CHARACTERISATION OF MUTATIONS IN THE GENE FOR NEUROFIBROMATOSIS TYPE 1

Smita Madhav Purandare

Thesis submitted to
The Faculty of Medicine
University of Glasgow
For the degree of Doctor of Philosophy

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

DECLARATION

I certify that this thesis does not contain material previously published or written by any other person except where referred to in the text and that the results in this thesis have not been submitted for any other degree or diploma.

SMITA PURANDARE

CONTENTS

ACKNOWLEDGEMENTS

I would like to express my gratitude to a number of people,

Professor J.M. Connor for his supervision and guidance throughout the project.
Dr George Lanyon, for his patience, supervision and helpful discussions.
All the clinicians who organised the collection of samples and the patients for their participation in this study.
To all the CMG staff, especially Dr Sandy Cooke.
To the staff of cytogenetics, especially Jim Colgan, Billy Baird, Isabel Boyd and Janet Stewart.
My appreciation also extends to all in the PGR lab and Dev Gen lab.
I would specially like to thank David Stevenson, Gordon Graham, Anne Theriault and Betty O'Hare.
Carol, who has always been a good friend and for all her help,
My family for having faith in me,
And finally my husband Sanjay, for his continuous support and encouragement without which none of this would have been possible.

ABBREVIATIONS

A	adenine
a.a	amino acid
APS	ammonium persulphate
ATP	2'-adenosine 5'-triphosphate
bp	base pair
C	cytosine
CCM	chemical cleavage of mismatches
Ci	Curie (3.7×10^{10} Becquerel)
cpm	counts per minute
cm	centimetre
cDNA	complementary deoxyribonucleic acid
CNS	central nervous system
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytosine 5'-triphosphate
ddNTP	dideoxyribonucleoside triphosphate
DEPC	diethyl pyrocarbonate
DGGE	denaturing gradient gel elctrophoresis
dGTP	2'-deoxyguanosine 5'-triphosphate
DNA	deoxyribonucleic acid
DTT	dithiotreitol
dTTP	2'-deoxythymidine 5'-triphosphate
G	guanine
g	gram
GAP	GTPase activating protein
GTP	Guanosine triphosphate
GDP	Guanosine diphosphate
GRD	GAP related domain
HA	heteroduplex analysis
IRA (1&2)	Inhibitory regulators of RAS
kb	kilobase
kDa	kilodalton
M	molar
min	minute
MOPS	3' (N-morpholino) propanesulfonic acid
mg	milligram
ml	millilitre
µg	microgram
µl	microlitre
$MgCl_2$	magnesium chloride
$MnCl_2$	manganese chloride
mM	millimolar
mRNA	messenger ribonucleic acid
MZ	monozygotic
NaI	sodium iodide
NF-1	neurofibromatosis type 1
NH_4OH	ammonium hydroxide

ng	nanogram
NNFF	National Neurofibromatosis Foundation
nm	nanometre
OD	optical density
OsO_4	osmium tetroxide
PCR	polymerase chain reaction
pmol/pm	picomole
PNS	peripheral nervous system
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RNAse	ribonuclease
r.p.m	revolutions per minute
RT-PCR	reverse transcription PCR
SDS	sodium dodecyl sulphate
SSC	standard saline citrate
SSCP	single stranded conformational polymorphism
T	thymidine
TBE	tris borate ethylenediamine tetra acetic acid
TE	tris ethylenediamine tetra acetic acid
TEMED	N,N,N,N-tetramethylethylenediamine
Tris	tris (hydroxymethyl) aminomethane
UV	ultraviolet
V	volts

LIST OF FIGURES

LIST OF TABLES

SUMMARY

Neurofibromatosis-1 is an autosomal dominant disorder with a prevalence of approximately 1 in 3000 individuals. Its manifestations involve tissues derived from the neural crest and include mainly cafe au lait spots, neurofibromas and Lisch nodules. The gene for NF-1 was identified in 1990 and found to encode a protein, neurofibromin, with sequence similarity to a family of GTPase activating proteins (GAP). The region of homology is called the NF-1 GAP related domain (NF-1 GRD). Its expression has been shown to complement yeast strains deficient in the yeast GAP homologues IRA1 and IRA2 and to interact with human ras proteins and accelerate the conversion of active GTP bound ras to inactive GDP bound ras. Neurofibromin is also known to associate with cytoplasmic microtubules and the connection between ras mediated signal transduction and the cytoskeleton suggests that neurofibromin may play multiple roles in the regulation of cell division.

Neurofibromatosis type-1 is caused by mutations in the NF-1 gene. Mutation analysis in NF-1 is complicated due to the large size of the gene, which extends for over 300 kb on chromosome 17 and is made up of more than 49 exons, due to its high mutation rate and due to the presence of NF-1 pseudogenes and homologous sequences.

The present study involved the characterisation of germline and somatic mutations within the NF-1 gene. Characterisation of germline mutations was carried out in 25 randomly selected, unrelated patients with neurofibromatosis type-1 from Scotland and included both inherited and sporadic cases. Characterisation of somatic mutations was carried out in tumours unrelated to NF-1, to assess the tumour suppressor function of the NF-1 gene.

The strategy for germline mutation analysis involved initial amplification of the NF-1 coding sequence by the polymerase chain reaction (PCR) using both DNA and RNA as templates. In order to do so, PCR primers were designed to amplify 78% the NF-1 coding sequence. Primers for amplification of selected individual exons from genomic DNA were also designed. After initial amplification by PCR, the products were electrophoresed on agarose gels in order to check for any abnormal alterations in size. If no alteration was identified, the segments amplified using DNA as a template i.e. individual exons of the NF-1 gene were analysed using single stranded conformational analysis and chemical cleavage analysis. The segments amplified using RNA as a template were larger in size (0.4-1 kb) and thus were directly subjected to chemical cleavage analysis, to precisely locate the presence of small alterations or point mutations within the NF-1 gene. Any mismatch detected by chemical cleavage was then fully characterised using direct sequencing by the dideoxy chain termination method, using single stranded DNA generated by asymmetric PCR amplification. Larger rearrangements within the NF-1 gene were screened by Southern blotting of genomic DNA.

Using the above strategy, 17 positive screening results were detected on analysis of 78% of the coding sequence in 25 patients. Of these, 13 have been characterised by direct sequencing. The mutations include three splice site errors responsible for exon skipping, two other gross abnormalities of the NF-1 mRNA resulting from a partial deletion in exon 16 and the complete deletion of exon 18, two insertions, a nonsense mutation, two missense mutations, three silent mutations and a novel intragenic polymorphism in intron 41.

The mutation 3113+1G to A affected the splice donor site of intron 18 and resulted in the skipping of exon 18 from the NF-1 mRNA. This loss of exon 18 did not cause a shift in the reading frame and is predicted to result in a loss of 41 amino acids from the protein product. The 41 amino acids lost include two cysteine residues at positions 1016 and 1036, whose loss may lead to altered conformation / stability of neurofibromin. This mutation was identified in a familial case of NF-1.

5749+2T to G is a splice site mutation that affects the invariant GT dinucleotide of the splice donor site of intron 30 and was identified in a sporadic case of NF-1. This mutation resulted in the skipping of exon 30 and a shift in the translational reading frame. This is predicted to result in a truncation of the protein product due to the translation of a single altered amino acid before the termination at a premature stop codon at position 1851.

1721+3A to G is a splice site mutation at position +3 of the donor site in intron 11 and was identified in a familial case of NF-1. This transition resulted in an error of splicing, leading to skipping of exon 11 from the NF-1 mRNA. This is predicted to cause a shift in the translational reading frame, resulting in the formation of 12 altered amino acids and the creation of a premature stop codon at position 560. This would result in the synthesis of a protein of 559 amino acids instead of the normal 2818 amino acids, which would lack the NF-1 GAP related domain.

Two other gross abnormalities of the NF-1 mRNA were identified by RT-PCR and confirmed by direct sequencing. These included a 229bp deletion in exon 16, identified in a sporadic case of NF-1 and a complete deletion of exon 18 from the NF-1mRNA, in a familial case of NF-1. Sequence analysis

of the region surrounding exon 18 revealed a novel homologous sequence to the NF-1 gene in this region.

Four small alterations were detected within the NF-1 gene on analysis of the cDNA segments. These were as follows: A missense mutation G1166D at codon 1166 in exon 21, caused by a G to A substitution at nucleotide 3497 was identified in a familial case of NF-1. K1419R is a missense mutation at codon 1419 in exon 24, caused by an A to G transition at nucleotide 4256. This mutation created a *Mnl* I site and was identified in a sporadic case of NF-1. G1404G is a silent mutation at codon 1404 in exon 24, caused by a G to A substitution at nucleotide 4212. This mutation created an *Apo* I site and was identified in a sporadic case of NF-1. S1311S is a silent mutation at codon 1311 in exon 23, caused by a C to T substitution at nucleotide 3933. This mutation resulted in the loss of a *Fok* I restriction site and was identified in a familial case of NF-1.

Exons 28-36 and 42-44 were amplified from genomic DNA and were subjected to SSCP and chemical mismatch cleavage analysis. Positive screening results were further characterised using direct sequencing. The mutations identified on screening these exons were as follows: A single base insertion of guanine, 6519insG was detected in exon 34 that resulted in a shift in the reading frame and premature termination of translation at codon 2220. This mutation was identified in a familial case of NF-1. R2496X is a nonsense mutation in exon 42, caused by the conversion of a C to T at nucleotide 7486. This transition, at a hypermutable CpG dinucleotide, converts an Arginine to a termination codon at position 2496. This mutation was identified in a sporadic case of NF-1. 7485insGG is a frameshifting two base insertion in exon 42, predicted to cause the translation of 5 altered amino acids before creating a premature stop codon at position 2502. A

consequence of this mutation will be the synthesis of a truncated protein product of 2501 amino acids instead of the normal 2818 amino acids. Interestingly, the insertion was found to affect the same codon (Arginine 2496) which was converted to a premature stop codon in the previous patient (R2496X) and was identified in a sporadic case of NF-1. N1776N is a silent mutation at codon 1776 in exon 29, caused by a C to T substitution at nucleotide 5328. This mutation created a *Hinf* I restriction site and was identified in a sporadic case of NF-1.

A novel intragenic polymorphism was identified in intron 41, the change being that of an A to G transition, 28 bases upstream of the first base of exon 42. 75 unrelated Caucasian individuals of Scottish origin were studied by heteroduplex analysis which showed three distinct electrophoretic patterns and an allele heterozygosity of 47% was calculated. This polymorphism will be useful in families where the disease causing mutation remains to be identified.

Four other chemical cleavage results were detected on analysis of the NF-1 cDNA segments and were localised to exons 2, 4, 16 and 27a. Two possible polymorphisms were detected, in exon 16 (as the same cleavage product was identified in 4 different patients) and in exon 28, where SSCP analysis revealed three distinct patterns.

An interesting feature presented in a case of segmental neurofibromatosis on RT-PCR of the NF-1 mRNA surrounding the alternatively spliced isoform of the NF-1 gene, NF-1 GRD II. Amplifcation and analysis revaled equal expression of both isoforms in this patient, while the other patients and normal controls showed expression of only the type I isoform. This may not be the cause of disease but was the only finding on screening 78% of the coding sequence.

The spectrum of the above 13 and other reported mutations within the NF-1 gene is discussed and their mechanisms of mutagenesis are presented. The distribution of mutations was random and in this study no additional mutation hotspots were identified on analysis of 78% of the coding sequence. However the remaining 22% of coding sequence and other regions important for efficient expression remain to be analysed. A genotype-phenotype correlation is presented, which will help in drawing conclusions regarding the function of the various regions within the NF-1 gene and may help in determining the molecular basis for the wide variation in clinical features.

Screening for somatic mutations (small alterations) in exon 24 (FLR exon) of the NF-1 gene (by SSCP and CCM) was carried out in the following nine tumour types unrelated to NF-1: colonic adenocarcinoma (2), colonic adenoma (2), pancreatic carcinoma (2), seminoma (2), melanoma (3), lung adenocarcinoma (2), thyroid carcinoma (2) and myelodysplastic syndrome (10). Screening for large alterations (by Southern and Northern blotting) was carried out in 4 neuroblastoma samples.

The role of NF-1 mutations was considered in the tumour types in which activated ras genes are frequently found. NF-1 mutations, considered to be the functional complement of ras mutations, were looked for in tumour samples lacking ras mutations. Initially the tumour samples were screened for activating mutations in the appropriate group of ras oncogenes (H-RAS, K-RAS, or N-RAS) depending on the tumour type. Two shifts on SSCP analysis were identified in the PCR products encompassing codons 12-13 of the K-ras oncogenes in a colonic adenoma sample and a colonic adenocarcinoma sample. No small or large somatic mutations in the NF-1 gene were detected in the tumour types analysed in this study.

Thus in this study a heterogeneous group of germline mutations was identified in the NF-1 gene and a mutational spectrum was established. The majority of germline mutations identified and characterised in this study are predicted to cause disruption of the protein product neurofibromin. Correlation of genotype with the phenotype in the cases with pathogenic mutations characterised in this study did not show a clear correlation and a larger number of mutations will have to be analysed.

The strategy used in this project for characterisation of mutations in the gene for neurofibromatosis type-1 has proved to be useful and may be applied to the detection of molecular pathologies in general.

Note: Some of the data presented in this thesis have been published. Reprints are bound at the back.

CHAPTER: 1

INTRODUCTION

CHAPTER 1: INTRODUCTION

1.1 HOW MUTATIONS CAUSE DISEASE

The gene is the unit of inheritance. Each gene is a nucleic acid sequence that carries information representing a particular polypeptide. A gene is a stable entity, but can acquire a change in the base sequence. Such a change is called a 'mutation'. When a mutation occurs, the new form of the gene is inherited just like the previous form. The organism carrying the altered gene is called a mutant and an organism carrying the normal (unaltered) gene is called the wild type. Mutations can be divided into two general classes: point mutations and rearrangements or length mutations. Point mutations are changes affecting a single position in a gene, rearrangements or length mutations affect larger regions and the simplest type are insertions or deletions. Mutations will be copied at subsequent replications unless a reverse mutation occurs.

By and large mutations that produce disease manifest in two ways: mutations which cause synthesis of an abnormal gene product and those which cause a reduction or absence of a protein product. Mutations which cause synthesis of an abnormal gene product comprise single base substitutions (missense mutations) which alter the structure of the mRNA such that an abnormal protein product is synthesised. The change in function or stability of the protein results in an abnormal phenotype. Mutations causing decreased output or absence of gene product are mutations affecting transcription or processing of mRNA or mutations that act at the translational level by interfering with initiation, elongation or termination.

Mutations which lead to a greatly reduced level or no protein product are classified by immunological protein assay methods as cross reacting

material negative, (CRM -ve), whereas those mutations which produce a protein product in normal amounts but with aberrant function are CRM +ve. Mutations can occur in the coding sequences, which are the exons, that comprise mature RNA and the introns or intervening sequences (IVS), which are transcribed but then cut out of the transcript during mRNA processing. Although the boundaries of the coding sequences can be precisely defined, other less well defined sequences such as the promoters, enhancers and 5' and 3' untranslated regions are required to constitute a unit which is transcribed efficiently in the appropriate tissues and at the proper time.

1.1.1 MUTATIONS CAUSING SYNTHESIS OF AN ABNORMAL GENE PRODUCT

A single base substitution or a more subtle rearrangement of a gene may alter the structure of the mRNA such that an abnormal protein product is synthesised. Point mutations can be divided into two types depending upon the nature of the change when one base is substituted for another. The more common class of mutation is the TRANSITION, comprising the substitution of one pyrimidine by the other or of one purine by the other. The other class is the TRANSVERSION, in which the purine is replaced by a pyrimidine or vice versa.

Mutations induced by a base substitution are often 'leaky' (the mutant has some residual function). This situation arises when the sequence change in the corresponding protein does not entirely abolish its activity. A point mutation that alters only a single base will change only the one codon in which that base is located so only one amino acid is affected in the protein. While this substitution may reduce the activity of the protein, it may not abolish it entirely.

Some single base substitutions have no apparent effect and they fall into two types: One group involves base changes in DNA that do not cause any change in the amino acid present in the corresponding protein and are called *silent mutations*. Others change the amino acid (usually with an amino acid of the same class-basic / acidic / neutral, hydrophilic / hydrophobic) but the replacement in the protein does not affect its activity and the phenotypic effects are small. These are called *neutral substitutions*, or *conservative changes*.

The results of single amino acid substitutions in proteins vary depending on the type of amino acid that is substituted and the site of substitution in a particular protein. However, many amino acid substitutions have no effect on function or stability. The primary amino acid sequence of a protein or its subunits determines the way it assumes its secondary structure. Some amino acid substitutions in the primary amino acid sequence go on to cause abnormalities of a tertiary structure and to reduce the overall stability of the protein or its subunit. For example, proline cannot participate in an alpha helix except as one of the initial three residues. Thus the substitution of proline by another amino acid can sometimes seriously disrupt helical conformation and result in protein instability. Most proteins are folded into a complex tertiary configuration, so that most of the charged amino acids such as lysine, arginine, glutamic acid and aspartic acid are found on the surface of the molecule, allowing their ionised groups to be in contact with water. On the other hand, residues oriented towards the interior of the molecule have non polar groups; thus the inside of the molecule is stabilised by hydrophobic interactions. The substitution of a charged for an uncharged residue can disrupt these important interactions and lead to molecular instability.

20 amino acids are used to synthesise proteins and are classified by their ionic charge into four groups (Lewin, 1993):

1)Basic: lysine, arginine and histidine 2)Acidic: aspartic acid and glutamic acid 3)Neutral and polar amino acids: amino acids that have no net charge are neutral. Some of the neutral amino acids are polar i.e. electrically charged because of the distribution of charges within the molecule. These are: glycine, serine, threonine, tyrosine, cysteine, glutamine, asparagine 4)Neutral and hydrophobic: the apolar amino acids are hydrophobic. These are: alanine, valine, leucine, isoleucine, proline, tryptophan, phenylalnine and methionine.

Occasionally, the substitutions may involve chain termination or initiation codons or may disrupt the genetic code such that either elongated or shortened peptide chain products are produced. Extended gene products result from the following mechanisms: base substitution in the chain termination codon or preservation of the initiator methionine residue. Shortened gene products are due to substitutions resulting in premature termination of translation.

1.1.2 MUTATIONS CAUSING DECREASED OUTPUT OF GENE PRODUCT

The second group of mutations are those that cause a reduction or absence of a particular protein product. These mutations can be divided into mutations that cause defective transcription, defective processing of messenger RNA and mutations that interfere with translation.

Mutations causing defective transcription

These comprise deletions, insertions, inversions, duplications, fusion genes and promoter mutations.

a)Deletions: e.g. DMD, cystic fibrosis etc. Deletions are the physical absence of DNA sequence. Deletions can result due to non-homologous recombination, from faulty repair of damaged DNA, or from a skip during DNA replication. The effects of gene deletions are as follows:

Table 1: Effects of gene deletions

Deletion of	Consequences
many genes	chromosomal abnormalities
one whole gene	no gene function
exon material	truncated protein, often unstable, may generate a frameshift
intron material	usually no phenotypic effect
splice site	usually no functional product/ or truncated product
promoter	reduced/absent gene function

b)**Insertions**: e.g. Lesch-Nyhan syndrome, Haemophilia B etc. Insertions may be large sequence insertions, or insertion of a few bases into gene coding regions causing disease. These involve insertion of either novel bases, or of a specific DNA sequence.

c)Inversions: e.g. Haemophilia A. The term inversion signifies that a region of DNA is back-to-front with respect to its normal orientation in the genome.

d)Fusion genes: e.g. red/green colour blindness. There are three separate genes for each of the cone pigments responsible for colour vision, blue (on chromosome 7) and red and green (Xq28). There is a single red gene on each X chromosome and 1-3 copies of the green genes. The red and green genes have 96% sequence homology and unequal crossing over in the area can result in the formation of hybrid genes or fusion genes, which produce pigments of altered function.

e)Promoter mutations: e.g. haemophilia B (Leyden), retinoblastoma.

The haemophilia B Leyden phenotype is seen in patients with different mutations within the promoter region of the factor IX gene.

f)Duplications: e.g. CMT1A, DMD/BMD. Duplications can be tandem or inverted. Duplication of a whole gene will cause dosage effects only, but partial duplications can result in disease in two general ways: In one, duplication of several amino acids of a protein may change the protein conformation, leading to an unstable / dysfunctional protein. In the other, the duplication may result in a reading frame shift in the mRNA, producing a truncated and usually dysfunctional protein.

Mutations causing defective processing of messenger RNA

These comprise splice junction and consensus site mutations, mutations creating cryptic splice sites in exons and introns and polyA signal site mutations.

a)Splice junction and consensus sequence mutations: e.g. Ehler Danlos syndrome, α and β thalasaemia. Splice site mutations alter the 5' donor (GT) or 3' acceptor (AG) sequences at the ends of an intron which are required for excision of the intron during mRNA processing and usually abolish gene function. In addition to the GT/AG junctional sequences, there are highly conserved sequences at the boundaries between introns and exons that also must be involved in splicing of mRNA. Several forms of β thalassaemia are described, which result from the production of alternatively spliced sites within these sequences.

b)Cryptic splice sites in introns and exons e.g. β thalassaemia. In β thalassaemia, single base substitutions within introns may result in preferential alternative splicing of the precursor β mRNA molecules at the

site of mutation. The mRNA produced as a result of the abnormal splicing contains intron sequences and therefore is useless as a template for globin chain synthesis, so more abnormal than normal mRNA is produced.

Mutations have been also found in the exons of the globin genes that seem to activate cryptic splice sites. e.g.: The structural haemoglobin variant haemoglobin E, due to a G to A transition at position 26 results in the activation of a cryptic spice site, which competes with the normal splice site and leads to a reduced output of β globin chains.

c)Polyadenylation/cleavage signal mutations: e.g. α and β thalassaemia. Polyadenylation signal site mutations also interfere with the normal processing of mRNA. For example, the single base change AATAAA to AATAAT, in the α globin genes of patients with a certain form of α thalassaemia. Due to the mutation, instead of the normal cutting and polyadenylation of the mRNA precursor, a long molecule is produced which does not appear in the cytoplasm. A small amount of polyadenylated mRNA is produced, but the overall effect is to almost entirely inactivate the affected globin gene.

Mutations causing defective translation

These comprise initiation codon mutations, termination codon mutations, nonsense mutations and frameshift mutations.

a)Initiation codon mutations: e.g. phenylketonuria, Tay-Sachs disease and α & β thalassaemia. Several mutations have been observed in patients with α thalassaemia which involve either the initiation codon itself or the sequences immediately adjacent to it and no α chains are produced from the affected α globin gene.

b)Termination codon mutations: e.g. α thalassaemia. Termination codon mutations have been observed only in the α globin genes. The base alteration results in the insertion of an amino acid instead of chain termination. Messenger RNA sequence that is not normally utilised is then translated until another in-phase stop codon is reached, resulting in elongated but stable α chains (Hb Constant Spring).

c)Nonsense mutations: e.g. Haemophilia A, B, Lesch-Nyhan, OTC deficiency, porphyrias. Nonsense mutations replace the codon for an amino acid by a stop signal (UGA, UAA, UAG) leading to premature chain termination, with the production of a shortened and physiologically useless peptide fragment.

d)Frameshift mutations: e.g. Duchenne muscular dystrophy. Frameshifts are a consequence of the way the genetic code in the mRNA is read in triplets. Any change in the number of nucleotides in a coding sequence which does not add or remove complete triplets alters the reading frame of all the message downstream of the change. Frameshifts usually abolish the function of the protein. Sometimes the altered base sequence generates a new termination codon leading to premature termination of translation of the abnormal RNA. If the normal stop codon is thrown out of sequence, the scrambled mRNA is translated until another stop codon is produced, leading to an elongated translation product.

1.2 MECHANISMS OF MUTAGENESIS

Cytosine methylation and the role of CpG dinucleotides in disease

The methylation of mammalian DNA at the cytosine residues, is one of the most common forms of DNA modification and is involved in a number of cellular and developmental processes (Cooper and Youssoufian, 1988). The

dinucleotide CpG is a hotspot for mutation in the human genome. This is as a result of the modification of the 5' cytosine by cellular DNA methyltransferases and the consequent high frequency of spontaneous deamination of 5-methylcytosine. Methylated cytosine residues are relative hotspots for mutations (25-35% of all point mutations) because with deamination they produce thymidine (with substitution of adenine for guanine on the complementary strand), which is not recognised by DNA repair mechanisms and thus results in mutation (Cooper and Krawczak, 1991).

Mechanisms of insertional mutagenesis

Insertions of either novel bases or of a specific DNA sequence into a gene coding region usually result in alteration of the reading frame of the encoded protein, which may lead to termination of translation at some distance downstream. Examples of the insertion of <10bp of DNA sequence into human gene coding regions causing genetic disease was analysed by Cooper and Krawczak (1991), in order to study the underlying mechanisms and the following observations were made: insertional mutation, involving the introduction of <10bp of DNA sequence into a gene coding region is not a random process and appears to be highly dependent on the local DNA sequence context. The majority of insertion type mutations are consistent with an explanation which involves an endogenous replication associated mechanism of mutagenesis and may be explained either in terms of a)direct repeats/runs of single bases causing slipped mispairing b)inverted repeats or c)symmetric elements facilitating the formation of secondary structure intermediates.

Mechanisms of deletional mutagenesis

Deletional events in human genes are at least in part sequence directed and the frequency of occurrence reflects underlying structural differences between genes. Human gene deletions appear to be caused by multiple mechanisms, whose relative importance is influenced by local primary and secondary DNA structure. Reports describing short (<20bp) gene deletions causing human disease were analysed by Krawczak and Cooper (1991) in order to study underlying causative mechanisms. Direct repeats are a feature of a number of recombination, replication or repair based models of deletion mutagenesis. Palindromes or inverted repeats could potentiate the looping out of single stranded DNA. Inverted repeats may promote instability by facilitating the formation of secondary structure intermediates. A significant excess of symmetrical sequence elements was found at the sites of single base deletions. These elements were seen to possess an axis of internal symmetry (e.g. CTGAAGTC, GGACAGG) and varied between 5bp and 11bp in length. In addition a consensus sequence proposed to be hot-spot for deletions was drawn up: (TGA/GA/GG/TA/C).

Mechanisms of gene duplication

Tandem duplication involving parts of genes is now recognised as a contributor to the mutational spectrum that results in genetic disease. The mechanisms of duplication formation were analysed by Hu and Worton (1992), with special emphasis on the molecular details of the nucleotide sequences at the duplication junctions. Partial gene duplication can result in disease in two general ways. In one, duplication of several amino acids of a protein may change the protein conformation, leading to an unstable / dysfunctional protein. In the other, the duplication may result in a reading frame shift in the mRNA producing a truncated and usually dysfunctional

protein. Two general mechanisms leading to duplications and deletions have been proposed: homologous and non-homologous recombination. In the former sequence homology between the two parental strands in the regions of cross-over is needed for strand exchange. Homologous recombination may involve recombination between repetitive sequences, i.e. the Alu elements. Non homologous recombination involves recombination between unrelated (non-homologous) sequences and probably the duplications may have been created by random chromatid cleavage and rejoining events. Unequal sister chromatid exchange, i.e. unequal crossing over between sister chromatids, rather than between non-sister chromatids of two X chromosomes, is the predominant event for producing duplications in the dystrophin gene.

Expansion of trinucleotide repeats as a mechanism of mutagenesis

The unstable expansion of trinucleotide repeats represents one of the previously unrecognised mechanism of mutagenesis and so far seven examples of triplet repeat diseases are known. These are as follows: fragile X syndromes (FRAXA and FRAXE) (Fu et al., 1991), myotonic dystrophy (DM) (Fu et al., 1992), Kennedy's disease (spinal and bulbar muscular atrophy SBMA) (La Spada et al., 1991), Huntington's disease (The Huntingtons Disease Research Collaborative Group, 1993), spinocerebellar ataxia (Orr et al., 1993) and the most recently discovered DRPLA (dentatatorubral and pallidoluysian atrophy) (Koide et al., 1994). All the diseases are due to a novel mutational mechanism by which normally polymorphic trinucleotide repeats expand beyond the normal size range and result in changes involving gene expression (FRAX), message stability (DM), or gain of function (SBMA and HD).

1.3 MUTATION SCREENING METHODS

The spectrum of mutations found may range from cytogenetically visible chromosome rearrangements to large and small alterations within the gene, including single base changes. The mutation screening methods can be divided into screening methods for detection of unknown mutations (small and large alterations) and methods for detection of known mutations (Loss or gain of an RFLP, ARMS, ASO and PCR directed mutagenesis). While several useful techniques for detection of sequence heterogeneity exist, no single method is applicable for all situations. The different methods used have complementing strengths and a combination of methods can be used successfully for mutation detection. The various methods are outlined below and the principles, advantages and disadvantages are discussed and compared in to order to evaluate the screening methods available and apply them to screen for mutations within the NF-1 gene.

Detection of small alterations

Small alterations include deletions, insertions and single base substitutions which comprise nonsense, missense, silent and splice mutations. All the methods used are PCR based detection methods and rely on PCR amplification of sample DNA or RNA prior to analysis. The methods discussed include: 1)Single stranded conformation polymorphism analysis (SSCP), 2)Heteroduplex analysis (HA) 3)Denaturing gradient gel electrophoresis (DGGE), 4)RNase cleavage and chemical mismatch cleavage analysis (CCM), 5)Protein truncation test (PTT), 6)Direct sequencing and its modifications.

Single Stranded Conformational Polymorphism (SSCP) Analysis

This procedure was first described by Orita et al. (1989). Single stranded DNA molecules assume a three-dimensional conformation which is dependent on the primary sequence. If a sequence difference exists between wild type and mutant DNA, this may result in differential migration of one or both the mutant strands when the products are denatured and then electrophoresed through a non denaturing polyacrylamide gel, under different sets of electrophoretic conditions. It is a simple and relatively sensitive technique, with no additional steps required after PCR. A large number of samples can be analysed simultaneously. However, SSCP analysis detects mutations but does not localise them within a fragment. SSCP analysis detects 70-95% of mutations in PCR products of 200bp or less and the sensitivity of the method is less than 50% when fragments >400bp are analysed (Grompe M., 1993). To increase the sensitivity of the technique, a radioactive label may be necessary, but silver staining of the gels also shows equally good results. Size limitations may be overcome by restriction digestion of a larger amplification product prior to electrophoresis. Occasionally there may be occurrence of additional bands which may correspond to different stable conformations (conformers) of the same sequence and whose presence may be sensitive to the temperature employed.

A modification of SSCP termed RNA conformation polymorphism analysis (rSSCP) has been described by Sarkar et al. (1992a). RNA is used instead of DNA, the RNA being generated by T7 RNA polymerase transcription, from a PCR amplified DNA template fragment. Sense and antisense RNA strands yielded different conformational patterns and this method has been claimed to be more successful than conventional SSCP

(Sarkar et al, 1992a). A second modification is the dideoxy fingerprinting (ddF) method, which combines SSCP analysis and direct sequencing and was described by Sarkar et al. (1992b). After amplification the PCR product is sequenced with dideoxyCTP to generate a C-ladder (fingerprint) of bands. The sequencing reactions from wild type and mutant samples are then electrophoresed through a non denaturing polyacrylamide gel. Mutations are detectable as shifts of individual bands in the ladder. Another modification of SSCP was proposed by Lazaro and Estivill, (1992). Their assay consisted of generation of single stranded fragments by asymmetric PCR, which were then subjected to SSCP analysis and visualised by ethidium bromide staining.

Heteroduplex analysis

Mutant and wild-type sequences, when present simultaneously in a PCR reaction, form heteroduplexes during the late cycles in a PCR reaction (Nagamine et al., 1989). Heteroduplex molecules with a single base pair variance may show different mobility from homoduplexes in regular polyacrylamide gels (White et al., 1992), as well as other newer polyacrylamide based gel matrices such as Hydrolink™ and MDE™, due to sequence dependent conformational changes in the dsDNA (Tassabehji et al., 1992). The technique is very simple and is equally effective using non-isotopic detection methods (ethidium bromide staining) and detects 80-90% changes in fragments less than 300bp (Grompe M., 1993). However it cannot be solely used as a mutation detection technique due to its inability to detect all types of mutations and due to the fact that there is decreased sensitivity when large products are analysed.

Denaturing gradient gel electrophoresis (DGGE)

DGGE is also a technique that allows the separation of DNA molecules differing by single base changes due to differential electrophoretic migration of wild type and mutant DNA (Fischer and Lerman, 1983, Myers et al., 1985). The separation is based on the fact that DNA molecules differing by a single base change have slightly different melting properties. These cause them to migrate differently into discrete sequence dependent domains of low melting temperature, in a polyacrylamide gel containing a linearly increasing gradient of DNA denaturants such as urea and formamide, from top to the bottom of the gel. There are two types of denaturing gradient gels, a)parallel gels which contain a linearly increasing gradient of DNA denaturants from top to bottom in the gel. b)perpendicular gels which contain a linear gradient of denaturants from left to right across the gel. The denaturing gradient can also be generated by temperature and this method is termed TGGE. The sensitivity of DGGE is greatly enhanced if heteroduplex DNA between wild type and mutant sequences is used for analysis (Sheffield et al., 1989). Wild type and mutant PCR products are denatured and reannealed to generate four species of hetero and homoduplexes. Differential melting behaviour of these heteroduplexes leads to altered migration in a gradient of denaturing gels compared to wild type homoduplex. Once the appropriate PCR primers and denaturant conditions (using computer programs to predict theoretical melting profiles and design PCR primers) have been developed for a specific region DGGE is a highly reliable and rapid method for mutation detection. Single base differences can be detected with about 95% accuracy in PCR products of up to 600bp in length and detection is usually carried out by non-radioactive means (Grompe M., 1993). However this method requires a special apparatus to

control gel temperature and long PCR primers which include a 30-40 nucleotide "GC clamp" at the 5' end of the primer to ensure that the amplified sequence has a low dissociation temperature (Sheffield et al., 1989, Myers et al., 1985). Also the location of the sequence difference within the fragment cannot be localised and has to be determined by sequencing.

RNase cleavage and Chemical mismatch cleavage analysis

The principle of heteroduplex analysis is applicable to both RNase and chemical mismatch cleavage and the two procedures were described by Myers et al. (1985) and Cotton et al. (1988), respectively.

In RNase A cleavage an RNA-DNA heteroduplex between a radioactive wild type riboprobe and mutant DNA generated by PCR is subjected to cleavage by RNase A. The enzyme will recognise and cleave single stranded RNA at the points of mismatch. The reaction is analysed by electrophoresis and autoradiography. The presence and location of the mutation is indicated by a cleavage band of a given size. However use of radiolabelled RNA is necessary and RNase A can detect only 50% of mismatches.

Chemical mismatch cleavage analysis (CCM) is based on the principle of creation of heteroduplexes between radiolabelled wild type and mutant DNA or RNA molecules. Chemical modification at the sites of mutation is then carried out using hydroxylamine and osmium tetroxide. This is then subjected to cleavage at the site of modification using piperidine followed by denaturing gel electrophoresis and autoradiography. CCM is very sensitive detecting > 95% of mismatches when only the wild type DNA is labelled and 100% when both wild type and mutant DNA are labelled (Forrest et al., 1991). PCR products of up to 1.7kb can be screened, permitting efficient

screening of amplified mRNAs (Grompe M., 1993). Finally, the sequence alteration is detected reliably and the precise localisation and nature of the change is also indicated by the size of the cleavage band and cleaving reagent. However it is a relatively complex and lengthy procedure involving use of radioactivity and toxic chemicals.

Protein truncation test (PTT)

The above techniques, used to recognise point mutations in genetic disease detect all sequence differences including phenotypically silent changes. Consequently the methods are not convenient to analyse mutations in large multiexonic genes where a large fraction of pathological point mutations produce early termination. The technique of PTT, or protein truncation test is based on the combination of RT-PCR, transcription and translation and selectively detects translation terminating mutations and was first described by Roest et al. (1993). This technique does not detect phenotypically silent alterations. The site of mutation is localised, so only a small part of the gene needs to be sequenced. Comparatively larger (2.4kb) stretches of coding sequence can be screened. However this method can only detect translation terminating mutations.

Direct sequencing

Direct DNA sequencing refers to the direct sequence analysis of PCR products without prior subcloning into sequencing vectors and can be a primary method of mutation detection, if the coding sequence of the gene to be screened is relatively small (Mgone et al., 1992). Direct sequencing involves the synthesis of a DNA strand by a DNA polymerase *in vitro* using a single stranded DNA template (Sanger et al., 1977). Template DNA is

purified and annealed to a synthetic oligonucleotide primer. The DNA synthesis is carried out in 2 steps, the first a labelling step and the second the chain termination step using dideoxynucleotides. A radioactively labelled nucleotide is included in the synthesis so the labelled chain of variation can be visualised by autoradiography after separation by high resolution electrophoresis on denaturing gels. DNA sequencing defines the location and exact nature of the change and therefore is the necessary final step of any mutation detection method.

In order to sequence PCR products successfully by the conventional dideoxy termination protocol, it is essential to convert the double stranded product into a single stranded template for sequencing. The methods described are:

1)Asymmetric PCR: This is a modified type of PCR to produce single stranded DNA of a chosen strand and utilises an unequal or asymmetric concentration of the 2 amplification primers.(Gyllensten and Erlich, 1988). In asymmetric PCR, during the initial 15-25 cycles, most of the product generated is double stranded and accumulates exponentially. As the low concentration primer becomes depleted, further cycles generate an excess of one of the two strands depending on which of the amplification primers is limited. The primer ratios used for asymmetric PCR are usually 50pmoles:1pmol or 100pmoles:1pmol. The single stranded DNA accumulates linearly and is complementary to the limiting primer.

2)Biotinylation of one of the primers: In this method, one of the PCR primers is biotinylated. After the reaction the double stranded PCR product is captured on an avidin coated magnetic bead. The nonbiotinylated strand is separated with NaOH and the sequencing reactions are carried out on the immobilised single stranded template (Gibbs et al., 1990).

3)Direct sequencing with phage promoters:

This can be done in two ways: RAWTS: RNA amplification with transcript sequencing and GAWTS: genomic amplification with transcript sequencing (Stoflet et al., 1988). In this method, the original PCR primers carry T7 RNA polymerase binding sites. In vitro transcription is then used to generate single stranded template for sequencing. These are methods of direct sequencing that utilise a phage promoter sequence 5' to at least one of the PCR primers. RAWTS is a four step procedure that involves:

a)cDNA synthesis with oligo dT random primers or an mRNA specific oligonucleotide primer.

b)A PCR in which one or both nucleotides contain a phage promoter attached to a sequence complementary to the region to be amplified.

c)Transcription with a phage polymerase and d)dideoxysequencing with a reverse transcriptase.

The procedure for GAWTS is identical except that genomic DNA is the input to step b). The advantages of this procedure are that the transcription step produces an additional level of amplification that obviates the need for purification subsequent to PCR; the amplification afforded by transcription can compensate for a sub optimal PCR; and the generation of single stranded template provides a more reproducible sequence than obtained from a double stranded template. The disadvantages include the fact that there is added expense of attaching phage promoters to the PCR primers.

Detection of large alterations:

Large gene alterations are mutations in which substantial portions of a gene are deleted, duplicated or otherwise rearranged. The techniques available for detection of large alterations can be divided into cytogenetic

techniques and molecular techniques. Cytogenetic techniques include: a)Cytogenetic analysis using conventional and high resolution cytogenetics, b)Flow cytometry using FACS (fluorescence activated cell sorter) and c)Fluorescent In Situ Hybridisation (FISH). Molecular techniques include a)Southern blot hybridisation and b)pulsed field gel electrophoresis (PFGE).

1)Using cytogenetic techniques: Flow cytometry can be used to measure the DNA content of individual chromosomes as they pass in a fluid stream through the laser beam of a FACS (fluorescence activated cell sorter). This technique is useful in identifying chromosome aberrations in particular microdeletions as its lower limit of resolution is 1-2Mb compared with 4Mb for the light microscope used in conventional cytogenetics. FISH utilises fluorescently labelled DNA probes hybridised to chromosome spreads, to detect the presence/absence of a chromosomal region corresponding to the probe and its position within the genome.

2)Using molecular techniques: a)Southern blot hybridisation: this technique was first described by E. Southern in 1975. This a quick method to screen for large alterations and offers a good first step in mutation analysis. By this method, large deletions and insertions may be detected by the presence of junction fragments, or changes in band intensities (for autosomal dominant conditions). Point mutations may be detectable if they alter restriction sites.

b)Pulsed field gel electrophoresis: In contrast to routine DNA analysis techniques which have an upper size limit of 40kb, PFGE can separate fragments in the size range of 1kb-10Mb and hence large areas of a gene can be screened for alterations in a single experiment. PFGE involves switching of the voltage potential relative to the gel during electrophoresis and as the shorter DNA molecules can reorientate to the potential more quickly than the longer molecules, they migrate further into the gel. The large fragments for

PFGE are generated using special restriction enzymes which have infrequent cutting sites such as *Not* I (8bp) and *Sfi* I (13bp) in length.

Methods used for detecting known mutations

Specific point mutations can be detected by several approaches including loss or gain of an RFLP, Allele-Specific Oligonucleotide probes (ASO) and the Amplification Refractory Mutation System (ARMS).

Loss or gain of an RFLP

A difference in the pattern of fragments after restriction enzyme digestion is called restriction fragment length polymorphism. The use of PCR has made it possible to amplify a short region of DNA surrounding the restriction site of interest which is then exposed to the relevant restriction enzyme. The presence of a mutation may destroy a restriction enzyme site or lead to creation of a site. This can be used for rapid detection of that particular mutation on digestion with the relevant restriction enzyme.

Allele specific oligonucleotide hybridisation (ASO)

In this method, ASO probes are used, which are short (17-30 mers) and have the complementary sequence to either the normal or the mutant DNA sequence at the point of interest. Under very stringent conditions, such probes will only hybridise to their perfect homologous sequences and not to those that vary by even a single nucleotide residue. Thus a normal gene can be detected using the wild type probe and the mutant gene with the probe containing the same mutation. In practice, for the routine analysis, dot-blot filter hybridisation is used and detection done by either radioactive or biotinylated and enzyme labelled probes. This method also allows

simultaneous amplification and analysis of multiple polymorphic sites by using ASO probes of identical length and tetramethylammonium chloride (TMACl). When washed in 3M TMACl, perfectly matched oligonucleotides of the same length tend to dissociate from the hybridised test sample at the same temperature, thus leaving behind the mismatched probes. Alternatively, a "reverse dot-blot" procedure can be used in which the oligonucleotide probe is immobilised on a membrane and hybridised to labelled PCR products. In this method a panel of different probes can be used to screen simultaneously for several mutations in a given PCR product.

Amplification refractory mutation system (ARMS)

This technique involves PCR amplification of the target DNA with a normal or mutant primer and a common end primer. The normal and mutant primer differ at their 3' ends only by a single base which corresponds to the normal and mutant alleles. The technique is based on the concept that template amplification by PCR in which one of the amplification primers is mismatched at the 3' end is not possible. This is due to the lack of 3' exonucleolytic proof reading activity of Taq DNA polymerase. The technique is performed by amplifying a test DNA with a mutant amplification primer and a common primer. As a control, internal primers are added in the same reaction. Similarly another set of reactions is performed using normal and common primers, with internal control primers also added to the reaction. After amplification, the products are size fractionated in agarose gels and visualised by ethidium bromide staining. With the normal sample, amplification will only be positive in the reaction with the normal primer, but not in the reaction with the mutant primer. In the case of heterozygotes, it will be positive on amplification with both the normal and mutant primer and in

case of homozygous mutant samples, there will be no amplification with the normal primer.

All the above methods can be used in the investigation of mutations that have already been characterised and are useful as a means of direct detection of the particular mutations.

PCR directed mutagenesis

All the mutations characterised in this study, were confirmed by digestion with the appropriate restriction enzyme site that was created or lost. However not all the single base substitutions characterised in this study altered restriction enzyme sites and therefore were not amenable to detection by this method. Detection of these mutations was therefore carried out using the method of PCR directed mutagenesis as described by Li et al. (1992).

1.4 CLINICAL FEATURES OF NEUROFIBROMATOSIS TYPE-1

Historical Perspective

Neurofibromatosis type-1 (NF-1), also known as von Recklinghausen's neurofibromatosis was first described by Freidreich von Recklinghausen in 1882, who gave the disease its full description, including the recognition that the tumours arose from fibrous tissue surrounding small nerves, leading to the recognition of these as neurofibromas. The autosomal dominant inheritance pattern was defined in the early twentieth century. A crucial diagnostic element, the Lisch nodule was defined by the Viennese ophthalmologist Karl Lisch in 1937. The landmark study of Crowe, Schull and Neel in 1956 brought together for the first time the salient clinical features of NF-1.

Incidence

Neurofibromatosis type-1 is inherited as an autosomal dominant disorder with a 98% penetrance and with an incidence of approximately 1 in 3000 individuals. About one third of cases are new mutations and the mutation rate is approximately 1 per 10,000 gametes per generation.

Diagnostic criteria

The diagnostic criteria for NF-1 were established at an NIH 'Consensus Development Conference' in 1988. The diagnostic criteria for NF-1 are tabulated in Table 2 and are met in an individual if two or more of the criteria are present.

Table 2: Diagnostic criteria for neurofibromatosis type-1

1.	Six or more cafe au lait macules over 5 mm in greatest diameter in prepubertal individuals and over 15 mm in greatest diameter in postpubertal individuals
2.	Two or more neurofibromas of any type or one plexiform neurofibroma
3.	Freckling in the axillary or inguinal regions
4.	Optic glioma
5.	Two or more Lisch nodules (iris hamartomas)
6.	A distinctive osseous lesion such as sphenoid dysplasia or thinning of long bone cortex with or without pseudarthrosis
7.	A first degree relative with NF-1 by the above criteria

The basic disturbance in neurofibromatosis appears to be an abnormality in development of the neural crest cells with resulting tendency to abnormal, excessive growth of affected tissues and the development of multiple tumours.

Defining features present in most patients

Cafe au lait spots

The cafe au lait spot is a flat evenly lightly pigmented macule that does not macroscopically differ from normal skin and appears in 95% of NF-1 patients. Ordinarily they are not apparent at birth, but become visible during the first year of life, when at least six spots measuring from 0.5 to 10 cm are apparent, although the occasional patient may have fewer. Melanocytes within a cafe au lait spot have an increased number of macromelanosomes, although this is not diagnostic for NF-1. Cafe au lait spots are present from early infancy and their numbers (in contrast to freckling which is also a part of this disorder) are relatively constant until late middle age (late 40s and 50s) when the total number of spots declines (Riccardi, 1994).

Peripheral neurofibromas

NF-1 derives its name from its hallmark feature neurofibromas. Neurofibromas may be intracutaneous and are of a violet colour and soft consistency, or subcutaneous and presenting as firm tumours along the trunk of peripheral nerves. Their diameter varies from a few millimetres to 3-4 cm. Pathologically, these lesions are made up of a mixture of cell types, including fibroblasts, Schwann cells, mast cells and vascular elements. These tumours make their appearance in adolescence and increase in size and number with age, although the rate can be extremely variable. Some affected females with NF-1 note an increase in rate of progression during pregnancy, suggesting these tumours may be hormone responsive.

Freckling

The occurrence of freckles in the axillae, groin and intertriginous areas is seen in NF-1. Such freckling is not apparent at birth, but often appears during childhood. The occurrence of such freckling in the inframammary areas and other skin folds may suggest that these lesions are modulated by the local environment.

Lisch nodules

Raised pigmented nodules of the iris pathologically representing hamartomas are called Lisch nodules and represent an extremely important diagnostic feature of NF-1. Lisch nodules are of melanocytic origin and are present in 30% of NF-1 patients by 6 years of age and in over 90% of adult patients.

Frequent but non diagnostic and non morbid features of NF-1

Macrocephaly and short stature are frequent accompaniments of NF-1. The macrocephaly reflects accompanying megalencephaly. Macrocephaly >98th centile occurs in 16-45% of patients. Individuals with NF-1 are on the average about 3 inches shorter than predicted by their family background. However more significant causes such as aqueductal stenosis due to diffuse or membranous gliosis of the aqueduct, leading to hydrocephalus and growth failure which occasionally may result from a hypothalamic involvement by an optic glioma should be ruled out.

Variable but significant complications

Learning disability

Frank mental retardation is uncommon in NF-1 and such patients are more likely to have the disease because of a large deletion that removes the entire NF-1 gene and considerable flanking DNA (Kayes et al., 1994). Other nearby genes reduced to hemizygosity by the deletion in these patients may contribute to the retardation. Learning disabilities are found in 30-50% of the children with NF-1 and this is an important complication. The learning disabilities in NF-1 have a specific pattern, often involving difficulties in with reading and fine and gross motor co-ordination. Significance of MRI abnormalities in children with NF-1 was evaluated by North et al. (1994). Forty children were evaluated with MRI, medical, psychometric, speech therapy and occupational therapy assessments. The mean full scale IQ scores showed a left shift in comparison with the normal population and the distribution of the IQ scores was bimodal, suggesting that there were two populations of patients with NF-1, those with and without a variable degree of cognitive impairment. Areas of increased T2 signal intensity on MRI have been reported in children with NF-1 called "UBOs" or unidentified bright objects. It was seen that children with UBO+ had significantly lower values for IQ and language scores and significantly impaired visuomotor integration and co-ordination. Children without increased T2 signal on MRI (UBO-) did not significantly differ from the general population. Areas of increased T2 signal on MRI represent dysplastic glial proliferation and aberrant myelination in the developing brain and may be associated with deficits in higher cognitive function. The presence of these signals thus can divide the population into two distinct groups anatomically and developmentally, (UBO+ and UBO-). These two groups should be considered separately in the

assessment and management of children with learning disabilities in NF-1 (North et al., 1994).

