 in 13q12. An ideogram of the somatic cell hybrid mapping panel is shown in figure 1.

Figure 1
An ideogram of a somatic cell hybrid mapping panel showing the position of the STS for FLT1 and the 3'13 Alu probes.



Linking the closest flanking markers by YACs: The markers which flanked the DGF breakpoint most closely were 3'13 Alu 66 and 77 above DGF and a STS for the oncogene, FLT1, below the breakpoint. In order to isolate the region of DNA in between these markers and, therefore, the breakpoint junction fragment, FLT1 and 3'13 Alu 66 and 77 were used to screen the Imperial Cancer Research Fund yeast artificial chromosome (YAC) library which contains YACs averaging 600kb in length and has a relatively low incidence of chimerism (Lehrach et al., 1990). By Southern blotting 3'13 Alu 66 and 77 recognised a 1.2Mb YAC which is not chimeric and maps exclusively to 13q12 by FISH. The ends of the human insert DNA have been isolated by vectorette PCR (Riley et al., 1990). Unfortunately, they are repetitive on Southern hybridisation and are being sequenced in order to develop STS which will allow their position relative to the DGF breakpoint to be assessed.

The FLT1 STS was used to screen a YAC library suitable for analysis by PCR and candidate YACs have been identified and are being characterised.

Once the breakpoint has been isolated in a single YAC clone or a YAC contig, genes close to the breakpoint can be isolated. When the gene is cloned the plan would be to compare its structure to known genes to determine its possible biological action. The importance of the gene in Nb development could be ascertained by comparing its structure and expression in normal adrenal and Nb tissue samples.

SUMMARY

We have constructed a detailed physical map around the 13q12 region containing the constitutional chromosome translocation breakpoint from a patient with

ganglioneuroblastoma. Further analysis will allow us to determine whether a gene vital for the development of sympathetic neural tissue has been disrupted by this genetic change.

ACKNOWLEDGEMENTS

AJM is supported by a grant from the Medical Research Council of Great Britain. We gratefully acknowledge the support of the UK Neuroblastoma Society for this project. The CF27 and CF25 hybrids were isolated by Dr. Mohandas and kindly provided by Dr D Warburton who also provided the NM hybrid. The FLT1 oligonucleotide sequence was sent to us by Dr D Shapiro who also localised 3'13 Alu 90 by FISH. We thank Sheila Giles for help in the manuscript preparation.

REFERENCES

Brodeur GM, Green AA, Hayes RA, Williams KJ, Williams DL , Tsiatis AP (1981):Cytogenetic features of hemaneuroblastomas and cell lines. Cancer Res. 41:4678-4686.

Brooks-Wilson AR, Goodfellow PN, Povey S, Nevanlinna HA, De Jong PJ , Goodfellow PJ (1990): Rapid cloning and characterisation of new chromosome 10 DNA markers by Alu element-mediated PCR. Genomics 7:614-620.

Call KM, Glaser T, Ito CY, Buckler AJ, Pelletier J, Haber DA, Rose EA, Kral A, Yeger H, Lewis WH et al (1990): Isolation and characterisation of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumour locus. Cell 60:509-520.

Christiansen H, Schestag J, Christiansen NM, Grzeschik K-H , Lampert F (1992): Clinical impact of chromosome 1 aberrations in neuroblastoma: a metaphase and interphase cytogenetic study. Genes, Chroms. Cancer 5:141-149.

Cole CG, Goodfellow PN, Bobrow M , Bentley DR (1991): Generation of novel sequence tagged sites (STSs) from discrete chromosomal regions using Alu-PCR. Genomics 10:816-826.

Cotter FE, Hampton GM, Nasipuri S, Bodmer WF , Young BD (1990): Rapid isolation of human chromosome-specific DNA probes from a somatic cell hybrid. Genomics 7:257-263.

Cowell JK. (1992) Manipulation of somatic cell hybrids for the analysis of the human genome. In Rooney DE, Czepulkowski BH, (eds) "Human Cytogenetics a Practical Approach" II ed. Oxford: IRL Press, : 235-292.

Cowell JK and Mitchell CD (1989): A somatic cell hybrid mapping panel for the regional assignment of human chromosome 13 specific DNA sequences. Cytogenet.Cell Genet. 52:1-6.