Plexiform neurofibromas

A group of nerves affected by numerous oval and irregular swellings is referred to as a plexiform neurofibroma. It may diffusely involve nerve, muscle, connective tissue, vascular elements and overlying skin and occurs in 10% of individuals with NF-1. Plexiform neurofibromas are usually congenital and combine cutaneous and subcutaneous elements to form tumours that may become huge and represent one of the most serious complications of NF-1. They may be continuous with intracranial or intraspinal tumours. Histologically these tumours are a mixture of Schwann cells, perineural cells and fibroblasts loosely arranged in a myxoid stroma with variable amounts of collagen. 3-6% of neurofibromas undergo malignant transformation and the plexiform neurofibroma has a greater tendency to malignant transformation. Severe plexiform lesions are almost invariably apparent by the age of 4-5, so it is possible to reassure older individuals without plexiform lesions, that they are not at significant risk for development of these lesions.

Malignancy

A particularly aggressive and often fatal malignancy is the neurofibrosarcoma, which commonly arises in a plexiform neurofibroma.
Optic gliomas are present in 4% of children and 2% of adults (Table 3). Histologically, these differ from other optic gliomas by the presence of an arachnoidal gliomatosis surrounding the optic nerve. They may be limited to the optic nerve, or involve the chiasma and the retrochiasmatic portion of the

visual pathway. Certain forms of malignant myeloid disorders such as juvenile chronic myelogenous leukaemia (JCML), monosomy 7 (Mo 7), chronic myelogenous leukaemia (CML) and acute myeloid leukaemia (AML) are seen in children with NF-1 and show a male preponderance. The other tumours seen in NF-1 are astrocytomas, rhabdomyosarcomas, gliomas, ependymomas, meningiomas, acoustic neuromas, phaeochromocytomas and thyroid carcinomas.

Seizures

A seizure disorder develops in approximately 5% of patients with NF-1 and the onset of this can occur at any time during life.

Scoliosis

Vertebral defects including scalloping from dural ectasia are extremely common in NF-1 and approximately 10% of affected individuals have scoliosis during late childhood and adolescence.

Pseudarthrosis

An uncommon complication of NF-1, whose pathological basis is unknown, is the involvement of long bones, noted first as bowing particularly of the tibia, in young children. This progresses to thinning of the cortex, pathological fracture and severe difficulties, with non-union of the fragments and may go on to form a false joint, rendering the limb severely compromised.

Hypertension

Hypertension is extremely common in adults with NF-1, affecting one third of the patients. This is usually essential hypertension with no underlying cause, but the new development of hypertension may be due to renal artery stenosis which is particularly common in children, or phaeochromocytoma in adults.

The frequencies of the complications of NF-1 in 1000 patients was summarised by Birch and Freidmann (1993) from the information obtained from the NNFF 'CLINICAL' database, which has contributions by 47 centres from 12 countries (Table 3).

Table 3: Frequencies of complications of NF-1

	Complication	Rate in children	Rate in adults
1)	Plexiform neurofibromas	18%	29%
2)	Proptosis	6%	3%
3)	Strabismus	10%	10%
4)	Pseudoarthrosis	2%	2%
5)	Scoliosis	17%	31%
6)	Malignant tumours (excluding brain tumours)	0.3%	2%
7)	Symptomatic optic gliomas	4%	2%
8)	Brain tumours excluding optic gliomas	1.5%	1.5%
9)	Learning difficulties	38%	-
10)	NF-1 related behavioural, emotional, psychological problems	33%	33%
11)	NF-1 related cosmetic problems	19%	43%

Variant forms of NF-1

NF-1 is also known to exist in association with other disorders (e.g. NF-1 in association with CMT1A), but most of these reports appear to represent the coincidental occurrence of two unrelated conditions. However variant forms of NF-1 have been described, which are:

Segmental NF-1

This is a spatially restricted form of NF, with symptoms localised to one region of the body (left or right upper quadrant and limb) not crossing the midline. Segmental NF patients do not have an affected parent but occasionally have had children with classical NF. The hypothesis for its causation is that these individuals may be mosaic as a result of a mutation in the NF-1 gene in early embryogenesis and if the mosaicism involves the germline, the disease can be transmitted.

Watson Syndrome

This variant of NF-1 involves multiple cafe au lait spots, dull intelligence, short stature, pulmonary valvular stenosis and only a small number of neurofibromas and Lisch nodules. Molecular analyses have shown at least two families that appear to fall in the Watson syndrome category and have mutations in the NF-1 gene (Tassabehji et al, 1993).

Neurofibromatosis-Noonan syndrome (NFNS):

The clinical features seen in Noonan's syndrome include triangular facies, downward slanting palpebral fissures, micrognathia, short stature and learning difficulty. Many of the features of Noonan syndrome seen in patients with NF-1, may be caused by dysgenesis or other developmental alterations

of the central nervous system resulting in muscular hypotonia. In the presence of hypotonia, development of craniofacial structures will be altered leading to hypoplasia of the midface and micrognathia. When coupled with craniofacial changes known to be common in NF-1, (prominent forehead, ptosis, ocular hypertelorism and broad nasal tip) the resulting facial phenotype may suggest Noonan's syndrome (Stern et al., 1992). The phenotype of NF-1 can include features that overlap, but these disorders are probably genetically distinct. The absence of linkage of Noonan's syndrome to the neurofibromatosis type-1 locus was shown by Sharland et al. (1992).

Spinal neurofibromatosis

Rare families have been identified with a predominance of spinal tumours and relatively few peripheral neurofibromas, some of them being linked, while the others showing no linkage to the NF-1 locus.

Other variant forms of NF-1

Familial intestinal neurofibromatosis (NF-3), neurofibromatosis-Phaeochromocytoma-Duodenal carcinoid syndrome (NPDC syndrome; Duodenal carcinoid syndrome), neurofibromatosis type III of Riccardi (NF-III; neurofibromatosis, mixed central and peripheral type; Palmar cutaneous neurofibromatosis included) and neurofibromatosis type IV of Riccardi, (NF-IV; neurofibromatosis variant forms of; neurofibromatosis-atypical) are some of the other described variant forms of NF-1.

1.5 GENETIC ASPECTS OF NF-1

Inheritance pattern

NF-1 is inherited as an autosomal dominant disease.

In NF-1, maternal transmission of the disease gene is associated with a more severe clinical phenotype (Miller and Hall, 1978) whilst the vast majority of new mutations are paternal in origin but do not exhibit a paternal age effect (Jadayel et al, 1990, Stephens et al, 1992).

In other genetic disorders that show a bias towards paternal origin of new mutations, there is a marked increase in the incidence of mutations with paternal age, consistent with the mutations arising from replication errors in mitosis of spermatogonial stem cells. In NF-1 however, such effects are slight or absent and therefore, most NF-1 mutations probably arise either at mitosis in a cell that is not a self renewing stem cell, or independently of mitosis, for example in a mature sperm. Also in NF-1, the paternal chromosomes are more susceptible to this mutation than maternal chromosomes and the reason for this is unclear, but could be related to the differences in maternal and paternal germ lines. The sex of the affected parent may have an impact on the severity of the disease a phenomenon which is referred to as parental imprinting. (Hall, 1990). Genomic imprinting is the term that has been used to refer to the differential expression of genetic material, at either a chromosomal or allelic level, depending on whether the genetic material has come from the male or female parent. Genomic imprinting may involve modifications of the nuclear DNA of somatic cells in order to produce these phenotypic differences. The term imprinting is also meant to imply that something happens during a critical or sensitive period in development. In case of genomic imprinting, the stage during which germ-line cells are formed may represent one critical period in which genetic information is marked, temporarily changing the genetic information to allow differential

expression. Because this marking is thought to occur during germline formation, the term germline imprinting is sometimes used. Genomic imprinting appears to be a form of regulation, allowing flexibility within the control and expression of the mammalian genome and may explain why mutations in some parts of the mammalian genome function differently depending on whether they come from the father or the mother (Hall, 1990).

Penetrance

The penetrance of NF-1 is 98% in individuals who have reached adulthood and been subjected to careful examination by an experienced physician. Rare cases of normal parents giving rise to affected children have been described and could be due to germline mutations in one of the parents, or rarely two independent mutations.

Variable expressivity

NF-1 is notable for its variable expression and an explanation for the variation in phenotypic expression was put forward by Riccardi, (1994) who suggested that within a family with NF-1 (for whom a single allele at the 17q11.2 locus was being dealt with), the key modifying factor is stochastic and the least critical is the epigenetic factor.

The phenotypic expression of NF-1 is to a large extent determined by the genotype at other modifying loci and the modifying genes are trait specific (Easton et al., 1993). The study was based on an analysis of variation in expression of NF-1, to determine whether the variation has an inherited component, using three quantitative traits and five binary traits. 175 individuals in 48 NF families were examined including 6 monozygotic (MZ) twin pairs. The quantitative traits showed a high correlation between MZ

twins and suggested a strong genetic component in the variation of expression. All five binary traits except plexiform neurofibromas showed significant familial clustering. There was no evidence of association between the different traits in affected individuals. Easton et al. (1993) concluded that the phenotypic expression of NF-1 is to a large extent determined by the genotype at other modifying loci and identification of putative genes when possible, would be candidates for evaluation in NF-1 families.

1.6 MOLECULAR BIOLOGY OF THE NF-1 GENE: CHARACTERISATION OF THE NF-1 LOCUS

Molecular genetic investigations into NF-1 began in the 1980s with the goal to localise and characterise the NF-1 gene.

Linkage and physical mapping

The first step was chromosomal localisation by family linkage studies. The NF-1 locus was mapped to chromosome 17 by linkage with DNA markers in families (Barker et al., 1987). Seizinger et al. (1987) presented evidence that the NF-1 gene is linked to the locus for nerve growth factor receptor (NFGR), a gene on the long arm of chromosome 17 (17q12-22). However the authors found cross-overs which indicated that NFGR was not the NF-1 gene. Goldgar et al. (1989) summarised the results reported by the International Consortium for NF-1 linkage. 142 families with more than 700 affected persons were studied using 31 markers in the pericentromic region of chromosome 17. The genetic analysis of these families indicated that the NF-1 gene lies on proximal 17q and markers on both sides of the gene were identified that were within 5 centimorgans of NF-1 (Goldgar et al., 1989).

In support of this localisation of the NF-1 gene, two unrelated patients with NF-1 having apparently balanced translocations involving chromosome 17 were identified. In both instances, the chromosome 17 breakpoint was in band q11.2, precisely where NF-1 mapped by genetic linkage analysis. Schmidt et al. (1987) described a patient with the balanced chromosome rearrangement [46XX,t(1;17)(p34.3;q11.2)]. Ledbetter et al. (1989) described a patient with a balanced translocation between chromosomes 17 and 22, [46XX,t(17;22)(q11.2;q11.2)]. This breakpoint was cytogenetically identical to the case with the 1;17 translocation reported by Schmidt et al. (1987). These results were suggestive that the translocation events disrupt the NF-1 gene and precisely mapped the NF-1 gene to 17q11.2. This provided a physical reference point for strategies to clone the breakpoint and therefore the NF-1 gene. The first step involved the construction of somatic cell hybrids from the cells of the two patients to add to the panel of hybrids that were in use for physical mapping. The hybrid derived from t(1;17) DCR-1 contained the translocation product carrying 17q11.2-qter as its only chromosome 17 material (Menon et al., 1989).The t(17;22) hybrid NF13 also contained the translocation product bearing 17q11.2-qter (Ledbetter et al., 1989). These two hybrids plus two other hybrids constructed by Van Tuinen et al. (1987) and Leach et al. (1989), provided starting material for construction of libraries in an effort to clone more of the NF-1 region. Progress in the physical mapping included localisation of most of the markers via the somatic cell mapping panel and physical connections were established between several probes with pulsed field gel electrophoresis (PFGE) (Fountain et al., 1989a). New probes were being developed by a number of approaches: Fountain et al. (1989a) mapped a series of chromosome 17 specific *Not* I linking clones to 17q and studied them by pulsed field gel electrophoresis. One clone 17LI

clearly identified the breakpoint in the patient with the t(1;17) translocation and it was found that 17LI lay closest to the NF-1 breakpoints on the centromeric side. Subsequently Fountain et al. (1989b) also detected the t(17;22) translocation from the centromeric side using the linking clone 17L1. O'Connell et al. (1989) generated probes from cosmid libraries made from microcell hybrids. Simultaneously O'Connell et al. (1989) detected the breakpoints from the telomeric side using the cosmid c11-1F10. The combined results narrowed the translocation breakpoints to a fairly well mapped physical region of 600kb. A 2.3Mb pulsed field map was constructed by Fountain et al. (1989b) and this indicated that the NF-1 breakpoint was 10-240kb away from 17LI. It was also seen that 17LI represented a CpG island. These CG rich hypomethylated regions are associated with the 5' regulatory sequences of active genes (Bird et al., 1986). Thus 17LI represented a potential candidate gene.

Cloning candidate genes

In attempting to clone the NF-1 gene, the goal was to clone cDNAs in the 600kb interval particularly around the translocation breakpoints and test each for its possibility of being the NF-1 gene. Random genomic fragments were also used to screen cDNA libraries including whole cosmids or YACs.

The genomic and cDNA cloning efforts in the area between 17LI and c11-IF10 were aided by the finding of a probe that lay between the two translocation breakpoints. The mouse gene evi2, involved in virally induced murine leukaemia was found to lie on mouse chromosome 11 (Buchberg et al., 1988) and showed synteny with the NF-1 region on chromosome 17 (Buchberg et al., 1989). Mapping of this gene and its human homologue EVI2A showed that EVI2A lay between the two translocation breakpoints. This was the first candidate gene for NF-1. EVI2A allowed the translocation

breakpoint region to be cloned in overlapping cosmids and mapped placing t(1;17) approximately 60kb closer to the centromere than t(17;22) (O'Connell et al., 1990a). EVI2A encoded a 1.7kb cDNA split between one small 5' exon and one larger 3' exon containing an open reading frame. The entire gene was contained completely between the two translocation breakpoints. However, inability to find mutations in NF-1 patients and lack of expression in neural crest tissues indicated that EVI2A was unlikely to be the NF-1 gene (Cawthon et al., 1991).

Efforts continued in trying to find additional candidate genes using cloned DNA from overlapping EVI2A cosmids (O'Connell et al., 1990a) and by chromosome jumping and YAC cloning in the EVI2A region (Wallace et al., 1990b). A contiguous 130kb genomic map of overlapping clones had been constructed by O'Connell et al. (1990a,b) that encompassed both translocation breakpoints. By screening cDNA libraries with these cosmids, two additional genes were identified, RC1 and HB36, which were similarly found to lie between the two NF-1 translocation breakpoints. A fourth locus, HB15 mapped distal to both breakpoints (Cawthon et al., 1990b). HB15 is a pseudogene of the adenylate kinase 3 multigene family. The predicted peptide from DNA sequence analysis indicated that RC1 was a transmembrane protein without specified function like EV12A (Cawthon et al., 1990b, Buchberg et al., 1990). The jump clone EH1 (Wallace et al., 1990b) detected the second candidate gene, EV12B from the beginning of the jump sequences (Cawthon et al., 1991). EV12B also lay between the translocation breakpoints several kilobases centromeric to EV12A (Cawthon et al., 1991). The cDNA sequence and genomic structure was described by the authors and was similar to EV12A in size, structure, direction of transcription and

tissue distribution. This gene too was eliminated as it failed to show mutations in NF-1 patients.

HB36 matched the sequence of the gene for oligodendrocyte myelin glycoprotein (OMGP) and OMGP was the third candidate gene. The sequence of OMGP was initially cloned by Mikol et al. (1990) and then by Viskochil et al. (1991). This gene also lay between the breakpoints, very close to the t(1;17) translocation. OMGP is a central nervous system cell surface peptide with properties of a cell adhesion molecule; potentially active in mediating proper cell motility and differentiation during brain development (Mikol 1990). However this gene, although expressed in oligodendrocytes was eliminated as a gene as it was not interrupted by either breakpoint and again did not show mutations in NF-1 patients.

Cloning the NF-1 gene

The NF-1 gene was discovered as the fourth candidate gene. The NF-1 gene was initially called the TBR gene (translocation breakpoint region) by the Utah group (Cawthon et al., 1990a, Viskochil et al., 1990) and NF1LT by the group in Michigan (Wallace et al., 1990a).

Wallace et al. (1990b) used chromosome jumping and yeast artificial chromosome technology to identify a 13kb ubiquitiously expressed transcript denoted as NF1LT. Two different strategies were used to derive cDNA clones that defined NF1LT. Initial experiments with the end of jump clone EH1 obtained by chromosome jumping showed that a single copy 1.4kb *Eco*R I-*Hind* III sub-fragment which lay just telomeric to the t(17;22) breakpoint was conserved across species. This probe was used to screen a human peripheral nerve cDNA library which resulted in the isolation of a clone P5, which had an insert of 1.7kb. Transcripts were also sought with the YAC clone A113D7, part of an overlapping contig of clones from this region.

This YAC contained the entire breakpoint region and on direct screening of a cDNA B lymphoblast library with this probe, clone B3A was isolated and P5 and B3A were seen to overlap. Using P5 as a probe, the authors showed that NF1LT was functionally disrupted by the translocation breakpoints. The above two approaches yielded 2 overlapping cDNAs that together represented 300bp of 3' untranslated region and 1.7kb of coding region.

The Utah group cloned a part of the same gene by initially screening a mouse cDNA library with a highly conserved human genomic fragment. They cloned the corresponding human cDNA and did several cDNA walks to obtain 4kb of cDNA sequence, including approximately the same portion of the 3' untranslated sequence. Viskochil et al (1990) used a 3.8kb (probe EE3.8) *Eco*RI fragment to screen several cDNA libraries. A 2.1kb TBR cDNA clone mDVI was obtained from a murine macrophage cDNA library. The *Eco*RI insert released from mDVI was then used to screen additional human cDNA libraries and many human foetal brain and peripheral blood TBR cDNA clones were isolated and partially characterised. Viskochil et al. (1990) showed that the TBR gene was interrupted by four NF-1 mutations, which included the previously characterised t(17;22) chromosome translocation. Three other deletions of 11kb, 190kb and 40kb were also shown to contain sequences from the TBR cDNA clones and the TBR conserved region.

Cawthon et al. (1990a) sequenced overlapping cDNA clones from the translocation breakpoint region (the TBR gene) which lay at the NF-1 locus and was interrupted by deletions and the t(17;22) translocation. The 4kb sequence of the transcript was compared with sequences of genomic DNA and a small number of exons were identified. PCR amplification of a subset of exons was followed by single stranded conformational polymorphism (SSCP) analysis and point mutations within the gene were identified.

Thus, further crucial support came from mutation analyses in NF-1 patients showing disruptions at the DNA level (Viskochil et al., 1990), point and nonsense mutations (Cawthon et al., 1990a) and a de novo insertion of an *Alu* repeat element which results in a splice error (Wallace et al., 1991). Both groups also determined that the cDNA originated from a large transcript of at least 11kb in length and that the gene is transcribed towards the telomere. The genomic organisation of nine exons (initially called exons 1-9) now known as exons 28-36 of the NF-1 gene, was described by Cawthon et al. (1990a).

Embedded genes within the NF-1 gene

The three previous candidate genes EV12A, EV12B and OMGP were shown to be embedded within an intron of the NF-1 gene in the opposite transcriptional orientation. The cDNA sequence and genomic structure of all the genes has been described; EV12A (Cawthon et al., 1990b), EV12B (Cawthon et al., 1991) and OMGP (Viskochil et al., 1991). An example of embedded genes, is a gene within intron 22 of the human factor VIII gene, also on the antisense strand (Levinson et al., 1990). There are very few previous examples of functional genes embedded within introns in higher eukaryotes and these examples raise the possibility that antisense RNA may regulate the expression of complex loci. Wallace et al. (1990a) proposed that the gene products of the embedded genes in the NF-1 gene may be responsible for some of the NF-1 phenotypic features.

1.7 THE NF-1 GENE AND ITS GENE PRODUCT (NEUROFIBROMIN)

The complete coding sequence of the NF-1 gene was cloned by an extensive cDNA walk using five different cDNA libraries (Marchuk et al., 1991). Walks proceeded sequentially by isolation of positive phage clones using the most 5' cDNA insert. The positive clones were characterised by restriction mapping using *Eco*RI and Southern blot analysis, using previously isolated inserts. The phage clones were sub-cloned and the ends were sequenced to anchor the position of the clones to the transcript map. As the walk neared completion, a very GC rich region of the transcript that contained an abnormally high concentration of the dinucleotide CpG, as well as rare cutting restriction endonuclease sites such as *Eag* I, *Nar* I and *Sac* I was encountered at the 5' end. These sites had been previously placed on the pulsed field map using the linking clone 17LI (Fountain et al., 1989a). The most 5' cDNA clone, KE-2 contained an in frame stop codon. Downstream from this codon, the first ATG fitted the rules for a proper translational start (Kozak, 1986). Marchuk et al. (1991) proposed that this ATG represented the authentic start codon, giving the protein a total of 2818 amino acids and a predicted molecular weight of 327kD. However, Marchuk et al. (1991) were unable to characterise the 3' end of the NF-1 transcript as a polyA tail had not been found in any of the clones. The authors were able to clone and sequence 9kb of the message and the remainder appeared to be the 3' untranslated region, as previous sequence analysis had shown the proper position of a stop codon (Wallace et al., 1990a).

Analysis of the sequences revealed an open reading frame of 2818 amino acids and it was shown that the gene extends for over 300kb on chromosome 17, with its promoter in a CpG island. The NF-1 gene is known to consist of at least 49 exons and the boundaries of some of the exons have

been made available via information from the NNFF (National Neurofibromatosis Foundation) Mutation Analysis Consortium.

Searches of the protein database with the known amino acids predicted from the sequence of neurofibromin revealed striking amino acid sequence similarities to IRA proteins of yeast (inhibitory regulators of the ras-cAMP pathway) and also to mammalian GAPs (GTPase activating protein). The similarities were initially identified between a 360 amino acid region of the NF-1 protein and the catalytic domains of mammalian GAP and the essential domain of yeast IRA proteins. These domains modulate the activity of the p21-ras protein in their respective hosts by accelerating the rate of hydrolysis of active rasGTP to inactive rasGDP (figure 1).

The region of homology was termed as the NF-1 GAP related domain or NF-1 GRD (figure 2). Within the GRD, four regions, termed box I-IV, contain residues that are conserved among different ras-GAP proteins which include NF-1, p120-GAP and yeast IRA1 and IRA2 proteins. Among these, two regions boxes III and IV, located towards the C-terminal side of the GRD, contain a high percentage of conserved residues. There are fourteen conserved amino acids within the GRD and ten of the fourteen are found in these two boxes. A sequence FLR...PA located in box III is highly conserved (Appendix 5).

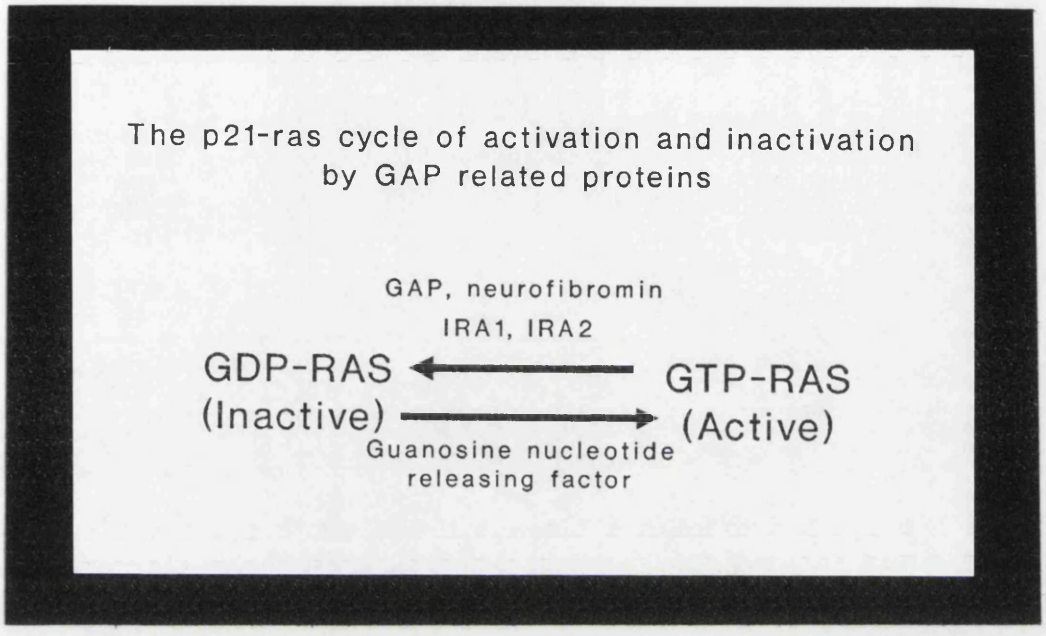


Figure 1

The p21-ras cycle of activation and inactivation by the GAP-related proteins

p21-ras is inactive in the GDP bound state and is converted into an active GTP bound state by guanosine nucleotide replacing proteins that substitute GDP for GTP. Interaction of GAP-like proteins with p21-ras accelerates the conversion of p21-rasGTP to p-21rasGDP by increasing the intrinsic GTPase activity of p21-ras and converting p21-ras to the inactive GDP bound form. In resting cells, the majority of the p21-ras is inactive and in the GDP bound form.

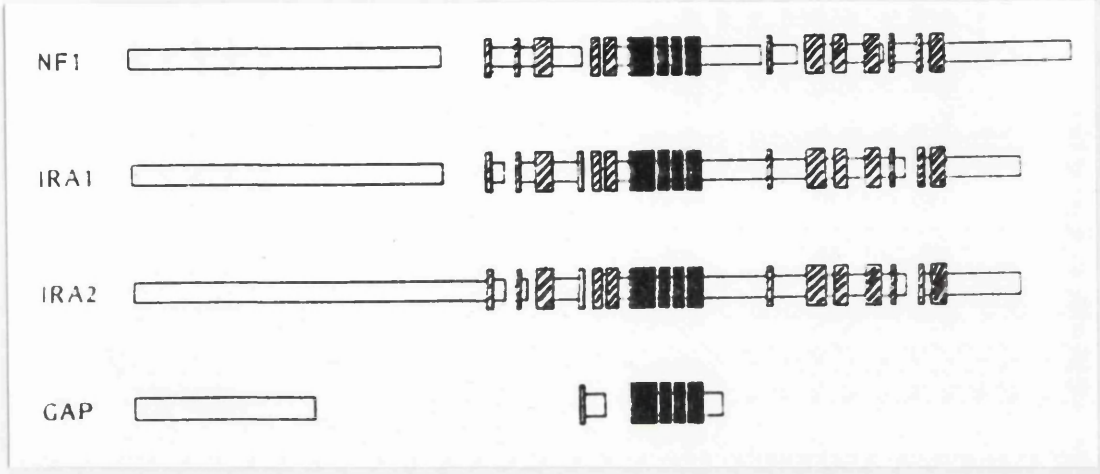


Figure 2

Regions of homology between the NF-1 protein and the IRA1, IRA2 and GAP proteins

Unfilled boxes represent regions with little or no homology and open spaces are used to align the most homologous region. Hatched boxes represent regions homologous among only three of the proteins and the black boxes show the regions conserved among all four proteins (which represents most of the catalytic domain or NF-1 GRD, Ballester et al., 1990).

(Adapted from Wallace and Collins, 1991).

Guanine nucleotide binding to ras proteins mediates signal transduction that regulates cell growth. Binding of ras to GTP activates signalling, while hydrolysis to GDP terminates signalling (Barbacid, 1987). A GTPase activating protein GAP was the first protein found to catalyse the hydrolysis to GDP and thereby mediate the signal termination event (Trahey and McCormick, 1987). In addition GAP may also function in signal propagation as a downstream effector of ras (McCormick, 1989; Hall, 1990). The GAP related domain of NF-1 (NF1-GRD), similarly stimulates the GTPase of ras (Ballester et al., 1990, Martin et al., 1990, Xu et al., 1990a) and possesses properties consistent with the functioning of neurofibromin as a downstream effector of ras (McCormick and Bollag, 1991). The GAP related domain of the NF-1 gene spans exons 20-27, exon 24 being the most highly conserved, also termed the FLR exon (Li et al., 1992).

The complete amino acid sequence of the NF-1 transcript showed three blocks of homology most conserved among the GAP family of proteins as seen above. There were no SH2 or SH3 (S RC-homology domains-elements that control interactions of cytoplasmic signalling proteins) in neurofibromin. SH2 and SH3 domains are homologous to the non-catalytic regions of the oncogene src and are thought to direct interactions with phosphotyrosine proteins involved in signal transduction. Their absence in neurofibromin implies that neurofibromin and GAP are not interchangeable in the cell and that neurofibromin is probably not directly modulated through tyrosine phosphorylation by activated growth factor receptors.

Six potential cAMP dependant protein kinase phosporylation sites and a single potential tyrosine phosphorylation site were found to be present in the NF-1 gene. The potential sites for tyrosine and serine threonine phosphorylation may mean that an intermediate between an activated

receptor and neurofibromin may modulate its activity, since there is evidence that neurofibromin is phosphorylated on serine and threonine residues (Marchuk et al., 1991). The sequence of neurofibromin showed no significant homology to the bcr-related GAP family or rap I GAP (Marchuk et al., 1991).

To find conserved, potentially functional domains of neurofibromin which would allow identification of potentially significant sequence alterations, the complete mouse NF-1 sequence was identified by Bernards et al. (1993). The sequence was predicted to be approximately 12kb with a 2841 amino acid protein, that is more than 98% identical to human neurofibromin, with the 3' segment being highly conserved. 45 amino acid differences were identified between mouse and human neurofibromin, of which all but 9 occur in the N-terminal half of the protein with 16 changes clustered just upstream of the IRA related segment. Given the high degree of sequence identity virtually any sequence alteration in NF-1 patients or tumours was thought to be potentially significant (Bernards et al., 1993).

Alternatively spliced transcripts of the NF-1 gene

Examination of the intron-exon organisation of the NF-1 gene revealed the presence of two alternatively spliced exons (Marchuk et al., 1991 Andersen et al., 1993a, Nishi et al., 1991).

The first alternatively spliced exon is exon 23a. It is located within the GAP related domain and inserts 63 nucleotides into the NF-1 mRNA (Nishi et al., 1991). These are inserted just upstream from the most conserved exon (exon 24) in the GRD and it encodes a basic stretch of amino acids including 6 lysines out of 21 amino acids. This isoform is termed GRDII for GAP related domain II and is expressed predominantly in adult brain and its expression is induced in neuroblastoma cells on retinoic acid treatment (Nishi et al., 1991). This isoform produces a protein with GAP activity but its

ability to down regulate ras is reduced. Thus differential regulation of neurofibromin isoforms may control the degree of influence p21-ras transduction pathways have on the overall signal for cell proliferation (Viskochil et al., 1993). Andersen et al. (1993a) demonstrated that because high lysine content could significantly alter the peptide conformation in a region presumably critical for p21-ras interaction, it was thought that the alternatively spliced form lacked function. To test this hypothesis a baculovirus *sf*9 expressed peptide termed (NF-GRDII), identical to the NF-1 GRD, now termed (NF-GRDI), except for inclusion of the 21 amino acid insert was purified and its functional properties were compared with NF-GRDI. NF-GRDII demonstrated decreased stimulation of H-ras as compared with NF-GRDI, however, its affinity for H-ras was about two fold greater. Thus the affinity of NF-GRDII for ras-GTP is the highest yet demonstrated between wild type p21-rasGTP and a GAP. This feature of the NF-GRDII could be significant in cells where limiting concentrations of rasGTP play a major role in signal transduction of physiological responses (Viskochil et al., 1993).

The other alternatively spliced exon (48a), is located in the extreme carboxy terminus of the NF-1 gene and inserts 54 nucleotides into the NF-1 mRNA (Marchuk et al., 1991), 18 amino acids between residues 2771 and 2772 (Cawthon et al., 1991). This isoform is called 3' ALT or Neurofibromin III for 3' alternatively spliced exon and this isoform was originally detected in a foetal brain cDNA library (DeClue et al., 1991). Examination of its expression by reverse transcribed RT-PCR demonstrated a high level of expression in cardiac muscle, skeletal muscle and smooth muscle (Gutmann et al., 1993). Additionally this isoform was also shown to be expressed in muscle tissues from other vertebrate species and the expression of this isoform in muscle suggested that the NF-1 gene may play additional tissue specific roles in

muscle development and signal transduction. Northern blot analysis of human adult tissues detected faint NF-1 expression in muscle at levels 10-20 fold lower than in brain. However it is not known whether this low level of expression (also seen in rat tissues) of neurofibromin in muscle compared with brain reflects the 3' ALT isoform or the uninserted neurofibromin protein. There are reports of patients with NF-1 and heart disease and although there is no consistent pattern of congenital heart defect in these patients, perhaps the timing of the 'second hit' (somatic mutation) could determine the phenotypic expression. There are a small number of neural crest derived cells that migrate into the heart during development. Some of these cells contribute to the orderly separation of the aorta and the pulmonary vessels, while others may give rise to the heart conduction system. Mice derived from homozygous targeted disruption of the NF-1 gene die at embryonic day 13.5 of generalised oedema secondary to a heart defect (Gutmann et al., 1993). This cardiac development defect known as double outlet right ventricle has been described to result from neural crest ablation in the developing chick. Further investigation of the role of neurofibromin in muscle is necessary to determine the tissue specific role of this GTPase activating protein in signal transduction pathways and development.

Expression and tissue distribution of neurofibromin

Neurofibromin is ubiquitiously expressed in most human tissues. The methods used to ascertain expression are: Northern blot analysis, RNA reverse transcribed PCR (RT-PCR), Antibody binding and NF-1 GAP activity. Northern blot analysis demonstrated an 11-13kb transcript from a neuroblastoma, 2 melanoma cell lines, brain frontal lobe tissue and kidney tissue (Wallace et al., 1990a). Viskochil et al. (1990) demonstrated transcripts from a choriocarcinoma cell line. Using RT-PCR, the

neurofibromin transcript was present in all tissues examined, including lymphoblastoid cell lines, skin fibroblasts, spleen, muscle, brain, kidney, liver and lung (Nishi et al., 1991, Wallace et al., 1990a, Suzuki et al., 1991).

To further elucidate the normal function of the NF-1 gene product and to determine the pathophysiologic basis whereby alterations in the gene give rise to neurofibromatosis, specific antibodies recognising the NF-1 protein product were developed. Identification of neurofibromin in cells by means of an antibody has been reported by DeClue et al., 1991, Gutmann et al., 1991 and Daston et al., 1992. Antibodies raised against various domains of the protein identified a 220-280kD protein by several techniques. Gutmann et al. (1991) generated polyclonal antisera (antibodies) in rabbits against 3 fusion proteins and 2 synthetic peptides. Initial characterisation of the two anti peptide antibodies and one fusion protein antibody demonstrated a specific protein of approximately 250kD by both immunoprecipitation, immunoblotting in HeLa cells, NIH 3T3 cells and various murine tissues. The authors also stated neurofibromin was detected in all cell lines and tissues examined. In another study by DeClue et al. (1991), identification of the NF-1 protein was carried out in mammalian cells. DeClue et al. (1991) raised rabbit antisera to a bacterially (E.Coli fusion expression construct) synthesised 48kD peptide corresponding to the GAP related domain of NF-1 (NF-1 GRD). The sera detected specifically, a 280kD protein in ^{35}S labelled cell lysates of mouse NIH 3T3, human HeLa and rat Schwannoma cell lines. The detected protein corresponded to the NF-1 gene product. Daston et al. (1992) used antibodies generated against two E.Coli fusion peptides that contained approximately 310 amino acids of neurofibromin to detect a 220kD band from total rat spinal cord on SDS/PAGE Western blots. Direct Western analysis on human spinal cord failed to detect the protein, but immunoprecipitation of human spinal

cord tissue lysate with one antibody followed by Western immunoblot analysis using the second antibody identified a 220kD protein.

Although the open reading frame of the NF-1 cDNA predicts a total protein of 2818 amino acids and a molecular size of 327kD, neurofibromin migrates at about 250kD. The size discrepancy was thought to result from anomalous migration due to protein folding, or the difference could reflect post translational modifications such as processing of a pro-protein species and the modification would necessarily involve cleavage of amino terminal sequences. These studies thus demonstrated a 220-280kD protein in cell and tissue lysates recognised by several different antibodies.

The expression pattern of neurofibromin during development, was characterised by Daston et al in 1992 and 1993, using the rat as a model system. Highest levels of neurofibromin using antibodies have been detected in brain, spinal cord, sciatic nerve and adrenal glands. Lower levels have been detected in liver, spleen, pancreas and cardiac tissue and none in skeletal muscle, lung kidney, skin (using antibody studies). Daston et al. (1993) also showed that neurofibromin immunoreactivity was enriched in neurons of the cortical plate, in the peripheral ganglia and in the developing CNS and PNS fibre tracts. Within the PNS the dorsal root ganglia and small axonal fibres associated with non-myelinating Schwann cells stain intensely with antineurofibromin antibodies. Transient expression of neurofibromin during development in many tissues suggested the importance of this protein in morphogenesis and organ growth. Daston and Ratner (1993) proposed a separate role for neurofibromin in growing axons and in the mature nervous system.

To determine whether the cellular and neuroanatomical distribution of neurofibromin revealed a possible function of neurofibromin in the brain,

Nordlund et al. (1993) stained rat brain tissue sections with antineurofibromin antibodies. They found that neurofibromin was present in cell bodies and in axons and was highly enriched in dendrites. Immunoelectron microscope analysis demonstrated that neurofibromin was associated with smooth vesiculotubular elements and cisternal stacks, but not with the plasma membrane, nucleus, nuclear envelope, golgi apparatus or rough endoplasmic reticulum. The preferential localisation of neurofibromin to the smooth endoplasmic reticulum, together with evidence that neurofibromin modulates rasGTPase activity suggested that some of the CNS manifestations of NF-1 may result from the altered expression of neurofibromin in neurons, perhaps through disruption of calcium signalling, translocation of organelles or endocyte pathways (Nordlund et al., 1993).

Analysis of the sequence and embryonic expression of the chicken neurofibromin RNA was carried out by Schafer et al., 1993. They isolated a partial cDNA for chicken neurofibromin. Sequence analysis revealed that the predicted amino acid sequence was highly conserved between chick and human. The chicken cDNA hybridises to a 12.5kb transcript on Northern blots, a molecular size similar to that reported for human and murine mRNAs. Ribonuclease protection assays indicated that the NF-1 mRNA was expressed in a variety of tissues in the chick embryo, which was confirmed by in situ hybridisation analysis. This pattern of expression suggested a role for neurofibromin during normal development including that of the nervous system.

mRNA expression of the NF-1 gene was studied in blood vessels in quiescent and proliferative conditions, by Ahlgrenbeckendorf et al. (1993) as vascular hypertrophic lesions are known to occur in NF-1. The expression and alternatively splicing pattern of the catalytic domain of NF-1 consistently

changed in proliferating cells, supporting a role for this gene in the regulation of vascular smooth muscle and the vascular pathology in NF-1. Expression of neurofibromin is varied on comparison of expression by the RNA and antibody methods. By RT-PCR detection, neurofibromin was ubiquitiously expressed, yet Northern blot analysis revealed restricted expression patterns. Also, antineurofibromin antibodies recognised a 220-280kD protein from all tissue and cell line lysates tested. However immunohistocytochemistry revealed a restricted pattern of expression. This may be explained by the relative sensitivities in signal detection (Viskochil et al., 1993). RT-PCR may allow detection of a small number of cellular NF-1 mRNAs which represent the leaky nature of housekeeping gene transcription, yet such a low level may not provide any specific contribution to cellular phenotypes. On the other hand, neurofibromin may be expressed at the lower limits of detection (e.g. antibody) in most cells, but still contribute to the cellular phenotype. The p21-rasGTPase activating function of neurofibromin provides an approach for further characterising expression of this protein with respect to cellular physiology (Viskochil et al., 1993).

The GAP related domain of neurofibromin is inhibited by some lipids, whereas p120GAP is not (Bollag and McCormick, 1991). Therefore total p21-rasGAP activity from various cell lysates can be divided into p120GAP like activity (non-inhibitable by lipids) and NF-1 GAP like activity (lipid inhibitable). Bollag and McCormick (1991) found that all mammalian cells tested contained NF-1 and p120GAP like activity in varying proportions. The relative proportions of lipid inhibitable (NF-1 GAP) and noninhibitable (p120GAP) rasGAP activities can be compared from each cell lysate. NF-1 GAP accounts for 75% of total rasGAP activity from rat phaeochromocytoma cells, (PC12 cell line), whereas p120GAP accounts for 90% of total activity in

human placental tissue. Both the proteins p120GAP and neurofibromin are widely expressed in mammalian tissues and this raised the question of whether or not they have different regulatory functions.

Cellular localisation of neurofibromin

Neurofibromin was localised to the non-nuclear particulate fraction of NIH3T3 cells by an immunodetection study that used differential centrifugation and no detergent in the cell lysis buffer (DeClue et al., 1991).

In functional studies in various cell lysates, Bollag and McCormick, 1991 found NF-1 GAP activity distributed between the supernatant (approximately 60% of total GAP activity) and the particulate fraction (approximately 40% GAP activity). Neurofibromin may be less active within the particulate fraction of the cell and immunologically undetectable protein in the cytosol may be responsible for 60% of the detectable GAP activity. Alternatively, detergent may release loosely bound neurofibromin from the particulate fraction into the cytoplasmic fraction.

DeClue et al. (1991) used detergent (NP-40) in the cell lysis buffer to demonstrate that neurofibromin was immunoprecipitated from NIH 3T3, HeLa and rat schwannoma cell lines, both as a monomer and as a complex with an unidentified 400-500kD protein. In tissue culture cell lines, neurofibromin could be part of a complex associated with the particulate fraction of cells. In addition to the above findings, immunostaining demonstrated localisation of neurofibromin in the cytoplasm of cells from various tissues (Daston et al., 1992).

Neurofibromin associates with cytoplasmic microtubules

In separate experiments, using double indirect immunofluorescent labelling, with antineurofibromin and antitubulin antibodies, Gregory et al. (1993)

demonstrated that neurofibromin associates with cytoplasmic microtubules. Immunoblotting of microtubule enriched cytoplasmic fractions using antibodies generated against neurofibromin, showed that neurofibromin co-purified with microtubules. When portions of neurofibromin are expressed in *sf* 9 insect cells, they associate with polymerised microtubules. Furthermore the critical residues for this interaction reside within the GAP related domain of neurofibromin. This association suggests that neurofibromin is involved in microtubule mediated intracellular signal transduction (Gregory et al., 1993).

The discovery that neurofibromin is a GAP-like molecule that associates with microtubules suggests several hypotheses to explain its function in cell growth and differentiation (Figure 3). One model which fits the upstream view of p21-ras-GAP interactions envisions that neurofibromin is regulated by serine/threonine kinases (Gregory et al., 1993). Neurofibromin would be active as a GAP while associated with microtubules keeping p21-ras in the inactive form and inhibiting cell division. After phosphorylation, neurofibromin would dissociate from microtubules and its GAP activity would be reduced or altered. Alternatively, neurofibromin could be compartmentalised in the microtubule compartment (perhaps performing some other function) until it was required for the control of p21-ras. Phosphorylation of neurofibromin on critical serine residues would release it from the microtubules to interact with p21-ras. Likewise the interaction of neurofibromin with microtubules may reduce its GAP like activity and its dissociation from microtubules may allow neurofibromin to associate with and downregulate p21-ras.

Further studies have demonstrated the involvement of neurofibromin in a in a B-lymphocyte signal transduction pathway involving microtubules and p21-ras. In this system neurofibromin and p21-ras co-localise during

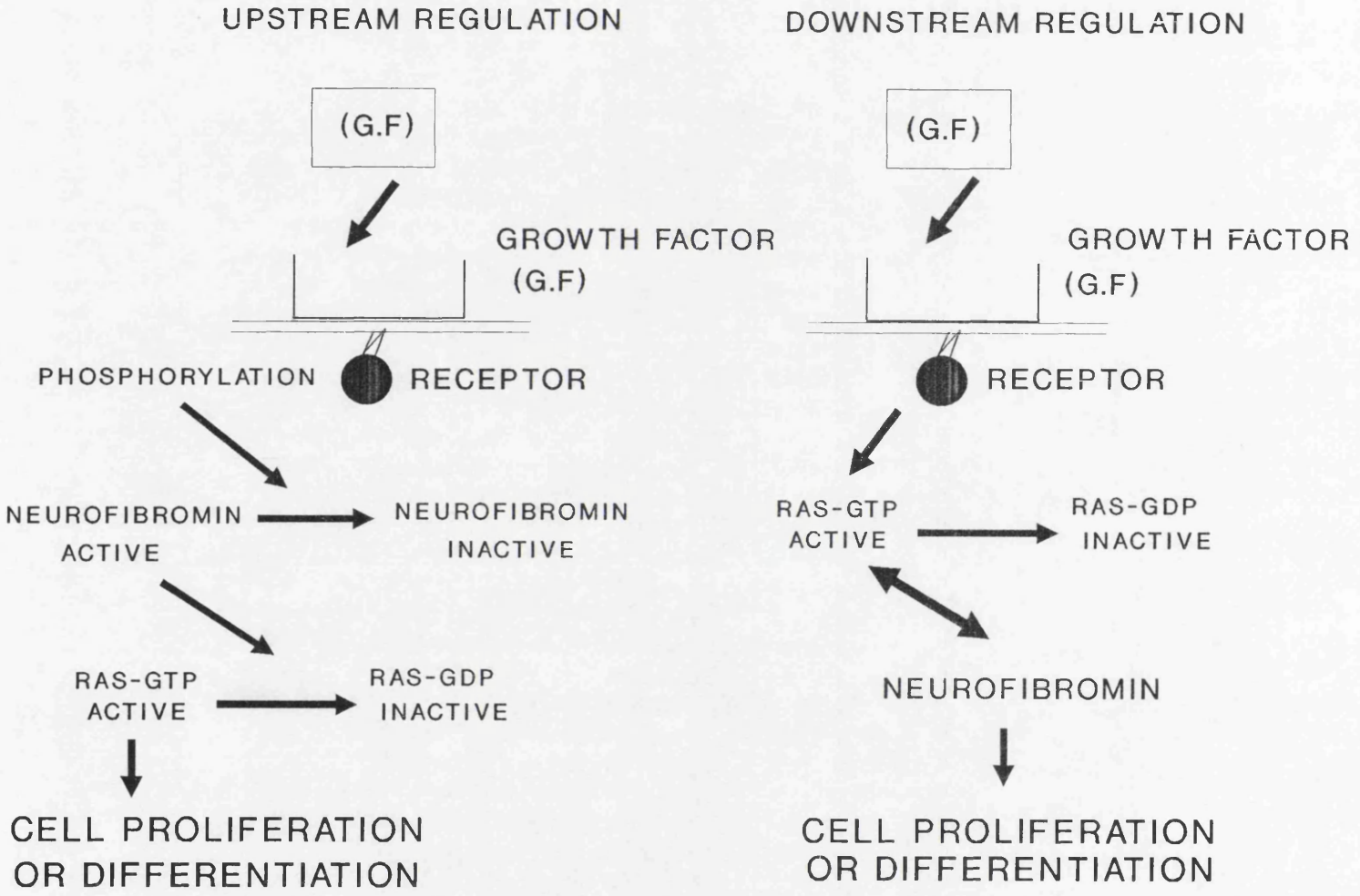
immunoglobulin receptor internalisation and neurofibromin becomes rapidly phosphorylated (Boyer et al., 1994). A second model, which falls into the category of a "downstream" hypothesis for p21-ras GAP interaction, is that neurofibromin is induced by the process of p21-rasGTP to p21-rasGDP conversion to transmit a signal through its influence on microtubule organisation.

Figure 3 (P.T.O)

Upstream versus downstream models of p21-ras neurofibromin interactions (proposed by Gutmann and Collins 1994).

In the upstream model, stimulation of appropriate cells expressing growth factor receptors leads to inactivation of neurofibromin perhaps through phosphorylation cascades. Inactivation of neurofibromin releases p21-ras from its downregulation, allowing p21-ras to predominate in the active, GTP bound form and signal other intracellular proteins to culminate in cell proliferation or differentiation. The downstream model envisions p21-ras to be a regulator of neurofibromin and transmits a signal via neurofibromin and p21-ras to culminate in cell proliferation or differentiation.

Upstream vs Downstream models of p21-ras-neurofibromin interactions



1.8 PUTATIVE FUNCTIONS OF NEUROFIBROMIN

Neurofibromin as a GTPase activating protein

Due to the homology between a predicted peptide moiety of neurofibromin and the catalytic domains of mammalian GAP and yeast IRA proteins it was suggested that at least one role of neurofibromin might be to stimulate rasGTPase activity. This was confirmed using both genetic and biochemical assays.

cDNA segments encoding the homologous catalytic domain were cloned into expression vectors for E.Coli, baculovirus / *sf* 9 insect cells and yeast (Ballester et al., 1990, Martin et al., 1990, Xu et al., 1990b). In genetic experiments, the peptide was expressed in ira-yeast strains and studied by *in vivo* assays based on heat shock resistance and glycogen accumulation. Transformation with the GAP related domain of neurofibromin NF-1GRD reverted the ira phenotype and cells survived the heat shock and accumulated glycogen (Ballester et al., 1990, Martin et al., 1990, Xu et al., 1990b).

In biochemical experiments, the *in vitro* rasGAP assays confirmed the hypothesis that the NF-1GRD can stimulate p21-ras hydrolysis of GTP. Ballester et al. (1990) used a cDNA construct encoding 412 amino acids that encompass the GRD to transform an ira-2 yeast strain. Cell lysates from that transformed strain stimulated hydrolysis of H-rasGTP but not of the oncogenic mutant H-rasGTPVal12. Xu et al. (1990b) purified a glutathione S-transferase/NF-1GRD protein expressed in E.Coli and demonstrated GTPase stimulating activity on both RAS2GTP and HRASGTP. The same investigators also showed that like the GAP and IRA2 proteins, the NF-1GRD fusion protein could not stimulate GTPase activity of H-rasVal12, Ras2val19, or RAS2Ala42 mutant protein.

Martin et al. (1990) purified an epitope tagged 474 amino acid peptide which encompassed the NF-1GRD from the baculovirus/*sf*9 expression system. This peptide stimulated the GTPase activity of N-ras, but not the oncogenic mutants of N-rasAsp12 and N-rasVal12. NF-1GRD also stimulated GTP hydrolysis of the N-rasAla38 effector mutant, but at a level less than 10% of the wild type activity.

In vitro assay results have been able to biochemically distinguish NF-1GRD from p120GAP. At high concentrations of N-rasGTP, GAP stimulates the GTPase to a much higher specific activity than does NF-1GRD. At low N-ras concentrations, the GAP and NF-1GRD activations are comparable (Martin et al., 1990). This demonstrated that NF-1GRD might have a higher affinity for N-rasGTP. In a competition binding assay, NF-1GRD demonstrated a 30 fold higher affinity for N-rasGTP than did GAP (Martin et al., 1990). Further studies demonstrated several differences between p120GAP and NF-1GRD, both in their affinities for various ras proteins and in their respective activations of p21-rasGTPase (Bollag and McCormick, 1991).

NF-1GRD and p120GAP also show differences with respect to their interaction with p21-ras in the presence of various lipids. The activity of lipid derived mitogenic agents depends on cellular p21 activity (Yu et al., 1988). Arachidonic acid at 100μg per ml inhibits stimulation of H-rasGTPase by GAP catalytic fragment and NF-1GRD, whereas high concentration of phosphatidic acid (β-arachidonoyl-γ-stearoyl) only partially inhibited NF-1GRD and did not inhibit GAP (Golubic et al., 1991). Other studies confirmed the inhibition by arachidonic acid and also demonstrated that GAP stimulation of H-rasGTP was increased by prostaglandins PGF2α, or PGA2 and decreased by PGI2, whereas NF-1GRD activity was not affected by the

prostaglandins (Han, 1991). Further comparisons of p120GAP and NF-1GAP with respect to lipid inhibition responses on N-rasGTP showed that phosphatidate, arachidonate and phosphatidylinositol 4,5 biphosphate inhibit NF-1GAP to a much greater extent than p120GAP (Bollag and McCormick, 1991). Furthermore, n-dodecyl-β-D-maltoside which did not inhibit p120GAP was identified as a stable inhibitor of NF-1GAP activity. This selective inhibition was used to differentiate the two GAP activities in cell lysates. Inhibitable GAP activity is due to NF-GAP and noninhibitable GAP activity is due to p120GAP. The potential physiological significance of the *in vitro* interactions of lipid moieties on GAPs raised the possibility that neurofibromin may play a key role in signal transduction pathways within the ras-dependant mitogenic response of many cell types (Viskochil et al., 1993).

Tumour suppressor function of the NF-1 gene

Neurofibromatosis type-1 represents a disease caused by disruption of a tumour suppressor gene, which encodes a protein involved in the proper regulation of cell differentiation and proliferation and its disruption predisposes to tumour formation. The NF-1 gene is established as a tumour suppressor gene (Legius et al., 1993 Andersen et al., 1993b, The et al., 1993, Johnson et al., 1993 and Legius, 1994 (NNFF newsletter), whose loss of function appears to be associated with typical NF-1 related tumours and with a variety of unrelated tumours in the general population. The genetic mechanisms through which aberrations in the NF-1 gene may contribute to tumourigenesis involve a two-hit model of tumourigenesis in which both copies of the tumour suppressor locus need to be deleted or inactivated in order to lead to tumour formation. The clinical variability of NF-1 symptoms affecting multiple organ systems and cell types suggested that the NF-1 gene

may play an important role in regulating growth and development not only in neural crest derived tissues.

The NF-1 gene has been shown to be a recessive tumour suppressor gene and the relationship between genomic imprinting and tumourigenesis is thus considered. Genomic imprinting has also been suggested as an important factor in the occurrence of certain cancers (Hall, 1990). Knudson (1991) presented a model for tumourigenesis based on two successive "hits" on the genome. This was speculated to result in inactivation of homologous alleles encoding a tumour suppressor locus, leading to cellular deregulation. In familial cancers, the first insult was inherited and the second event occurred somatically. In sporadic cases of cancer, it was suggested that both events occurred somatically. Sapienza (1991) was the first to incorporate genetic imprinting into this model. If an imprinted gene was involved, the first hit may actually be explained by non-expression of one of the alleles due to the imprinting process. An imprinted gene by definition is not expressed, leaving a hemizygous phenotype. The second hit may be mutational, or may result from a loss of all or part of the chromosome carrying the remaining functional suppressor allele. Loss of function is only one of the mechanisms proposed for tumourigenesis. Hochberg et al. (1994) suggested that the activation of imprinted protoncogenes and growth factors may also play a vital role in tumourigenesis. This is proposed to occur when the gene is removed (translocated) from its imprinted sequence area, it escapes repression and overexpression then leads to abnormal cellular proliferation and cancer.

The NF-1 gene however may be mutant in typical NF-1 tumours as well as in a variety of unrelated neoplasms. Li et al. (1992) have detected amino acid substitutions altering the codon Lys1423 of neurofibromin in an

NF-1 family as well as in a sporadic anaplastic astrocytoma, a sporadic colon carcinoma and myelodysplastic syndrome, the latter two of which are not known to be associated with NF-1. Lys1423 is one of the 14 amino acids in the NF-1 GRD which are absolutely conserved across all members of the GAP family of proteins, suggesting the mutations were significantly functional alterations.

The mechanism of action of neurofibromin may be distinct for the different types of tissues involved. Dramatically elevated levels of active GTP bound ras were found in malignant neurofibrosarcomas from NF-1 patients by DeClue et al. (1992) and Basu et al. (1992). To the contrary, The et al. (1993) found only moderately elevated rasGTP levels in neurofibromin deficient sporadic neuroblastoma cells.

Johnson et al. (1993) saw no correlation at all between the level of neurofibromin and the proportion of active GTP bound ras in neuroblastomas and sporadic melanoma cell lines. Previously it was seen that in Schwannoma cell lines from patients with neurofibromatosis, loss of neurofibromin was shown to be associated with impaired regulation of rasGTP on overexpression of cHras. However the above neuroblastoma and melanoma cell lines have greatly reduced or absent levels of neurofibromin with appropriate regulation of GTP ras, even when cHras was overexpressed in these lines. These results suggest that some neural crest tumours not associated with neurofibromatosis have acquired somatically inactivated NF-1 genes and suggest a tumour suppressor function of neurofibromin independent of its rasGTPase activation i.e. its GAP like activity. Neurofibromin may be involved in ras mediated differentiation pathways and loss of neurofibromin may inhibit this pathway (Seizinger, 1993). Indeed in some neural crest derived cell types p21-ras activation may actually result in

cell cycle arrest rather than in mitogenic stimulation depending on the physiological status of the cells.

The ras group of oncogenes, activating mutations and relationship with NF-1 somatic mutations

The ras gene family encodes membrane associated, guanine nucleotide binding proteins (p21) that are involved in the control of cellular proliferation and differentiation. Similar to other guanine binding proteins, the ras proteins cycle between the active guanosine triphosphate bound form (GTP) and an inactive guanosine diphosphate bound form (GDP). The weak intrinsic GTPase activity of ras proteins is greatly enhanced by the actions of GTPase activating proteins (GAPs). Both p120ras (protein product of the ras genes) and neurofibromin greatly stimulate p21 mediated GTP hydrolysis.

Point mutations in the ras genes, (activating or oncogenic mutants) have been detected in many human tumours and mice transgenic for mutant ras proteins develop tumours at a higher frequency than control mice. The ras genes acquire transforming potential when due to a point mutation, a single amino acid of the protein product at one of the critical codon positions 12, 13, or 61 is altered. The ras genes appear to be finely balanced at the edge of oncogenesis. Almost any mutation at codon position 12 or 61 can convert them to active oncogenes. Position 12 is occupied by glycine in exon 1. If it is replaced by any other of the 19 amino acids except proline it can transform cultured cells. Position 61 is occupied by glutamine in wild type genes in exon 2. Its change to another amino acid usually creates a gene with transforming potential. However conversion of the glutamine to proline or glutamic acid has no effect. Ras mutations have been detected in many

human malignancies and their incidence in various tumour types is shown in Table 4:Incidence of ras gene mutations in human malignancies

Tumour type	% of tumours harbouring a mutant ras gene	ras gene preferentially activated
Colon adenocarcinoma	50	K-ras
Colon adenoma	50	K-ras
Pancreatic adenocarcinoma	90	K-ras
Seminoma	40	K-ras, N-ras
Melanoma	20	N-ras
Neuroblastoma	<5	
Lung adenocarcinoma	30	K-ras
Thyroid follicular carcinoma	50	N,K, & H-ras
Myelodysplastic syndrome	30	N-ras

Biochemically, nearly all of these activating mutations decrease the intrinsic GTPase activity of the encoded ras and are insensitive to stimulation by GAPs (Ballester et al., 1990). Normally only a small proportion of cellular p21-ras is GTP bound. Mutations in ras can dramatically increase its biological activity by encoding mutant p21-ras, that accumulates much higher levels of the GTP bound form. These observations led to the hypothesis that at least some of the transforming activity of mutant ras is conferred by the ras protein being constitutively activated in the GTP bound state. Mutations in ras that render it insensitive to GAP regulation, result in tumour formation. Therefore mutations in GAP (C-terminal portion or the catalytic domain) and NF-1 (NF-1GAP related domain) that alter their ability to down regulate ras might result in a similar phenotype. Thus if a tumour phenotype can be related to the ras protein being in the active, GTP bound state, then perhaps a similar phenotype can be induced by an inactivation of the catalytic domain of GAP, or NF-1 which are the contributors to ras-GTP hydrolysis.

This was the hypothesis that was tested by undertaking somatic mutation analysis in the GAP related domain of the NF-1 gene in various tumour types. Before screening for mutations within the NF-1 gene, the same panel of mutations was analysed for mutations in the appropriate ras group of oncogenes. Somatic mutations that activate ras genes may contribute to tumourigenesis in up to 30% of human tumours (Bos, 1989). It was proposed that NF-1 mutations should be considered in any tumour type in which activated ras genes are frequently found and perhaps in these tumour types, NF1 mutations are the functional complement of ras mutations and will be found in those tumour samples lacking ras mutations. NF-1 mutations were also screened for in tumour types in which activated ras is shown to inhibit growth (e.g. medullary thyroid carcinoma) and it was proposed that in these tumour types, mutations would be found in those tumour samples composed of cells that continue to proliferate when activated ras is introduced (Li et al., 1992).

Treatment of NF-1

The cloning of the NF-1 gene has helped in understanding the pathobiology of neurofibromatosis and the eventual goal is to design specific nonsurgical treatments for affected patients. The finding of elevated p21-ras-GTP levels in tumours from NF-1 patients suggests that drug therapies directed at upregulating neurofibromin GAP activity or down regulating p21-ras activity might have some beneficial growth on neurofibromas. A number of groups have been studying the lipid sensitivity of neurofibromin GAP activity and find that specific lipids preferentially alter neurofibromin GAP activity as opposed to mammalian p-120 GAP activity. The discovery of a compound capable of up-regulating or replacing neurofibromin, or drugs that interfere with p21-ras activity, such as pharmaceutical events that block

farnesylation (a reaction necessary for p21-ras membrane localisation) might have a therapeutic potential in NF-1.

Animal models of NF-1

Three early models of NF-1 were reported. In the bicolour damselfish, spontaneous neurofibromas and hyperpigmented spots develop, but the disorder appeared to be transmissible and these tumours were more aggressive and malignant than human neurofibromas (Schmale et al., 1988). Hinrichs et al. (1987) reported a murine model resulting from overexpression of the HTLV-tat gene in mice. However other phenotypic features of NF-1 were absent and the neurofibromas lacked Schwann cells unlike human NF-1 neurofibromas. Also the relationship between HTLV-I and human NF-1 disease is unclear, as there is no increased incidence of HTLV-I expression or infection in NF-1 patients. Nakamura et al. (1989) described a third model of NF-1 by injecting N-nitros-N-Ethylurea into pregnant Syrian golden hamsters. The progeny developed neurofibromas histologically identical to those observed in NF-1 lesions, as well as pigmented lesions similar to cafe au lait spots. However these hamsters also had Wilm's tumours and other malignancies not seen in typical NF-1 patients. Point mutations in the n-protoncogene were identified in these tumour cells but no mutations were identified in the NF-1 gene.

Recent efforts have been directed at generating a mouse NF-1 knockout by homologous recombination. Brannan et al. (1994) used gene targeting in ES cells to generate mice carrying a null mutation at the mouse NF-1 locus. However heterozygous mutant mice did not exhibit obvious abnormalities. The homozygous mutant embryos died in utero, due to severe malformation of the heart and also displayed hyperplasia of neural crest derived sympathetic ganglia.

Jacks et al. (1994) created a mouse model by successfully disrupting the mouse homologue of exon 31. This was done by the insertion of a neomycin cassette to produce a truncated NF-1 mRNA. Mouse embryonic stem cells carrying this abnormal NF-1 gene, were introduced into blastocysts and viable offspring were selected that carried a copy of the mutated NF-1 gene. The heterozygous animals did not exhibit the classic symptoms of the disease, but were highly predisposed to the formation of different tumour types, notably phaeochromocytoma and myeloid leukaemia, both of which occur with increased frequency in human NF-1 patients. The wild type NF-1 allele was lost from approximately half of the tumours from heterozygous animals. In addition homozygosity for the NF-1 mutation led to abnormal cardiac development and mid-gestational embryonic lethality.

1.9 GERMLINE AND SOMATIC MUTATIONS IN NF-1

IMPORTANCE OF MUTATION DETECTION WITHIN THE NF-1 GENE

Identification of mutations in this large and complex gene is important for the following reasons:

1. By isolating a large number of independent mutations within the NF-1 gene a mutational spectrum for the NF-1 gene can be established. By determining the type and site of each mutation, delineation of mutational hot-spots, if any, within the gene can be carried out.
2. Identification of mutations within the gene will allow many variant forms of the protein to be compared. A detailed analysis of the mutations can be used to identify regions of neurofibromin responsible for specific functions.

3. Mutation analysis is necessary to assess the role of the three embedded genes EV12A, EV12B and OMGP within the NF-1 gene and to determine unique molecular mechanisms for disease. An example for this is seen in Haemophilia A, where the embedded gene (F8A) within intron 22 of the factor VIII gene was found to be responsible for the causation of mutations (inversions) in 50% of the severe cases. (Naylor et al., 1993, Lakich et al., 1993).
4. The existence of mutations allows comparison of properties of the wild-type gene with a defective gene. By analysing the type of changes that occur, an effort can be made to correlate the phenotype with the genotype and perhaps, associate certain domains of the gene with specific manifestations of the disorder.
5. Mutation detection within the NF-1 gene may help in creating a molecular classification for the clinically variant forms of NF-1 e.g. Watson syndrome, segmental NF, etc.
6. Detection of somatic mutations within the NF-1 gene in NF related and unrelated tumours will further establish the tumour suppressor function of the NF-1 gene.

DIFFICULTIES IN MUTATION ANALYSIS IN THE NF-1 GENE

Mutation analysis within the NF-1 gene is complicated due to the following:

1)The large size of the gene:

The NF-1 gene extends for over 300 kb on chromosome 17 and has at least 49 exons, whose intron exon boundaries have not yet been fully characterised. The NF-1 gene produces a transcript of 11-13 kb, of which 8.5 kb is the coding sequence, encoding 2818 amino acids. Screening for

mutations in a large gene, for which the coding sequence of only 9 exons (28-36) had been initially reported (Cawthon et al., 1990a) and whose complete characterisation is still awaited, created initial difficulties.

2)The high mutation rate:

When mutations are considered in terms of inactivation of the gene, most genes within a species show rates of mutation relative to their size. However it is seen that the spontaneous mutation rate in NF-1 is even higher (Huson et al., 1989), estimated at 1×10^{-4} mutations per gamete per generation.

Up to 30-50% of cases are due to new mutations. The prevalence and mutation rate for NF-1 from various studies are listed in Table 5.

Table 5: Prevalence and mutation rate in NF-1

Study (year)	Methods of ascertainement	Prevalence	Mutation rate
Crowe et al. (1956)	Surveys of general hospital admissions & state mental institutions	1/2500-3000	$1.4\text{-}2.6 \times 10^{-4}$
Sergeyev (1975)	Population sample of 16 year old Russian youths	1/7800	$4.4\text{-}4.9 \times 10^{-4}$
Samuelsson & Axelson (1981)	Population based	1/4600	4.3×10^{-5}
Huson et al. (1989)	Population based	1/2500-4590	$3.1\text{-}10.5 \times 10^{-5}$

Adapted from Huson et al. (1989).

3) Presence of the normal allele:

NF-1 is an autosomal dominant condition and during mutational analysis, the presence of the normal allele has always to be taken into account, on analysis of the directly sequenced PCR products. This may sometimes complicate analysis of complex sequences. This can be seen in the case of insertions and deletions within the NF-1 gene (e.g. sequence analysis of a single base insertion 6519insG -Figure 39).

4) Presence of NF-1 homologous sequences and pseudogenes

The presence of NF-1 homologous sequences has been detected (Marchuk et al 1993) and some of these have been characterised by Legius et al (1993) and Gasparini et al (1993). The presence of homologous sequences should be taken into account during amplification and analysis of mutations.

NF-1 mutations that have been detected in the constitutional DNA of NF-1 patients include translocations, medium to large sized deletions and insertions and nonsense and missense mutations. Most of the above mutations would be expected to inactivate neurofibromin by truncating the protein and/or drastically changing its overall shape. Mutant alleles with mutations lying 3' to the NF-1 GRD might produce stable neurofibromin with intact functioning GRD. For mutations within the GRD of NF-1 patients the mutant neurofibromin predicted even if stable, would not be expected to bind to rasGTP and stimulate the GTPase. It was seen that most NF-1 germline mutations lead to neurofibromatosis by inactivating the gene and suggested that NF-1 was a tumour suppressor gene. If so, the remaining allele would be expected to undergo inactivation during development of at least some of the tumours which arise in NF-1 patients.

Loss of allele is one mechanism of inactivation and one way of detecting loss of an allele is to examine benign and malignant tumours from

NF-1 patients for loss of heterozygosity (LOH) on chromosome 17 in the region of the NF-1 gene (Figure 4). Extensive analyses by various groups have failed to detect LOH along chromosome 17 in the benign tumours of NF-1 neurofibromas. This suggested that events other than inactivation of the second NF-1 allele, or inactivation of the second NF-1 allele by mechanisms more subtle than allele loss i.e. point mutations, were involved in the development of these tumours.

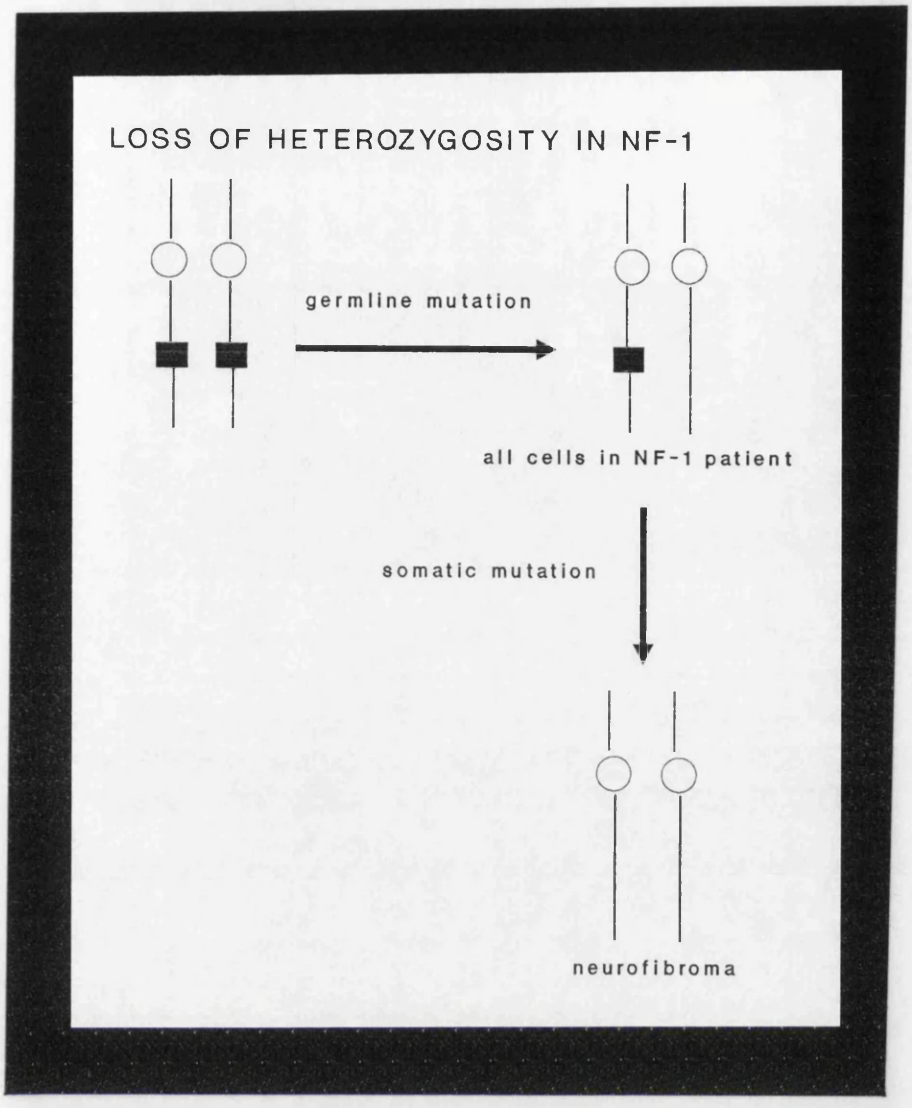


Figure 4: Loss of heterozygosity in NF-1

A given DNA marker polymorphism, denoted by the filled squares, is present in a normal NF-1 chromosome 17. A germline mutation found in all cells in an NF-1 patient would alter the NF-1 gene to result in loss of one of the DNA markers. Because the patient has one normal chromosome 17 DNA polymorphism and one mutated chromosome 17 DNA polymorphism, the patient is heterozygous with respect to that DNA marker. Mutation of the one remaining normal NF-1 gene in a tumour would result in loss of both copies of the gene and loss of heterozygosity with respect to that DNA marker polymorphism.

Analysis of some malignant tumours from NF-1 patients have shown LOH for DNA markers on 17p or along the entire length of chromosome 17 (Skuse et al., 1989, Glover et al., 1991, Menon et al., 1990). These findings were ambiguous because loss of a normal allele of the p53 tumour suppressor on 17p could contribute to tumour progression and drive chromosome loss. However other malignant tumours from NF-1 patients have shown LOH specific to 17q. Xu et al. (1992) reported allele losses with markers flanking the NF-1 region in each of the 7 phaeochromocytomas they examined and in 3 tumours the loss involved the wild-type chromosome. These results provided evidence that in cells of the adrenal medulla at least, the NF-1 gene may act as a tumour suppressor.

In total allele losses that were shown unequivocally to span the NF-1 region have been reported in 6/19 neurofibrosarcomas from NF-1 patients in the series of Xu et al. (1992), Glover et al. (1991) and Menon et al. (1990). Studies of neurofibrosarcomas in NF-1 patients by Basu et al. (1992) and DeClue et al. (1992) showed drastically diminished levels of neurofibromin by immunoassay with elevated levels of rasGTP. This suggested that the development of the tumours was due to nearly complete loss of neurofibromin function.

Andersen et al. (1993b), The et al. (1993) and Johnson et al. (1993) have shown aberrations in the NF-1 gene in sporadic malignant melanomas and neuroblastomas. Some of the melanomas and neuroblastomas were found to be associated with homozygous deletions at the NF-1 locus, i.e. both copies of the NF-1 gene were aberrant or deleted. Legius et al. (1993) also detected a homozygous deletion in a neurofibrosarcoma from an NF-1 patient. Loss of marker alleles in the vicinity of NF-1 or encompassing the NF-1 locus leaves open the possibility that some other gene in the region

and not NF-1 itself may be a relevant contributor to tumour progression in NF-1 associated malignancies.

The other mechanism of inactivation of the second allele could be due to small alterations within the NF-1 gene. Li et al. (1992) used the polymerase chain reaction and single stranded conformational polymorphism method to screen rapidly for point mutations and small alterations within the NF-1 gene. They initially screened exon 24 or the FLR exon, which is the most conserved exon in the NF-1 GRD. The three consecutive amino acids FLR are completely conserved across all 4 proteins and FLR occurs only once in the entire 2818 amino acid sequence of the predicted NF-1 protein. The relatively high degree of conservation of amino acid sequence across these different proteins and between species suggested that somatic mutations in this region of the gene would be likely to alter amino acid residues crucial to the function of the gene and so to the clinical pathology. The initial screening revealed an amino acid substitution altering Lys1423, that occurred in three tumour types: colon adenocarcinoma, myelodysplastic syndrome and anaplastic astrocytoma and in one family with neurofibromatosis-1. The GAP activity of the mutant GRD was shown to be 200-400 fold lower than that of wild type, whereas binding affinity was unaffected. Thus it was shown that point mutations seen in the germline and causing NF-1 also occur in somatic cells and contribute to the development of sporadic tumours, including tumours not associated with NF-1.

The above study reported the conversion of lysine 1423 to glutamine in a colon cancer sample and conversion to glycine in an anaplastic astrocytoma sample. An alteration of this lysine residue to methionine has been reported to affect stability of the protein (Weismuller and Wittinghoffer, 1992).

To further understand the significance of the lysine 1423 residue of neurofibromin and its role in NF-1 function, Poullet et al. (1994) mutated it to all possible amino acids. Functional assays using yeast ira complementation revealed that lysine was the only amino acid that produced functional neurofibromin. Quantitative analysis of different proteins suggested that their GAP activity is drastically reduced due to decrease in their ras affinity. Such a requirement for a specific residue was not observed in the case of other conserved residues within the GRD. The authors also reported that another residue phenylalanine 1434 plays an important role in NF-1 function. This was indicated by the finding that the defective neurofibromin due to an alteration of lysine 1423 to other amino acids could be rescued by a second mutation at residue 1434. The mutation partially restored GAP activity in the lysine mutant. Thus the above work demonstrated the importance of the lysine 1423 for NF-1 function.

Legius (NNFF newsletter-June 1994) recently identified a point mutation (L2317P) in exon 38 of the NF-1 gene. This mutation was detected in a patient whose neurofibrosarcoma was studied and found to show loss of heterozygosity for all of chromosome 17 (Glover et al., 1991) and lack of neurofibromin expression (DeClue et al., 1992 and Basu et al., 1992). With the finding of the germline mutation in this patient, it was shown that the normal allele was the one lost in the tumour, in accordance with the tumour suppressor hypothesis (NNNF International NF-1 Genetic analysis newsletter vol. 2, no 3). Thus the results further contributed to the fact that inactivation of both NF-1 alleles may contribute to the development of several NF-1 associated and non associated malignancies and that the NF-1 gene appears to follow the pattern of a recessive tumour suppressor gene.

1.10 AIMS OF THE STUDY

1. To amplify the NF-1 coding sequence by designing oligonucleotide primers for amplification using both DNA and RNA as templates for PCR.
2. To develop an efficient screening strategy for the detection of mutations, utilising techniques such as single stranded conformational polymorphism (SSCP) and chemical cleavage of mismatches (CCM).
3. To completely characterise the mutations identified by the above techniques, by direct sequencing of single stranded template generated by asymmetric PCR amplification.
4. To identify the mutational spectrum of the germline mutations affecting the NF-1 gene.
5. To analyse the distribution of mutations detected within the NF-1 gene and detect areas with a high mutation frequency (mutational hot-spots).
6. To assess the tumour suppressor function of neurofibromin by analysis of somatic mutations in tumours unrelated to NF-1.
7. To correlate the genotype with the phenotype by predicting the effect of the characterised mutations on the protein product.

CHAPTER: 2

MATERIALS AND METHODS

CHAPTER 2: MATERIALS AND METHODS

2.1 PATIENTS AND TUMOUR TISSUE USED FOR ANALYSIS

Twenty-five Caucasian patients including both sporadic and familial cases were screened for mutations in the NF-1 gene. 78% of the coding sequence was screened, of which 50% was screened using RNA as a template and 28% using DNA as a template.

The tumour samples studied were:

Melanoma: 3 samples (F-2564, B-5530, P-2852)

Neuroblastoma: 4 samples (S-1, B-2, C-3, M-4).

Myelodysplastic syndrome (MDS): 10 samples (922363, 921492, 921406, 921508, 921822, 922161, 922404, 921919, 921821, 921572)

Pancreatic carcinoma: 2 samples (9020, 2063)

Colonic adenocarcinoma: 2 samples (8496, 8278)

Colonic adenoma: 2 samples (8549, 8284)

Lung adenocarcinoma: 2 samples (0019, 2575)

Seminoma: 2 samples (0187, 0907)

Thyroid follicular carcinoma: 2 samples (7830, 3505)

2.2 EXTRACTION OF DNA

DNA was extracted from peripheral leukocytes, paraffin embedded tissue and cultured cell lines.

Extraction of DNA from whole blood

Genomic DNA was extracted from peripheral blood leukocytes (fresh and frozen) using a variation of the method described by Kunkel et al. (1977).

10ml of whole blood was put in a 50ml Falcon centrifuge tube, to which 40ml of lysis mix was added. Tubes were left on ice for 10 minutes and then centrifuged at 2500 r.p.m. for 10 minutes at 4^{o}C in an IEC DPR-6000 centrifuge. The resulting pellet was resuspended in 3ml of nuclei lysis mix, 200µl of 10% SDS and 100µl proteinase K (100 mg/ml). Tubes were incubated overnight at 37^{o}C. After incubation, 1ml of 6M sodium chloride was added with vigorous shaking and the tubes were spun at 2500 r.p.m. for a further 10 minutes. The supernatant was then transferred to a fresh 50ml Falcon centrifuge tube and 3ml of phenol:chloroform was added and spun at 2500 r.p.m. for 15 minutes. The upper aqueous phase was transferred to a fresh tube and DNA was precipitated by addition of two volumes of ethanol. The DNA was then spooled out, washed in 70% ethanol, allowed to air dry and suspended in 500µl of T.E buffer. DNA extractions were also carried out from bone marrow samples (in patients with myelodysplastic syndrome) that were stored at -20^{o}C in fixative (methanol:acetic, acid 3:1). In order to get rid of the fixative the cells were given 3 washes with PBS and the method used was the same as for the extraction from blood samples, except that the initial lysis step was omitted since the starting material was solely nucleated cells.

Extraction of DNA from paraffin embedded tissue

DNA extraction was carried out from tumour tissue embedded in paraffin blocks (Rolfs et al., 1992). 10µm sections were cut from the paraffin blocks, taking care to exclude surrounding normal tissue and only sections having tumour tissue were taken as far as possible. One 10µm section was placed in a 1.5ml microfuge tube. To remove the paraffin, 400µl of xylene was added. The tube was vortexed vigorously and centrifuged at full speed in a bench top centrifuge (IEC Centra 4X) for 5 minutes at room temperature.

The xylene was carefully removed with a pipette and the residual xylene was removed by gently mixing the tissue pellet with 400µl of absolute ethanol. This was again centrifuged at full speed in a bench top centrifuge for 5 minutes at room temperature. The ethanol was removed with a pipette and the remaining ethanol was allowed to evaporate off by incubating in an oven at 37°C for 10-15 minutes.

The tissue pellet was resuspended in 50µl of proteinase K buffer and this mixture was incubated at 50°C in a water bath for 24 hours. An additional aliquot of proteinase K was added after 12 hours, to ensure more complete protein digestion. The tubes were removed from the water bath and centrifuged briefly to remove any condensed water in the lid. To the tubes was added one volume phenol/chloroform/isoamyl alcohol (25:24:1). The tube was vortexed vigorously and then centrifuged at full speed in a bench top centrifuge for 10 minutes at room temperature, to separate the organic phase from the aqueous phase. The upper aqueous phase was then transferred to a clean 1.5ml microfuge tube. The organic phase was re-extracted by addition of 100µl T.E buffer. This was vortexed and spun at full speed in a bench top centrifuge. This aqueous phase was then added to the previously collected aqueous phase. The DNA from the aqueous phase was precipitated by addition of 1/10 volume of 3M sodium acetate (pH 5.2) and 2.5 volumes of pure ethanol. The mixture was mixed well and incubated at -20°C overnight. This was then centrifuged for 30 minutes at 4°C at full speed in a bench top centrifuge. The supernatant was removed and the pellet was washed with 200µl of 70% ethanol. The pellet was then allowed to dry and resuspended in 20µl of T.E buffer. Approximately 400ng of DNA can be recovered from a 10µm section and 5µl (100ng) of the above extracted DNA dissolved in T.E buffer was used for a PCR reaction.

Extraction of DNA from cultured cell lines

DNA was extracted from cultured cell lines using a modified procedure described by Laird et al. (1991).

Cells were grown in $80cm^2$ flasks at 37^oC (LTE incubator), in a closed system using Ham's F10 with Hepes (Gibco BRL) supplemented with 10% foetal calf serum, 10% new-born calf serum and 1% penicillin streptomycin until confluent (Boyd, I.). An 80% confluent flask contained about $2X10^6$ cells. The flasks were washed with 10ml of PBS (Gibco BRL) thrice to get rid of the medium. To the flask was added 5ml of lysis buffer. The flasks were incubated at 37^oC for 3-4 hours. The flasks were then taken out of the incubator and an equal volume of isopropanol was added. This was kept on a shaker and agitated for 15-20 minutes. The DNA obtained was spooled out and washed in 70% ethanol. This was then dissolved in 200µl of T.E buffer and stored at -20^oC until further use.

Determination of concentration of DNA

The concentration of each DNA sample was determined by measuring the optical density at 260nm in a dual beam spectrophotometer. An O.D. reading of 1 corresponds to 50µg/ml of DNA.

2.3 EXTRACTION OF RNA

RNA was extracted from peripheral lymphocytes and cultured cells by the modified acid-guanidinium thiocyanate method described by Chomczynski and Sacchi (1987).

Extraction of RNA from peripheral lymphocytes

The extraction of RNA from peripheral lymphocytes involved the initial separation of lymphocytes from whole blood, followed by the acid

guanidinium thiocyanate phenol-chloroform extraction. All the solutions used (except organic solutions) and the tubes and tips required for handling of the RNA were treated with a solution of 0.1% DEPC (diethyl pyrocarbonate), to prevent degradation of the RNA by RNases.

a) Separation of lymphocytes from whole blood:

5ml of whole blood was layered over 5ml of Histopaque (Sigma) in a Universal. The tube was then centrifuged for 30 minutes at 1400 r.p.m. at room temperature. The opaque interphase (buffy coat) was transferred into a fresh tube. The cells were then washed twice with 5ml of cold PBS (Gibco BRL). The supernatant was discarded after centrifugation at 1400 r.p.m. for 10 minutes.

b) The Acid guanidinium thiocyanate phenol chloroform extraction:

To the cells in the Universal, 500μl of solution D was added and the cells were thoroughly mixed with the solution D until viscous. This was then transferred to an autoclaved and DEPC treated 1.5ml microfuge tube. To this was added 50μl of 2M sodium acetate (pH 4.0), 500μl phenol and 100μl of chloroform-isoamyl alcohol (49:1). The microfuge tube was then vortexed for 10 seconds and placed on ice for 15 minutes. The tube was then centrifuged in a bench top centrifuge for 20 minutes. After centrifugation there was separation into an upper aqueous phase. This was transferred to a fresh microfuge tube, to which an equal volume of isopropanol was added. The tube was then vortexed and precipitated at -20^{o}C for a minimum of one hour. The tube was then centrifuged at full speed in a bench top centrifuge and the supernatant was discarded. The resultant RNA pellet was then dissolved in 300μl of solution D to which an equal volume of isopropanol was added. The mixture was kept at -20^{o}C for one hour. The tube was then centrifuged at full speed for 10 minutes. The supernatant was discarded and the pellet

was washed in 70% ethanol and air dried. The pellet was dissolved in 50μl of DEPC water and stored at -20°C until required.

Extraction of RNA from cultured cells

Human tumour (neuroblastoma) lines were cultured in 80cm^2 flasks until confluent. The cells were washed twice in PBS (Gibco BRL) and the flask drained. 1ml of solution D was added to the flask and the flask was agitated. A further 2ml of solution D was added and the viscous mixture was collected in a sterile 30ml polypropylene tube. The following reagents were added to this: 1/10th volume 2M sodium acetate (pH 4.0), 1ml unequilibrated phenol and 0.6ml chloroform:isoamyl alcohol (49:1). The sample was gently mixed after the addition of each reagent and incubated on ice for 15 minutes. The RNA was pelleted down by centrifugation at 10,000 r.p.m. for 20 minutes at 4°C in a Sorvall rotor. The aqueous phase was transferred to a fresh tube and 3ml of isopropranol was added. The solution was mixed and incubated at -20°C for 60 minutes. The RNA was pelleted at 10,000 r.p.m. for 20 minutes at 4°C and the supernatant discarded. The pellet was resuspended in 0.6ml solution D and an equal volume of isopropanolol was added. The RNA was precipitated by incubation at -20°C for 60 minutes followed by centrifugation at 10,000 r.p.m. for 10 minutes. The pellet of RNA was washed twice in 70% ethanol and resuspended in DEPC treated water. The method yielded 200-300μg of RNA per 5X10^6 cells.

Visualisation of the RNA

The RNA checking gel was prepared by boiling 0.45g agarose in 22ml dH_20 and cooled to 55°C. To this was added 5ml of 37% formaldehyde and 10 x MOPS to make up a total volume of 30ml. The gel was poured in a fume

hood. The RNA sample was prepared as follows: 1μl total RNA, 1.3μl dH_20, 5 μl formamide, 1.65μl (37%) formaldehyde solution and 1μl 10XMOPS made up to a total volume of 10μl. This was heated in a 55-60°C water bath for 10 minutes and quenched on ice. To this was added 2μl 10XRNA loading buffer and the sample was loaded immediately. The gel was run in 1XMOPS buffer at 70-80 volts for 40-60 minutes. After electrophoresis was complete, the gel was rinsed in water to wash out the formaldehyde and stained for 5 minutes in 5μg/ml ethidium bromide. The gel was destained in water overnight and viewed under UV light to check for integrity of ribosomal bands and to estimate concentrations. The concentration was also determined by measuring the optical density at 260nm. The integrity of the RNA was assesed by comparing the ratio of the O.D at 260/280. A good quality preparation should give a value of 2 (Gurr and McPherson, 1991).

2.4 DESIGN OF OLIGONUCLEOTIDES

The following guidelines were followed for primer design:

1)Where possible, primers were selected with an average G+C content of around 50% and a random base distribution.

2)The rule of thumb calculation of 2°C for A or T and 4°C for G or C (Thein and Wallace, 1986) was used in calculation of the Tm of both the primers such that they ranged from 50-75°C.

3)Most primers were designed to be about 20-30 bases in length and runs of C's or G's at the 3' ends of the primers and palindromic sequences within primers were avoided.

The primers used in this study were designed using the program OLIGO™ version 3.4 and in addition to the above, the following rules were considered:

-Self complementarity within the primers was avoided.

-3' end complementarity at the end of both primers was avoided, as this reduces the incidence of primer-dimer artefacts.

-Primers were designed such that there was not much difference between their melting temperatures (Tm) as calculated by the program.

-Duplex stability i.e. the strength of the bond the primer makes with the template was usually kept to >400 kcal/mol and was similar in both primers.

2.5 PURIFICATION OF OLIGONUCLEOTIDES

Oligonucleotides used in this study were synthesised on an ABI 391-DNA synthesiser. These oligonucleotides were phosphoramidite synthesised and require a final reaction step with concentrated ammonia to effect their complete deprotection. Ammonia was used for primer elution and deprotection and this needed to be removed from the oligonucleotides before the latter was used in various enzymatic processes. Primers were stored in their ammonia eluant at -20°C. Before use 10µl of this stock was heated to 95°C for 3 minutes to drive off the ammonia (Taylor G. 1991) cooled, diluted and used for amplification.

Another method described by Sawadago and VanDyke (1992) involved oligonucleotide purification through extractions with n-butanol. This method obviated the need for prior NH_3 removal and resulted in a high yield of oligonucleotides for use in PCR amplification and sequencing.

The procedure was carried out as follows: 100µl of deprotected oligonucleotide solution in NH_4OH was vortexed vigorously in a 1.5ml microfuge tube with 1000µl n-butanol for 15 seconds and then centrifuged for 10 minutes (minimum 1 minute) at 12,000 r.p.m. The H_2O containing n-butanol phase was removed and discarded. Following n-butanol extraction, the pellet was dried and resuspended in 50µl distilled H_2O and used for PCR

amplification. In this study both methods described were used as appropriate for the oligonucleotides used for PCR amplification or sequencing.

2.6 PCR USING GENOMIC DNA AS A TEMPLATE (DNA-PCR)

PCR reactions were performed in a 100µl PCR reaction mix containing 50mM KCl, 10mM Tris-HCl (pH 8.4), 1.5mM $MgCl_2$, 100µg/ml gelatin, 200µ mol of each dNTP and 500 ng of each primer This mixture was then subjected to 10 minutes of UV irradiation to minimise contamination. To this mixture was added 2.5 units of *Taq* DNA polymerase and 1µg of DNA. The mixture was vortexed, centrifuged, overlaid with 40µl of mineral oil and subjected to PCR. The primers used for amplification of the individual exons and the conditions for amplification were as described in Tables 6 and 7.

2.6.1 Primers and conditions used for amplification of exons

Exons 2, 4, 11, 18, 28-36, 42, 43 and 44 were amplified from genomic DNA. The primer sequences used are listed in Table 6.

TABLE 6: Primer sequences for amplification of exons

EXON	PRIMER SEQUENCE	REFERENCE
2F	5'TTT TAA GGA TAA ACT GTT TAC GTG 3'	designed for this study
2R	5'GAA AAG AAA GCA AAT TCC CCA AAA C 3'	"
4F	5'AGG TGG TGT GTA TGT AAG GTG TTC 3'	"
4R	5' TTT TAA TAC CAG AGA GGT TGT AAC 3'	"
11Fe	5'ATA AGT ACT CCA GTG TTA TGT 3'	"
11Ri	5'CAA AAT AAA ATT TAA AGT TGA AAG 3'	"
11Re	5'AAT GAA TAC AAG TTT AGA GAA GCT 3'	"
18Fe	5' CCC CTA CAC ACT GAT ACT GGT AGT 3'	"
18F	5'TGC CTT CTC TTT TGT CTA TAT CTG 3'	"
18R	5' ATT AAA ATT TAC AAA ACC CTA CAT 3'	"
28 f-a	5'AAT GAA TCC AGA CTT TGA AGA ATT G 3'	Cawthon (pers.comm)
28 r-a	5'GCT ACC TTT GAG GCC AGT CAG 3'	"
28 f-b	5'GGG AGT ACA CCA AGT ATC ATG AG 3'	"
28r-b	5'GCC AGG ATA TAG TCT AGT TAG TC 3'	"
29F	5'TTC CTT AAA TGG CAT AGT GTT TTG T 3'	"
29R	5'AAA CCC CAA ATC AAA CTG AAG AGA 3'	"
29Fi	5' AGC AGA GCG AAC AAA AGT CC 3'	designed for this study
29Ri	5' GGG GAT AGA GTC GGG CTG TG 3'	"

Table 6 (continued)

EXON	PRIMER SEQUENCE	REFERENCE
30F	5'AAC TTA TAA GGA TCG TTT AAA ACA ACT 3'	Cawthon (pers. comm)
30R	5'AAA GGG TTT TCT TTG AAT TCT CTT AG 3'	"
31F	5'ATA ATT GTT GAT GTG ATT TTC ATT G 3'	"
31R	5'AAT TTT GAA CCA GAT GAA GAG 3'	"
32F	5'ATC TAG TAT TTT TGA GGC CTC AG 3'	Cawthon et al. (1990)
32R	5'CAG ATA TGC TAT AGT ACA GAA GG 3'	"
33F	5'CAT ATC TGT TTT ATC ATC AGG AGG 3'	"
33R	5'AAG TAA AAT GGA GAA AGG AAC TGG 3'	"
34F	5'CAA AAT GAA ACA TGG AAC TTT AGA 3'	"
34R	5'TAA GCA TTA AGT ACA AAT AGC ACA 3'	"
35,36F	5'TAT GTA TTC AGA GTA TCC CCT T 3'	designed for this study
35,36R	5'ATA AAT GCT AGA ATG ATT TCT CAT GCT 3'	Cawthon et al. (1990)
42F	5'TAG GAC AGC CAC TTG GAA GGA 3'	Shen et al (1993)
42R	5'ACA TAA TTT AGA TAA ACT AAC 3'	"
43F	5' ATA GAC ACT GTA GTT AAT GAA 3'	"
43R	5'ATC AAT ACT TTG TTT CTC AC 3'	"
44F	5' CAC GTT AAT TCC CTA TCT TGC 3'	"
44R	5' TGA GAA GTA GAA GAC TGT ATC C 3'	"

F/f-forward, R/r-reverse, e=external reaction, i=internal reaction, 35,36f/r are the primers for amplification of exons 35 and 36 together

Primers for exons 2 and 4 were designed in order to characterise the positive chemical cleavage results obtained on CCM analysis of segment I which spans exons 2-4. Primers for exons 11 and 18 were designed to characterise the splice site mutations.

TABLE 7:Conditions for amplification of exons

EXON	DENATURATION	ANNEALING	EXTENSION	CYCLES
2	94°C- 1 min	50°C- 1 min	72°C- 1 min	35
4	94°C- 1 min	56°C- 1 min	72°C- 1.30 min	35
11	94°C- 1 min	48°C- 1 min	72°C- 1 min	35
18	94°C- 1 min	54°C- 1 min	72°C- 1 min	35
28	94°C- 1 min	58°C- 1 min	72°C- 1 min	30
29	94°C- 1 min	58°C- 1 min	72°C- 1 min	30
30	94°C- 1 min	58°C- 1 min	72°C- 1 min	30
31	94°C- 1 min	58°C- 1 min	72°C- 1 min	30
32	94°C- 1 min	58°C- 1 min	72°C- 1 min	30
33	94°C- 1 min	58°C- 1 min	72°C- 1 min	30
34	94°C- 1 min	58°C- 1 min	72°C- 1 min	30
35-36	94°C- 1 min	60°C- 1 min	60°C- 2 min	35
42	94°C- 1 min	50°C- 1 min	72°C- 1 min	35
43	94°C- 1 min	50°C- 1 min	72°C- 1 min	40
44	94°C- 1 min	54°C- 1 min	72°C- 1 min	40

Identification of PCR products

After completion of the 100µl PCR, 10µl of the PCR product was resolved on a 1% agarose minigel prepared in 1X TBE buffer. Electrophoresis was carried out at 100 volts for 45-60 minutes and the products were visualised on a UV transilluminator. After confirmation of size, each band was excised from the gel and covered with 50µl of T.E. buffer and stored in a labelled microfuge tube at 4°C, after overnight incubation at 37°C. The DNA thus eluted into the buffer and this was the template for

asymmetric PCR reactions (if necessary for that product). The remaining 90μl of PCR product was utilised as follows: 50μl was subjected to 'Geneclean' and the 'Genecleaned' products were stored separately at 4°C. 40μl of the remaining PCR product was stored at 4°C.

2.7 REVERSE TRANSCRIPTASE PCR: RT-PCR

RNA cannot serve as a template for PCR so RNA has to be converted to a complementary DNA which is amenable to PCR. The combination of reverse transcription and PCR was referred to as RT-PCR. Reverse transcription was carried out using MoMLV reverse transcriptase and using specific oligonucleotides to prime selectively the mRNA of interest. The procedure was carried out as follows: 1-2μg of RNA diluted in DEPC treated water was heated at 95°C for 5 minutes for denaturation. 10-50pM of the downstream primer used to prime reverse transcription was added and the mixture was heated at 65°C for 5 minutes to open up the regions in the RNA template that were rich in secondary structures. To the mixture was added reverse transcriptase buffer, dNTPs (1μl of a 10mM solution), 2μl 0.1M DTT (dithiothreitol) and 200U of MoMLV reverse transcriptase. Reverse transcription was carried out at 42°C for 45-60 minutes. After reverse transcription the mixture was heated at 95°C for 5 minutes to inactivate the enzyme, as reverse transcriptase is known to inhibit *Taq* polymerase, (Sellner et al., 1992) and to disrupt the mRNA:cDNA hybrid.

Table 8 lists the primers used for amplification of the seven RNA segments.