Gessler M, Poustka A, Cavenee W, Neve RL, Orkin SH , Bruns GAP (1990): Homozygous deletion in Wilms' tumours of a zinc-finger gene identified by chromosome jumping. Nature 343:774-778.

Hogg A, Onadim Z, Baird PN , Cowell JK (1992): Detection of heterozygous mutations in the RB1 gene in retinoblastoma patients using single strand conformation polymorphism (SSCP) analysis and polymerase chain reaction sequencing. Oncogene 7:1444-1451.

Jelinek WR and Schmid CW (1982): Repetitive sequences i eukaryotic DNA and their expression. Ann.Rev.Biochem. 51:813-844.

Knudson AG and Strong LC (1972): Mutation and cancer: neuroblastoma and pheochromocytoma. Am.J.Hum.Genet. 24:514-532.

Kushner BH, Gilbert F , Helson L (1986): Familial neuroblastoma: Case reports, literature review and etiologic considerations. Cancer 57:1887-1893.

Laureys G, Speleman G, Opdenakker G, Benoit Y, Leroy J (1990): Constitutional translocation t(1;17)(p36;q12-21) in a patient with neuroblastoma. Genes, Chroms.Cancer. 2:252-254.

Lehrach H et al., (1990) Genetic and Physical Mapping. In Davies KE,Tilghman SM (eds) "Genome analysis Vol 1" Cold Spring Harbor Cold Spring Harbor Laboratory Press: 39-81/

Mannens M, Slater RM, Heytig C, Bliek J, De Kraker J, Coad N, De Pagter-Holthuizen P, Pearson PL (1988): Molecular nature of genetic changes resulting in loss of heterozygosity of chromosome 11 in Wilms' tumours. Hum.Genet. 81:41-48.

Michalski AJ, Cotter FE , Cowell JK (1992): Isolation of chromosome-specific DNA from an Alu polymerase chain reaction library to define the breakpoint in a patient with a constitutional translocation t(1;13)(q22;q12). Oncogene 7:1595-1602.

Michalski AJ and Cowell JK (1993): Assignment of 4 sequence tagged sites to 3 sub-regions of 13q12 using a somatic cell hybrid mapping panel. Genomics In press.

Moyzis RK, Torney DC, Meyne J, Buckingham JM, Wu JR, Burks C, Sirotkin CM, Goad WB (1989): The distribution of interspersed repetitive DNA sequences in the human genome. Genomics 4:273-289.

Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989): *Alu* polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. Proc.Natl.Acad.Sci. USA 86:6686-6690.

Olson MV, Hood L, Cantor C , Botstein D (1989): A common language for physical mapping of the human genome. Science245:1434-1435.

Riccardi VM, Sujansky E, Smith AC, Francke U (1978): Chromosome imbalance in the aniridia-Wilms' tumour association: 11p interstitial deletion. Pediatrics 61:604-610.

Riley J, Butler R, Ogilivie D, Finniear R, Jenner D, Powell S, Anand r, Smith JC , Markham AF (1990): A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucl.Acids Res. 18(10):2887-.

Rosnet O, Stephenson D, Mattei M-G, Marchetto S, Masabumi S, Chapman VM , Birnbaum D (1993): Close physical linkage of the FLT1 and FLT3 genes on chromosome 13 in man and chromosome 5 in mouse. Oncogene 8:173-179.

Seeger RC, Brodeur GM, Salter H, Dalton A, Siegel SE, Wong KY , Hammond D (1985): Association of multiple copies of the Nmyc oncogene with rapid progression of neuroblastomas. N.Eng.J.Med. 313:1111-1116.

Varesco L, Thomas HJW, Cotrell S, Murday V, Fennell SJ, Williams S, Searle S, Sheer D, Bodmer WF, Frischauf A-M et al (1989): CpG island clones from a deletion encompassing the gene for adenomatous polyposi coli. Proc.Natl.Acad.Sci.USA 86:10118-10122.

Wadey RB, Pal NP, Buckle B, Yeomans E, Pritchard J , Cowell JK (1990): Loss of heterozygosity in Wilms' tumour involves two distinct regions of chromosome 11. Oncogene 5:901-907.

Yunis JJ and Ramsay N (1978): Retinoblastoma and subband deletion of chromosome 13. Am.J.Dis.Child. 132:161-163.