2.7.1 Primers and conditions for amplification of RNA segments I-VII

TABLE 8: Primers used for amplification of RNA segments I-VII

SEGMENT	PRIMER SEQUENCE
I-F*	5' GCA GCC GCT TCG ACG AGC A 3'
I-R*	5' CTA ACT GCG AAC CTT CTT TAG G 3'
II-F	5' TGC CTT GTT TCT TGC TTT CG 3'
II-R	5' TTC TAT CTG CCT GCT TAT TT 3'
III-Fe	5' CCT GAT GCT CCT GTA GAA AC 3'
III-Re	5' ATT TTG GCT TTT GGA TAG TT 3'
III-Fi	5' AAC ATT TTG GGA GAT TAG CT 3'
III-Ri	5' GCT TTG TTG CTT GTT CCC AT 3'
IV-Fe	5' GAA ACA CTG AGG CTT GGG AA 3'
IV-Re	5'AAA AAC TTG CTG ATG GTA TT 3'
IV-F-i	5' GCA AAA TGG GAA CAA GCA AC 3'
IV-Ri	5'ACT TGC TGA TGG TAT TCT TC 3'
V-Fe	5' CCT GCT CTG TAT CCA ATG CT 3'
V-Re	5' CCC ACT GAG AAC AAG GAA CCA CAT 3'
V-Fi	5' GCT CTG TAT CCA ATG CTA TTT AAC 3'
V-Ri	5' TTC CAC CAA TCT CTC AAA CC 3'

SEGMENT	PRIMER SEQUENCE
VI-Fe	5' AGT ATT GGC TGA TCG GTT TGA GAG 3'
VI-Re	5' AGG TCG TCT TCC AAC AGC TTT ATG 3'
VI-Fi	5' CGG TTT GAG AGA TTG GTG GA 3'
VI-Ri	5' CTG TTG CTG GAA AGA TAC TG 3'
VII-Fe	5' GTG CTT GCT TTA CAT CGT CTA CTC 3'
VII-Re	5' GTC GTA AGC AAA GCC AGG AAA AAC 3'
VII-Fi	5' ACA TCG TCT ACT CTG GAA CAA TCA 3'
VII-Ri	5' CAG GAA AAA CAA CAA ACC ACT TAG 3'
29-Fe *	5' TCA GCA GAG CGA ACA AAA GTC CTA 3'
33-Re*	5' GCG GAC CTG TGG CTA CTA AGA AAG 3'
29-Fi*	5' TCA CAG CCC GAC TCT ATC CCC CAA 3'
33-Ri*	5' AAA GTA GGA GTT GGA GAT AAG CAT 3'

'e' denotes the external reaction and **'i'** the internal reaction of the nested PCR reaction.

All the primers described in this table, except the primers for segment-I were designed for this study using the sequence information supplied by the NNFF International Mutation Analysis Consortium. The primers used for amplification of segment I were described by Guttmann et al. (1993).

TABLE 9: Conditions for amplification

SEGMENT	DENATURATION	ANNEALING	EXTENSION	CYCLES
I	94°C for 1 min	56°C for 1 min	72°C for 2 min	30
II	93°C for 1 min	53°C for 1 min	72°C for 2 min	35
III-e	94°C for 1 min	52°C for 1 min	72°C for 2 min	30
III-i	94°C for 1 min	56°C for 1 min	72°C for 2 min	30
IV-e	94°C for 1 min	52°C for 1 min	72°C for 2 min	30
IV-i	94°C for 1 min	56°C for 1 min	72°C for 2 min	30
V-e	94°C for 1 min	52°C for 1 min	72°C for 2 min	30
V-i	94°C for 1 min	56°C for 1 min	72°C for 2 min	30
VI-e	94°C for 1 min	52°C for 1 min	72°C for 2 min	30
VI-i	94°C for 1 min	56°C for 1 min	72°C for 2 min	30
VII-e	94°C for 1 min	56°C for 1 min	72°C for 2 min	30
VII-i	94°C for 1 min	56°C for 1 min	72°C for 2 min	30
29-33-e	94°C for 1 min	52°C for 1 min	72°C for 2 min	30
29-33-i	94°C for 1 min	58°C for 1 min	72°C for 2 min	30

'e' denotes the external reaction and **'i'** the internal reaction of the nested PCR reaction.

Other primer sets were designed internal to Segment IV, V and Segment VI for specific further characterisation or as primers for sequencing. They were as follows:

Segment IV was subdivided into 2 parts and 2 new primer sets (A-R and B-F) were designed for use with IVF-e and IVRe: (IVF-e and IVA-R) and (IVB-F and IVRe).

IVA-R 5' CCT GAT GCT CCT GTA GAA AC 3'

IVB-F 5' CGG AAG GGA AAA GGG AAC TC 3'

A forward sequencing primer internal to IV (IV-Fseq) was as follows:

IV-Fseq 5' TGT TAA GAG GCG AAT GTC CCA T 3'

Segment VI was further divided into two segments (a and b) after amplification using VI-Fe and VI-Re. The primer sets used to amplify segment a and b were:

VI-Fa 5' TTA CTC TAC CAA CTG CTC TG 3'

VI-Ra 5' GAG GAG GAA CTG ATG ATG GC 3'

VI-Fb 5' ATC ATC AGT TCC TCC TCA GA 3'

VI-Rb 5' ATG AAG GAA AGA CTA TGA TT 3'

PCRs using these primers was performed for 30 cycles of denaturation for 1 min at 94^{o}C, annealing of 56^{o}C for 1 min, extension of 72^{o}C for 1.30 minutes and a final extension of 72^{o}C for 10 minutes.

Identification of PCR products

The products of amplification (10μl) were electrophoresed on 1% agarose gels and visualised by ethidium bromide staining, to confirm the size and detect any abnormal alterations in size. The bands were then excised from the gels and covered with 40μl of T.E. buffer and stored at 4^{o}C. In the case of RT-PCRs with nested segments all the RT-I products (external reactions) were stored for later use and the remaining 90μl of RT-PCR product (RT-2) (internal reaction) was subjected to 'Geneclean' and stored.

2.8 AMPLIFICATION OF THE RAS GROUP OF ONCOGENES

Point mutations in codons 12, 13 and 61 in the ras group of oncogenes were screened for using primers for amplification of the sequence surrounding codons 12 and 13 (in exon 1) and a second set of oligonucleotides for amplification of the region surrounding codon 61 (in exon 2). The primers used for amplification of these regions in the H-ras (H), K-ras (K) and N-ras (N) group of oncogenes were as follows (Lyons, 1990).

Table 10: Primers used for amplification of the regions surrounding codons 12-13 and codon 61 of the H,K,and N ras group of oncogenes

Primer	Sequence 5' to 3'
N (12-13) F	CTT GCT GGT GTG AAA TGA CT
N (12-13) R	ACA AAG TGG TTC TGG ATT AG
K (12-13) F	TTT TTA TTA TAA GGC CTG CT
K (12-13) R	GTC CAC AAA ATG ATT CTG AA
H (12-13) F	GAG ACC CTG TAG GAG GAC CC
H (12-13) R	CGT CCA CAA AAT GGT TCT GG
N 61-F	GTT ATA GAT GGT GAAA ACC TG
N 61-R	AAG CCT TCG CCT GTC CT CAT
K 61-F	ACC TGT CTC TTG GAT ATT CT
K 61-R	TGA TTT ACT ATT ATT TAT GG
H 61-F	CCG GAA GCA GGT GGT CAT TG
H 61-R	ACA CAC ACA GGA AGC CCT CC

Amplification conditions for all 6 PCR reactions were as follows: 30 cycles of 94°C for 1 min, 52°C for 1 min and 72°C for 1 min followed by a final extension of 72°C for 5 min.

2.9 PCR OF THE TETRANUCLEOTIDE REPEATS (AAAT)

A highly polymorphic (AAAT) Alu-repeat sequence located in intron 27 of the NF-1 gene (Xu et al., 1991) was analysed in 5 families, where the causative mutation had not been characterised. This repeat sequence (AAAT) was used to track the NF-1 gene through the family. The primer sequences used were:

Alu I : 5' CAA GAA AAG CTA ATA TCG GC 3'

Alu II :5' GGA ACC TTA AGT TCA CTT AG 3'

PCR was carried out with 300ng of DNA, 200μmol each dNTP, 20pM of primer *Alu* I and 1.5pM of end labelled and 18.5pM of cold primer *Alu* II in a total volume of 25μl.

20pM of *Alu* II was end labelled using 20μCi [γ-^{32}P] ATP (3000Ci/mmol) and 5U of polynucleotide kinase (nbl).

2μl of amplification products were mixed with 10μl of formamide stop solution and loaded onto a 6% polyacrylamide gel and electrophoresed at 1500V for 5 hours. The gel was dried and exposed to radiographic film for 10-15 hours.

2.10 METHODS FOR MUTATION DETECTION

2.10.1 SINGLE STRANDED CONFORMATION POLYMORPHISM (SSCP)

SSCP analysis allows the detection of single base changes in short DNA fragments, due to mobility differences of single stranded molecules (Orita et al., 1989).

In PCR-SSCP analysis, the DNA sequence of interest was first amplified by PCR. For efficient detection of the product, isotopic label was coupled to the PCR primers or was incorporated into the PCR product itself. The PCR product was then heated to dissociate the strands and analysed by non-denaturing polyacrylamide or MDE™ gel electrophoresis.

Radiolabelled PCR

Radio labelled PCR was carried out using two different methods:

1)The first method involved labelling of oligonucleotides by T4 polynucleotide kinase and either one or both the primers were radio labelled and used for amplification. Synthetic oligonucleotides were synthesised without a 5' phosphate group and therefore can be marked quickly at the 5' OH end by the transfer of [γ-^{32}P]. This was supplied by [γ-^{32}P] ATP and catalysed by the T4 kinase enzyme. The procedure for labelling was as follows:

1)50pM of the oligonucleotide was taken in a microfuge tube, to which was added 1.0µl of T4 kinase buffer (500mM Tris-HCl pH 7.6, 0.1M DTT, 0.1M $MgCl_2$), 1.0µl of [γ-^{32}P] ATP (10µCi of 3000 Ci/mmol, 10 mCi/ml) and 1.5µl (15U of 10U/µl) of T4 polynucleotide kinase enzyme and the mixture was made up to 10µl with sterile distilled water.

2)The mixture was mixed thoroughly centrifuged and incubated at 37°C for 30-60 minutes. The 10μl mixture was used for a 50μl PCR which was made up as follows:

10 μl of the above mixture containing the labelled oligonucleotide, 50pM of the other primer, 5μl of 10XPCR buffer, 1μl dNTPs (250μM), 0.5μl of *Taq* polymerase and 0.5μl DNA (20-200ng) was made up to a total of 50μl, overlaid with mineral oil and subjected to PCR under the appropriate conditions.

For the radio labelling using [α^{32}-P] dCTP incorporation, a 10μl PCR was performed, using a master mix prepared for 40 tubes which contained: 1Xreaction buffer (50mM KCl, 10mM Tris-HCl pH 8, 1.5mM $MgCl_2$), 20μCi [α-^{32}P] dCTP, 100pM of the appropriate primers and 20U *Taq* polymerase in a total volume of 380μl. Reactions were assembled on ice by mixing 9.5μl master mix with 0.5μl DNA (20-200ng) and overlaid with mineral oil and subjected to PCR.

Larger PCR products were analysed by SSCP after digestion with restriction endonucleases to obtain DNA fragments of 100-400 bp. The advantage of this approach was that the length of the primary PCR product was not limitative as long as appropriate restriction sites were available. In addition, using small restriction fragments for SSCP, the sensitivity of the technique may be enhanced. The PCR product containing exons 35 and 36 of the NF-1 gene was also amplified from genomic DNA as a single fragment. This PCR product was first non radioactively amplified by PCR. This was then digested using two restriction enzymes *Rsa* I and *Xba* I and the products were electrophoresed on a 1% agarose gel to identify the sizes of the digested products. 10μl of the digested product was radioactively end-labelled and subjected to SSCP analysis.

Preparation of samples and conditions for gel electrophoresis

1μl of the PCR product was diluted in 10μl loading buffer (95% formamide, 10mM NaOH, 0.1% bromophenol blue, 0.1% xylene cyanol) and denatured for 5 minutes at 90°C. After quick cooling on ice, 5-10μl of each sample was loaded onto non denaturing polyacrylamide gels, or on a MDE™ gel, which is a modified polyacrylamide based vinyl polymer solution.

For preparation and electrophoresis of the SSCP gels, the Sequigen sequencing gel apparatus (Biorad) was used. The spacers and well forming comb used were 0.4 mm in thickness. The glass plates were thoroughly cleaned using soapy water, dried and wiped with 100% ethanol. The spacers were kept in place and the apparatus assembled. The bottom of the gel was sealed using the sealing apparatus, a wick of 3MM Whatmann paper and 30ml of a 10% polyacrylamide solution, after addition of TEMED and ammonium persulfate (20%) which allowed polymerisation and sealing. The gel forming solution using either acrylamide or MDE™ was prepared and 100 μl of ammonium persulphate (20%) and TEMED was added. The gel was carefully poured between the two glass plates and allowed to polymerise for a minimum of 1 hour before electrophoresis. After the gel had set, the sealing apparatus and the well forming comb were removed. The buffer chambers were filled with 0.6XTBE or 1XTBE depending on the type of gel matrix used and the wells were washed with buffer. The prepared samples were then loaded onto the gel and electrophoresis was carried out under the following conditions:

Non denaturing polyacrylamide gels:

1)At room temperature, 6% polyacrylamide, 10% glycerol at 8 watts for 14-16 hours in 1X TBE buffer. 2)At 4°C 6% polyacrylamide, without glycerol, at 30 watts for 4-5 hours.

MDE™ gels: only one electrophoretic condition was used, 10% glycerol at 6-8 watts for 16-18 hours in 0.6XTBE buffer at room temperature.

Autoradiography

After electrophoresis, the glass plates were separated and the gel was transferred to a pre-cut piece of Whatmann 3MM paper. The gel was covered with Saran wrap or cling film and dried under a vacuum for 20-40 minutes at 80^{o}C on a gel dryer (model 583 Biorad) and subjected to autoradiography. Alternatively, the gel was autoradiographed without drying for 10-16 hours, using X-Omat diagnostic AR imaging film.

2.10.2 CHEMICAL CLEAVAGE OF MISMATCHES (CCM)

The technique for chemical mismatch cleavage identifies sequence mismatches in heteroduplexed molecules, by chemical modification of the single strand causing a break or cleavage at the site of mismatch (Cotton et al., 1988).

The protocol involved five steps; (i)preparation of labelled probe and unlabelled test DNA, (ii)formation of heteroduplexes, (iii)chemical modification, (iv)chemical cleavage of mismatches and (v)electrophoresis and autoradiography.

Preparation of the labelled probe and unlabelled test samples

The DNA or cDNA required for CCM was obtained by amplification of the required segment by PCR. PCR was carried out for all the test samples and a couple of control samples, against which the sequence variants were detected. The PCR products were checked on a 1% agarose gel. If non-specific products of amplification were seen, the entire PCR product was

electrophoresed on a low melting (NuSieve) agarose gel and the required band was cut out after visualisation on a UV transilluminator. The DNA from the gel slice was obtained by subjecting the band to purification using the 'Geneclean' kit. If the PCR product was well defined and did not require separation from other non-specific products of amplification, Geneclean was carried out directly on the PCR product (without excision). The 'Geneclean' kit was used to remove and purify DNA from agarose, thus removing ethidium bromide, primers and dNTPs from PCR reaction products. By this method, 15-20μl of a DNA solution with a concentration of 100ng/μl was obtained. The purified PCR product served as the unlabelled test DNA to be used in the preparation of heteroduplexes. The 'Geneclean' procedure was carried out as follows:

To the required band (excised from the agarose gel) or the PCR product was added 3 volumes of NaI stock solution. This was incubated at 45-55^{o}C to dissolve agarose. To this was added 10μl glassmilk and the tubes were vortexed and kept on ice for 10 minutes. The glassmilk-DNA complex was pelleted for 5 seconds. The pellet was then washed 3 times in NEW wash buffer. The DNA was then eluted into T.E. buffer by incubation at 65^{o}C for 5 minutes. The tube was then centrifuged and the supernatant was transferred to a fresh tube.

Preparation of radiolabelled probe

50pM of each PCR primer was end labelled at the 5' ends using T4 polynucleotide kinase and [γ^{32}-P] ATP. This was carried out in two separate reactions of 10μl each for each primer.

50pM of the oligonucleotide was taken in a microfuge tube, to which was added 1.0μl of T4 kinase buffer. 1.0μl of [γ-^{32}P] ATP (10μCi of 3000 Ci/mmol, 10 mCi/ml) and 1.5μl (15U of 10U/μl) of T4 polynucleotide kinase enzyme were added and the mixture was made up to 10μl with sterile distilled water. The mixture was mixed thoroughly, centrifuged and incubated at 37°C for 30-60 minutes.

The PCR reaction was prepared containing all the components excluding the primers and including 0.5μl of the 'Geneclean' product of the control DNA PCR (as previously described) as the template. 5μl of each end labelled reaction mixture was directly added to the PCR mix. The mixture was layered with mineral oil and subjected to PCR under the appropriate conditions. The wild-type PCR product, which was the radiolabelled probe was directly mixed with 3 volumes of the NaI solution and the same method (Geneclean) was followed.

Formation of heteroduplexes

15ng of labelled probe (10^5 cpm) was mixed with 150-200ng of each test DNA in a 20μl volume of 1Xannealing buffer. Heteroduplex formation was carried out in 0.5ml microfuge tubes. 50μl mineral oil was layered onto the mixture and it was heated to 99°C for 10 minutes and then placed in a 65 °C water bath for two hours. The heteroduplexes were transferred to 1.5ml siliconised microfuge tubes to be precipitated. To each tube was added 3μl of glycogen (20 mg/ml) and 750μl of precipitation mix and the mixture was incubated at -70°C for 15 minutes. The tubes were centrifuged at 12,000 r.p.m. for 10 minutes, the pellet was rinsed with 70% ethanol, air dried and resuspended in 14μl T.E. buffer.

Chemical modification using hydroxylamine and osmium tetroxide

Each heteroduplex was split between two tubes (7μl each) i.e. one tube each for the hydroxylamine and osmium tetroxide reactions.

Preparation of hydroxylamine and the hydroxylamine reaction: 1.39G of hydroxylamine hydrochloride was mixed in 1.6ml of pre-warmed (37°C) distilled water. The pH of the solution was adjusted to 6.0 by the addition of diethylamine (about 1.5ml) and the solution was stored at 4°C for upto 7-10 days. 20μl of this solution was added to the heteroduplex and incubation carried out at 37°C for 2 hours.

Preparation of osmium tetroxide (OsO_4) and the OsO_4 reaction: A commercially available 4% solution (which was stored at 4°C for up to 3 months) was used. A 10Xreaction mix was prepared by adding 1.5μl osmium tetroxide, 6.75μl pyridine and 154μl T.E. buffer on ice. 18μl of this mix was then added to the other half of the heteroduplex. The mixture was incubated at 37°C for 1 - 2 hours.

Both the hydroxylamine and osmium tetroxide reactions were stopped by transferring the tubes onto ice which was followed by the addition of 750μl of precipitation mix. The DNA was precipitated at -70°C for 15 minutes and centrifuged at 12,000 r.p.m. to obtain a pellet, which was rinsed and briefly air dried.

Piperidine cleavage of the chemically modified mismatch

50μl of a 10% piperidine solution was added to each pellet and the tubes were vortexed for up to a minute. They were then incubated at 90°C for 30 minutes. The tubes were chilled on ice and the DNA was precipitated as before and the pellet was resuspended in 15μl formamide loading buffer

(95% formamide, 10mM EDTA, 10mg/ml bromophenol blue, 10mg/ml xylene cyanol).

Gel electrophoresis and autoradiography

The resuspended pellet was heated to 95°C for 10 minutes and applied to an 8% denaturing polyacrylamide gel containing 7M urea. The gel mixture was prepared and polymerised in between the plates of a sequencing gel apparatus (Sequigen; Biorad) as previously described. A labelled 1Kb DNA ladder (Gibco BRL) was denatured and electrophoresed on the gel, in order to permit accurate size estimation of the cleavage products. Electrophoresis was carried out at 35 watts constant power till the bromophenol blue reached the end of the gel. The gel was transferred onto a pre-cut piece of Whatmann 3MM paper and autoradiographed using Kodak X-Omat AR film for a period of 12-16 hours.

2.11 DIRECT SEQUENCING

Direct sequencing of the NF-1 DNA and cDNA was carried out using dideoxy chain termination method using Sequenase version 2.0 (USB) (Sanger et al., 1977).

Generation of single stranded template

To perform sequencing under optimal conditions, it was necessary to obtain a well purified single stranded DNA template, to avoid the reassociation of the two amplified strands. Two methods were used:

1)Unequal primer ratios were used right from the beginning (Mgone et al., 1992). The template used for asymmetric amplification using unequal ratios of primers, was the PCR product of the corresponding region. 50μl of

the PCR product was resolved on a 1.0 % agarose gel and the appropriate ethidium bromide stained band was excised and eluted by soaking in 100µl T.E. buffer overnight. A 10µl aliquot of the eluted DNA or cDNA was used in a 100µl asymmetric PCR mixture containing 50mM KCl, 10mM tris-HCl (pH 8.3), 1.5mM $MgCl_2$ and a primer ratio of 100:1 (50pM:0.5pM). A second asymmetric PCR was performed with the primer ratio reversed, so that both strands could be generated. Another method of elution was electrophoresis on a 8% polyacrylamide gel and visualisation and excision of the appropriate band. The band was then covered with 100µl of T.E buffer and the tube with the excised band was incubated at 37°C overnight. 10µl of this was then subjected to asymmetric amplification as above.

2)One primer only was used to reamplify one DNA strand, of the pre-existing double stranded PCR fragments (Rolfs et al 1992). This was the other method used for production of single stranded DNA and involved taking a small amount (1-2%) of a PCR product (0.5-1 µl of PCR product from a 50µl reaction) and repeating the temperature cycling process using one primer only (50-100pM) for between 15-30 cycles.

Visualisation of single stranded DNA

At the end of the asymmetric PCR, 10µl of the PCR product was fractionated on a 1.0% Seakem agarose gel, along with the double stranded control and a suitable marker to monitor the amount of single stranded products. When a high proportion of single stranded products were observed these were used in dideoxysequencing reactions.

Purification of the template for sequencing

Prior to sequencing, the PCR products were purified by removing the excess dNTPs, salts and primers. This was accomplished by precipitation at room temperature for 10 minutes, by addition of an equal volume of 4M ammonium acetate and 2 volumes of isopropanolol. The template was recovered by centrifugation in a microcentrifuge at room temperature for 10 minutes, washed in 70% ethanol and allowed to dry. The pellet was then dissolved in 7μl of distilled water and used in the sequencing reaction.

Annealing of template and primer

7μl of purified asymmetric PCR product was mixed with 2μl 5Xsequencing buffer giving a final concentration of 40mM Tris-Hcl pH 7.5, 20mM $MgCl_2$ and 50mM NaCl and 1pM of sequencing primer, either limiting or nested and complementary to the synthesised single stranded DNA. 1.0pM of primer was used for each set of sequencing reactions. The concentration of the primer was determined by reading the optical density at 260 nm. If the primer has N bases, the approximate concentration (pM/μl) was given by the following formula:

Concentration (pM/μl) = O.D. at 260 / (0.01 x N)

Formula obtained from 'step-by-step protocols for sequencing with Sequenase version 2.0' 5th edition USB.

The template-primer mixture was annealed by heating to 65^{o}C for 2 minutes and then allowed to cool slowly, to less than 35^{o}C over a period of 30 minutes.

Labelling reaction

The mixture was then placed on ice and 1μl of 0.1M DTT, 2μM of each dNTP except dATP, 0.5μl [α^{35}-S] dATP (1000 Ci/mmol;10 μCi/μl) and 2μl (1.5u/μl) of a 1:8 dilution of T7 DNA polymerase (Sequenase version 2.0) were added and incubated at room temperature for 2-5 minutes.

Termination reaction

The termination mixes contain 80μM of each dNTP and 8μM of the appropriate dideoxyribonucleoside triphosphate. The 4 labelled (A,T,G,C) microfuge tubes containing 2.5μl of the appropriate termination mix were prewarmed at 37^{o}C. Once the labelling reaction was complete, 3.5μl was transferred to each tube and the incubation reaction was continued for a further 5 minutes at 37^{o}C. The reactions were terminated by addition of 4μl of formamide dye stop solution. The samples were stored at -20^{o}C until further required.

Sequencing gel electrophoresis and autoradiography

The samples were heated at 75^{o}C, quenched on ice and loaded on a 8% polyacrylamide gel containing 7M urea. The sequencing gels were poured and run on a Sequigen sequencing gel apparatus (Biorad) as previously described. The gels were run at a temperature of 50^{o}C for 2-4 hours. On completion of electrophoresis the apparatus was dismantled and the plates were gently pulled apart. The gel was then treated with a solution of 10% methanol/10% acetic acid for 10 minutes and then dried in a vacuum gel dryer. The gel was then exposed for 16-18 hours to Kodak X-Omat AR film in a cassette using intensifying screens.

2.11.1 Variation in sequencing protocols

The sequencing reactions and electrophoretic conditions were usually carried out as described. However some variations were carried out in the following cases:

1)Elimination of compressions: Some DNA sequences, especially those with dyad symmetries containing dG and dC residues, were not fully denatured during electrophoresis. When this occurs, the regular pattern of migration of DNA fragments was interrupted; the bands were spaced closer than normal (compressed together), or sometimes further apart than normal. In this case the substitution of a nucleotide analogue for dGTP, dITP was used. Using dITP, the bands were sharper and most compressions were eliminated. The reactions containing dITP were run in conjunction with reactions containing dGTP, since the overall quality of the reactions containing dITP was lower.

2)Reading sequences close to the primer: For reading sequences close to the primer, two methods were used: One method was to use less dinucleotide in the labelling step. This meant a further dilution of the labelling mix to a 1:10 or 1:20 of the stock reagent instead of the usual 1:4 and both reaction times were 3-5 minutes. When reading sequences within 20 nucleotides of the 3' end of the primer, it was essential that sufficient template was present and so the amount of DNA template used and the primer for annealing was doubled. The other alternative was to use Mn buffer, which affects reactions in the termination step. The addition of Mn buffer which was a buffered solution of $MnCl_2$, reduces the average length of DNA synthesised in the termination step, intensifying bands corresponding to sequences close to the primer.

3)Extending sequences farther away from the primer: For extending sequences farther from the primer, two different methods were used: One

method involved using more nucleotide in the labelling step. This meant use of undiluted labelling mix and the labelling reaction was lengthened to 5 minutes. Excess dATP was also added and this resulted in a sequencing ladder that emphasised the bands in the 200-800 nucleotide range, at the expense of bands in the 20-100 nucleotide range. The other method involves alteration of the nucleotide mixture used in the termination step. This was achieved by using the sequence extending mix and the usual volume of termination mix was replaced by a mixture of termination mix (1μl) and sequence extending mix (1.5μl). The gels on which the sequences (sequenced using the extension mix) were loaded were electrophoresed for 4-10 times longer and cooler at 40°C instead of 50°C for the best resolution.

2.12 PLASMID TECHNIQUES

The cDNA probe P5 was obtained as a plasmid DNA sample from ATCC and was transformed, grown up and stored as a glycerol stock. The details of the probe were as follows: The probe contains part of the 3' region of the NF-1 cDNA. The vector used was pBluescript SK-, the total size being 4.7 kb and the insert size being 1.7 kb. Restriction digests of the clone with *Eco*RI give the following sizes: 3.0, 1.2 and 0.5 kb (ATCC), which were confirmed on restriction digestion. The complete insert detects human genomic *Eco*RI fragments of 15, 6.5, 4.0, 2.3, 2.0, 1.8 and 1.3 kb (Wallace et al., 1990).

Transformations

The competent cells used were subcloning efficiency DH5α™ (Gibco BRL). The competent cells were removed from the freezer and thawed on ice. The required number of autoclaved 1.5ml microcentrifuge tubes were

placed on ice. The cells were gently mixed and 50µl were aliquoted into each chilled tube. 1-10ng of DNA was added to the tube and the tube was gently tapped to mix. The cells were incubated on ice for 30 minutes. The cells were then heat shocked for 20 seconds at 37°C and then placed on ice for 2 minutes. 950µl of prewarmed L-Broth was added to the cells and the cells were incubated in Universals at 37°C in a shaker at 225 r.p.m. for 1 hour. The control Universal contained only 950µl of L-broth. 200µl of the reaction was spread onto LB plates (1.5% L-agar with ampicillin added to a final concentration of 100 µg/ml) and the plates were incubated at 37°C.

Preparation of Glycerol stocks

The agar plates were removed from the incubator and a single colony was picked and grown overnight in 15ml of L-broth containing ampicillin at 37°C. After incubation glycerol stocks were made in 15% glycerol and stored at -20°C.

Preparation of probe

a)Using Circle-prep method:

i)Plasmid growth and cleared supernatant preparation:

Cultures were set up with 100µl glycerol stock in 50ml L-broth containing ampicillin and incubated overnight with vigorous shaking. The next day, the cultures were centrifuged at 1200 r.p.m. at 4°C for 10 minutes. The supernatant was removed and the tube was drained for 1-2 minutes. 4ml of prelysis buffer from the kit was added and the pellet resuspended. 4ml of alkaline lysis reagent was then added and the solution mixed well for 1-5 minutes, until the cells were uniformly lysed. The cell suspension at this point was very viscous. The suspension was then centrifuged at 1200 r.p.m. for 5

minutes at 4°C. The supernatant was then transferred through a sieve to a fresh tube. One volume of isopropyl alcohol was added and mixed well producing a fine precipitate. The tube was centrifuged for a further 5 minutes at 4°C at 1200 r.p.m. The supernatant was removed to a fresh tube. The pellet obtained, was dissolved in 0.5ml of water.

ii)Lithium chloride precipitation of RNA and single stranded DNA:

The cleared supernatant, which was transferred to a microcentrifuge tube, was incubated at 100°C in a hot block for 3-5 minutes and cooled in an ice bath for at least 1 minute. 300µl of lithium chloride was added and after mixing was allowed to stand at room temperature for 5 minutes. The solution was then centrifuged for 2 minutes. The supernatant was transferred to a new tube and 600-700µl of isopropanolol was added, the solution was mixed and then centrifuged for 2 minutes. The drained pellet was then dissolved in 0.5ml of water and 300µl of lithium chloride was added.

iii)Purification of plasmid DNA on Circle-prep glassmilk:

The Circle-prep glassmilk vial was vortexed until homogenous and then 75µl was added to the tube. After mixing, this was incubated for 5 minutes and then centrifuged for a few seconds. The supernatant was removed and the pellet washed twice with 1ml of binding buffer. The tubes were spun for a few seconds to remove traces of liquid. The pellet was then washed with wash solution, the tube spun for a few seconds and the supernatant discarded. The pellet was resuspended in 150µl of water and incubated for 5 minutes in a 60°C water bath. The tube was then spun for 30 seconds to form a tight pellet and the supernatant containing plasmid DNA transferred to a fresh tube.

b)Using the Insta-Mini-Prep ™:

The unopened Insta-prep-mini tube was centrifuged at 12000g for 10 seconds to pellet the gel. The bacterial culture was prepared as previously described and 1.5ml was transferred to a separate microfuge tube (not insta-prep mini tube) and spun at 12000g. After centrifugation all the supernatant was aspirated and care was taken not to disturb the bacterial pellet. To the pellet was added 50μl of STE buffer and the pellet was vortexed briefly to resuspend the bacterial cells. 300μl of shaken PCI solution was added to the tube containing the resuspended cells and the tube was mixed by repeated inversion. The entire contents of this tube were then carefully transferred to the pre-spun Insta-prep-mini tube. This tube containing the bacterial lysate was then centrifuged at 12000g for 1 minute. To the spun tube, 300μl of CI solution was added and the two upper liquid phases were mixed by repeated inversion. The tube was then centrifuged at 12000g for 1 minute. The plasmid DNA was recovered by pipetting the topmost phase to a fresh microcentrifuge tube and stored at -20^{o}C until further use.

c)Using RT-PCR:

The PCR product was visualised on a 1% agarose gel and the size and specificity was confirmed. The PCR product was then purified using 'Geneclean' (BIO 101) as previously described. The optical density of the 'Genecleaned' product was measured and depending on the concentration, 2-4μl was used for labelling.

Releasing insert from plasmid

Plasmid concentrations were determined by measuring the optical density of each sample at 260nM in a dual beam spectrophotometer (LKB

Biochem Ultrospec 4050). An O.D. reading of 1 corresponds to approximately 50μg/ml of DNA. 50μl digests were set up containing 7.5μl of the plasmid DNA obtained using Circleprep and 15μl of plasmid DNA obtained using Insta-prep respectively, along with 4μl of the enzyme *Eco*RI and 5μl of the enzyme buffer. The digests were incubated at 37°C for 2 hours. A 1% low melting agarose gel (Seakem agarose, FMC bioproducts) was made and the digests were run at 100 volts for 2 hours, alongside a 1 kb ladder (Gibco BRL). The gel was visualised on a UV illuminator and the 1.2 kb band (insert) was excised using a sterile scalpel blade and placed in a weighed Sarstedt tube. The weight of the insert was determined and a volume of sterile distilled water 3 times the weight of the insert was added. The probe was boiled for 10 minutes and stored at -20°C until further use.

2.13 SOUTHERN BLOTTING

Digestion of genomic DNA

A 10μg aliquot of genomic DNA was digested using the appropriate restriction enzyme in a concentration of 4U/μg of DNA in a total volume of 40 μl. The digests were incubated overnight at the suppliers recommended temperature, for optimal enzyme activity.

Agarose gel electrophoresis

The digests were resolved on 0.8% agarose gels (FMC bioproducts) containing 0.05% ethidium bromide (10 mg/ml). 4μl of loading mix containing 0.05% bromophenol blue was added to each digest which was loaded onto the gel. The gels were electrophoresed at 60-70mA for 12 hours. The gels were visualised on a UV transilluminator and photographed. After photographing the gel, the gel was treated in the depurination solution

(0.25M HCl) for 20 minutes, then transferred to denaturation solution (0.5M NaOH, 1.5M NaCl) to denature the DNAs for 20 minutes and finally transferred to neutralisation solution (1M Tris-HCl pH 8.0, 1.5M NaCl) for 30 minutes.

Southern Transfer

The blotting apparatus comprised a plastic tray containing about 500-1000ml of 10-20XSSC with a raised platform for the gel. The platform was covered with 3MM Whatmann paper whose ends had been dipped into the 10XSSC to form a wick. The gel was removed from the neutralisation solution and placed on the wick on the platform. Any air bubbles were smoothed out. The edge of the gel was surrounded by Saran-Wrap. The blotting membrane (Hybond-N) and 2 sheets of 3MM paper which have been cut to the size of the gel were wet in 2XSSC and placed on the gel one by one, the Hybond-N being laid first on the gel without any air bubbles. Finally a pack of dry paper towels was laid on the top and covered with a plate and some weight. The transfer was carried out overnight at room temperature. After the transfer was complete, the Hybond-N membrane was dried between two layers of 3MM paper and the DNA was fixed onto the membrane by placing it on the UV box with the DNA side down for 4 minutes. The membrane was then wrapped in cling film (or Saran wrap) and stored at -20^{o}C until hybridisation.

Prehybridisation

Prehybridisation was carried out to prevent non-specific binding of radiolabelled probe to the filter. The filters were soaked in 2XSSC and placed onto a mesh. The mesh was rolled up and placed in a hybridisation bottle containing 30ml 2XSSC. The bottle was laid flat and gently rolled to unfold

the mesh and filter against the inner surface of the bottle. The 2XSSC was poured off and replaced with 15ml of prewarmed prehybridisation solution containing 100μl of denatured salmon sperm DNA (10 mg/ml). The bottle was placed in the hybridisation oven for a minimum of 6 hours at 65°C.

Labelling of the probe

The probe was labelled using the random primed DNA labelling kits (Boehringer Mannheim) according to the method of Feinberg and Vogelstein (1984) and were used for both Southern and Northern Blotting. 25μl of the probe (using Circleprep and Insta-prep) or 2μl of the RT-PCR product subjected to 'Geneclean' and made up to 25μl with distilled water was boiled for 10 minutes and incubated at 37°C for 2 minutes. 2μl each of dATP, dGTP and dTTP, 4μl of the reaction mix, 1.5μl of Klenow enzyme and 50μCi (5μl) of [α-32P] dCTP (Amersham International) were added. The mixture was incubated for 1-2 hours and the reaction was stopped by addition of 2μl of 0.5M EDTA.

Separation of the labelled probe from unincorporated [α-^{32}P] dCTP

This was done using NICK™ columns from Pharmacia Biotech, which were prepacked columns containing Sephadex G-50 DNA grade. The top cap of the column was removed and the excess liquid poured off. The tube was rinsed with 1XSSC and the bottom cap was removed. The column was supported over a suitable receptacle and to the column was added 3ml of 1XSSC which was allowed to drip through. The labelled probe (40μl) was then added to the column and immediately 400μl of 1XSSC was added to it, which was allowed to drip through. A further 400μl of 1XSSC was added to the column and this was collected.

Hybridisation of filters

The collected probe was denatured by boiling for 10 minutes and placed on ice for 2 minutes. The prehybridisation solution from the bottle was poured into a Universal and the probe added to it. The contents of the Universal were mixed gently and transferred back into the bottle. Hybridisation was carried out at 65^{o}C in a Hybaid oven overnight.

Post-hybridisation washing

The hybridisation solution was poured off from the bottles and the filters were washed in 2XSSC/0.1%SDS for 20 minutes at 65^{o}C. The filters were then transferred to a box and washed for a further 20 minutes with shaking in a 65^{o}C water bath with fresh 2XSSC/0.1%SDS for 20 minutes. The filters were monitored using a series 900 minimonitor and if a signal greater than 10cpm was seen the filters were washed at an increased stringency at 65^{o}C. The solutions used were: 1XSSC/0.1%SDS, 0.5XSSC/0.1%SDS, 0.1XSSC/0.1%SDS.

Autoradiography

After washing the filters, they were placed between a layer of Saran Wrap and placed in autoradiography cassettes with intensifying screens and exposed to Kodak Diagnostic AR Imaging film. The cassettes were stored at -70^{o}C and the films were developed after 5-7 days.

2.14 NORTHERN BLOTTING

Northern blotting consists of separation of total cellular RNA by size on an agarose gel. The RNA molecules in the gel are transferred to a nitrocellulose or nylon membrane and detected using an appropriate probe.

In this study, Northern blotting was used to examine neurofibromin expression from 4 neuroblastoma cell lines.

Gel and sample preparation

To make a 20 cmX20 cm gel, 4.5g agarose was boiled in 220ml dH_2O and cooled to 55°C. To this was added 50ml of formaldehyde solution (37%) and 30ml 10XMOPS to make a total volume of 300ml. The gel was then poured in the fume hood.

10-20μg total RNA was used in a volume up to 4.7μl, to which was added 10μl formamide, 3.3μl formaldehyde solution (37%) and 2.0μl 10XMOPS in a total volume of 20μl. This was heated at 55-60°C in a water bath and quenched on ice. To this was added 5μl 6XRNA loading buffer, or 2 μl 10XRNA loading buffer. The gel was run in 1XMOPS buffer at 50-100 volts for 4-8 hours.

Northern transfer

After electrophoresis, the gel was washed in dH_2O for 20 minutes, to get rid of the formaldehyde. The marker and control RNA lanes were cut for staining before the gel was blotted onto Hybond-N membrane, as described in the procedure for Southern transfer.

Northern hybridisation

After the transfer was complete, the membrane and the hybridisation mesh were wet in 2XSSC. The membrane was laid onto the mesh and these were rolled up together. This was then inserted into a Hybaid bottle which contained 30ml 2X SSC. The bottle was then laid flat on a flat surface and gently rolled, such that the membrane and mesh become gently unrolled

against the inner surface of the bottle. The 2XSSC was then poured off and replaced with hybridisation solution. The bottle was then placed into the hybridisation oven and prehybridised for at least 1 hour at 42°C. The prehybridisation solution was then poured off and replaced with the hybridisation solution containing the labelled probe and hybridised overnight at 42°C.

Posthybridisation washing

Posthybridisation washing was carried out using 2XSSC, 0.1% SDS in the bottle at 42°C for 10 minutes. The next wash was carried out at 0.2XSSC, 0.1% SDS in a sandwich box at 65°C for 10-15 minutes with shaking. The membrane was wrapped in Saran wrap and autoradiographed at -70°C in a cassette with an intensifying screen and exposed initially for 2 days.

CHAPTER: 3

RESULTS

CHAPTER 3: RESULTS

3.1 STRATEGY FOR AMPLIFICATION OF THE NF-1 CODING SEQUENCE

25 unrelated patients with NF-1, including both sporadic and inherited cases were analysed for mutations within 78% of the NF-1 coding sequence. NF-1 cDNA obtained by reverse transcription of mRNA was amplified by RT-PCR to yield a set of seven segments (RNA segments I-VII) spanning mRNA positions 39-632 and 1135-4773. Genomic DNA was used for amplification of selected exons of the NF-1 gene, which spanned positions 4773-6756 and 7395-7806 of the coding sequence (Figure 5).

Amplification of exons from genomic DNA

Exons 28-36 and 42-44 were amplified from genomic DNA using flanking intronic primers. The sizes of all the DNA-PCR products ranged from 200-400bp. Exons 35 and 36 were amplified together (due to the small size of exon 36).

Amplification of the NF-1 cDNA by RT-PCR

Part of the NF-1 coding sequence was amplified to yield a set of 7 segments, which were amplified from NF-1 mRNA. The amplified segments ranged from 0.4 - 1.0kb and were amplified either as a single amplification product, or by using nested PCR.

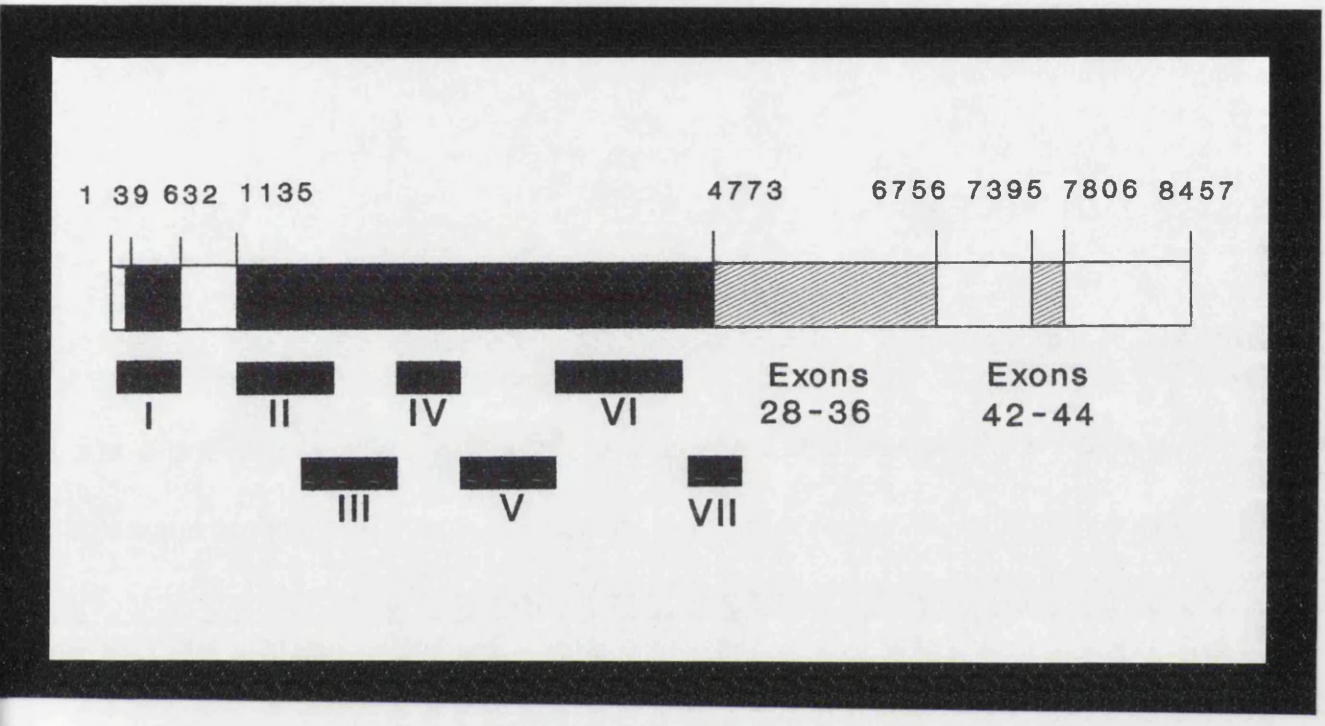
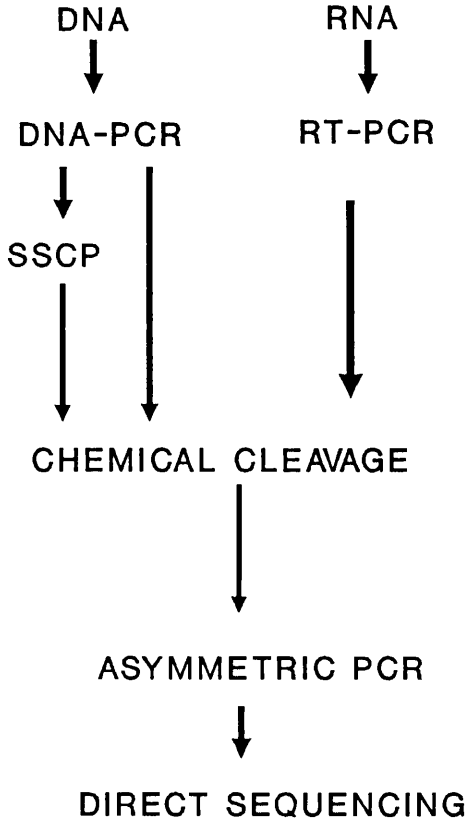


Figure 5a Diagrammatic representation of the NF-1 coding sequence. The filled box spans cDNA positions 39 - 632 (Segment I) and 1135 - 4773 (Segments II-VII) and is amplified in seven segments using the NF-1 mRNA as a template for RT-PCR. The two hatched boxes span exons 28-36 and exons 42-44 respectively and are amplified from genomic DNA.

STRATEGY FOR DETECTION OF MUTATIONS



3.2 STRATEGY FOR DETECTION OF GERMLINE MUTATIONS

After initial amplification by PCR, the products were electrophoresed on 1% agarose gels to check for any abnormal alterations in size. If no alteration was seen, the amplification products of DNA-PCR were analysed using SSCP analysis. The amplification products of RT-PCR and eventually all DNA-PCR products were subjected to chemical mismatch cleavage analysis, to detect and locate the presence of small alterations or point mutations within the NF-1 gene. Any mismatch detected was then subjected to asymmetric PCR and direct sequencing by the Sanger dideoxy chain termination method. Each change found on sequencing was confirmed by sequencing the complementary strand. If a change was found on sequencing the cDNA, amplification of the corresponding exon from genomic DNA followed by direct sequencing was carried out. Every sequence change seen was analysed to see if it created or abolished a restriction enzyme site. Alteration of a restriction site was checked for using the GCG package "MAP" program. Digestion of the PCR product with the appropriate restriction enzyme, was then used as a further confirmation of the change in sequence. For missense and silent mutations, the restriction digests were used to screen normal controls, to find if the mutations were polymorphic. Restriction analysis with the appropriate enzyme was also used to analyse other affected and unaffected members of the family.

In cases where the mutation did not affect a restriction site, assays using PCR directed mutagenesis were designed to confirm the presence of the sequence change. A simple non-radioactive method of detection, such as heteroduplex analysis, was also made use of in screening for already characterised mutations (for segregation analysis & population studies).

Using the above strategy, 17 positive screening results were obtained, of which 13 were characterised. In addition a novel intragenic polymorphism in intron 41 was identified. The mutations and the methods they were detected and further screened by are tabulated below and each mutation is discussed in detail in the following pages.

Table 11:Summary of mutations characterised within the NF-1 gene

Mutation	Location	Method detected by	Further screened by
1721+3A to G	intron 11	RT-PCR, sequencing	PCR-directed mutagenesis
patial del exon 16	exon 16	RT-PCR, sequencing	RT-PCR
3113+1G to A	intron 18	RT-PCR, sequencing	direct sequencing
deletion exon 18	exon 18	RT-PCR, sequencing	RT-PCR
G1166D	exon 21	CCM, sequencing	CCM
S1311S	exon 23	CCM, sequencing	restriction analysis (*Fok* I)
G1404G	exon 24	CCM, sequencing	restriction analysis (*Apo* I)
K1419R	exon 24	CCM, sequencing	restriction analysis (*Mnl* I)
5749+2T to G	intron 30	SSCP, sequencing, RT-PCR	SSCP
6519insG	exon 34	SSCP,sequencing	SSCP
-28 A to G	intron 41	CCM, sequencing	heteroduplex
7485insGG	exon 42	CCM, sequencing	heteroduplex
R2496X	exon 42	CCM, sequencing	CCM

All the sequence numbering is in accord with that provided by the NNFF Mutation Analysis Consortium.

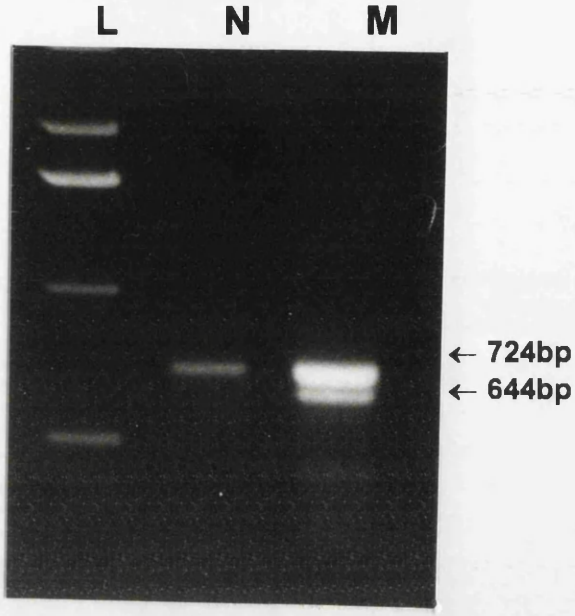
3.3 LARGE ALTERATIONS OF THE NF-1 CODING SEQUENCE

Upon electrophoresis of the amplification products of RT-PCR five abnormal sized products were seen in addition to the normal sized product in segments II, IV, V (in two cases) and VI. The altered sized fragments could either be due to deletions or insertions of the NF-1 coding sequence or errors of splicing. Each alteration was characterised and is now considered separately.

Splice donor mutation in intron 11: 1721+3A to G

The first altered segment was detected upon amplification of segment II in a 30 year old female patient, who is a familial case of NF-1 with an affected son. Upon electrophoresis on a 1% agarose gel, followed by visualisation by ethidium bromide staining, a smaller sized fragment of 644bp was seen in addition to the normal sized fragment of 724bp (Figure 6). Segment II covers exons 8-12a of the NF-1 gene. To confirm the specificity of the altered band, the gel was subjected to Southern blotting using segment II as a probe (Figure 7). The altered fragment lacked 80bp, which corresponds to the size of exon 11. This was confirmed on sequencing the NF-1 mRNA (Figure 8). The loss of 80bp from the NF-1 mRNA could either be due to a deletion or an error of splicing causing the skipping of exon 11. In order to rule out a deletion the patient's genomic DNA was digested with five different restriction enzymes and subjected to Southern blotting using segment II as a probe. No altered fragments or dosage differences were seen. Primers were designed to flank exon 11 (using intronic sequence information provided by the NNFF Mutation Analysis Consortium), the upstream primer in intron 11 and the downstream primer in the coding region of exon 12, to rule out a deletion. This allowed the determination of the size of intron 11 which had not previously been reported/published. After ruling out the presence of a

deletion, this mutation was suspected to be splice site mutation. Using primers immediately flanking exon 11 for amplification and sequencing the splice sites, an A to G transition was seen at the splice donor site of exon 11 (nt 1721+3A to G). This transition resulted in an error of splicing, which caused the skipping of exon 11 from the NF-1 mRNA (Figure 9). This mutation was shown to be present in the affected son, who showed the same pattern of cleavage products as his affected mother, on CCM analysis of his NF-1 mRNA. (Figure 10).



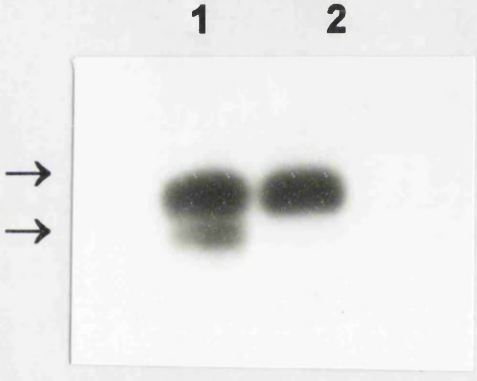


Figure 6

Amplification of NF-1 cDNA (segment II) showing a smaller sized transcript of 644bp in addition to the normal sized fragment of 724bp (marked by arrows) in lane M = mutant. Lane N = normal control and lane L = 1kb ladder (Gibco BRL).

Figure 7

Southern blotting of the product of amplification of RNA segment II, in order to confirm specificity of the altered band, using segment II as a probe. Both the normal and altered bands in lane 1 are marked by arrows, lane 2 is the control.

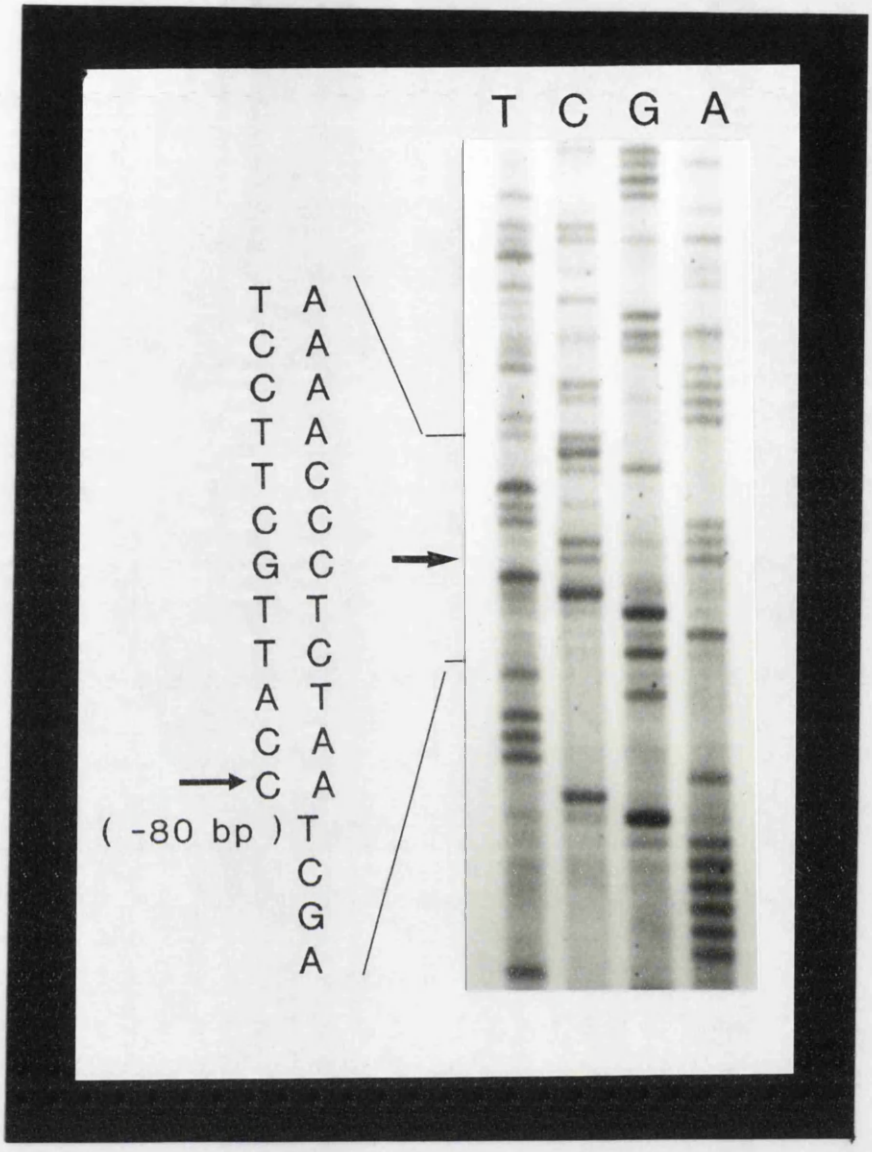


Figure 8

mRNA sequence analysis surrounding exon 11. Beyond the point marked by an arrow, a parallel sequence is seen along with the normal sequence. After analysis, this sequence was shown to correspond to the sequence from the normal allele and the mutant allele which lacks exon 11 running together.

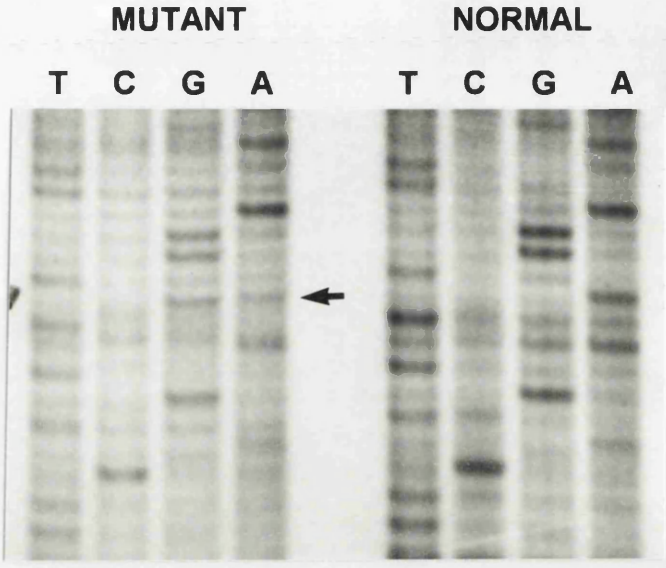


Figure 9

Sequence of the mutant DNA across the intron-exon boundary, showing the intron 11 splice donor site mutation, A to G at position +3, marked by an arrow.

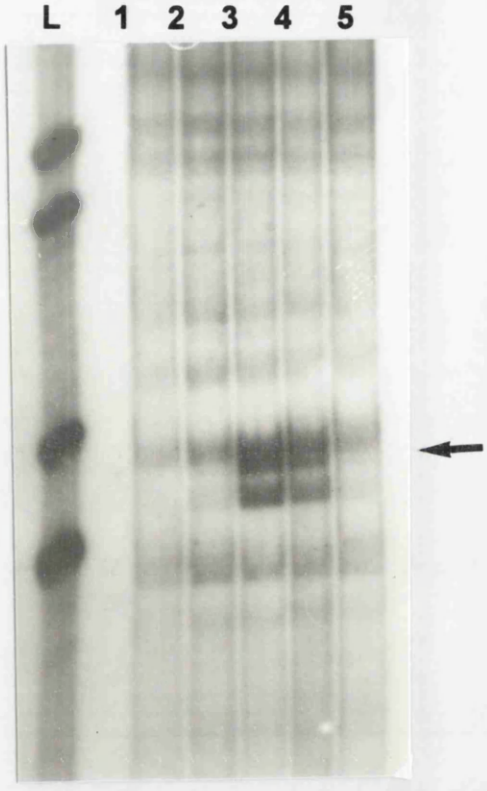


Figure 10

Deletion mapping using chemical mismatch cleavage analysis of RNA segment II. The proband and her affected son show cleavage products in lanes 3 and 4. L indicates the 1kb ladder. The remaining lanes show CCM analysis in controls.

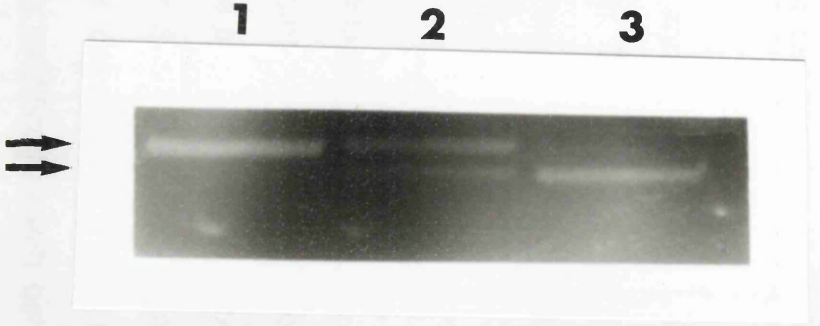
In cases where the mutations did not affect a restriction site and where it was necessary to screen normal controls and the remaining members of the family, PCR directed mutagenesis was used to introduce part of the recognition sequence of the enzyme in the primer used for amplification (Li et al 1992). In this study this strategy was used in the case of the above splice donor mutation in intron 11, (nt 1721 +3A to G). The presence of a normal adenine (A) nucleotide at the site of transition (1721+3A) allows cutting on digestion with the restriction enzyme *Acc* I. The appearance of any other nucleotide destroys the recognition site of this restriction enzyme. A PCR primer covering the sequence immediately preceding the site of transition was used in combination with an upstream intron based primer for amplification. By designing the downstream primer to have a non-template C nucleotide at its 3' terminus, the PCR product derived from the normal allele had a part of its sequence to be **GTA|TAC**, in which the **T** comes from the 3' base in the downstream primer and the **A** comes from the adjacent base in the template (Figure 11). This is the recognition sequence for the enzyme *Acc* I. Conversion of A to G causes loss of the restriction site for *Acc* I and prevents cutting in the mutant allele, whose sequence is **GTGTAC** instead of **GTATAC.**

Figure 11: The schematic representation of the mutant and normal PCR products is presented and the *Acc* I restriction site and its loss due to the mutation is indicated. Visualisation of *Acc* I digested PCR products by electrophoresis on a 3% MetaPhor™ (FMC) gel is shown in the lower half of the figure. Lane 1: Undigested control, lane 2: PCR product from the patient who is heterozygous for the mutation and shows the digested normal allele (185bp) and the undigested mutant allele (161bp). lane 3: Normal control showing complete digestion of the PCR product.

```
                    EXON 11       INTRON 11
NORMAL SEQUENCE 5'  TAG    GTA TAT GTA CTT  3'

MUTANT SEQUENCE 5'  TAG    GTG TAT GTA CTT  3'
PRIMER                      3' TAC GTA CTT TAA TTT TTA AAT CTT  5'


NORMAL PCR PRODUCT  TAG GTA TAC GTA CTT
                          AccI
MUTANT PCR PRODUCT  TAG GTG TAC GTA CTT
                      Loss of AccI
```



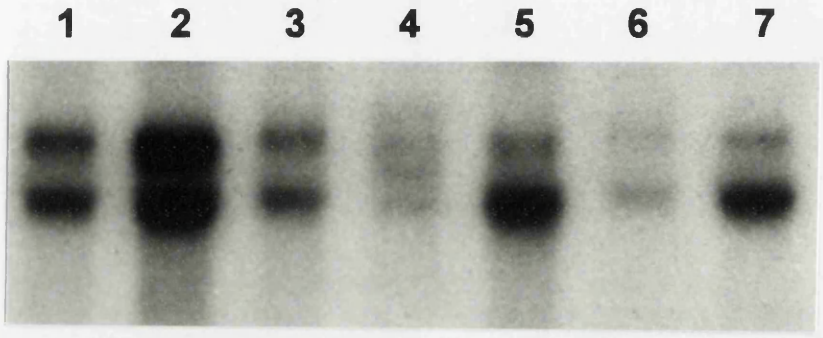
Therefore on digestion of the amplification product of the patient, the presence of the digested normal allele (161bp) and the undigested mutant allele (185bp) is seen, while the normal control shows complete digestion of the PCR product.

This analysis was used to confirm the mutation and to screen 30 controls, none of which were shown to have this mutation. Thus segregation of the mutation with disease and absence of this mutation in controls, thus shows that mutation is responsible for disease in this family.

Splice donor mutation in intron 30: 5749 +2T to G

The second splice site mutation was detected on analysis of exon 30 and the flanking sequences, in a 29 year old male patient, who is a sporadic case of NF-1. The mutation was seen as a shift on SSCP analysis of exon 30 (Figure 12), which was confirmed on chemical cleavage analysis. CCM analysis showed an altered cleavage product on modification with osmium tetroxide (Figure 13). DNA sequencing of the sequence flanking exon 30 revealed a T to G substitution at position +2 (nt 5749 +2T to G) at the invariant GT dinucleotide of the splice donor site of intron 30 (Figure 14). After confirmation of the mutation by sequencing the opposite DNA strand, analysis of the corresponding mRNA region by RT-PCR of transcripts derived from peripheral lymphocytes was carried out. This amplification showed the presence of a smaller sized transcript of 504bp in addition to the normal sized PCR fragment of 707bp. This corresponds to the skipping of exon 30, which is approximately 203bp in size (Figure 15). Sequence analysis of the amplified cDNA fragments confirmed that the transcript lacked exon 30. This patient is a sporadic case of NF-1 and analysis of his unaffected parents and brother was carried out and none of them had this mutation. In addition 30

normal controls were screened and none of them were found to have this mutation.



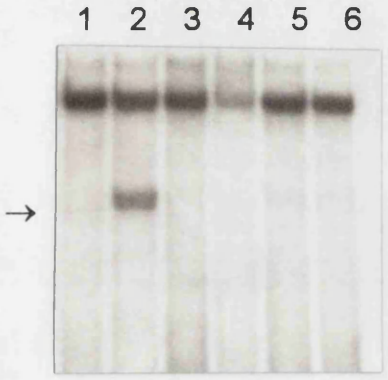


Figure 12

SSCP analysis of exon 30 showing an altered pattern in lane 4, as compared with lanes 1-7.

Figure 13

Chemical mismatch cleavage analysis of exon 30, showing the presence of a cleavage product marked by an arrow in lane 2, on modification with osmium tetroxide. In normal controls only the uncleaved probe is seen.

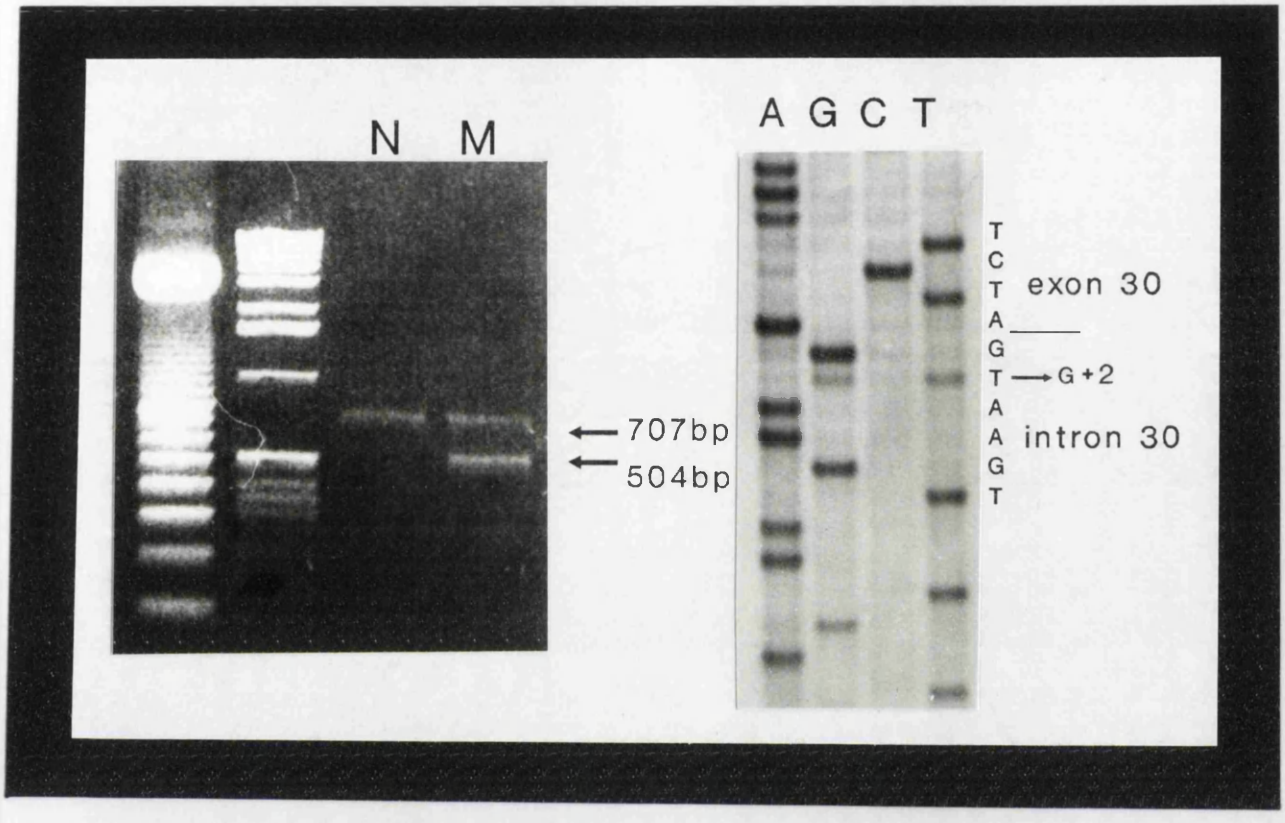


Figure 14 and 15

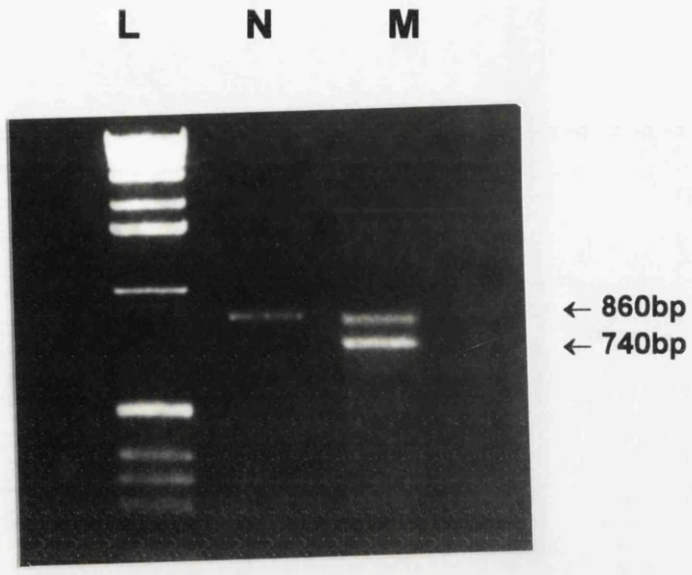
Amplification of a segment of the NF-1 cDNA encompassing exon 30. A smaller transcript (504bp) is seen in addition to the normal sized transcript (707bp) in lane M (mutant). Lane N = normal control showing only the normal sized product and lane L = 1kb ladder (Gibco BRL).

Sequence of the mutant DNA across the intron-exon boundary, showing the intron 30 splice donor site mutation, T to G at position +2 marked by an arrow. The intron-exon boundary is indicated.

Splice donor mutation in intron 18: 3113+1G to A

The third splice site mutation was identified when amplification of segment V showed the presence of an altered fragment of approximately 750bp in addition to the normal sized fragment of 860bp (Figure 16a, 16b). This mutation was seen in a 41 year old female patient, who has an affected son. Chemical mismatch cleavage analysis was carried out to confirm the boundaries of the deletion (Figure 17a). Upon analysis, the altered fragment appeared to be approximately 120bp shorter than the normal sized fragment. To rule out the presence of a deletion, Southern blotting of the patient and control DNA was carried out using five different restriction enzymes using segment V as a probe. The results did not show the presence of a deletion, which would be seen as a junction fragment or a dosage effect. Exonic primers were designed such that a forward primer was in exon 17 (17F) and the reverse primer in exon 19 (19R). This product of amplification would span intron 17, exon 18 and intron 18, thus confirming or ruling out a deletion if any. Analysis of this product of amplification showed no deletion. Therefore the splice sites of intron 17 and 18 were sequenced. On sequencing, a G to A transition was seen at the splice donor site of intron 18 (nt 3113+1G to A) (Figure 17b). The splice mutation was characterised using primers immediately flanking exon 18 (18F, 18R), as well as the primers in exon 17 and 19 (17F, 19R).

Using a combination of primers 17F, 19R, 18F and 18R also allowed the determination of the size of introns 17 and 18 which were shown to be 0.2kb and 0.7kb respectively.



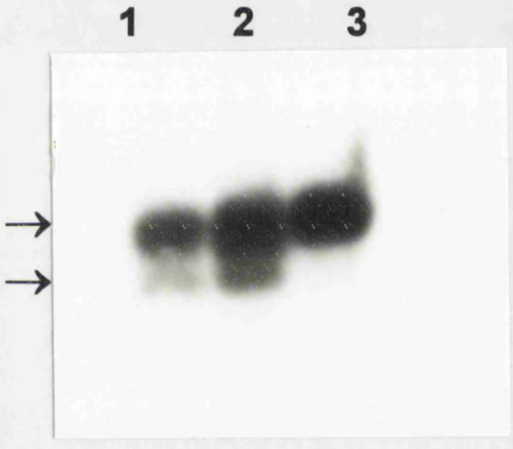


Figure 16a

Amplification of NF-1 cDNA (segment V) showing a smaller sized transcript of 750bp in addition to the normal sized fragment of 860bp (marked by arrows) in lane M = mutant. Lane N = normal control and lane L = 1kb ladder (Gibco BRL).

Figure 16b

Southern blotting of the product of amplification of RNA segment V, using segment V as a probe, in order to confirm specificity of the altered band. Both the normal and altered bands are marked by an arrow in the patient and her affected son, in lanes 1 and 2, in comparison with only the normal band in lane 3 (control).

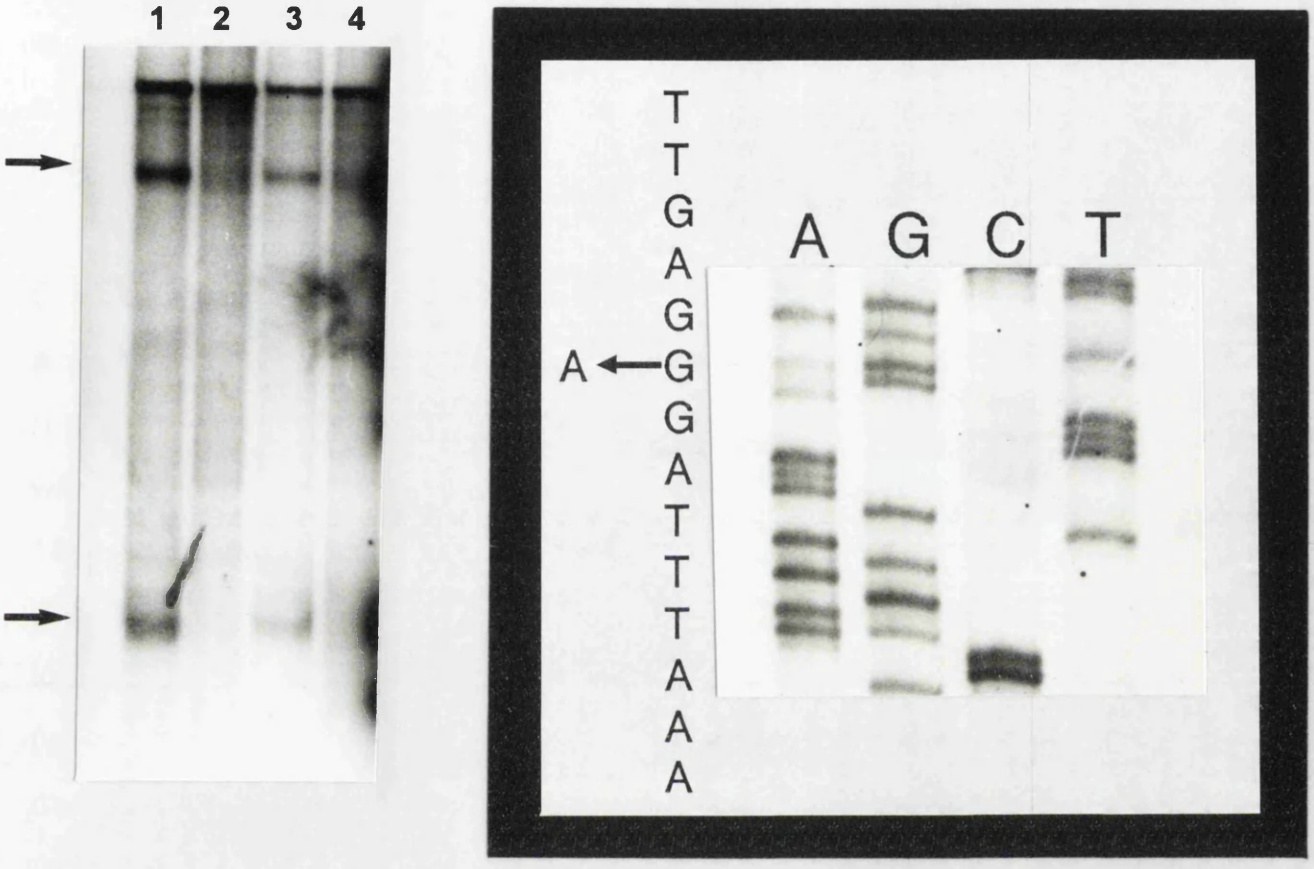


Figure 17a

Deletion mapping using chemical cleavage analysis in segment V in the patient in lane 1 and her affected son in lane 3. Lanes 2 and 4 are the controls showing only the presence of the uncleaved product. Lanes 1 and 3, show the presence of the two cleavage products marked by arrows.

Figure 17b

Sequence of the mutant DNA across the intron-exon boundary, showing the intron 18 splice donor site mutation, G to A at position +1 marked by an arrow.

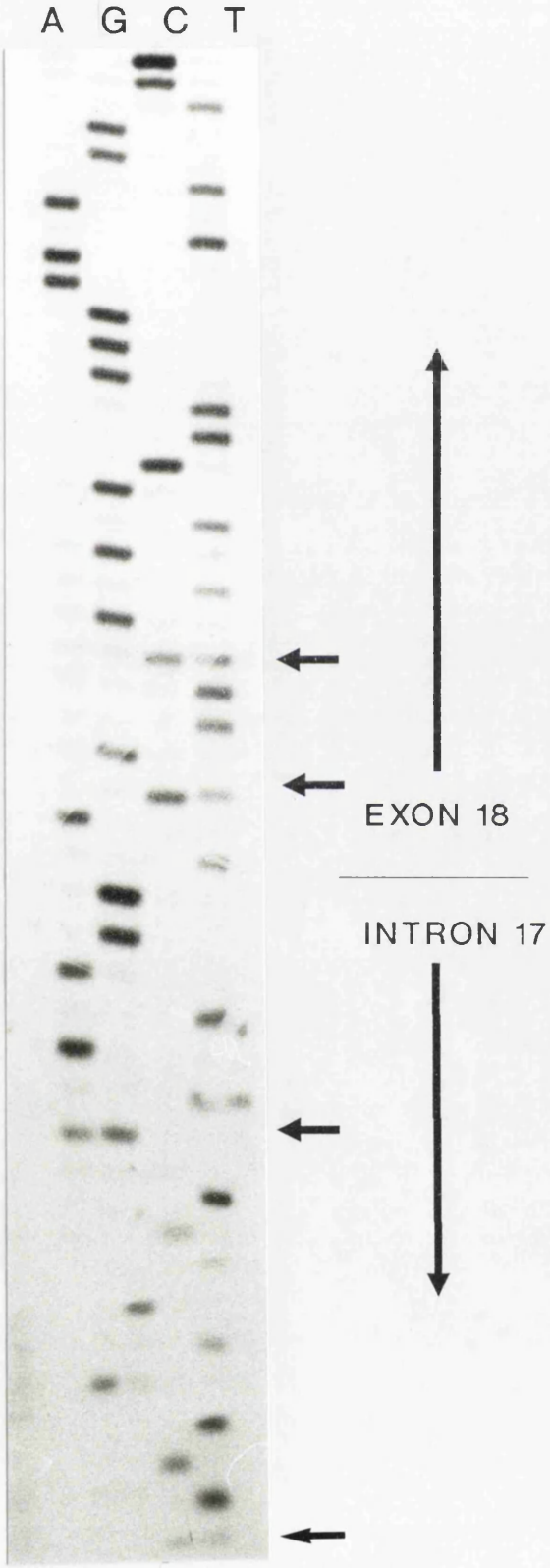
This mutation caused an error of splicing and resulted in the skipping of exon 18 from the NF-1 transcript. The skipping did not alter the translational reading frame, but resulted in the exclusion of 41 amino acids from the predicted protein product. This mutation was seen in a familial case of NF-1 and was shown in her affected son. 10 other controls were also screened by sequencing their splice sites and none of them were shown to have this mutation.

An interesting feature was seen on amplification of the region using the primers in exon 17 and exon 19 (17F, 19R). Upon sequencing this region, a number of sequence variants were identfied in 10 individuals in the coding region of exon 18 and in intron 17 (Figure 19). The changes were not seen when the same region was amplified using primers immediately flanking exon 18. These changes, thus did not appear to be solely amplified from the NF-1 locus. Using primers 17F and 19R, amplification occurred from the NF-1 locus as well as a homologous locus. But by using primers 18F and 18R, (which were used for amplification and characterisation of the splice mutation), the altered sequence was not seen. Therefore, this means that the regions on which 18F and 18R prime were not conducive to PCR amplification on the homologous region. This appeared to be an as yet unreported sequence which was homologous to exon 18 of the NF-1 gene.

Figure 18

Partial sequence analysis of exon 18, amplified using the primers 17F and 19R which are in exons 17 and 19 respectively. Four of the seven changes detected in the coding sequence and intronic sequence are shown in this figure, the remaining lie at the other end of the exon. The positions marked by arrows show the substitutions. These were not seen on amplification of the same region using primers immediately flanking exon 18 (18F, 18R), indicating dual amplification from the NF-1 gene and a homologous locus.

NF-1 HOMOLOGOUS SEQUENCE
(INTRON 17/EXON 18)



Partial loss of exon 16 from the NF-1 mRNA

Amplification of segment IV in a 12 year old male patient, who was a sporadic case of NF-1 showed an altered sized band in addition to the normal sized amplification product. The altered band appeared to be shorter by approximately 220-230bp (Figure 19). Direct sequencing of the mRNA revealed a deletion of 229bp, which began exactly at the first base of exon 16 and an altered sequence ran in parallel with the normal sequence beyond this point (figure 20).

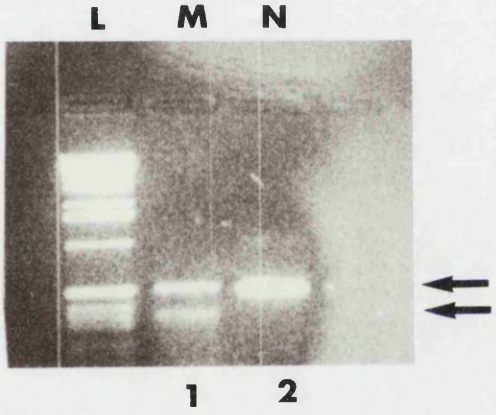


Figure 19

Agarose gel electrophoresis showing an altered band shorter by 229bp in addition to the normal sized product in lane 1 (mutant). The control shows only the normal sized product of approximately 500bp in lane 2. L indicates the 1kb ladder.

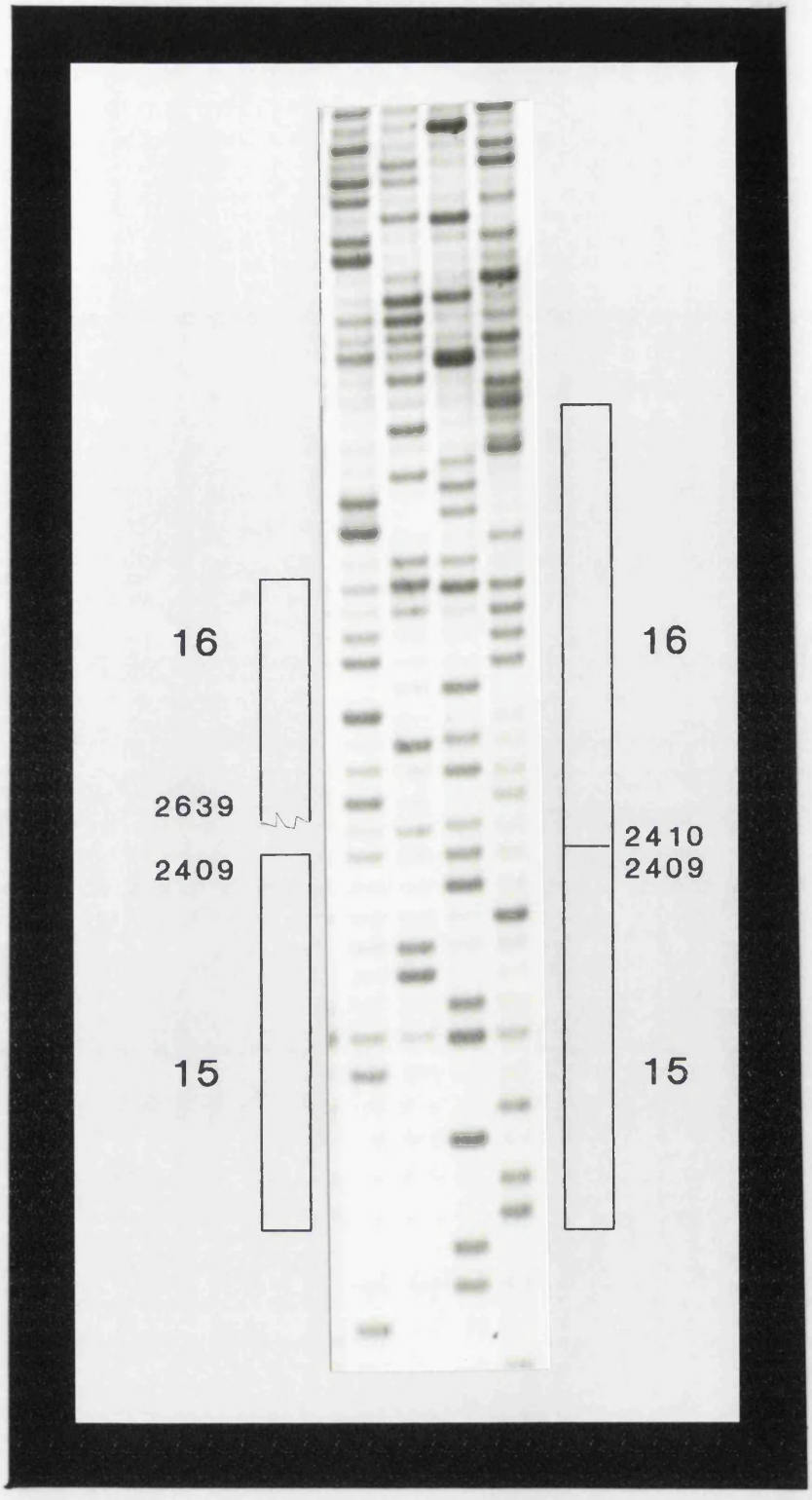


Figure 20

Partial sequence analysis of the NF-1 mRNA surrounding the junctions of exon 15 and 16 in the mutant (Segment IV). Up to the junction between exons 15 and 16 (2409/2410), normal sequence is seen and beyond this point, a complex sequence is seen. This corresponds to a sequence 229bp away, indicating that the mutant allele running in parallel with the normal allele has 229 nucleotides missing.

Insertion in exon 23

The second large alteration was seen on amplification of a part of segment VI, using internal primers in a case of segmental neurofibromatosis. Upon amplification, a larger fragment of approximately 60bp was seen in addition to the normal sized fragment. This region spans the alternatively spliced form of the NF-1 gene, called NF-GRD II, which has a 63bp insertion. This patient therefore is expressing equal amounts of the two neurofibromin isoforms NF-GRD-I and II (with the 63 base insertion), in comparison with the normal control who expresses only the NF-GRD I, which is the 390bp product.

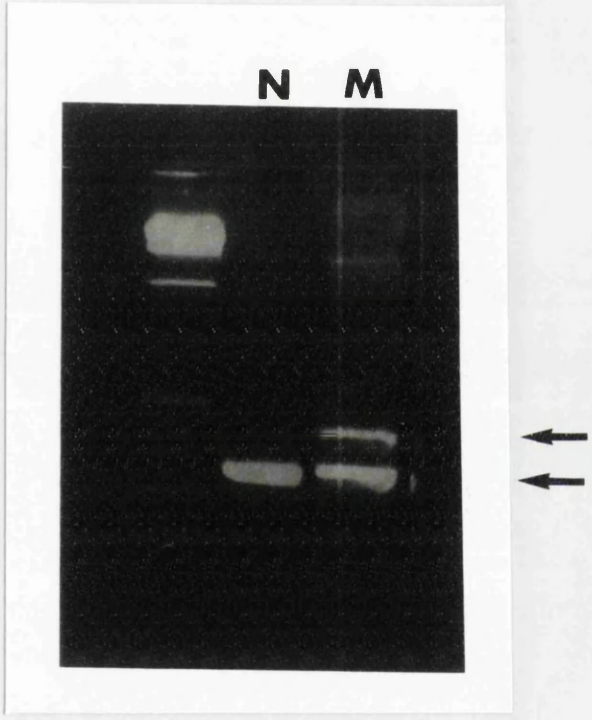


Figure 21

Agarose gel electrophoresis of the RT-PCR surrounding exon 23 (segment VI). The presence of a larger product of 450bp is seen in addition to the normal sized product of 390bp in lane M (Mutant), in comparison with only the normal sized product in lane N (Normal wild-type). The unmarked first lane is the 1kb ladder.

Loss of exon 18 from the NF-1 mRNA

Another interesting alteration was seen in a familial case who showed an altered fragment in the region surrounding exon 18 (segment V) on RT-PCR, which corresponded to shortening of approximately 120bp of the NF-1 mRNA (Figure 22). On sequencing of the same region this was seen to be due to loss of exon 18 (123bp) from the NF-1 mRNA (Figure 23).

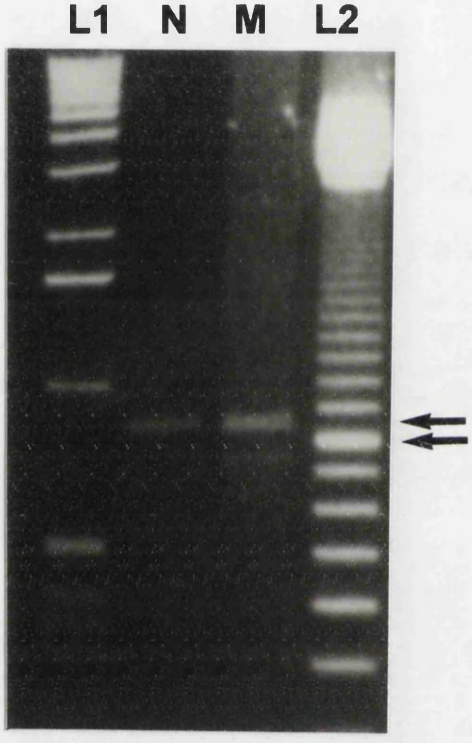


Figure 22

Agarose gel electrophoresis of amplification product of segment V, showing the presence of an altered product smaller than the normal sized product by about 120bp in the mutant lane (marked by arrows). Lanes L1 and L2 are the 1kb ladder and 100bp ladder respectively.

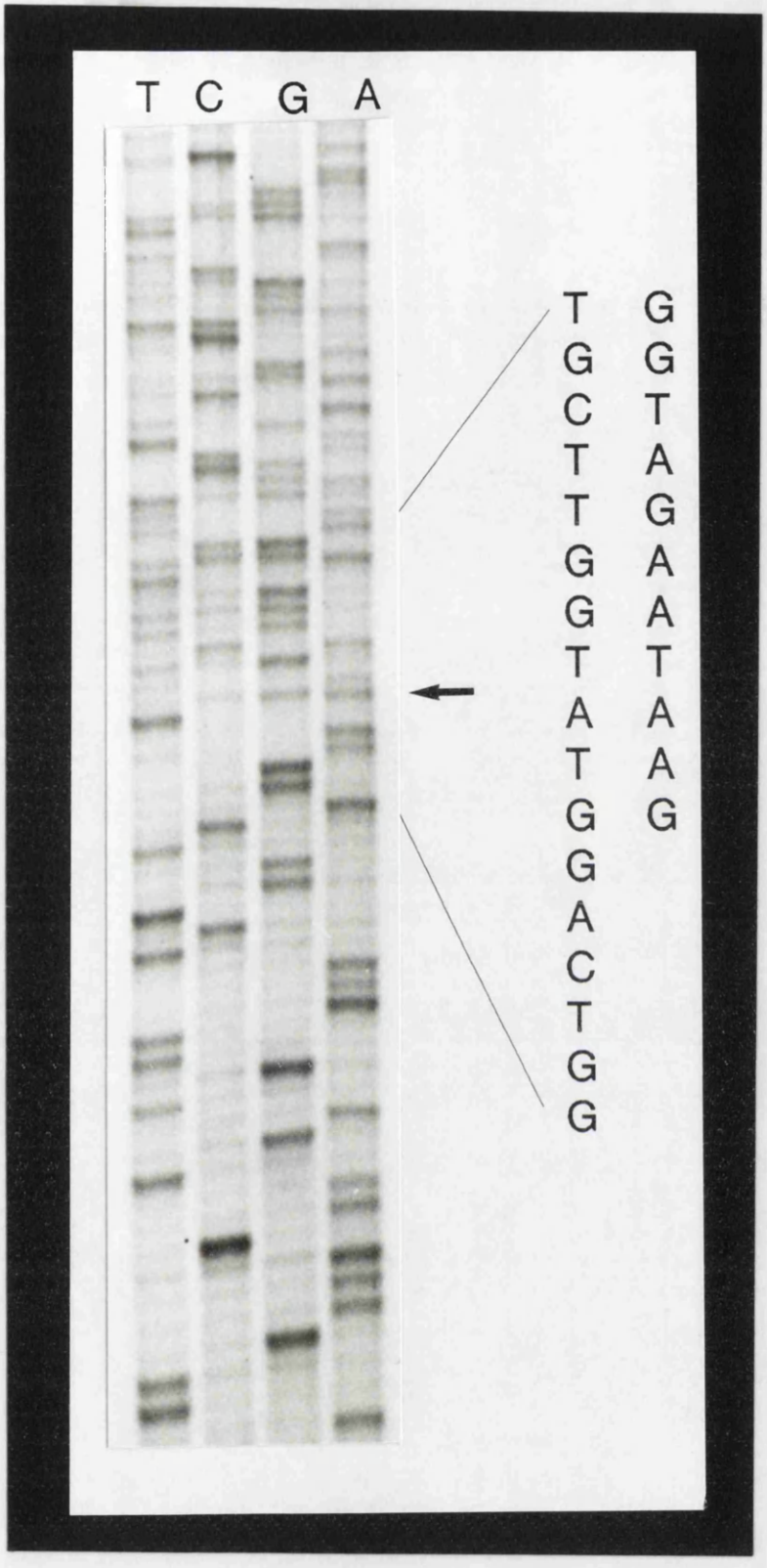
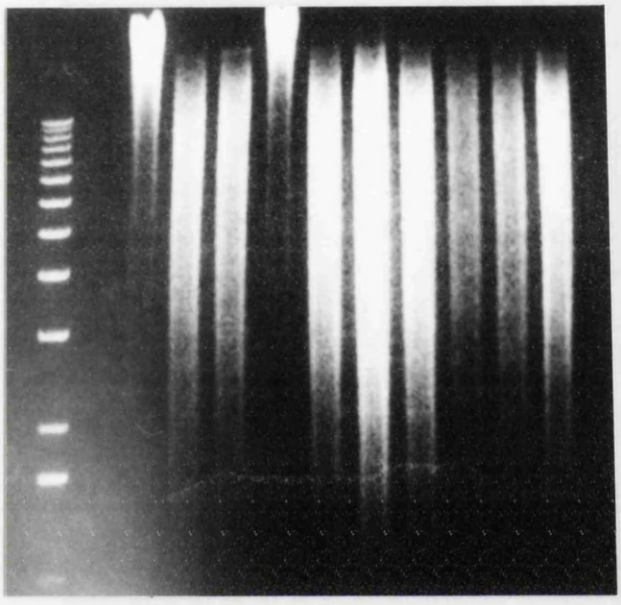


Figure 23

Partial mRNA sequence analysis surrounding exon 18. The normal sequence is seen up to the point marked by an arrow (beginning of exon 18), beyond which a complex sequence is seen. On analysis, the parallel sequence is due to a mutant allele which has skipped exon 18 and the normal allele with exon 18, which run together.

Large alterations within the NF-1 gene were screened for using Southern blotting with restriction enzyme *EcoR* I and probe p5 (ATCC). Analysis showed no dosage effect or junction fragment indicative of a large alteration using this probe-enzyme combination. However detection of large alterations using other probe-enzyme combinations have not been carried out.

24a



24b

Figure 24a

Agarose gel electrophoresis of *Eco* RI restriction digested DNA. The smear indicates digestion of the DNA into various sized fragments. Lanes 1 and 4 show incomplete digestion.

Figure 24b

Southern analysis using probe P5.

3.4 SMALL ALTERATIONS WITHIN THE NF-1 GENE

3.4.1 Characterisation of single base substitutions and small rearrangements in products of RT-PCR, using chemical mismatch cleavage (CCM) and direct sequencing.

Amplification of the RNA segments in the remaining patients showed PCR products of the expected size after electrophoresis on 1% agarose gels. However chemical cleavage analysis revealed the presence of cleavage products in these segments. The mutations identified on analysis of these RNA segments I-VII are presented below.

Missense mutation in exon 21: G1166D

A cleavage product was seen on modification with hydroxylamine in RNA segment V (Figure 25a) in an 11 year old male patient, who was a familial case of NF-1. This mutation was localised to exon 21, which is the first exon of the GAP related domain. Direct sequencing revealed a G to A transition at nucleotide position 3497 (Figure 25b), which is the first base of exon 21. This mutation caused the conversion of Glycine to Aspartate at amino acid position 1166 (G1166D). The mutation was seen in the proband and in his mother who also has NF-1.

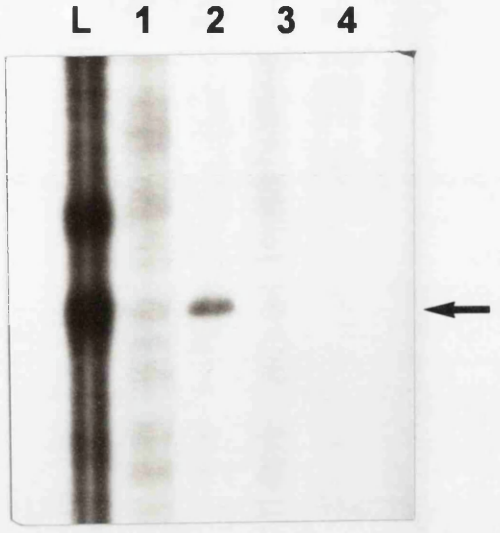
Silent mutation in exon 24: G1404G

Two different sized cleavage products on modification with both hydroxylamine and osmium tetroxide were seen in Segment VI in a 19 year old male patient who is a sporadic case of NF-1 (Figure 26a, Figure 26b). These were localised to exon 24 of the NF-1 gene. Direct sequencing demonstrated the hydroxylamine modification to be to be a G to A transition

in nucleotide 4212 (Figure 27) and the mutation creates an *Apo* I restriction site. This transition maintains a Glycine at position 1404 (G1404G) and is thus a silent mutation. Digestion of the 400bp PCR product by *Apo* I results in two bands of 220 and 180bp from the mutated allele, in addition to the undigested 400bp band from the normal allele in the patient (Figure 29a). Digestion of 30 controls was carried out to find out whether this mutation was polymorphic, but the mutation was not seen in any other control. Thus this is a rare sequence variant, within the coding region of the NF-1 gene.

Missense mutation in exon 24: K1419R

The cleavage product seen on modification with osmium tetroxide was due to a missense mutation. Direct sequencing demonstrated this mutation to be a A to G transition at nucleotide 4256 (Figure 28), causing the conversion of Lysine to Arginine at position 1419. This mutation creates an *Mnl* I restriction site (Figure 29b). The product of amplification contains another normally present *Mnl* I site which serves as an internal control. Digestion of the mutant PCR product clearly shows a different pattern from normal, which corresponds to the bands created due to the already present and the newly created *Mnl* I enzyme site.



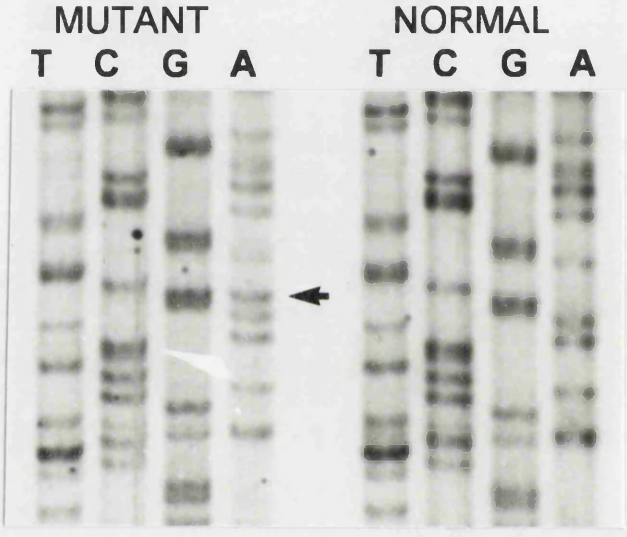


Figure 25a

Chemical cleavage analysis of RNA segment V on modification with hydroxylamine. The cleavage product due to the missense mutation G1166D, caused by the G3497 to A substitution is seen in lane 2, indicated by an arrow.

Figure 25b

Partial mRNA sequence analysis surrounding exon 21 in wild type (normal) and mutant, showing the G3497 to A substitution in exon 21, indicated by an arrow.

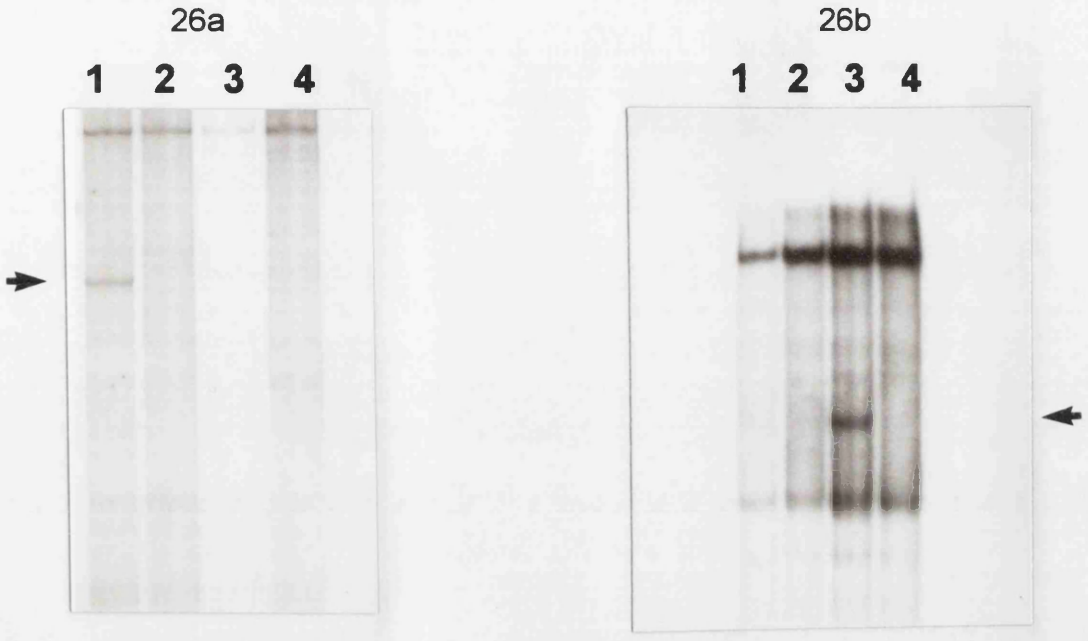


Figure 26a

Chemical cleavage analysis of RNA segment VI, on modification with hydroxylamine. The cleavage product due to the silent mutation G1404G, caused by the G4212 to A substitution, is indicated by an arrow.

Figure 26b

Modification of segment VI with osmium tetroxide showing a cleavage product due to the missense mutation K1419R, caused by an A4256 to G substitution and is indicated by an arrow.

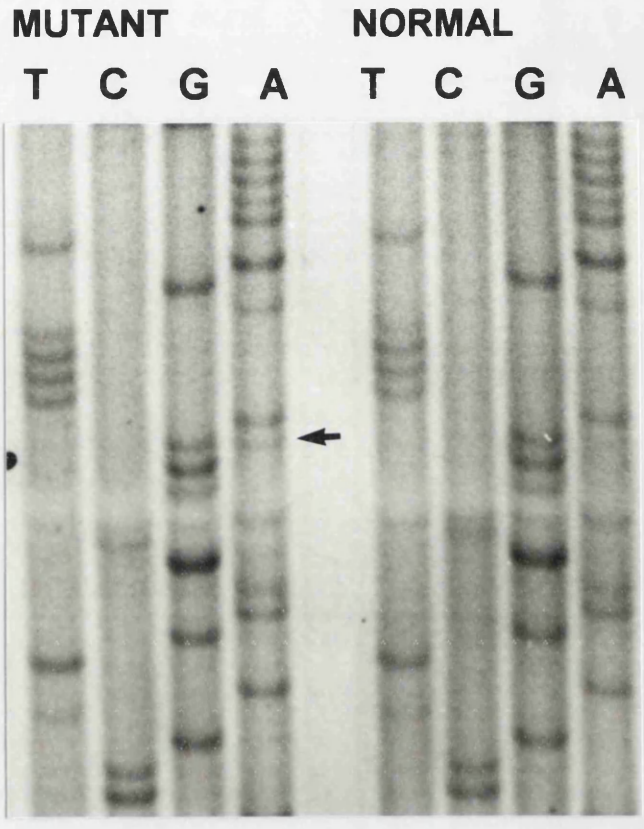


Figure 27

Partial mRNA sequence analysis of segment VI (exon 24) in wild type and mutant. The G4212 to A substitution, due to the silent mutation G1404G is indicated by an arrow.

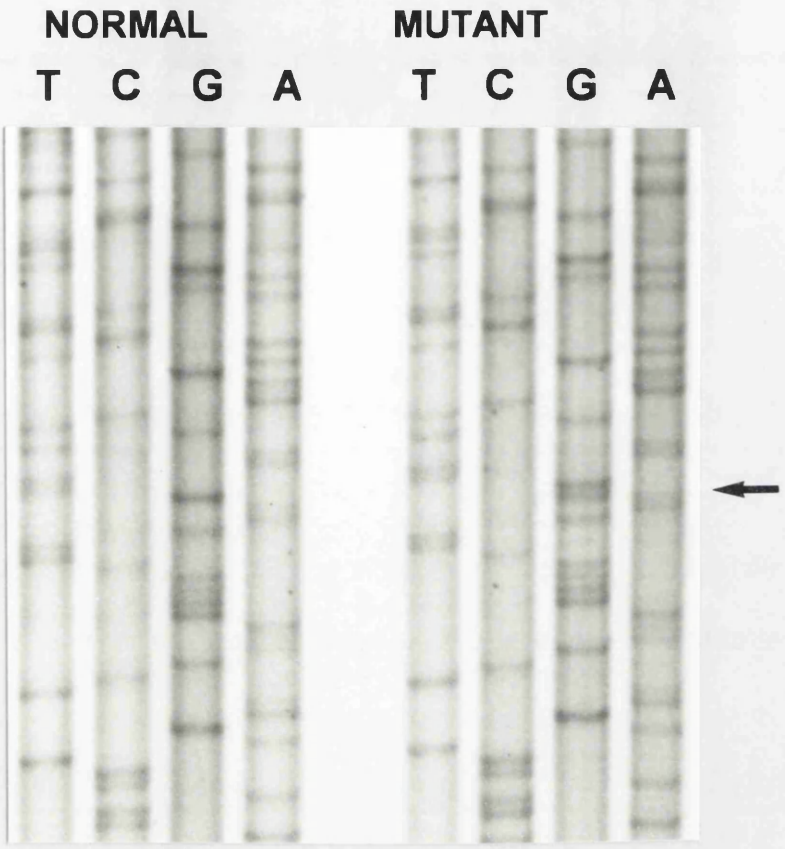


Figure 28

Partial mRNA sequence analysis of Segment VI (exon 24) in wild type and mutant, showing the A4256 to G substitution due to the missense mutation K1419R, indicated by an arrow.

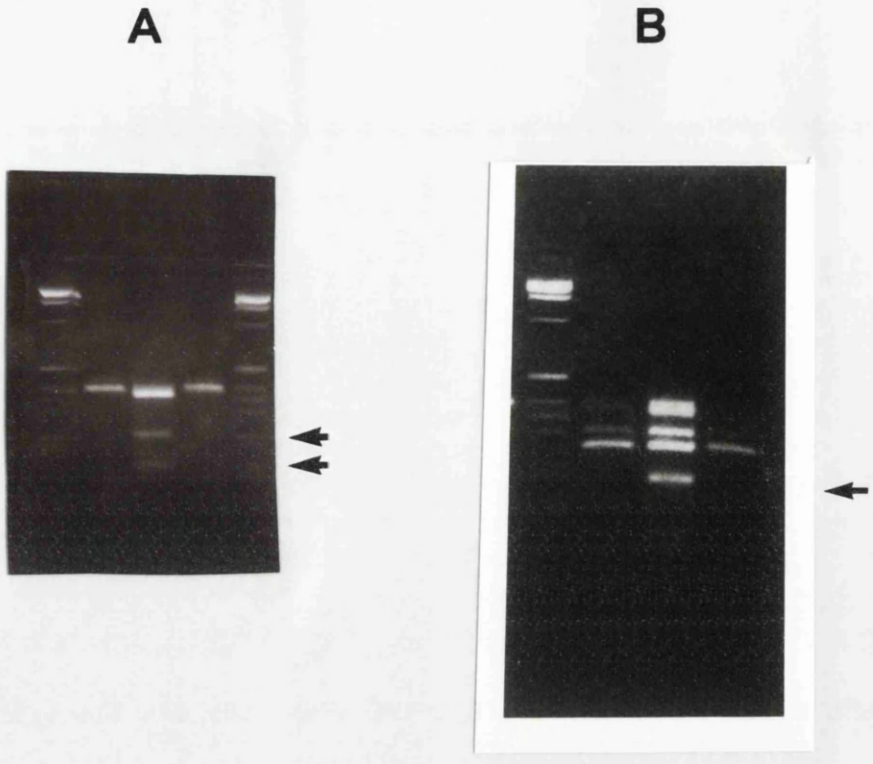


Figure 29a

Restriction analysis with the enzyme *Apo* I, the site being created due to the mutation G1404G in exon 24. Agarose gel electrophoresis showing digestion of the mutant allele into two smaller products of 240 and 160bp, indicated by arrows.

Figure 29b

Restriction analysis with the enzyme *Mnl* I. An additional *Mnl* I site is created due to the missense mutation K1419R in exon 24. Agarose gel electrophoresis showing digestion of the product in normal controls (N) due to a normally present *Mnl* I site. Lane M in the mutant shows a different pattern in the digested product, due to creation of an additional *Mnl* I site at 200bp due to the mutation and is indicated by an arrow.

Silent mutation in exon 23: S1311S

Analysis of segment VI in a 30 year old female who is a familial case of NF-1, showed a cleavage product on treatment with hydroxylamine (Figure 30). This mutation is a C to T substitution at nucleotide 3933, maintaining a Serine at amino acid position 1311 (S1311S). This mutation causes loss of a *Fok* I restriction site and restriction analysis was used to screen 30 controls, none of whom were shown to have this mutation.

Upon direct sequencing, only the altered allele (T) was seen instead of both C from the normal allele and T from the mutant allele at the same level (Figure 31a).

Similarly, only the presence of the allele with the silent mutation (undigested, due to loss of a the *Fok* I site) was seen on digestion of the product of amplification (Figure 31b). Partial or incomplete digestion was ruled out as the same result was seen on three different attempts at digestion.

In order to rule out a deletion, Southern analysis was performed on the patients genomic DNA, using segment VI as a probe and 5 different restriction enzymes. However the results from Southern analysis were inconclusive and a deletion could not be ruled out. Amplification of this region using PCR did not show an altered product, probably because the deletion boundaries extended outside the amplified PCR fragment.

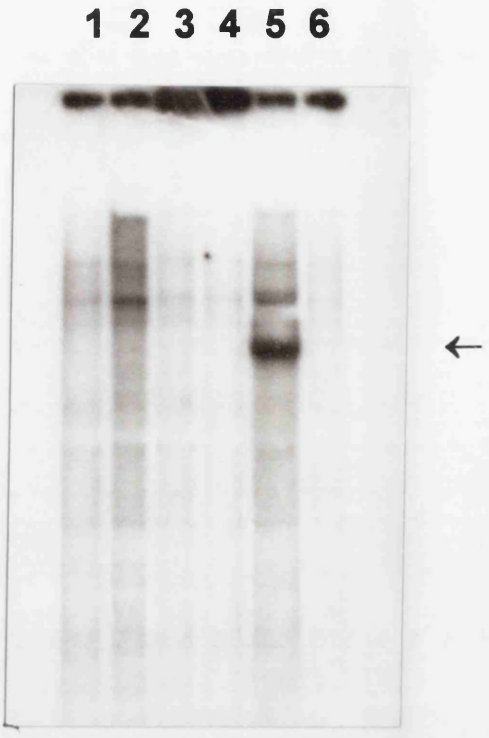


Figure 30

Chemical cleavage analysis of RNA segment VI on modification with hydroxylamine. The cleavage product due to the silent mutation S1311S, caused by a C3933 to T substitution is seen in lane 5 and is indicated by an arrow.

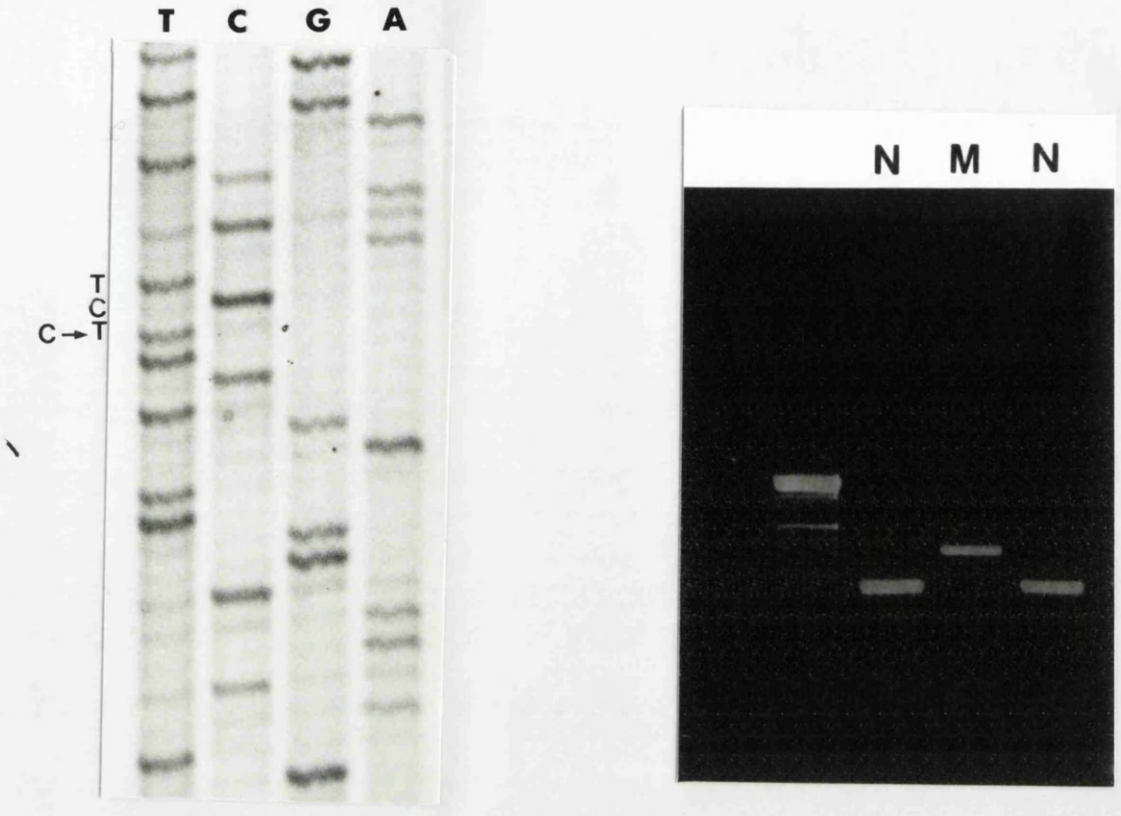


Figure 31a and b

a)Direct sequence analysis showing only the mutant allele T at the point of alteration marked by an arrow, instead of a C and T at the same level.

b)Restriction enzyme analysis using enzyme *Fok* I. Agarose gel electrophoresis of digested PCR products. Lane M (mutant) shows the undigested mutant product (S1311S abolishes a *Fok* I restriction site), while the other lanes N (normal wild type) show complete digestion in normal controls. The unmarked lane represents the 1kb ladder. It is seen that in lane M, only the undigested mutant allele is seen, when there should have been a digested band from the normal allele and an undigested band from the mutant allele.

Positive chemical cleavage results yet to be characterised by direct sequencing

The remaining positive results of chemical cleavage analysis were seen in the following RNA segments: (These have not yet been characterised by direct sequencing).

In Segment I, a cleavage product of 500bp in addition to the normal sized fragment of 600bp (Figure 32), was seen in a 22 year old female patient, who is a sporadic case of NF-1.

Segment I also showed two cleavage products in a 12 year old female patient who is a sporadic case of NF-1, at 150bp and 450bp respectively (Figure 33). These mutations were mapped to lie in either exon 2 or exon 4 of the NF-1 gene.

A cleavage product was seen on modification with hydroxylamine in a 3 year old female sporadic patient. The cleavage product appeared to be 220bp in addition to the normal sized fragment of 500bp (Figure 34). This was seen in segment IV.

A cleavage product of 180bp was seen in addition to the normal sized fragment of 500bp (Figure 35) in a 36 year old female patient of NF-1, with a family history of NF-1 and cancer. The cleavage product was seen on modification with osmium tetroxide, in segment VII spanning exons 26 and 27a and 27b, which covers part of the GAP related domain of the NF-1 gene.

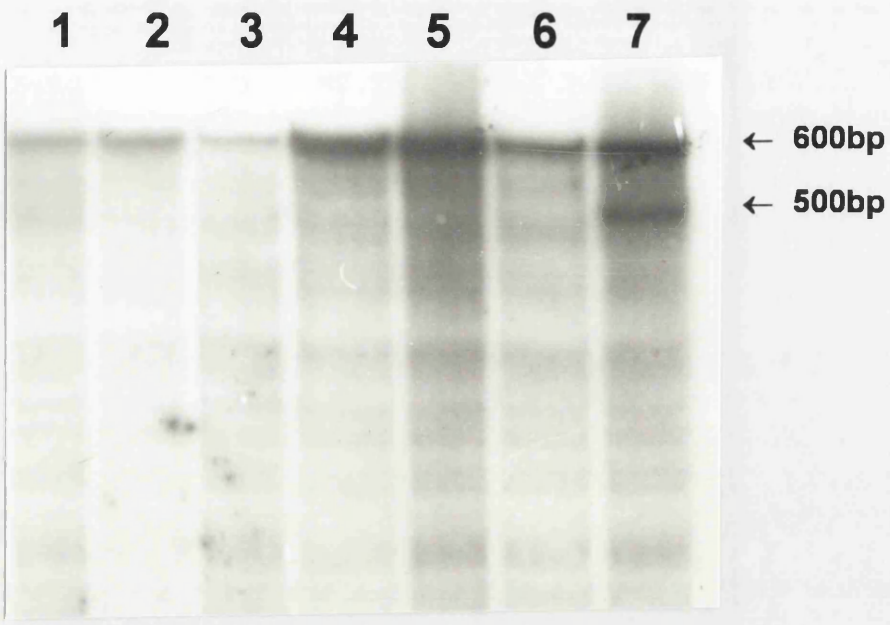


Figure 32

Chemical cleavage analysis of segment I, showing a cleavage product of 500bp in addition to the normal sized fragment of 600bp, on modification with osmium tetroxide, in lane 7 marked by an arrow.

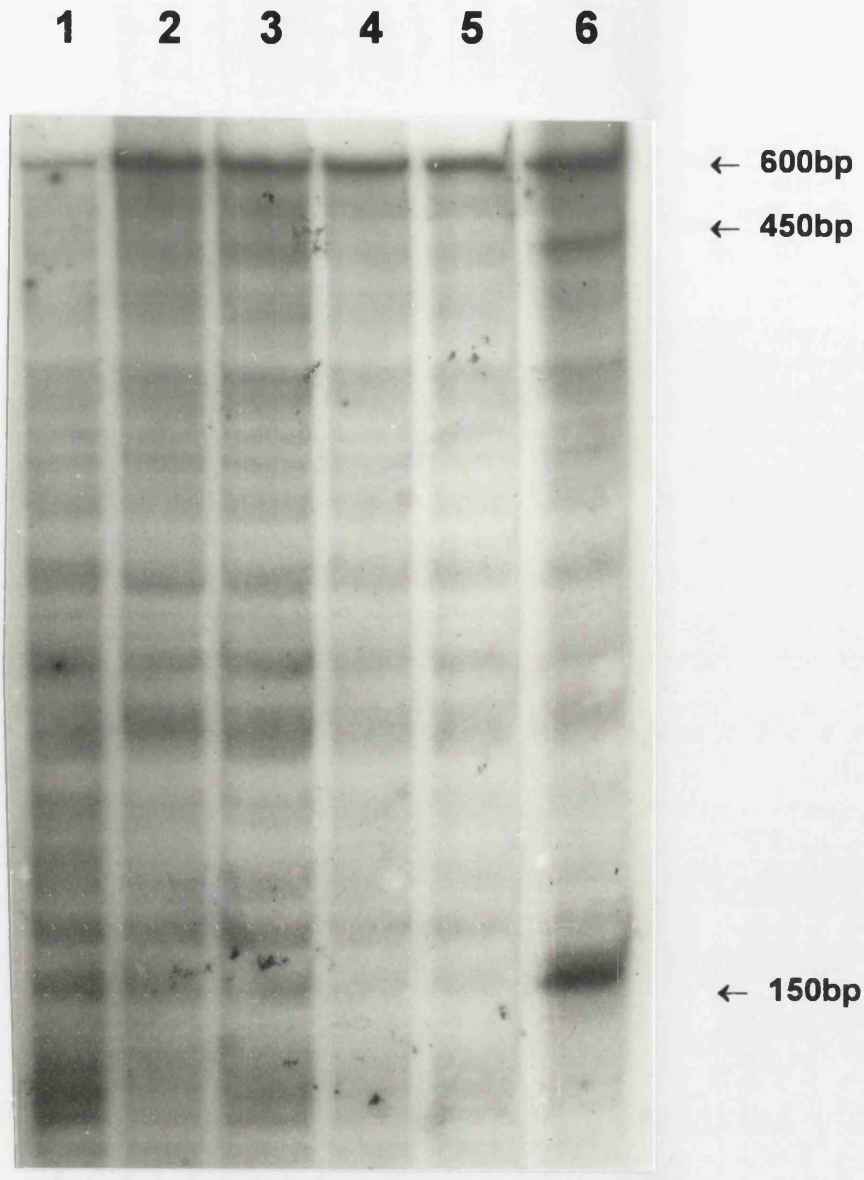


Figure 33

Chemical cleavage analysis of segment I, showing cleavage products of 150bp and 450bp in addition to the normal sized fragment of 600bp, on modification with osmium tetroxide, in Lane 6 marked by an arrow.

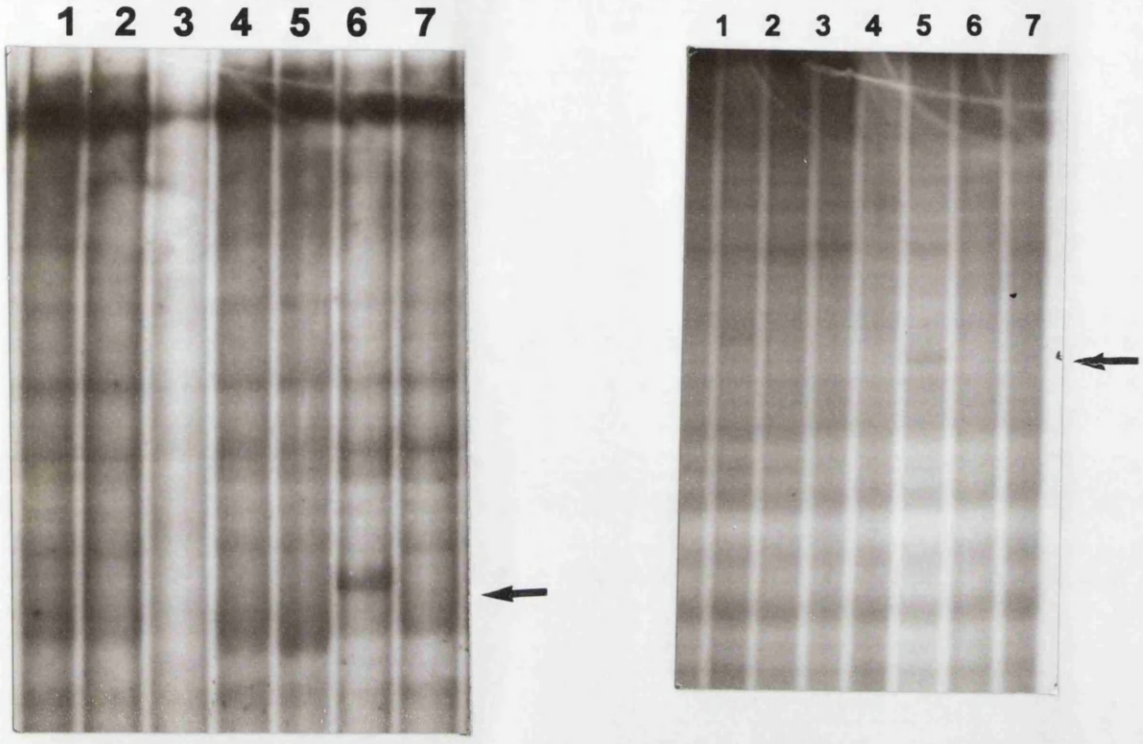


Figure 34

Chemical mismatch cleavage analysis of segment IV, showing a cleavage product of 220bp in addition to the normal sized fragment of 500bp, upon modification with hydroxylamine, in lane 6 marked by an arrow.

Figure 35

Chemical mismatch cleavage analysis of segment VII showing a cleavage product of 180bp in addition to the normal sized fragment of 500bp on modification with osmium tetroxide, in lane 5 marked by an arrow.

3.4.2 Characterisation of single base substitutions and small rearrangements in the products of DNA PCR using SSCP analysis, chemical cleavage of mismatches and direct sequencing.

Silent mutation in exon 29: N1776N

Chemical mismatch cleavage analysis of exon 29 in a 19 year old male, sporadic case of NF-1 showed a cleavage product on modification with hydroxylamine (Figure 36a). Direct sequencing revealed a C to T transition at position 5328, changing the third base of codon AAC to AAT (Figure 37). This transition maintains an Asparagine at position 1776 and is thus a silent mutation (N1776N).

The C to T transition results in creation of a *Hinf* I restriction site which upon digestion of the 230bp PCR product produces two bands of 190 and 40bp (Figure 36b). The digested PCR products were resolved on an agarose gel and visualised by ethidium bromide staining. 100 controls were screened using restriction analyses and none were found to have this mutation. This is therefore a rare sequence variant within the coding region of the NF-1 gene.

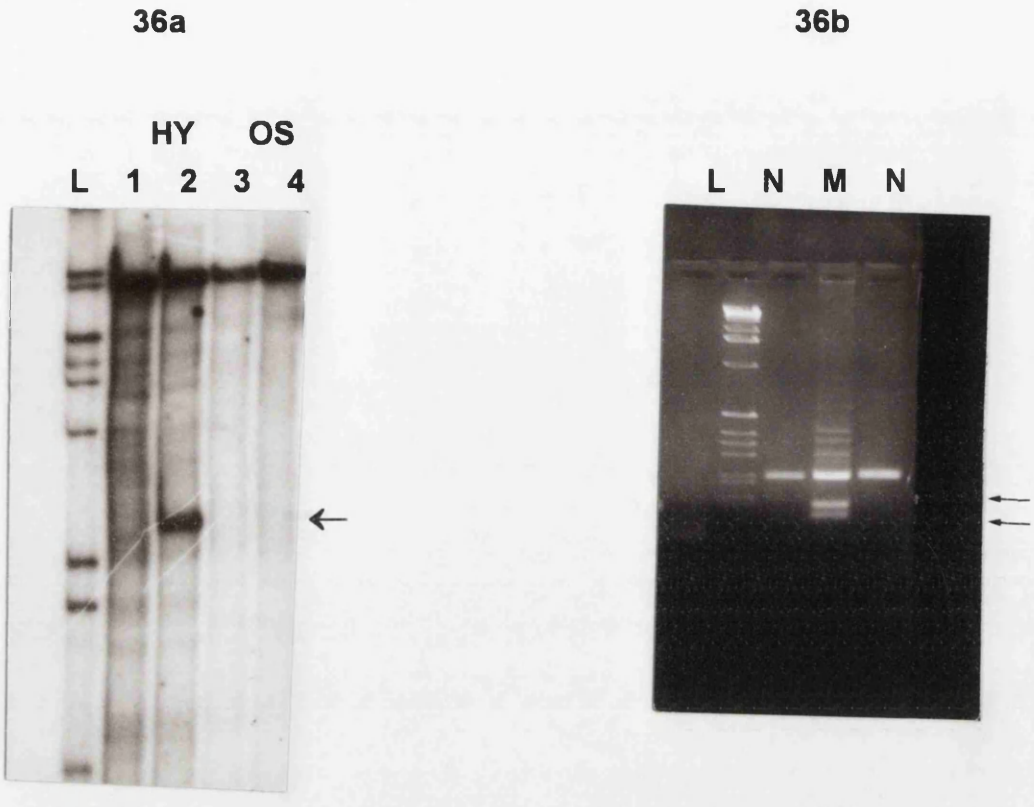


Figure 36a

Chemical mismatch cleavage analysis of exon 29. The presence of a cleavage product is seen on modification with hydroxylamine in lane 2. Lanes 1 and 3 are the control lanes on modification with hydroxylamine and osmium tetroxide respectively. Lanes 2 and 4 are the patient lanes on modification with hydroxylamine and osmium tetroxide respectively. L indicates the 1kb ladder.

Figure 36b

Restriction analysis with the enzyme *Hinf* I. The C5328 to T transition creates a *Hinf* I restriction site, showing digestion of the 230bp PCR product into two bands of 190 and 40bp resolved on a 1% agarose gel and visualised by ethidium bromide staining.

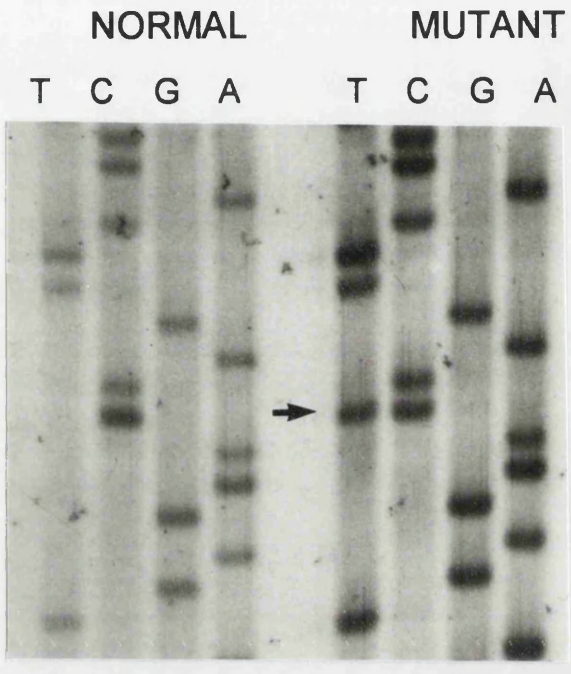


Figure 37

Partial sequence analysis of exon 29 showing the C5328 to T transition marked by an arrow in the mutant lane. Only the normal allele (C) is seen in the wild type.

Single base insertion of guanine in exon 34: 6519insG

The next mutation was identified on SSCP analysis of exon 34 in a 34 year old female patient with four affected children and a family history spanning three generations. An altered pattern was seen on SSCP analysis of exon 34 (Figure 38a). This product was then subjected to chemical mismatch cleavage, which showed a cleavage product of the same size (165bp) on modification with both hydroxylamine and osmium tetroxide in comparison with the uncleaved probe at 400 bp (Figure 38b). Direct sequencing demonstrated a single base insertion of guanine at position 6519 (6519insG) (Figure 39).

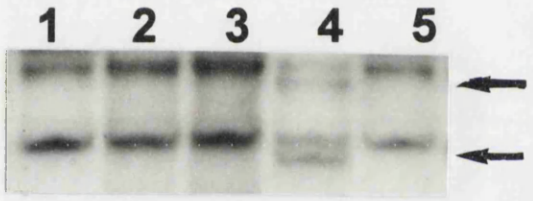


Figure 38a

SSCP analysis of exon 34, showing a shift at the level of the single strands in lane 4, compared with the level of the single strands in the control lanes.

Figure 38b

Autoradiograph of the chemical cleavage analysis of exon 34 with hydroxylamine (Hy) and osmium tetroxide (Os) in patient (lanes 1,3) and control (lanes 2,4). The control shows only the presence of the uncleaved probe at 400bp, while the patient shows a cleavage product of 165bp on modification with both hydroxylamine and osmium tetroxide. L is the 1kb ladder.

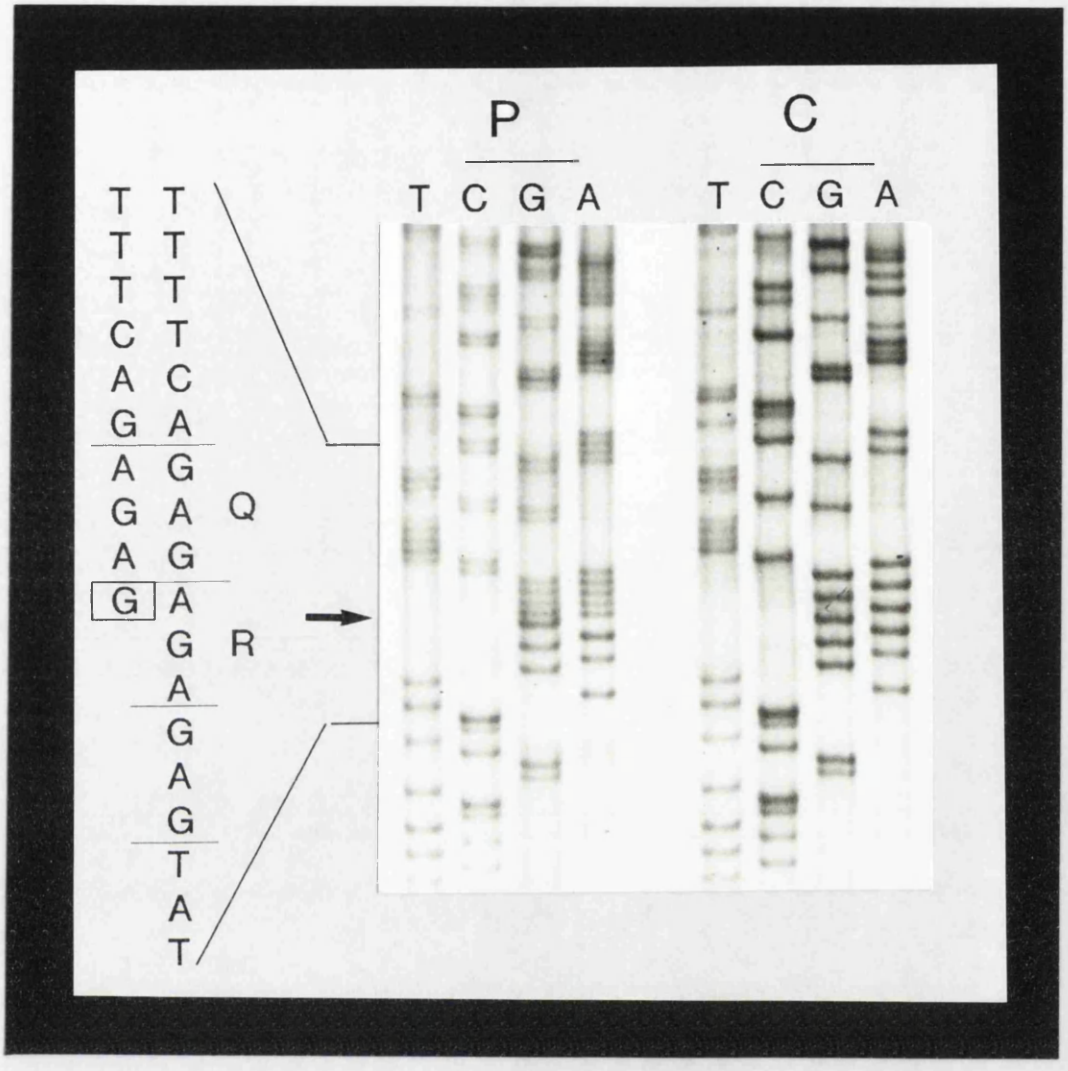


Figure 39

Sequence analysis of exon 34 in the patient and a normal control. The sequence surrounding position 6519 and the point of insertion of guanine (indicated by an arrow) is shown. As the patient is heterozygous for the mutation, both the normal and the mutated sequences run together beyond the point of insertion.

Insertion of two guanine bases in exon 42: 7485insGG

Amplification of exon 42 and flanking intronic sequences, followed by chemical cleavage analysis showed a cleavage product of the same size on modification with both hydroxylamine and osmium tetroxide (Figure 40). This went in favour of the mutation being a small insertion or deletion and was detected in a 8 year old male patient, who is a sporadic case of NF-1.

Direct sequencing of this exon demonstrated the presence of an insertion of 2 guanine bases at position 7485 of the published sequence (Figure 41). This mutation creates the following restriction enzyme sites: *Msp* I, *Sma* I, *Nla* IV, *BsaJ* I and *Nci* I. Electrophoresis on MDE™ (BIOCHEM) gel matrices revealed an altered pattern due to slower migrating heteroduplexes (Figure 42).

Nonsense mutation in exon 42: R2496X

Chemical cleavage analysis of exon 42 also detected a cleavage product on modification with hydroxylamine (Figure 43a) in a 23 year old female patient, who is a sporadic case of NF-1. Direct sequencing showed this to be due to a C to T conversion at position 7486 (Figure 43b). This transition at a hypermutable CpG dinucleotide converts the codon for Arginine (CGA) to a premature stop codon (TGA) at position 2496 (R2496X). This nonsense mutation was identified in the same codon (Arg 2496) in which the 2 base insertion 7485insGG was identified.

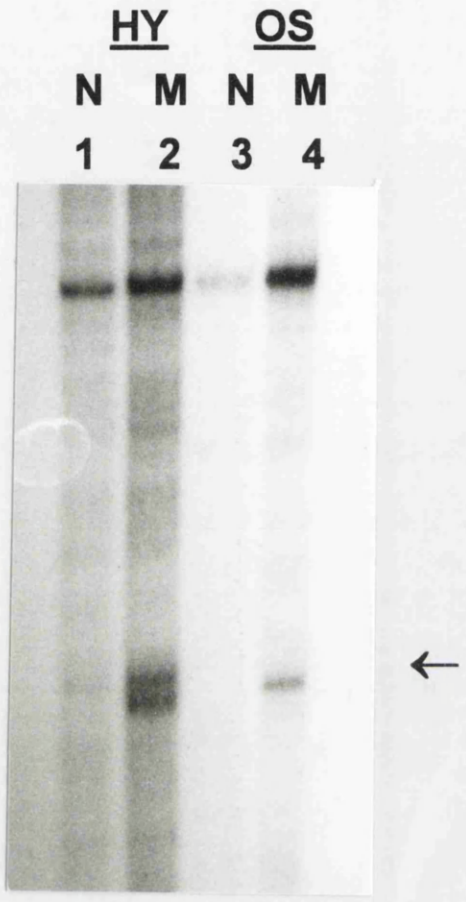


Figure 40

Chemical cleavage analysis of exon 42, showing cleavage products in lanes 2 and 4 in the mutant on modification with hydroxylamine (HY) and osmium tetroxide (OS) respectively (marked by arrows). Lanes 1 and 3 are the respective controls.

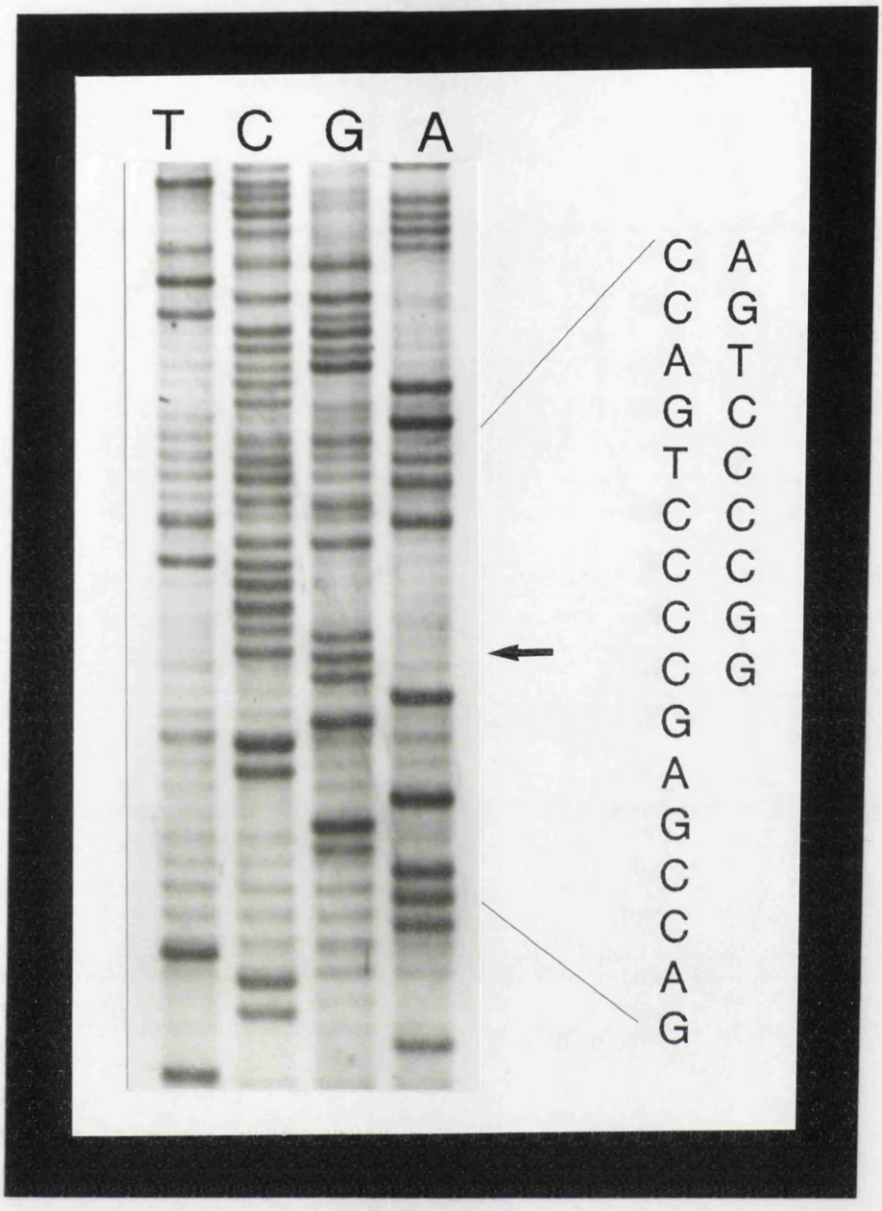


Figure 41

Partial sequence analysis of exon 42 showing normal sequence up to a point marked by an arrow. Beyond this point a complex sequence is seen, due to the insertion of two guanine bases at position 7485. Beyond the level of insertion, the sequence from the normal allele and the sequence from the mutated allele (whose sequence has been pushed forward two bases due to the insertion) are seen running in parallel.

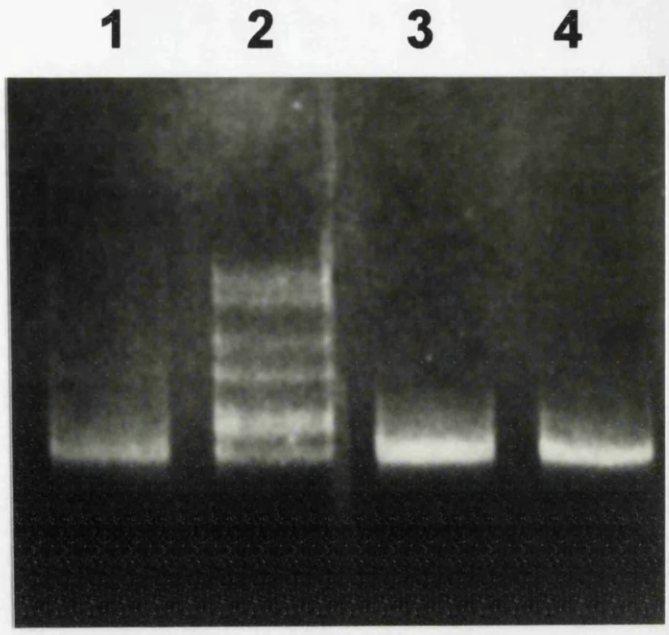


Figure 42

Heteroduplex analysis of exon 42. Lane 2 shows a distinct pattern with the formation of slower migrating heteroduplexes. More than one heteroduplex pattern is seen and a possible explanation for this is as follows: This patient was shown to be heterozygous for a polymorphism in intron 41, which itself generates a specific heteroduplex pattern (Figure 44). This product of amplification thus has two changes, heterozygosity for the intron 41 polymorphism and a two base insertion. The various combinations of sequence alterations within this product could be responsible for formation of more than one heteroduplex.

Figure 43a

Chemical cleavage analysis of exon 42 on modification with hydroxylamine showing the presence of a cleavage product in the mutant lane (lane 1).

Figure 43b

Partial sequence analysis of exon 42, showing the C to T transition at position 7486 marked by an arrow in the mutant lane.

3.5 IDENTIFICATION OF A NOVEL POLYMORPHISM IN INTRON 41 OF THE NF-1 GENE

An intragenic polymorphism was detected in intron 41 by chemical cleavage analysis. While screening for mutations in exon 42, this was seen as a constant cleavage product. DNA sequence analysis demonstrated this to be a transition of Adenine 'A' to Guanine 'G', in intron 41, 28 bases upstream of the first base of exon 42 (Figure 44). The sequence analysis was carried out initially in 4 individuals and it was seen that this region was polymorphic (2/4 individuals were heterozygous on initial sequence analysis).The mutation results in the loss of the following restriction enzyme sites: *Mse* I, *Pac* I, *Tsp*509 I and *Vsp* I. Analysis on an MDE™ gel matrix was carried out in an attempt to develop an easy screening strategy to find out the frequency of this polymorphism. It was seen that three distinct patterns were seen on electrophoresis (Figure 44). Correlation of the sequence analysis with the patterns detected on the gel permitted the following conclusions to be made: Pattern 1 is produced by alleles homozygous for A, pattern 3 is produced by alleles homozygous for G and pattern 2 heterozygous for A and G. 75 unrelated Caucasian individuals of Scottish origin were studied and the genotype prevalences for the total group (n=75) were as follows: 36% (27/75) for A allele homozygotes, 12% (9/75) for G allele homozygotes and 52% (39/75) for A/G heterozygotes. The individual frequencies of the A and G alleles were 62% and 38% respectively. Autosomal codominant segregation was observed in one family with three generations and two families with two generations.

When the observed heterozygosity (52%) was compared with that calculated from the homozygous allele frequencies (42%), a discrepancy is apparent

due to the small sample number. The average of these numbers gives an allele heterozygosity of 47%.

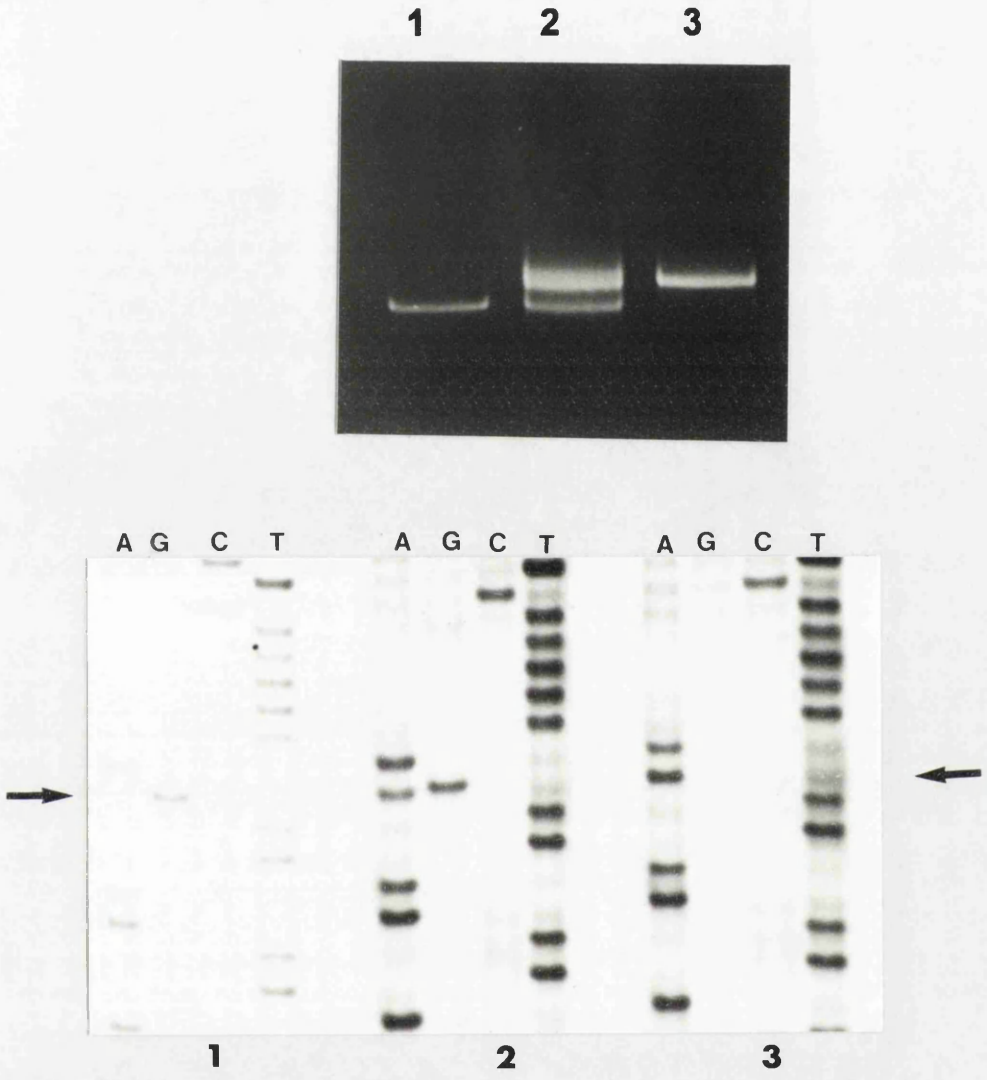


Figure 44

Heteroduplex analysis of the intron 41 polymorphism 7395 -28 A to G, showing three distinct heteroduplex patterns. Lane 1 pattern for alleles homozygous for A, lane 2 pattern for alleles heterozygous for A/G and lane 3 pattern for alleles homozygous for G.

Sequence analysis surrounding the intron 41 polymorphism 7395 -28 A to G

Tracks 1: Alleles homozygous for G

Tracks 2: Alleles heterozygous for A/G

Tracks 3: Alleles homozygous for A

The other two possible polymorphisms which have not been characterised by direct sequencing are as follows: Three distinct patterns were seen on SSCP analysis of exon 28 and could be due to a polymorphism in exon 28.

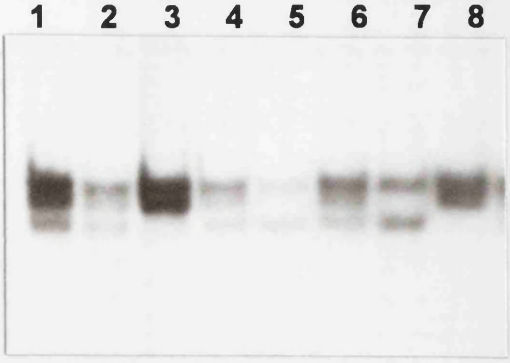


Figure 45 shows lanes 1,2,4 and 6 with one pattern, lanes 3 and 8 with another pattern and lanes 5 and 7 with a third pattern.

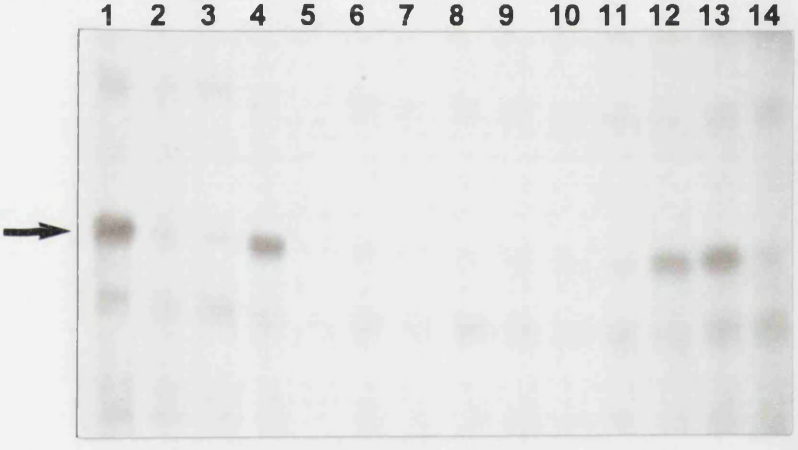


Figure 46

Also seen on chemical cleavage analysis of the mRNA (segment IV) surrounding exon 16 was the observation that 4 out of 14 samples showed the same cleavage product, which could be due to a polymorphism within this region. Figure 46 shows lanes 1,4, 12 and 13 with the same sized cleavage product on modification with hydroxylamine.

3.6 SEGREGATION ANALYSIS IN NF-1 FAMILIES USING A TETRANUCLEOTIDE REPEAT IN INTRON 27 OF THE NF-1 GENE

Analysis was carried out using an Alu-repeat polymorphic sequence AAAT located in intron 27 of the NF-1 gene (Xu et al 1991).

Lazaro et al. (1993) analysed 294 independent normal chromosomes and reported the distribution of alleles which were as follows: allele 0 (411bp:11 repeats), allele 1 (407bp:10 repeats), allele 2 (403bp:9 repeats), allele 3 (399bp:8 repeats) and allele 4 (395bp:7 repeats). This locus was shown to have a heterozygosity of 0.50. In this study the different alleles seen on analysis of affected patients and their families are shown in Figure 47a.

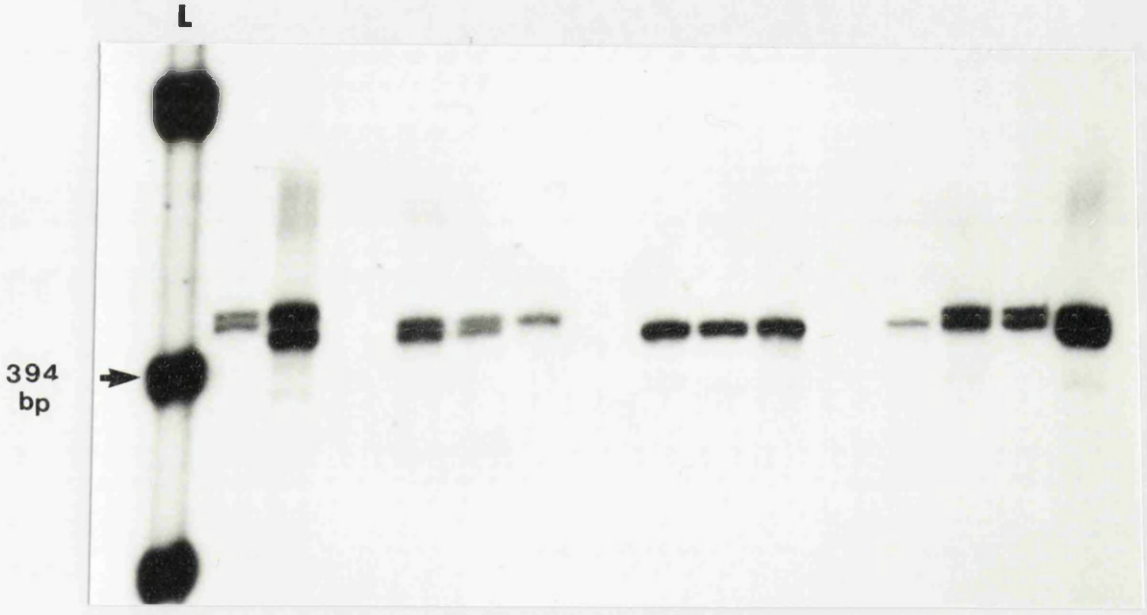


Figure 47a

Autoradiograph showing the different alleles in the Scottish population on tetranucleotide (AAAT) repeat analysis in intron 27. L represents the radiolabelled 1kb ladder.

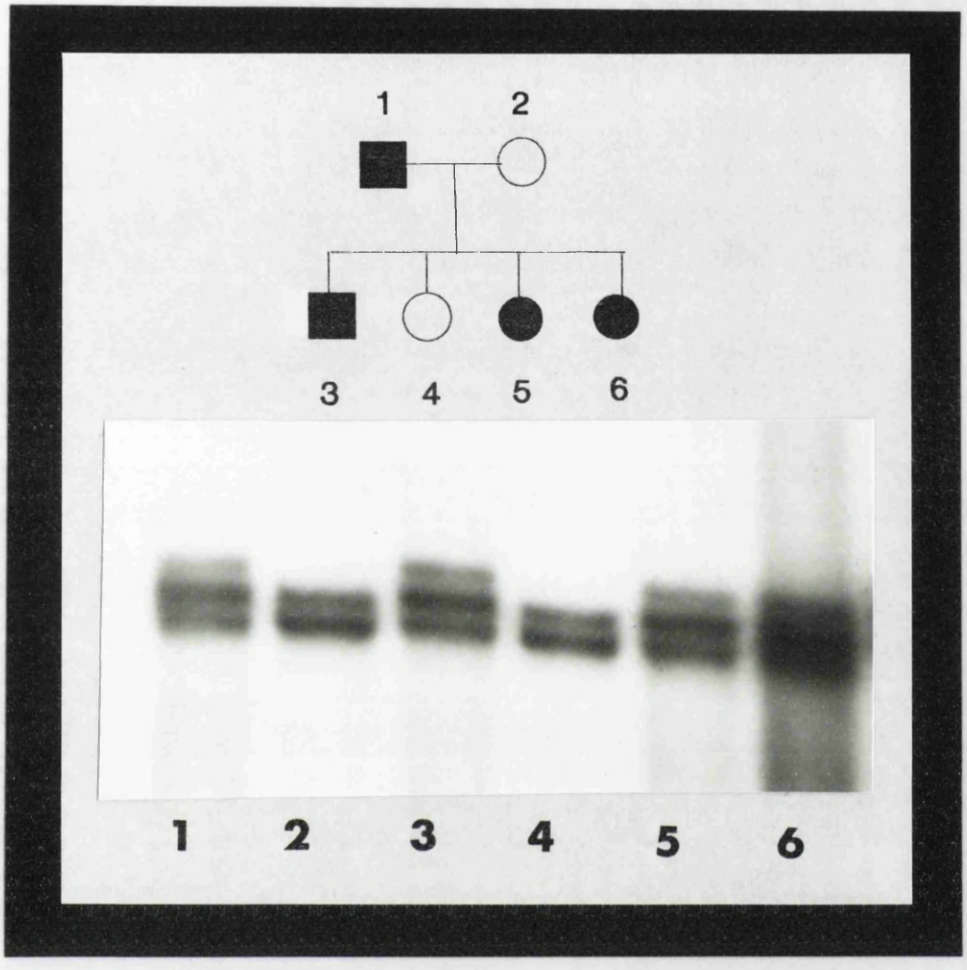


Figure 47b

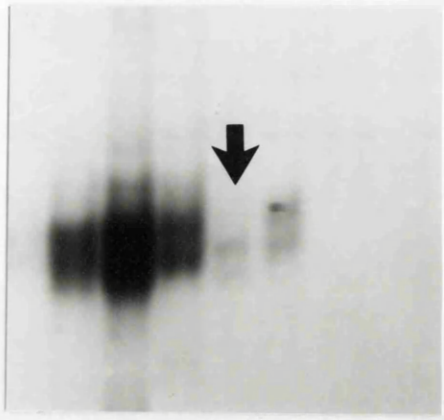
Analysis of the intron 27 tetranucleotide repeats. The unaffected members of the family (lanes 2 and 4) show a different pattern from the affected members.

3.7 SCREENING FOR SOMATIC MUTATIONS

Prior to screening for mutations in the NF-1 gene, the tumour samples were analysed for mutations within the appropriate ras group of oncogenes using SSCP analysis. Two tumour types showed an altered pattern on SSCP analysis in the regions surrounding codons 12 and 13 of the K-ras group of oncogenes. The tumour types with the alteration were a colonic adenoma sample (8549) and a colonic adenocarcinoma sample (8496) (Figure 48a, 48b). This is in accordance with the fact that up to 40% of colon carcinomas harbour a mutant K-ras gene (Bos 1989).

Screening for small alterations within the NF-1 gene was carried out by initial screening of exon 24 of the NF-1 gene (Li et al 1992). A pair of intron based primers were used to generate a 236bp PCR product, containing the FLR exon (159bp). Initial screening was carried out by an assay (FLR-*Taq* assay) designed by Li et al (1992). To detect sensitively and rapidly any change in the first nucleotide of Lysine 1423 in exon 24 of the NF-1 gene, in which 2 mutations affecting the first base had already been reported, Li et al designed a restriction enzyme based assay using the restriction enzyme *Taq* I. Any mutation altering the first A in the AAG Lys-1423 codon prevents cutting by *Taq* I (Figure 49). Using this assay, tumour DNA samples were analysed and all the samples showed complete digestion with the *Taq* I restriction endonuclease which indicated no alteration at the AAG Lysine 1423 codon. To analyse for mutations in other parts of exon 24, these tumour samples were subjected to chemical cleavage analysis, which also did not reveal any abnormalities (Figure 50). A positive control was done with each chemical cleavage analysis to ensure that the procedure was working. A cleavage product was seen in the control, but not in any of the tumour samples analysed.

48a



48b

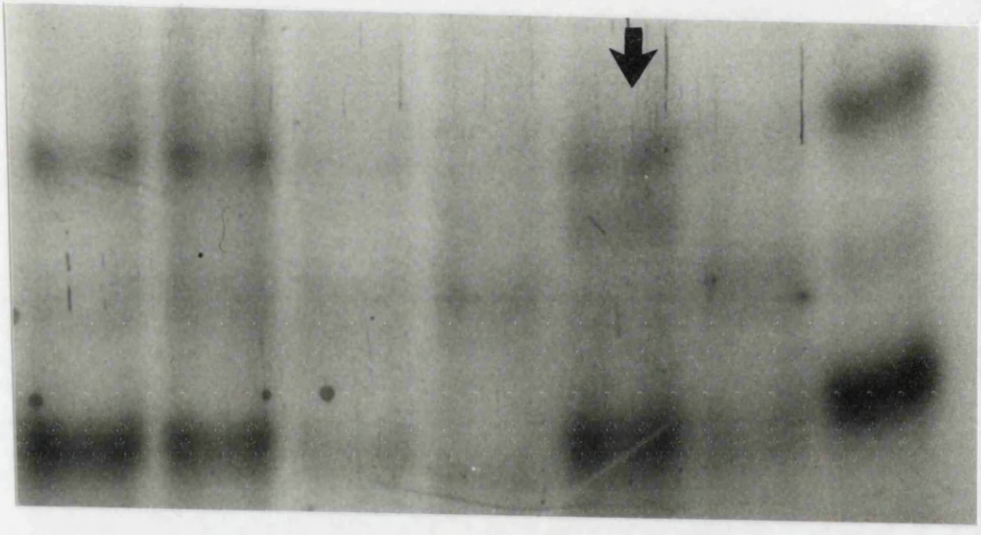
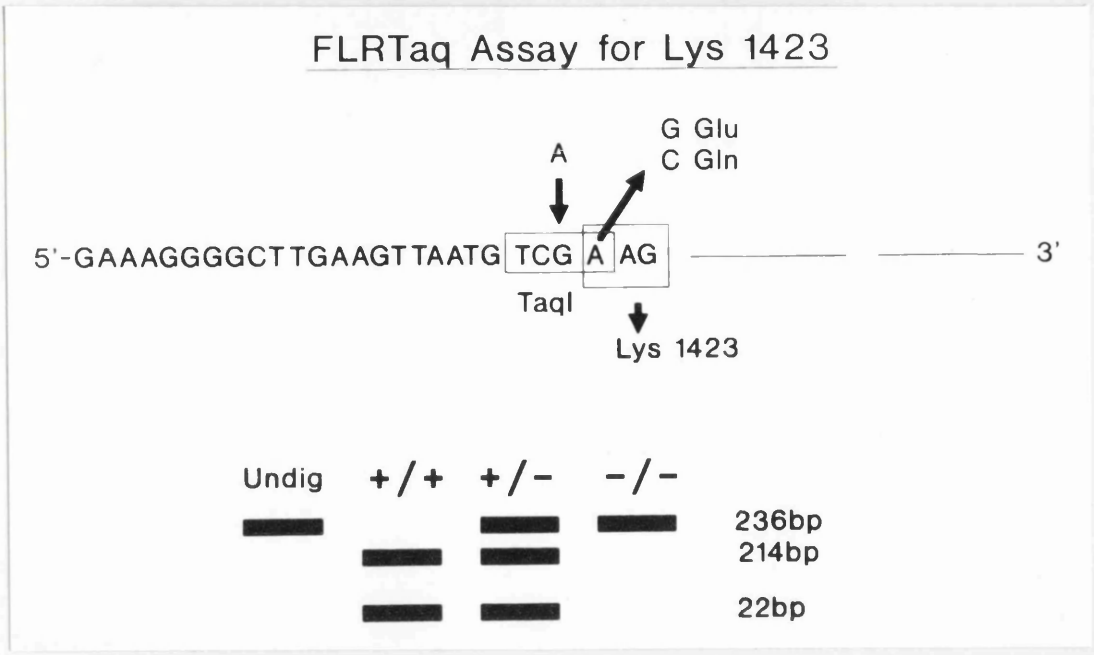


Figure 48

SSCP analysis of the K-ras group of oncogenes (codons 12-13).

a)An altered pattern is seen at the level of the single strands (marked by an arrow) in tumour sample 8549, which was a colonic adenoma. b)An altered SSCP pattern in tumour sample 8496 (marked by an arrow), which is a colonic adenocarcinoma.



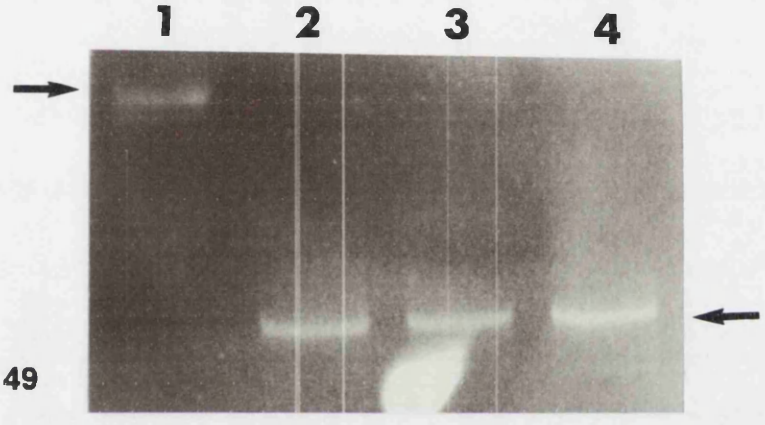


Figure 49

FLR-*TAQ* Assay for Lys 1423 (Lys 1389): The figure shows the upstream primer used for amplification, with the modification of one base A to G, in order to create the restriction site for *Taq* I (TCGA), the A being from the first base of the codon for Lysine (1423/1389). Modification of the codon for Lysine by a mutation in the first base of the codon e.g. A to G transition changing Lysine to Glutamine, or an A to C transition changing Lysine to Glycine would result in loss of the *Taq* I site.

The undigested allele of 236bp is shown below. On digestion with an intact Lysine codon, the product would show digestion into a 214bp and a 22bp fragment (+/+). However, if the codon for Lysine was mutated, it would result in digestion into two fragments of 214bp and 22bp from the normal allele and an undigested product of 236bp from the mutant allele (+/-). If both alleles were mutated, a pattern of no digestion (-/-) would be seen. Electrophoresis on an 8% polyacrylamide gel showing lane 1 with the undigested control (236bp). Lanes 2,3 and 4 show complete digestion indicating absence of a mutation (only the 214 bp band is shown, the 22bp band being much smaller runs faster and is not shown). All the tumours were screened using this assay and no mutations affecting the codon Lysine (1423/1389) were seen.

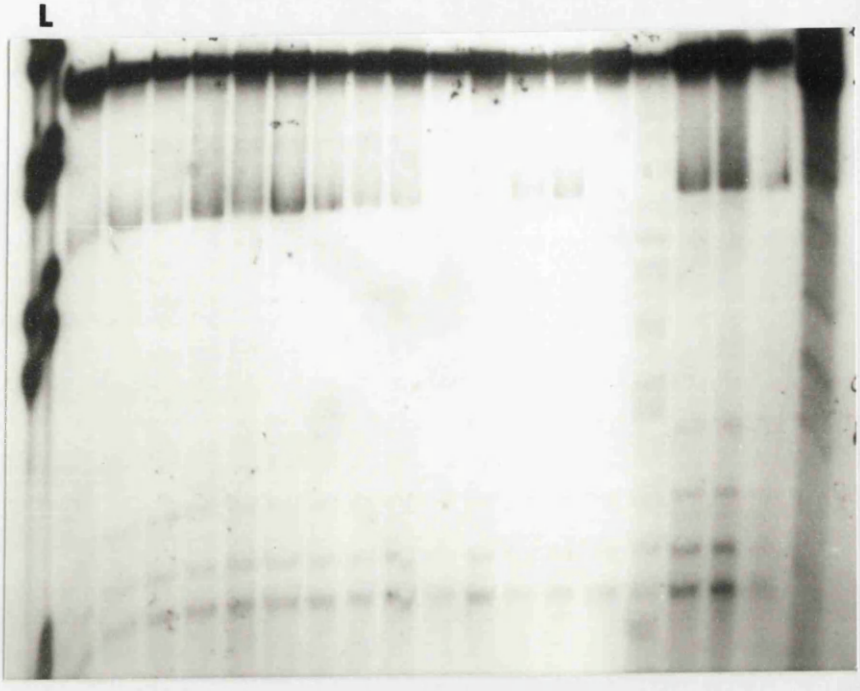


Figure 50

Chemical mismatch cleavage analysis of exon 24 (FLR exon) showing no cleavage product in the tumour samples analysed.

Four neuroblastoma samples were screened for large alterations using Southern (p5/*Eco*R I) and Northern blot analysis, none of which revealed any abnormalities.

CHAPTER: 4

DISCUSSION

CHAPTER 4: DISCUSSION

4.1 SPECTRUM OF MUTATIONS

The identification and characterisation of mutations within the NF-1 gene is the first step in addressing a number of important features of NF-1, including the possibility of a mutational hotspot, variability of expression and the correlation between the genotype and clinical symptoms.

The NF1 gene was cloned in 1990 and until recently very few mutations were identified. This could be due to a number of reasons which include the large size of the gene, the high mutation rate (up to 50% of cases being due to new mutations), the as yet uncharacterised intron-exon boundaries and due to the fact that the screening methods used so far may have missed mutations.

So far, a total of 68 mutations (June 1994 newsletter-NNFF Mutation Analysis Consortium) have been reported in the NF-1 gene. The mutations characterised include large and small deletions and insertions, single base substitutions which include nonsense, missense, splice junction and silent mutations, of which 13 mutations have been characterised in this study (Figure 51). Due to the compiling of the mutations by the NF-1 consortium, a group of mutations can now be analysed for their spectrum, distribution and recurrence.

In this study, a positive screening result was detected in 17 out of 25 patients analysed, of which 13 have been characterised. Of these, 10 mutations are potentially pathogenic. The mutations described in the NF-1 gene so far appear to be heterogeneous and include gene deletions and insertions of various sizes as well as single base substitutions. Thus a spectrum of changes has been identified and this offers an opportunity to

compare the relationship between the structure and function of neurofibromin.

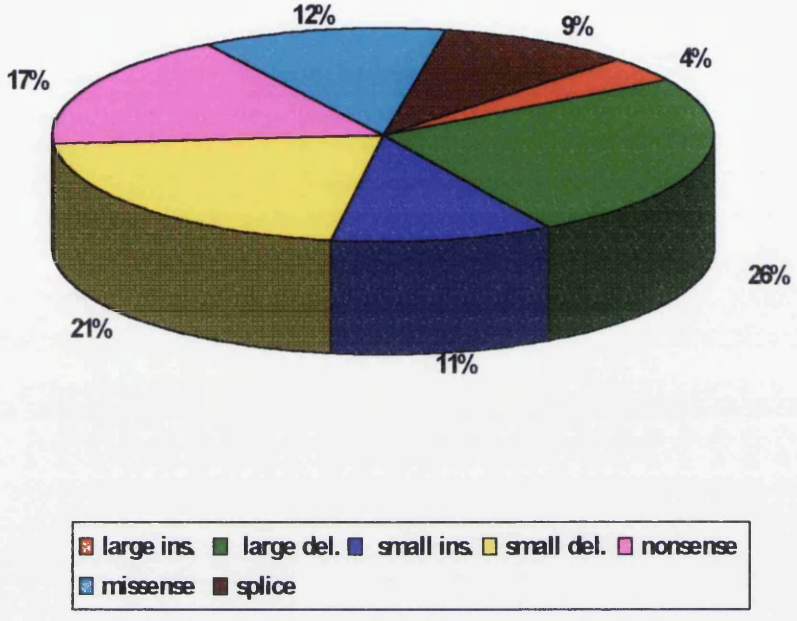
Table 12 shows the various germline mutations tabulated by their positions within the NF-1 gene, including the mutations detected in this study, which are marked by an *. Mutations that are underlined have not been reported to the Consortium, but have been published elsewhere (Valero et al., 1994). Mutations in italics lie within the NF-1 GAP related domain, which extends from mRNA positions 1125-1536.

Table 12: Mutations reported within the NF-1 gene

EXON	MUTATIONS
3	Deletion of exon 3
11	**1721+3G***
13	Q682X
16	**partial del of exon 16***
17	2970delAAT, 2875insA, 2851delCTTTT
18	M1035R, **3113+1A***, **del exon 18***
21	***G1166D****
22	*R1276X*
23	*3974delGTTA,* ***S1311S****, *L1339R*
24	*4152delA, 4183delC, 4190delT, K1432E,* ***G1404G****, ***K1419R****
27a	*R1513X*
28	5123delCCAACC, 5077delAAACTGGCTGAGC, 5010delG, 5108delAG, 4969delTCTATA
29	**N1776N***, S1745X, 5451insC, 5468insT, R1748X, Q1754X
30	5678delAC, **5749+2G***
31	5843delAA, R1947(7), 5849insTT, 5816insG, del 5787-5814
32	ins (Alu), 5949delA, 5944-5G (2)
33	del (ex 33), P2056R, 6364+4G
34	**6519insG***, L2143M, Y2171R
37	del (ex 37)?
38	6922insGATGAGGTCA
39	7096delAACTTT
42	**R2496X*, 7485insGG***
44	7745delTTCTCTTGGA

FIGURE 51

SPECTRUM OF MUTATIONS



This chart shows the spectrum of mutations reported within the NF-1 gene in the literature and via the NNFF Mutation Analysis Consortium.

A heterogenous group of mutations is seen and 88% of the mutations (i.e. excluding the missense mutations) are predicted to cause a disruption of neurofibromin.

From analysis of all the types of mutations identified so far, 88% of mutations are predicted to cause disruption of the protein product neurofibromin (Figure 51). This study has shown that 75% of the mutations characterised are predicted to cause disruption of the protein product neurofibromin. The reasons for this discrepancy (75% / 88%) could be, the technique of PCR amplification of the NF-1 mRNA used in this study is only capable of detecting large alterations if the primer pairs used entirely encompass the alteration. Also Southern blotting was carried out using only one probe-enzyme combination and therefore large alterations may have been missed.

However from all the mutations identified so far and from the spectrum of mutations detected in this study, it appears that the majority are small sequence alterations causing disruption of the NF-1 gene. This study has also identified 3 silent mutations, in addition to the remaining potentially pathogenic mutations. A mutation detection method PTT (protein truncation test) has been described for rapid detection of translation terminating mutations (Roest et al., 1993). This technique is unable to detect missense mutations (which may be pathogenic), but has the advantage of not detecting silent mutations. This method would be useful in analysing mutations in large complex multi-exonic genes, where a large fraction of pathological mutations arise from termination of translation. One such example, the dystrophin gene was analysed for mutations using PTT (Roest et al., 1993) and this technique may be a useful screening method for mutation analysis of the NF-1 gene.

The mutations characterised within the NF-1 gene that are predicted to cause disruption of neurofibromin, i.e. insertions, deletions, duplications, splicing errors and nonsense mutations are discussed below.

Insertions

There have been two reports of large insertions, by Wallace et al., (1991) who reported insertion of an *Alu* element in the NF-1 gene and by Upadhyaya et al., (1992) who reported insertions of approximately 10kb. Small insertions have been reported in the literature by Upadhyaya et al., (1992), Zhong et al., (1993) and Ainsworth et al., (1993). Two small insertions have been identified as part of this study (6519insG and 7485insGG).

6519insG: An insertion of a single guanine base was identified in exon 34 at position 6519. The mutation causes a shift in the reading frame and is predicted to result in the formation of 46 altered amino acids, before producing a premature stop codon at position 2220. Thus the predicted protein product will contain 2219 amino acids instead of 2818.

7485insGG: An insertion of 2 guanine bases was identified in exon 42 at position 7485. The mutation produces a shift in the reading frame and is predicted to cause the translation of 5 altered amino acids, before creating a premature stop codon at position 2502. A consequence of this mutation will be the synthesis of a truncated protein product of 2501 amino acids instead of the normal 2818.

Deletions

The deletions reported so far in the NF-1 gene are as follows: Viskochil et al., (1990) reported 3 deletions of 190, 40 and 11kb. Upadhyaya et al., (1990), (1992) reported three deletions of 90kb, 40kb and 80kb. Kayes

et al., (1994) reported 5 patients with deletions of the entire NF-1 gene. Weiming et al., (1992) reported a comparatively smaller deletion of 571bp. Small deletions within the NF-1 gene have been reported in the literature by Stark et al., (1991), Xu et al., (1992), Zhong et al., (1993), Shen et al., (1993a), Shen et al., (1993b), Colman et al., (1993) and Valero et al., (1994).

In this study two deletions of the NF-1 mRNA were detected, the cause of which at the DNA level has not yet been characterised.

The first was a complete deletion of exon 18, detected by RT-PCR and direct sequencing. The exclusion of exon 18 from the NF-1 mRNA could be due to a splice mutation, a deletion at the DNA level or a nonsense mutation. The cause at the DNA level of this mutation remains to be determined.

The second deletion was a partial loss of exon 16 (229 bases) from nucleotide positions 2410 to 2639. This loss of 229 residues is predicted to result in a shift in the reading frame and creation of a premature stop codon. A protein product of 803 residues instead of the normal 2818 amino acids is predicted, with loss of the NF-1 GRD. This exclusion of 229 nucleotides from exon 16 in the NF-1 mRNA could be due to a deletion at the DNA level or due to aberrant splicing (discussed in splice site mutations).

Duplications

A 10 base pair duplication (6922insGATGAGGTCA) was reported in a sporadic case of NF-1 with the complication of a neurofibrosarcoma in exon 38 of the NF-1 gene by Legius et al., (1994b).

A duplication within the NF-1 gene was also reported by Tassabehji et al., (1993) in a family with Watson syndrome and features of Noonan's syndrome.

Splice junction mutations

The first splice junction mutation reported within the NF-1 gene was reported by Hutter et al., (1994). The mutation was an A to G transition at +4 of the intron 33 splice donor site and caused exon skipping. The other splice mutation was at position - 5 of intron 31, reported independently by Rodenheiser et al., (1994) and Legius E. (1994) (NNFF Mutation Analysis Consortium newsletter-June' 94).

In this study 3 splice mutations were identified at the splice donor sites of introns 11, 18 and 30.

1721+3A to G: The first splice mutation detected was an A to G substitution at position +3 (nt 1721 +3) of the splice donor site of intron 11. The mutation resulted in an error of splicing leading to the skipping of exon 11 from the NF-1 mRNA and a shift in the translational reading frame. This is predicted to result in the formation of 12 altered amino acids and the creation of a premature stop codon at position 560. This would result in the synthesis of a protein of 559 amino acids instead of the normal 2818 amino acids. This product of translation would lack the NF-1 GAP related domain.

3113+1G to A: The second splice mutation detected was a G to A substitution at position +1 (nt 3113 +1) of the invariant GT dinucleotide of the splice donor site of intron 18. It resulted in an error of splicing which led to the skipping of exon 18 from the NF-1 mRNA. The skipping would result in the loss of 41 amino acids from the normal protein product of 2818 amino acids, including two cysteine residues at positions 1016 and 1032.

5749+2T to G: The third splice mutation detected was a T to G substitution at position +2 (nt 5749 +2) of the invariant GT dinucleotide of the splice donor site of intron 30. The mutation resulted in an error of splicing which led to skipping of exon 30 from the NF-1 mRNA. The skipping of exon 30 shifts the translational reading frame and is predicted to lead to the formation of

one altered amino acid, before creating a premature stop codon at position 1851. This would result in a protein of 1850 amino acids instead of the normal 2818 amino acids, assuming the predicted neurofibromin is stable.

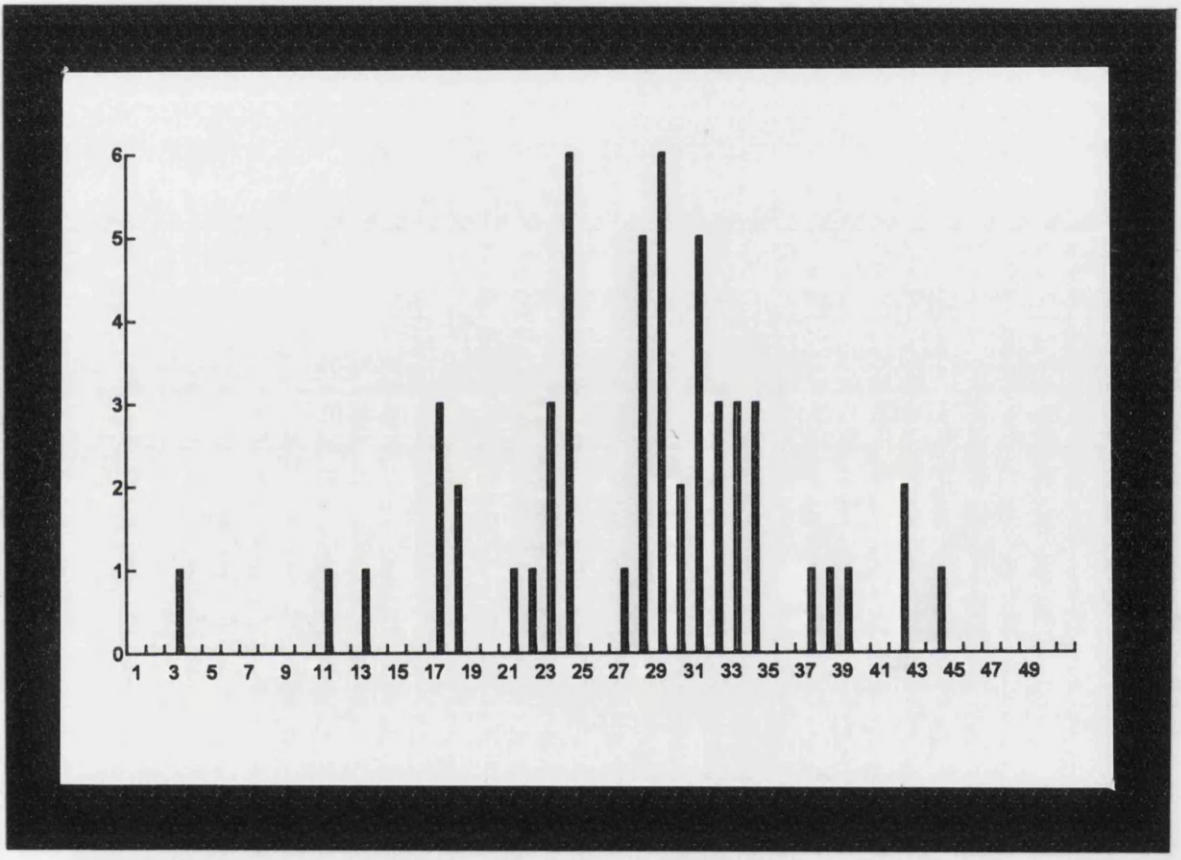
Nonsense Mutations

Thirteen nonsense mutations have been reported so far within the NF-1 gene. Seven of these are the same mutation (R1947X) in exon 31 of the NF-1 gene reported by different investigators. The remaining nonsense mutations are R1267X, R1523X, R1748X, Q628X, S1745X and Q1745X (Cawthon et al., 1990a, Ainsworth et al., 1993, Estivill et al., 1991, Horiuchi et al., 1994, Valero et al., 1994 and NNFF Mutation Analysis Consortium newsletter, June 1994).

In this study one nonsense mutation was detected in exon 42, R2496X. This nonsense mutation resulted due to conversion of C to T at position 7486, altering the codon for arginine (CGA) to a premature stop codon (TGA). It occurred at a hypermutable CpG dinucleotide and is predicted to produce a truncated protein product of 2495 amino acids.

4.2 DISTRIBUTION OF MUTATIONS

Figure 52 shows the distribution of all the germline mutations identified within the NF-1 gene. From analysis of the graph, it appears that most mutations are concentrated in exons 20-35 of the NF-1 gene. However this is a biased finding, as exons 24 and 28-36 were among the first exons to be screened, since the intron sequence at their boundaries was the first to be reported (Cawthon et al., 1990a).



4.3 RECURRENCE OF MUTATIONS

There are several examples of recurrent point mutations in disease genes. The most striking example is the independent origin of the sickle cell mutation in Africa (Antonarakis et al., 1984, Pagnier et al., 1984) and in Asia (Kulozic et al., 1986) and the multiple origins of the beta-E-gene (Kazazian et al., 1984). Another example is the G6PD Mediterranean C to T at 563 mutation which has originated independently in Europe and Asia (Beutler and Kohl 1990). Within the NF-1 gene, 4 different examples of recurrent mutations have been reported which include a nonsense mutation (R1947X) which occurs at a CpG dinucleotide reported in seven different studies, a single base insertion (5451insC), a three base deletion (2970delAAT) and a splice site mutation (5944-5G) reported by two different investigators.

The mutation R1947X in exon 31 of the NF-1 gene has been independently reported by seven investigators. The first report was by Cawthon et al., (1990a) who detected this mutation on screening 72 families, followed by Estivill et al., (1991) who screened 38 Spanish families and found one such mutation. The same mutation was reported by Ainsworth et al., (1993) on screening 48 families from South Western Ontario, which is an ethnically diverse region. The cases reported above are sporadic cases of NF-1 (origin of the mutation reported by Cawthon et al., (1990a) is unknown). Horiuchi et al., (1994) reported two familial cases of NF-1 with this mutation, on analysis of 25 Japanese families. These cases were the first described in an ethnic group other than Caucasian. Valero et al., (1994) also reported the R1947X mutation in a familial case of NF-1 on screening 70 unrelated NF-1 patients.

Therefore of the 253 families screened so far five mutations at this particular nucleotide have been detected. The same mutation was also reported by

Korf et al. (1994) and Vidaud et al. (1994) via the NNFF Mutation Analysis Consortium.

Since exon 31 represents 2% of the NF-1 coding region, seven independent reports of the same mutation in different populations has established this to be a hotspot for mutation in the NF-1 gene and shown that this specific cytosine is particularly prone to mutation. The identification of the same mutation in seven independent cases also suggests that a specific search for this mutation (R1947X) should be made when studying NF-1 familial or sporadic cases for genetic analysis.

Upadhyaya et al., (1992) reported an insertional mutation 5451insC in exon 29 in two unrelated individuals. The other example is a 3bp deletion, 2970delAAT, also reported by Upadhyaya et al., (1993) in exon 17 of the NF-1 gene in two unrelated individuals. A recurrent splice mutation has been reported by Rodenheiser (1994) and independently by Legius (1994), the mutation being 5944-5G in intron 31, which results in the shift in the reading frame (NNFF newsletter, June 1994). In the above case however, it is not known whether the splice mutation represents true recurrence or identity by descent.

In the present study, the absence of exon 18 from the NF-1 mRNA was detected in two unrelated individuals. In one patient, a familial case of NF-1 the absence of exon 18 from the NF-1 mRNA was due to a splice site mutation at the donor site of intron 18 (3113+1G to A). In the other patient, the absence of exon 18 (shown by RT-PCR and direct sequencing) could be due to a splice site mutation or a deletion and this remains to be characterised.

4.4 NF-1 HOMOLOGOUS SEQUENCES AND PSEUDOGENES

NF-1 related loci have been identified on chromosomes 2, 12, 14, 15, 20, 21 and 22 (Marchuk et al., 1991). All these show greater than 90% homology with some NF-1 exons and intron sequences. A number of the above regions have been sequenced and most, if not all represent non-processed pseudogenes. Most of the related loci reside on acrocentric chromosomes and it was hypothesised that their origin was due to acrocentric chromosome pairing during meiosis. These pseudogenes may act as a reservoir of mutations that can be crossed on to the NF-1 locus on chromosome 17 by gene conversion, which might explain the high mutation rate for the disorder (Cummings et al., 1991).

Two groups have further characterised some of the NF-1 homologous regions. One of them lies on chromosome 15 and is called the "NF-1 related locus" (Legius et al., 1992). Another sequence homologous to exon 24 of the NF-1 gene, called the "NF-1 highly homologous sequence" (NF-1HHS) has been reported by Gasparini et al. (1993).

The presence of these (and as yet undetected) pseudogenes may complicate the genetic analysis of NF-1. In the present study two mutations were detected in exon 24, a missense mutation K1419R and a silent mutation G1404G. These mutations are likely to be present within the NF-1 gene and not in the pseudogenes. The region in which these mutations exist is deleted in the "NF-1 related locus" on chromosome 15 and neither of these mutations correspond to codon changes in the "NF-1HHS sequence". Furthermore, various combinations of intron and exon based primers were used to rule out the possibility of amplification from other as yet uncharacterised pseudogenes.

The presence of pseudogenes also had to be considered during amplification of the regions surrounding exon 18 from genomic DNA.

Amplification was carried out using a forward primer in exon 17 (17F) and a reverse primer in exon 19 (19R). This region was amplified in the patient who had previously shown an altered fragment corresponding to the skipping of exon 18 on RT-PCR amplification of the NF-1 mRNA. The above primers were used for amplification from genomic DNA to rule out a deletion. Upon sequencing the product of amplification, 6 single base substitutions were detected in the patient as well as in 14 normal controls (Figure 18). On amplification and sequencing of the same region using intronic primers immediately flanking exon 18 (18F and 18R), these changes were not detected in the patient nor in the controls. Thus on the amplification using primers 17F and 19R probably a part of a NF-1 homologous sequence was being amplified. This appeared to be an as yet unreported sequence which was homologous to exon 18. This homologous sequence was also detected by another group who used somatic cell hybrids to map the sequence to chromosome 14 (Legius E., personal communication). The presence of pseudogenes has thus to be borne in mind when amplifying segments of the NF-1 coding sequence and mutations identified should be confirmed to exist at the NF-1 locus. Every mutation detected within the NF-1 gene was therefore confirmed by repeat amplifications using various combinations of intron and exon based PCR primers.

4.5 MECHANISMS OF MUTAGENESIS

Mutations in human genes may arise either as a consequence of endogenous error-prone processes such as DNA replication and repair or as a result of exposure to exogenous factors such as chemical mutagens or ionising /UV irradiation.

Regarding the endogenous mechanisms of mutagenesis, nucleotide substitutions are thought to result from either chemical (e.g. deamination of 5-methylcytosine or depurination), physical (e.g. DNA slippage) or enzymatic (e.g. post-replicative mismatch repair or exonucleolytic proof reading) processes. Since the efficiency of all these processes is sequence dependent, the distribution of point mutations is seen to be non-random, giving rise to "hot spots" of base substitutions. The mutations characterised in this study were analysed to understand their mechanism of causation and the mechanisms of mutagenesis are discussed.

4.5.1 SINGLE BASE SUBSTITUTIONS

A) Role of CpG dinucleotides

It has been estimated that the CpG dinucleotide within human coding sequences is up to 42 times more mutable than is predicted from random mutation (Cooper and Krawczak, 1991). One important consequence of the hypermutability of CpG dinucleotides is the fact that a substantial proportion of intragenic single base pair mutations are the result of C→T or G→A transitions. A total of 880 point mutations were studied by Cooper and Krawczak, (1993) and their findings showed that 38% of mutations occurred in CG dinucleotides, of which 32.8% were C to T or G to A changes due to a methylation mediated event.

Evidence of DNA methylation in the NF-1 region was reported by Rodenheiser et al. (1993). Southern analyses using isoschizomeric methylation sensitive restriction enzyme pairs revealed DNA methylation in areas flanking the NF-1 gene, while similar PCR methylation assays showed that methylation occurred both on genomic sequences flanking the NF-1 gene and within the coding region itself. The identification of methylated

CpGs in and around the NF-1 gene suggests that these are potential sites for spontaneous germline and somatic mutations within this gene. In humans DNA methylation represents one of the examples of epigenetic modification by which several features of NF-1 may be accommodated, such as the fact that up to 50% of mutations are the result of new mutations in the NF-1 gene and phenotypic expression of the disease is highly variable, even within members of the same family.

To find out whether methylated CpG dinucleotides within the highly mutable NF-1 gene served as a reservoir, within which C→T or G→A transitions contributed to a high frequency of spontaneous germline mutations associated with the disease, all the mutations characterised in this study and all the single base substitutions reported so far (NNFF Mutation Analysis Consortium newsletter-June 1994) were analysed to see if they occurred at CpG sites.

In this study 9 single base substitutions were identified which included 3 splice site mutations. It was seen that only the nonsense mutation at position 2496 (R2496X) was due to a C to T transition at a CpG dinucleotide (CGA to TGA).

Thirty-one single base substitutions have been reported so far within the NF-1 gene which include 13 nonsense mutations, 12 missense and silent mutations and 6 splice site mutations. Of these, none of the missense or splice site mutations occurred at CpG sites. Of the 13 nonsense mutations, 7 are the same mutation, (R1947X) which occurred at a CpG site. On analysis of the 6 remaining mutations, 3 nonsense mutations occur at CpG sites (R1276X, R1513X and R1748X), while the remaining three Q682X, S1745X and Q1745X did not.

B) Other mechanisms responsible for single base substitutions

In this study, 9 single base substitutions were identified. Three of them were splice site errors and one was a nonsense mutation that occurred at a CpG site. The remaining 5 single base substitutions which did not affect CpG sites were analysed for possible mechanisms of causation (Cooper and Krawczak, 1993).

The sequence surrounding the 5 single base substitutions is as follows: (The affected codon is underlined and the mutated nucleotide is in lower case).

1) G1166D in exon 21: GGC to GAC (G to A)
   GCACTCCATAGgCTTAGGTTACCA
2) G1404G in exon 24: GGG to GGA (G to A)
   TATGAAGCAGGGgATTTTAGAT
3) S1311S in exon 23: TCC to TCT (C to T)
   TGTGATCACATCcTCTGATTGGCAA
4) N1776N in exon 29: AAC to AAT (C toT)
   CTAGTAGATGAGAAcCAGTTCACCTT
5) K1419R in exon 24: AAG to AGG (A to G)
   GGGGCTTGAaGTTAATGTCAAA

Deoxyadenosine has been found to be the most frequently misincorporated base by both prokaryotic and eukaryotic polymerases. Besides deamination, if depurination is also a major cause of spontaneous mutation in human genes, the observed mutational spectrum would be expected to be heavily biased towards substitutions by deoxyadenosine. Consistent with this theory, Cooper and Krawczak (1993) found that in their sample of 880 point mutations, G to A transitions were the most frequently

observed, excluding those produced by deamination of CpG sites. In this study, 2/5 mutations were G to A transitions.

Analysis of the nearest neighbour (Cooper and Krawczak, 1993) showed that after excluding those mutations that were methylation mediated at CpG sites, (a)G residues were clearly over represented as 3' flanking nucleotides when T was mutated and (b)G residues were clearly over represented as 5' flanking nucleotides when G was mutated. So probably the relative rates of mutations, are influenced by the nucleotides flanking the mutation site. This was detected in both examples of the G to A substitutions in the five non-CpG mutations in this study, i.e. G was the flanking 5' nucleotide when the adjacent G was mutated.

Analysis of molecular mechanisms of mutagenesis carried out in cultured mammalian cells (Kresnick and Davidson, 1992) showed that C to T / G to A transitions occurred preferentially at the 3' guanine residue of a run of 2 or more guanines. This was detected in example 2, when the third guanine of the codon GGG was mutated to GGA. In fact in example 4, if we consider the C to T to be a G to A on the opposite strand, the 3' G in a set of 2 G's is affected, resulting in a C to T change on the sense strand. This example if considered would make the frequency of G to A mutations 3 out of 5 in the examples detected in this study.

4.5.2 SPLICE JUNCTION MUTATIONS

Mutations that affect mRNA splicing can be assumed to fall into 3 main categories:

(a)Mutations within a 5' or 3' splice site: Such lesions usually either abolish or at least reduce the amount of mature full length mRNA generated with/without activating cryptic (alternative) splice sites in the vicinity of the

wild-type sites. Exons are recognised and defined as distinct units by the initial binding of protein factors to the 3' end of the intron followed by a search for a downstream 5' splice site (Robberson et al., 1990). Thus a splice site mutation can result in the utilisation of an alternative site, so that an exon no longer is recognised as such and as a result be excluded from the mature mRNA transcript (exon skipping). The utilisation of an alternative (cryptic) splice site results in the production of mRNAs that lack a portion of the coding sequence or which contain additional sequence of intronic origin.

(b)Mutations within an intron, which may serve to activate cryptic splice sites, leading to the production of aberrant mRNA species.

(c)Mutations within a branch point sequence.

In a review of 101 splice mutations (Krawzcak and Cooper 1991), it was observed that of the 5' splice site mutations, 60% involved the invariant GT dinucleotide at positions +1 and +2. Of the 3' splice junction mutations, 87% were seen to involve the AG dinucleotide with an excess at position -2.

In this study, three splice mutations 1721+3A to G in intron 11, 3113 +1G to A in intron 18 and 5749+2T to G in intron 30 of the NF-1 gene were identified (Figure 53). The relative efficiencies of the 5' splice sites (5' SS) of normal and mutant introns 11, 18 and 30 was calculated on the basis of conservation of bases at the eight positions most likely to contribute 5' splice recognition determinants (position -2 through to +6). This was done by the method of Shapiro and Senapathy (1990), in order to calculate if the mutant intron sequences showed a reduction from that of the normal sequence.

Consensus values (CV) were calculated by the formula below, which gave a score between 0 and 100.

CV = 100(t-min t) / (max t-min t)

where 'min t' and 'max t' are the sums of the lowest and highest possible total of percentages in each of the eight positions (-2 to +6) at the splice donor site. 't' is the total of percentages for the same eight positions observed in the splice donor sequence in question.

Two sets of calculations were done for estimating the consensus values of the three splice mutations found in this study:

1)By calculating the 'min t' and 'max t' values for the NF-1 gene from the published splice donor sequences of 49 of its introns (Appendix 2 & 3).

2)Using the 'min t' and 'max t' values used by Shapiro and Senapathy for the various Genebank files (vertebrates) (Appendix 4).

1) Using the values calculated for the NF-1 gene (positions -2 to +6),

min t= 2+0+0+0+4+2+2+13=23

max t= 74+86+100+100+74+81+67+50=632

The consensus sequence determined for positions -2 to +6 of the splice donor sites of the NF-1 gene is (Appendix 3):

-2	-1	+1	+2	+3	+4	+5	+6
A_{74}	G_{86}	G_{100}	T_{100}	A_{74}	A_{81}	G_{67}	T_{50}/A_{22}

1)Consensus values (CV) for the three splice mutations 1721+3A to G in intron 11, 3113 +1G to A in intron 18 and 5749+2T to G in intron 30, using the calculations for the NF-1 consensus sequence are as follows:

't' values for the **normal** (wild type) introns 11, 18 and 30 (the mutated bases are underlined)

INT	-2	-1	+1	+2	+3	+4	+5	+6	't'
11	A74	G86	G100	T100	**A74**	T15	A20	T50	519
18	A74	G86	**G100**	T100	G11	A81	G67	T50	569
30	T16	A10	G100	**T100**	A74	A81	G67	T50	498

't' values for the **mutant** introns 11, 18 and 30 (the mutant bases are underlined)

INT	-2	-1	+1	+2	+3	+4	+5	+6	t
11	A74	G86	G100	T100	**G13**	T15	A20	T50	458
18	A74	G86	**A0**	T100	G11	A81	G67	T50	469
30	T16	A10	G100	**G0**	A74	A81	G67	T50	398

Therefore the consensus Values the various normal (CVn) and mutant (CVm) 5' splice sites were as follows:

NORMAL	CVn	MUTANT	CVm
intron 11	81.4	intron 11	71.4
intron 18	89.7	intron 18	73.2
intron 30	78.0	intron 30	62.0

2)Similarly values were calculated using the 'min t' and 'max t' values calculated for 542 vertebrate 5' SS sequences by Shapiro and Senapathy (1990), which were 'min t'= 47 and 'max t' = 595.

The 't' values were calculated using consensus values in vertebrates for positions -2 to +6 and were as follows:

	Normal t	Mutant t
intron 11	454	436
intron 18	577	477
intron 30	514	414

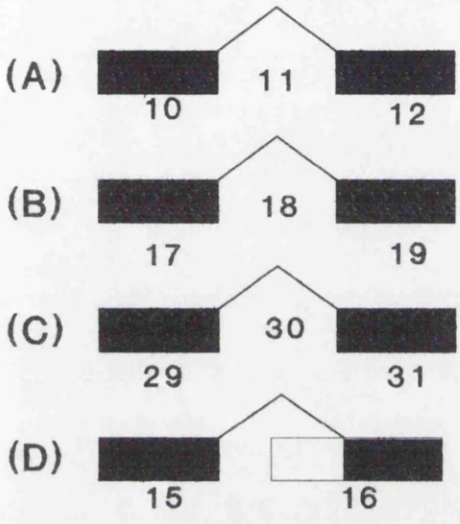
Using the above 't' values, the consensus value (CV) for the normal and mutant introns was as follows:

NORMAL	CVn	MUTANT	CVm
intron 11	74.0	intron 11	70.0
intron 18	96.0	intron 18	78.0
intron 30	85.0	intron 30	66.0

Thus it can be seen that using the consensus values described by Shapiro and Senapathy (1990) the mutations at the splice sites of introns 11 18 and 30 reduce the consensus values by 4%, 18% and 19% respectively. This is similar to the values calculated specifically for the NF-1 gene, where the mutations in introns 11, 18 and 30 cause reduction in consensus values of 10%, 17% and 16% respectively. The relatively larger difference in CV for intron 11 using the values calculated for the NF-1 gene may indicate the importance of calculating values specifically for the NF-1 gene, as this is a more significant reduction (10%) in comparison with the reduction observed using general values for vertebrates (4%).

Therefore these results show that the mutations at the donor splice sites in introns 11, 18 and 30 would be expected to reduce their efficiency, thus causing the skipping of exons 11, 18 and 30 respectively, as was shown earlier.

Figure 53



Mutation	CVn *	CVm *
1721+3A→ G	81.44	71.42
3113+1G→ A	89.65	73.23
5749+2T→ G	77.99	61.57
?Intron 15 splice acceptor mutation	79.3 (15)	84.9

*CV as Shapiro & Senapathy (1990)

The three splice donor mutations (A,B,C) and one possible utilisation of a cryptic splice site (D) are depicted. The splice mutations are listed along with the normal and mutant consensus values (CVn and CVm). In (D) the CVn represents the value for the splice acceptor of intron 15 and CVm that of the deletion junction situated 229bp into exon 16, as a potential splice acceptor. The exons skipped from the NF-1 mRNA are indicated in each case in between filled boxes that represent intact exons.

Possible utilisation of a cryptic splice site due to a splice acceptor mutation

An interesting feature was identified in a patient who showed partial loss of exon 16 from the NF-1 mRNA. The nucleotides missing included positions 2410 to 2639 i.e. from the first base of exon 16 extending for 229 nucleotides into exon 16. When the sequence around the deletion junction was analysed, it appeared to be that of a splice acceptor site.

The sequence at the deletion junction was:

T T C T A T G A T T T C A G | T

where | denotes the deletion junction.

b)The consensus sequence at a splice acceptor site, derived by Padgett et al., (1986) who scanned approximately 400 vertebrate genes was as follows:

Y Y Y Y Y Y Y Y Y Y N C A G G

from positions -14 to +1 where Y is a pyrimidine and N is any base.

The sequence of the normal splice acceptor site of intron 15 (from -14 to +1) was as follows:

A A T T T T G T G T T T A G G

On detailed analysis this was confirmed to be a potential splice acceptor site by calculating the consensus values for splice acceptor sites by the method of Shapiro and Senapathy (1990). On analysis, the sequence at the deletion junction had a consensus value of 84.9. This was found to be higher than the consensus value of the normal intron 15 splice acceptor site which was 79.3 which is normally utilised and may reduce even further if this normal intron 15 site was mutated. Hence the utilisation of a potential cryptic splice site in this case was considered to be a possibility.

The choice between cryptic splice site usage and exon skipping may be visualised as a decision whether to utilise the next available splice site or,

the next best although illegitimate sequence in the immediate vicinity. Cryptic splice site usage is favoured under conditions where a number of such sites are present in the vicinity and these sites exhibit sufficient homology to the splice consensus sequence for them to be able to compete successfully with the mutated splice site (Cooper and Krawczak 1991, 1993).

The partial loss of exon 16 thus could have two explanations at the DNA level. One possibility could be a deletion which involves intron 15 and part of exon 16. The other possibility could be a mutation at the splice acceptor site of intron 15, in which case the next most suitable splice acceptor site, which happens to lie in the first half of exon 16 (seen at the deletion junction) would be utilised leading to partial skipping of exon 16. As DNA from the patient was unavailable for analysis characterisation at only the RNA level was possible. Whatever the cause of this loss of 229 nucleotides from exon 16, the mutant allele has a premature termination codon following the deletion junction and so a truncated protein product of 803 amino acid residues instead of the normal 2818 amino acids is predicted.

4.5.3 INSERTIONS

The majority of insertion type mutations are consistent with an explanation which involves an endogenous replication associated mechanism of mutagenesis. This may be explained either in terms of (a)direct repeats/runs of single bases causing slipped mispairing, (b)inverted repeats or (c)symmetric elements facilitating the formation of secondary structure intermediates (Cooper and Krawczak, 1991).

In this study two insertions, 6519insG in exon 34 and 7485insGG in exon 42 were identified within the NF-1 gene.

The mechanism of insertion 6519insG may be explained by the theoretical model of 'slipped mispairing mediated by direct repeats' (Cooper and Krawczak 1991). The sequence surrounding the insertion is shown and the five direct GA repeats surrounding the site of insertion are underlined (A). The insertion is indicated in lower case (B).

(A) **TCTCCTGGCTCCTATGAGAGAGAGACTTTTGCTTTTGA**

(B) **TCTCCTGGCTCCTATGAGAGAgGAGACTTTTGCTTTTGA**

In principle, slipped mispairing at the replication fork can account for deletion and insertion type mutations. An insertion takes place when the newly synthesised strand disconnects from the primer strand during replication synthesis and slips or folds back so that pairing between different direct repeat copies becomes possible. If synthesis is resumed so as to stabilise this mispairing, DNA sequence from between the direct repeats is inserted. In 6519insG, the single guanine base has been inserted in a series of five GA repeats and the insertion of G has occurred at the level of A in the third GA repeat.

On the other hand, the insertion of 7485insGG can be explained by a more complex mechanism. Cooper and Krawcak (1991) have seen in their studies of the sequence context of short gene deletions and insertions, a significant excess of symmetric elements (e.g. CTGAAGTC). On analysis of the sequence surrounding the 2 base insertion, the following was identified (the insertion is indicated in lower case letters):

AAAGGACCGAGCCCCTGACC

AAAGGACCGAGggCCCCTGACC

A symmetric element CCGAGCC (underlined) is seen within the sequence where the insertion has occurred. Its relevance can be explained by predicting a stemloop formation within the sequence. Imperfect self

complementarity among the stemloops can mediate the formation of partially misaligned secondary structures. The non-palindromic portions of this structure then provide a template putatively for insertions by gap repair. The DNA sequence flanking the insertion correctly predicts the insertion of the appropriate base at the appropriate site.

The stemloop formed in the sequence surrounding the insertion was predicted using a sequence analysis program "STEMLOOP" (Genetics Computers Group GCG), which is shown in Figure 54.

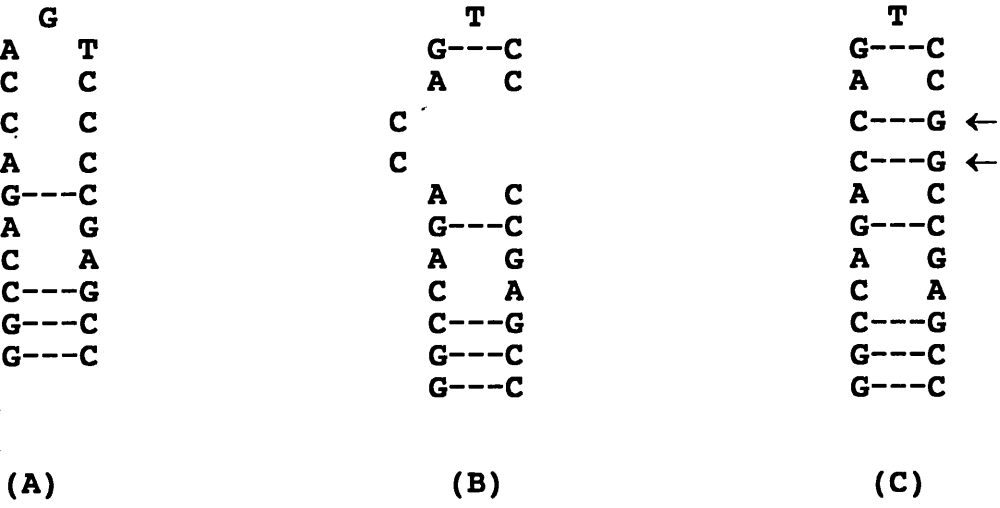


Figure 54

(A) shows the predicted stemloop pattern prior to the GG insertion. The insertion of GG is templated by the base sequence within the hairpin (stemloop).

(B) shows the formation of a partially misaligned secondary structure due to imperfect self complementarity within the stem loop.

(C) shows the prediction of 2 bases (GG) complementary to the two C residues. These sequence changes (insertion of GG) (indicated by arrows) would be predicted to result in increased stability of the hairpin loop.

4.6 CORRELATION OF THE GENOTYPE WITH THE PHENOTYPE

Inherited traits are defined by their ability to be passed from one generation to the next in a predictable manner. Visible or otherwise measurable properties are called the phenotype, while the genetic factors responsible for creating the phenotype are called the genotype. The phenotypic results of mutations vary from the undetectable (no defect appears to result from the absence of the protein), to lethal (the organism cannot survive without the protein).

Table 13 compares the clinical features and complications of NF-1 in all the patients in this study whose mutations were fully characterised and predicted to be pathogenic, thus allowing a comparison of the genotype with the phenotype.

Figure 55 shows a diagrammatic representation of the predicted products of translation in these mutations.

TABLE 13 :GENOTYPE-PHENOTYPE CORRELATION

Age, sex, type	Mutation	1	2	3	4	5	6	7	Complications
30 yrs, female familial	1721+3A→G intron 11	+	+	+	+	+	-	-	plexiform neurofibroma left arm, parasthesia L-arm, ↓ power R-arm, ↑ reflexes R-leg, abnormal gait
41 yrs, female familial	3113+1G→A intron 18	+	+	-	-	+	-	-	grand mal epilepsy, sciatica, macrocephaly
29 yrs, male sporadic	5749+2 T→G intron 30	+	+	+	-	-	-	-	sensorineural deafness, abnormal EEG, rt sided deviation of palate, left convergent squint, pes cavus
11yrs, male familial	G1166D exon 21	+	-	+	-	+	-	-	short stature
19yrs, male sporadic	K1419R exon 24	+	+	+	-	-	-	-	none
34yrs, female familial	6519insG exon 34	+	+	+	-	+	-	-	none
8yrs, male sporadic	7486insGG exon 42	+	+	+	+	-	-	-	learning difficulties
23yrs, female sporadic	R2486X exon 42	+	+	+	+	-	-	-	none
12yrs, male sporadic	partial loss of exon 16 ? splice/del	+	-	+	-	-	-	-	macrocephaly, precocious puberty,CT scan showing intracranial calcification and dysplastic cerebellum
59yrs, male familial	exon 18 ? splice/del	+	+	+	-	+	-	-	leiomyosarcoma of colon at 34 yrs

1=cafe au lait patches, 2=neurofibromas, 3=axillary freckling, 4=Lisch nodules, 5=NF-1 relative 6=glioma 7=osseous dysplasia

Figure 55

PREDICTION OF PRODUCTS OF TRANSLATION

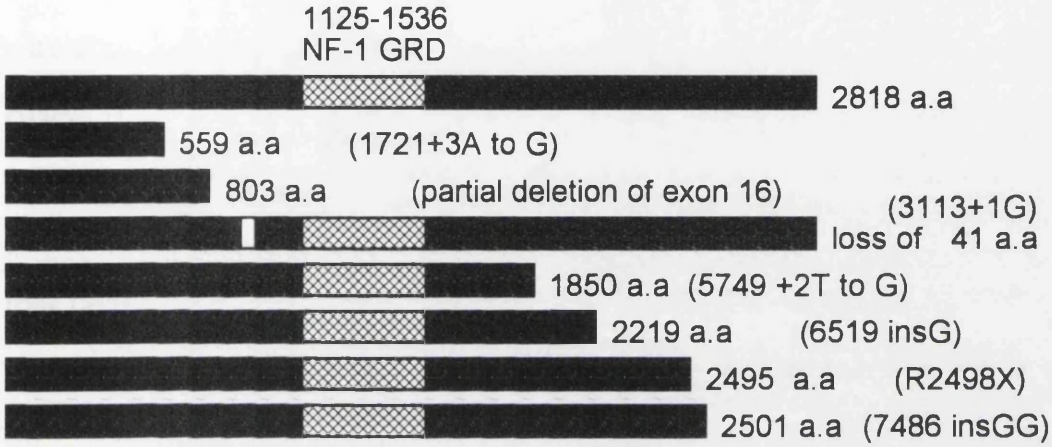


Figure 55 shows the prediction of products of translation in each of the mutations, predicted to cause disruption of the product of translation. A diagrammatic representation of normal and mutant neurofibromins is seen in relation to the NF-1 GAP related domain which extends from positions 1125-1536 (indicated by the hatched box). The sizes of the products of translation are predicted presuming the predicted mutant neurofibromins are stable.The causative mutation and the size of the product of translation is indicated on the right

4.6.1 PREDICTION OF THE PRODUCTS OF TRANSLATION

Seven patients were identified with mutations that were predicted to result in either loss of amino acids or truncation of neurofibromin (Figure 55). The mutations were as follows:

The mutation at the splice donor sequence of intron 11 (1721+3A to G) caused skipping of exon 11 from the NF-1 mRNA in the proband, who was the first case of NF-1 in her family. However, she had an affected son, who was shown to inherit the mutation from his mother by analysis of his DNA and RNA. The skipping of exon 11 caused a shift in the reading frame and a premature termination of translation, resulting in the synthesis of a shorter protein product of 559 amino acids. The mutant neurofibromin predicted in this case would not be expected to stimulate the GTPase, as it would not possess the NF-1 GAP related domain. As the microtubule binding domain also lies distal to the site of premature termination, this mutant neurofibromin would also not have the putative microtubule mediated signal transduction function. The patient does not have any evidence of malignancy, but has neurological complications of NF-1 which include parasthesia of the left hand, decreased power and increased reflexes in the right arm and leg and abnormal gait. The patient had plexiform neurofibromatosis of the left arm for which surgery was undertaken. Her affected son has pseudarthrosis of the left tibia.

The partial loss of exon 16 was detected in a sporadic 12 year old male with NF-1. The loss of 229 bases from exon 16 is predicted to shift the translational reading frame and result in the formation of a premature stop codon and a shortened protein product of 803 amino acids which would lack the NF-1 GRD. In addition to NF-1, the patient has precocious puberty, an abnormal CT scan showing intracranial calcification, dysplastic cerebellum and complications of NF-1 including an optic glioma.

Thus, two patients have been detected in whom the loss of the NF-1 GRD is predicted due to truncation of their protein products. Due to predicted loss of the NF-1 GRD, malignancies would be expected in these patients as the NF-1 GRD is involved in ras regulatory pathways controlling cell proliferation and differentiation. The patient with the intron 11 splice donor mutation has, as yet no evidence of malignancy. However it should be taken into account that this patient had a plexiform neurofibroma and it is known that some plexiform neurofibromas may progress to malignant neurofibrosarcoma. This risk could not be assessed as the plexiform neurofibroma had been operated upon and removed in this patient. The patient with the partial loss of exon 16 does have a malignant complication of NF-1 (optic glioma).

Disease causing mutations within the GAP related domain of the NF-1 gene have been reported in patients with NF-1 and include an A to G transition at codon 1423 in a familial case of NF-1 (no details about the presence/absence of malignancy were published) reported by Li et al., (1992). The same mutation was also identified in two different tumour types unrelated to NF-1. A single base deletion of thymidine at position 4190 has also been reported by Anglani et al., (1993) in a sporadic case of NF-1 (no details about the presence/absence of malignancy in this patient were published).

In the present study, analysis of the GAP related domain revealed four characterised and two partially characterised mutations. The 4 characterised mutations are all single base substitutions, (two missense G1166D & K1419R and two silent mutations G1404G & S1311S) which are not predicted to cause truncation of neurofibromin. The two patients with the

missense mutations within the NF-1 GRD (G1166D and K1419R) did not have any evidence of malignancy.

The other two mutations include a chemical cleavage result in exon 27a of the NF-1 gene and a possible deletion of exon 23 of the NF-1 gene. The patient with the chemical cleavage result in 27a (which lies within the GAP related domain) herself did not have any malignancy, but had a positive family history of NF-1 and cancer. The proband's mother with NF-1 died of oesophageal cancer and her maternal uncle with NF-1 died of a brain tumour.

The patient with a possible deletion of exon 23 within the NF-1 GRD also has a positive family history of NF-1 and malignancy. The patient is a 17 year old familial case of NF-1. The deletion was suspected due to loss of heterozygosity at a newly detected rare sequence variant in exon 23. The patient does not as yet have any evidence of malignancy, but had an affected mother who died of malignant complications of NF-1 (brain tumour).

It is interesting to note that the 5 patients in the study reported by Kayes et al. (1993, 1994) have complete deletions of the NF-1 gene (which would include loss of its GAP related domain) and have no evidence of malignancy.

The loss of exon 18 from the NF-1 mRNA was detected in two unrelated patients. In one the cause was identified to be a splice mutation causing skipping of exon 18 and in the other case, the cause at the DNA level (splice mutation/deletion) remains to be determined. Both patients would have an identical predicted protein product with the loss of 41 amino acids including two cysteine residues at positions 1016 and 1032. Cysteine residues are important in covalent bonding by the formation of S-S disulphide bridges, which are important for the maintenance of conformation of

polypeptide chains. Loss of the two cysteine residues (1016, 1032) may result in an altered conformation of neurofibromin, which could result in reduced stability/function.

The splice mutation (3113+1G to A) in intron 18 causing skipping of exon 18 was detected in a familial case of NF-1 and in her affected son. The proband is a 41 year old Caucasian female, who was first diagnosed with neurofibromatosis at the age of 28 years, when she was admitted to the hospital for grand mal seizures. The patient has short stature and several cafe au lait spots and widespread neurofibromas on her trunk and limbs which appeared in early childhood and a head circumference of 56cm (+1S.D.). The proband has seven unaffected sisters and one son who is affected. He has cafe au lait spots and neurofibromas which were noticed on his back, at the age of 2 years. The son has a head circumference of +2 S.D., his height is around the 25th centile for his age and has been treated for squint.

The other patient who was shown to have exon 18 missing from his mRNA by RT-PCR and direct sequencing, is a 59 year old Caucasian male, with cafe au lait macules, neurofibromas, axillary freckling and has a history of two unrelated malignancies in the family. The proband had a leiomyosarcoma of the colon which was diagnosed at 31 years of age and an affected son who died of gliobalstoma, which is a malignant complication of NF-1.

A splice mutation was identified in the splice donor site of intron 30 (5749+2T to G) in a sporadic case of NF-1. The mutation caused an error of splicing which resulted in the skipping of exon 30. The patient's unaffected mother, father and brother were screened and shown not to have the same mutation. Thirty normal controls were similarly screened by SSCP and were

shown not to have this mutation. The patient has neurological complications of NF-1 which include sensorineural deafness, right sided deviation of the palate, left convergent squint and an abnormal EEG. In addition the patient also has skeletal deformities which include pes cavus.

A single base insertion of guanine (6519insG) was identified in exon 34 of the NF-1 gene. The proband is a 40 year old Caucasian female with four affected children and a family history of NF-1 involving four generations. The patient shows typical skin stigmata of NF-1 and has no complications of NF-1.

An interesting feature presented in two cases involving exon 42, where a nonsense mutation and an insertion of 2 guanine bases were seen to affect the same codon (Arginine 2496) in two unrelated sporadic cases. The patient with the nonsense mutation had a predicted protein product of 2495 amino acids and the patient with the two base insertion had a predicted product of 2501 amino acids. The clinical features in the two cases were compared. Both patients showed typical features of NF-1 including multiple cafe au lait patches, neurofibromas, axillary freckling and Lisch nodules. The patient with the nonsense mutation had no complications of NF-1, but the patient with the 2 base insertion had specific learning difficulties. He has an IQ of 110, but has problems with complex visual data. This was a cognitive problem, as there was no abnormality in visual acuity. The patient has marked grapho-motor difficulties, high activity impulsivity and a short attention span.

A study of the neuropsychological profile in 45 children with NF-1 was carried out by Legius et al., (1994a). Mental retardation (total IQ<70) was noted in 5% of the children. The mean IQ of the entire group was 90, with a significantly higher verbal IQ than performal IQ. A weaker score was seen on

perceptual organisation as compared to verbal comprehension. Legius et al., (1994a) concluded that learning disabilities were frequent in the children analysed and nonverbal learning disabilities were the predominant type identified in 40% of the children. The cognitive phenotype in NF-1 was also assessed by Hofman et al., (1994) who studied 12 families each comprising an affected child with NF-1 (6-16 years), an unaffected sibling of the same age and the biological parents using neurocognitive testing and magnetic resonance imaging (MRI). The full scale IQ ranged from 70-130 among children with NF-1 and 99-139 in unaffected siblings. Children with NF-1 were seen to have significant learning disabilities in written language, reading and neuromotor dysfunction. Hofman et al., (1994) concluded that NF-1 is associated with a significantly lower full scale IQ, multifocal cognitive deficits, reading disability and neuromotor deficit and the cognitive differences correlated with the number of brain lesions on magnetic resonance imaging.

4.6.2 PATHOGENICITY OF MISSENSE MUTATIONS

While it is clear how insertions and deletions cause disease, missense mutations have to be evaluated for their pathogenicity. This can be done in the following ways:

1. Segregation analysis: the novel appearance and subsequent co-segregation of the gene lesion and disease phenotype through a family pedigree.
2. Population studies: the failure to observe such a mutation in a sufficiently large sample of normal controls.

3. Occurrence of the mutation in a region of known structure or function: by analysis of commonly occurring functional domains.
4. Occurrence of the lesion in an evolutionarily conserved residue: conservation among other homologous proteins and across species.
5. The previous independent occurrence of the mutation in an unrelated patient, the mutation being the only change on screening the entire coding sequence and other regions important for efficient expression.
6. Functional studies to determine the effect of the mutation on the protein: by demonstrating the pathological phenotype in the cultured cells by expression of the mutant gene/protein and its wild-type counterpart.

In the case of the two missense mutations detected in this study, G1166D in exon 21 and K1419R in exon 24, the following analyses were carried out:

1)G1166D: This mutation was identified in a familial case of NF-1 where the mutation in the proband was transmitted from his affected mother. The mutation affected a rare restriction site and this single base substitution could not be detected by heteroduplex analysis, hence CCM was used to screen normal controls, none of whom showed this mutation.

The mutation results in the conversion of glycine, which is a neutral and polar amino acid to aspartate which is acidic. The amino acid glycine lacks a β-carbon and can assume many backbone dihedral angles not possible by other amino acids. Due to this, glycine can be used in certain types of reverse turns where other amino acids are not suitable. Replacing glycines in such turns with any other residue would be expected to be destabilising, unless the protein could form an alternative type of turn (Pakula and Sauer, 1989). However due to lack of structural information for the NF-1 protein, it is not possible to determine if the glycine 1166 residue is used in

this manner, hence the effect of the substitution of this residue by aspartate cannot be predicted.

2)K1419R: This mutation was identified in a sporadic case of NF-1. Fifty normal controls were screened by restriction analysis with the restriction enzyme *Mnl* I (the mutation created a new *Mnl* I site) and none of the controls were shown to have this mutation. The mutation causes conversion of lysine to arginine, both amino acids being basic in nature and this change would normally be considered to be a conservative one. However, there are known examples of conservative missense mutations that can result in a disease phenotype if they occupy sites in the protein that are key determinants of stability or function (Pakula and Sauer, 1989).

The NF-1 GRD spans exons 20-27a of the NF-1 gene and both mutations lie within this region. These missense mutations were also analysed for their involvement with common functional domains mapped on the human NF-1 sequence, using the 'MOTIF' sequence analysis program (GCG) (Smith et al., 1990). This program searches for sequence motifs by scanning through a protein sequence for patterns catalogued and defined in the PROSITE dictionary of protein sites and patterns (Bairoch et al., 1991). The two missense mutations were found not to lie within important functional sites. All the important sites within the NF-1 gene according to the 'MOTIF' sequence analysis program, are described in Appendix 5.

Attribution of these missense mutations to disease can thus only be confirmed by (a)screening the remainder of the coding sequence and other regions important for gene expression to exclude other causative mutations and (b)by functional studies.

4.6.3 OVEREXPRESSION OF THE NEUROFIBROMIN ISOFORM NF-1 GRD-II

An interesting feature was detected in a case of a variant form of neurofibromatosis, in a 44 year old female patient with segmental neurofibromatosis affecting the right upper arm. The mutation arose *de novo* in this patient, as there was no family history of the disease. The patient had a plexiform neurofibroma affecting her right ulnar nerve. She also has some permanent loss of function of that nerve and several small neurofibromas along the line of the ulnar nerve. Two neurofibromas in her right upper arm were confirmed histologically to be neurofibromas. The patient had no other cutaneous manifestations and was also examined for the presence of Lisch nodules which she did not have.

Upon RT-PCR amplification of RNA segment VI spanning part of the NF-1 GRD agarose gel electrophoresis showed the presence of an approximately sixty base pair insertion. This was confirmed by CCM analysis and shown to be an insertion of 63bp which has been described by Marchuk et al., (1991). This 63 base pair insertion represents the alternatively spliced isoform of the NF-1 gene called the NF-GRD II. This insertion is present in block 2 of the NF-1 GRD which is a region of strong homology between GAP, IRA1, IRA2 and neurofibromin (Andersen et al., 1993a). This insertion significantly weakens the statistical significance of this homology block. (Andersen et al., 1993a).

A study by Nishi et al., (1991), demonstrated that the type I transcript (NF-GRD I) is predominantly expressed in undifferentiated cells, while the type II transcript (NF-GRD II) predominates in differentiated cells.

Suzuki et al., (1991) have shown that the type I transcript (NF-GRD I) is predominantly expressed in normal brain tissue, while the type II transcript (NF-GRD II) predominates in brain tumours.

Using RT-PCR analysis Andersen et al., (1993a) were not able to detect evidence of consistent tissue specificity differences of NF-GRD I and NF-GRD II, in a variety of normal and abnormal tissues. The authors found at least some evidence of both forms in all tissues analysed.

The additional 63bp insertion in the type II transcript reflects the insertion of a separate exon by alternative splicing. The evolutionary conservation of the type II form supports the conclusion that both the forms are important for neurofibromin function. However in *in vitro* biochemical assays, type II transcripts lack apparent GTPase stimulating activity with H-ras. This suggests that the 21-amino acid insertion results in the reduction of the catalytic activity of NF-GRD II.

Using primers designed in this study, only amplification of the type I transcript was detected in the remaining patients with NF-1. Expression of both type I and type II transcripts (with the 63bp insertion) was detected in only the above mentioned patient. This would mean that in this assay, using these sets of primers, either only the smaller product (NF-GRD I) was being preferentially amplified in all other patients examined, or that the amplification from the type II transcript was in very small quantities, which would not be visible on agarose gel electrophoresis. However, the reproducible (three RT-PCR amplifications) detection of NF-GRD II transcript in the above patient would indicate that at least for this assay, the patient was producing equal quantities of type I and type II transcripts in comparison with other patients who produced only type I transcripts.

Thus this observation is an interesting finding in a patient with segmental neurofibromatosis. In this patient, a somatic mutation in that particular tissue which would lead to the manifestations of localised neurofibromatosis is expected instead of a germline mutation. This finding cannot yet be attributed to disease but was the only finding detected in this patient after screening 78% of her NF-1 coding sequence.

4.6.4 UNEQUAL EXPRESSION OF NF-1 ALLELES

An interesting finding was detected as a chemical cleavage result of RNA segment VI in a familial case of NF-1. Direct sequencing revealed this to be a C to T substitution, maintaining serine at position 1311 (S1311S). Upon direct sequencing, only the altered allele was seen. This was detected on sequencing both strands and on restriction analysis. The mutation also causes loss of a *Fok* I restriction site. Upon digestion of the PCR product, only the allele with the silent mutation (undigested due to loss of the restriction site was seen). This could be accounted for by one of the following:

(a)One possibility is that this could be a common polymorphism and this patient is homozygous for it. This could not be confirmed by family studies, as the father of the patient was unavailable for analysis and her affected mother had died due to a malignant brain tumour. Thirty other individuals screened in this region did not show a similar chemical cleavage result, i.e. this was probably a rare sequence variant and not a common polymorphism. Thus homozygosity due to a polymorphism does not appear likely.

(b)The second possibility and probably the most likely, is that the patient has a large deletion in this region on one allele and a silent mutation on the other allele, hence there was no amplification from the allele with the deletion. This

mutation causes the loss of a *Fok* I site. On amplification and digestion with *Fok* I and electrophoresis on an agarose gel, only the allele with the silent mutation, which was undigested due to loss of the restriction site was detected. Ideally there should have been digestion of the PCR product from the normal allele (showing complete digestion) and the mutant allele showing no digestion due to loss of the restriction site (Figure 31). Thus in this case it appeared that the other allele in this patient had a large deletion that involved at least one primer binding site and therefore was not being amplified.

(c)Another reason could be unequal expression of the two NF-1 alleles: Disproportionately low amounts of mRNA from mutant alleles in patients heterozygous for stop mutations has been reported in the literature, (Jones et al., 1992). A transcribed polymorphism in exon 32 of the NF-1 gene was reported by Hoffmeyer et al. (1994). Using the polymorphic *Rsa* I restriction site in exon 32 of the NF-1 gene (in patients heterozygous for the polymorphism) the relative abundance of mRNAs transcribed from both alleles was evaluated by Hoffmeyer et al., (1994). Putative mutations lying outside the coding region but in regions essential for gene expression will not have a direct effect on the NF-1 protein, but may affect the expression of the mutant allele or the stability of the mutant messengers. This may therefore lead to the unequal representation of wild type and mutant messengers in the cells of NF-1 patients.

Thus S1311S is a silent mutation on the allele being expressed. The other allele may not be expressed due to a deletion or a mutation causing decreased expression or increased instability of the NF-1 mRNA.

In conclusion, the strategy of using both DNA and RNA as templates for PCR amplification and use of chemical mismatch cleavage analysis as the main screening method, has proved to be a useful screening strategy for detection of mutations within the NF-1 gene.

The advantages of using both RNA and DNA as templates for PCR in mutation analysis in the NF-1 gene are as follows:

The NF-1 gene is ubiquitiously expressed and the NF-1 mRNA can be easily obtained from peripheral blood leukocytes and used for analysis after conversion to cDNA by reverse transcription. Using NF-1 mRNA as a template allowed the amplification of large areas of the coding sequence. Using mRNA as a template it was possible to begin screening for mutations in the NF-1 gene before the complete genomic organisation was known (initially intron-exon boundaries for only 9 of the 49+ exons were published, Cawthon et al., 1990a). Amplification using RNA segments readily allowed the detection of large alterations within the NF-1 coding sequence (five large alterations were detected in this study).

Amplification of DNA allowed the easy and rapid screening of relatively smaller fragments comprising individual exons of the NF-1 gene. Genomic DNA amplification was necessary for the characterisation of splice site mutations after detection of an altered transcript on RT-PCR. Genomic DNA was useful in the complete characterisation of deletions or insertions which were also seen as alterations on amplification of the NF-1 cDNA. Genomic DNA allows equal amplification of both alleles. This is useful in those cases where mutations cause unequal levels of mRNA production on amplification from mRNA.

In this study most mutations were confirmed by amplification of both DNA and cDNA. This was done in order to rule out amplification from NF-1

homologous sequences. Various combinations of intron and exon based primers were used for this purpose. In this study, a novel NF-1 homologous sequence was identified. Direct sequencing revealed the sequence to be homologous to exon 18 and the flanking introns.

The screening methods used in this study included SSCP and chemical cleavage analysis. Initially, SSCP analysis was used to screen exons 28-36. However the sensitivity of SSCP falls to < 50% in products of 400bp or higher, (Grompe M., 1993) and it is possible that using this technique some mutations were missed. Therefore all the products of DNA PCR were eventually analysed using chemical cleavage analysis. The segments amplified using RNA as a template ranged from 400bp - 1kb and chemical mismatch cleavage was used directly for screening these fragments.

Chemical mismatch cleavage analysis allows the analysis of relatively large fragments of up to 1.7 kb in length (Grompe,M., 1993). CCM detects greater than 95% of mutations when only the probe is labelled and almost 100% of mutations when both the wild-type probe and the mutant samples are labelled in separate reactions (Forrest et al., 1992). Chemical cleavage allowed accurate localisation of the mutation, as mismatches reduce the size of the end labelled probe after cleavage, thus enabling sequencing of a limited area instead of the whole product. Chemical cleavage also detects the nature of the altered base depending on the modifying chemical i.e. C and G residues using modification with hydroxylamine and A and T residues using osmium tetroxide, which further confirms the nature of the mutation on sequencing. In this study chemical cleavage analysis was also useful in determining the boundaries of deletions and insertions in the case of gross RNA rearrangements. CCM analysis also has to some extent the advantage

of exclusion of mutations in a region that has been screened and found negative for mutations, provided suitable controls are done with each analysis. Thus CCM analysis was a useful screening method in detection of mutations within the NF-1 gene and was used as the main screening method.

Nine tumour types unrelated to NF-1 (a total of 29 samples) were analysed for somatic mutations within the NF-1 gene. No small alterations were detected on screening the tumours for mutations in exon 24 (FLR exon) within the NF-1 GRD by CCM analysis. Four neuroblastoma samples were screened for large alterations by Southern (p5/*Eco*R I) and Northern blotting and they did not reveal any alterations.

This study has been able to identify a positive screening result in 17 out of 25 patients on analysis of 78% of the NF-1 coding sequence. Thirteen of these have been characterised, demonstrating a range of mutations including deletions, insertions and single base substitutions. Ten of these mutations are potentially disease causing and eight of them are predicted to cause disruption of neurofibromin.

Most of the mutations detected within the NF-1 gene (this study as well as other reported mutations) appear to be small sequence alterations resulting in disruption of the protein product neurofibromin.

Nine single base substitutions were detected including 3 splice site mutations, 1 nonsense mutation, 2 missense and 3 silent mutations. Only one of these 9 mutations (R2486X) affected a CpG site.

Except R1947X in exon 31 (seven independent reports of the same mutation) no mutational hotspot has been observed. However since 22% of the coding sequence and regions important for efficient gene expression have not yet been analysed in this study, a clustering of mutations in the yet unanalysed part of the NF-1 gene cannot be ruled out.

This study compares the genotype with the phenotype in a heterogeneous group of mutations affecting the NF-1 gene. No definitive correlation can be made based on these observations and a multicentre study, facilitating the analysis of a large number of mutations would be helpful.

APPENDIX 1

Solutions used for extraction of DNA from peripheral blood

a)Lysis mix	0.32 M Sucrose 10mM Tris-HCl (pH 7.5) 5mM $MgCl_2$ 1% Triton X-100
b)Nuclei lysis mix (pH 7.8)	10mM Tris-HCl 0.4M NaCl 2mM EDTA
c)T.E. buffer	10mM Tris-HCl (pH 7.5) 1mM EDTA
d)Proteinase K (10mg/ml)	Boehringer Mannheim

Solutions used for extraction of DNA from parrafin embedded tissue

a)Proteinase K buffer	50mM Tris-HCl pH 8.7 10mM NaCl 0.1% Triton-X 100

Solutions used for extraction of DNA from tissue culture cells

a)Lysis buffer	100mM Tris-HCl pH 8.5 5mM EDTA 0.2% SDS 200mM NaCl
b)Proteinase K (100μg/ml)	Boehringer Mannheim

Solutions for extraction of RNA

a)Solution D	4M guanidinium thiocyanate (Fluka) 25mM sodium citrate 0.5% sarcosyl (Sigma) 0.1mM β-mercaptoethanol (Sigma)
b)2M sodium acetate (pH 4.0)	
c)Phenol	water saturated
d)Chloroform:isoamyl alcohol	49:1
e)Histopaque '1077'	Sigma

f)Phosphate buffered saline (pH 7.2)	0.14M NaCl 3mM KCl 0.1mM sodium hydrogen phosphate 1mM potassium dihydrogen phosphate alternatively PBS (Gibco BRL) was used

Solutions for the preparation of the RNA sample

a)10 X MOPS	200mM MOPS 50mM sodium acetate pH 7.0 10mM EDTA
b)Formamide	
c)37% formaldehyde solution	
d)Loading mix	2g Ficoll 1ml 2.5% bromophenol blue 1ml 2.5% xylene cyanol 0.2ml 0.5M EDTA 8 ml water

Solutions used in PCR

(All reagents were obtained from Boehringer Mannheim)

a)10 X PCR buffer	500mM KCl 100mM Tris-HCl (pH 8.4) 15mM $MgCl_2$ 1000μg/ml gelatin
b)dNTP (dTTP, dCTP, dGTP, dATP)	100mM each
c) Taq DNA polymerase	5U/μl

Solutions used for reverse transcription

(All solutions were from Gibco BRL)

a)Reverse transcriptase buffer (5X)	250mM Tris-HCl (pH 8.0) 375mM KCl 15mM $MgCl_2$
b)DTT (dithiothreitol 0.1M)	
c)MMLV reverse transcriptase	200 units/μl

Solutions used for chemical mismatch cleavage analysis (CCM)

a)Geneclean II kit (BIO 101)	1)6M sodium iodide (NaI) 2)NEW wash (pH 7.5-8.0). 14 mls of the NEW concentrate, 280 mls of distilled water, 310 mls of 100% ethanol. 3)Glassmilk
b)T4 kinase buffer (10 X)	500mM Tris-HCl pH 7.6 0.1M DTT 0.1M $MgCl_2$
c)Annealing buffer (10 X)	3M NaCl 1M Tris-HCl pH 8.0
d)Stop/precipitation mix	63mM sodium acetate 20μM EDTA 80% ethanol

Details of chemical reagents used for CCM analysis

1)T4 polynucleotide kinase	8U/μl (NBL)
2)Glycogen	Boehringer Mannheim
3)Hydroxylamine hydrochloride	BDH
4)Diethylamine	BDH
5)Osmium tetroxide	4% solution in water, Aldrich
6)Pyridine	BDH
7)Piperidine	Fluka

Solutions used for sequencing

(All solutions were supplied by USB - Sequenase version 2.0 kit)

a)Sequenase buffer (5 X)	200 mM Tris-HCl pH7.5 100mM $MgCl_2$ 250mM NaCl
b)Dithiothreitol (DTT)	0.1M
c)Labelling mix (dGTP) (5 X)	7.5μM dGTP 7.5μM dCTP 7.5μM dTTP
d)Labelling mix (dITP) (5 X)	15μM dITP 7.5μM dCTP 7.5μM dTTP
e)Termination mix (ddG,T,C or A)	8μM ddG,T,C or ATP 80μM dNTP 50mM NaCl

f)Enzyme dilution buffer	10mM Tris-HCl pH 7.5 5mM DTT 0.5mg/ml BSA
g)Stop solution	95% formamide 20mM EDTA 0.05% bromophenol blue 0.05% xylene cyanol

Plamid solutions

a)Luria broth (L-broth 1 litre pH 7.5)	10g Bactotryptone 5g Bactotryptone yeast extract 5g NaCl alternatively L-Broth (BRL) was used
b)Circleprep kit	Pre-lysis buffer: Tris/EDTA/glucose lysis reagent: 0.2%NaOH, 1%SDS. Neutralisation solution: 3M potassium acetate Lithium chloride solution glassmilk Binding buffer: KBr/NaI/Tris mixture Wash solution Sieve material
c)Instaprep kit	Insta-Mini-Prep tubes Phenol:Chloroform:Isoamyl alcohol (50:49:1) Chloroform:Isoamyl alcohol (49:1) TE (10mM Tris, 1mM EDTA, pH 8)
d)STE buffer	100mM NaCl 10mM Tris 1mM EDTA pH 8

Solutions used for Southern blotting

a)Denaturation solution	0.5M NaOH 1.5M NaCl
b)Neutralisation solution (pH 7.5)	0.5M Tris-HCl 3M NaCl
c)Depurination solution	0.25M HCl
d)Prehybridisation solution	50% Formamide 5 X Denhardt's solution 5 X SSC 2% SDS

Solutions used for Northern blotting

a)Pre hybridisation solution

5 X SSC
1 X Denhardts solution
100μg/ml salmon sperm DNA
0.5% SDS
50% formamide

b)Hybridisation solution

5 X SSC
1 X Denhardt's solution
10% Dextran sulphate
100μg/ml salmon sperm DNA
50% formamide
0.5% SDS
2 X 10^6 counts/ml labelled probe

APPENDIX 2

SEQUENCE AT 5'SS POSITIONS -4 TO +6 IN INTRONS OF THE NF-1 GENE

exon	-4	-3	-2	-1	+1	+2	+3	+4	+5	+6
1	G	C	A	G						
2	T	A	T	G	G	T	G	A	G	T
3	T	G	G	G	G	T	A	A	G	T
4	A	A	A	G	G	T	A	A	G	T
5	G	C	T	G	G	T	A	A	G	G
6	T	A	A	G						
7	T	A	A	G	G	T	A	A	T	A
8	T	A	A	G	G	T	G	A	G	G
9	C	A	A	T	G	T	A	A	G	T
10	G	G	A	G						
11	T	T	A	G	G	T	A	T	A	T
12A	T	A	A	G	G	T	A	A	G	C
12B	T	A	T	G	G	T	C	A	G	C
13	A	C	A	G	G	T	C	A	A	A
14	T	G	A	G	G	T	A	T	G	C
15	C	C	A	G	G	T	A	A	G	T
16	A	C	A	G	G	T	A	A	A	G
17	T	C	A	G	G	T	A	A	G	C
18	T	T	A	G	G	T	G	A	G	T
19	T	T	A	A	G	T	A	A	A	T
20	A	T	A	G	G	T	G	A	G	A
21	G	T	G	G	G	T	A	A	G	T
22	C	A	A	G	G	T	T	T	G	T
23	C	C	A	G	G	T	A	T	G	C
24	A	A	A	G	G	T	G	A	A	T
25	G	C	A	G	G	T	A	A	T	T
26	A	C	A	G	G	T	A	A	G	A
27A	C	T	A	G						
27B	G	G	A	G	G	T	A	A	G	A
28	T	A	A	A	G	T	A	A	G	T
29	T	A	C	G	G	T	A	G	G	T
30	T	C	T	A	G	T	A	A	G	T
31	G	C	A	G	G	T	A	T	T	G
32	C	A	A	G	G	T	A	A	T	C
33	A	G	T	G	G	T	A	A	G	T
34	G	G	A	G	G	T	A	T	A	G
35	A	A	A	G	G	T	A	T	G	T
36	C	A	A	G	G	T	A	C	C	T
37	T	A	A	G	G	T	A	A	T	T
38	C	A	A	G	G	T	A	A	G	C
39	A	A	A	G	G	T	A	A	A	A
40	G	C	A	G	G	T	A	A	A	A
41	A	T	A	G	G	T	A	A	G	T
42	C	T	T	G	G	T	T	A	G	T
43	A	T	G	G	G	T	G	A	G	A
44	T	C	T	A	G	T	A	A	G	G
45	T	T	G	T	G	T	A	A	G	T
46	C	A	A	A	G	T	C	A	A	A
47	A	A	A	G	G	T	A	A	G	A
48	C	C	A	G	G	T	C	A	G	T

APPENDIX 3

CALCULATION OF CONSENSUS SEQUENCES FOR POSITIONS -4 TO +6 OF THE NF-1 GENE

POSITION	TOTAL NUMBER OF T/C/G/A	%	CONSENSUS
-4	**T**=18/51	36%	**T**=36%
	C=12/51	23%	**C/A**=23%
	G=9/51	18%	THEREFORE
	A=12/51	23%	**T/C/A** OR **N**
-3	**T**=10/51	19%	**A**
	C=14/51	27%	
	G=7/51	14%	
	A=30/51	59%	
-2	**T**=8/51	16%	**A**
	C=1/51	2%	
	G=4/51	8%	
	A=38/51	74%	
-1	**T**=2/51	4%	**G**
	C=0/51	0%	
	G=44/51	86%	
	A=5/51	10%	
+1	**T**=0/46	0%	**G**
	C=0/46	0%	
	G=46/46	100%	
	A=0/46	0%	
+2	**T**=46/46	100%	**T**
	C=0/46	0%	
	G=0/46	0%	
	A=0/46	0%	
+3	**T**=2/46	4%	**A**
	C=4/46	9%	
	G=6/46	13%	
	A=34/46	74%	
+4	**T**=7/46	15%	**A**
	C=1/46	2%	
	G=1/46	2%	
	A=37/46	81%	
+5	**T**=5/46	11%	**G**
	C=1/46	2%	
	G=31/46	67%	
	A=9/46	20%	
+6	**T**=23/46	50%	**T/A**
	C=7/46	15%	
	G=6/46	13%	
	A=10/46	22%	

APPENDIX 4

Consensus nucleotide frequency patterns at 8 positions within 5' and 3'splice sites (Data from Shapiro and Senapathy, 1990).

Position	A	C	G	T
5' splice sites				
−2	0.58 (24)	0.12 (9)	0.15 (4)	0.15 (10)
−1	0.10 (5)	0.04 (4)	0.78 (33)	0.08 (5)
+1	0.00 (0)	0.00 (0)	1.00 (47)	0.00 (0)
+2	0.00 (0)	0.00 (0)	0.00 (0)	1.00 (47)
+3	0.57 (26)	0.02 (1)	0.39 (16)	0.02 (4)
+4	0.71 (35)	0.08 (1)	0.12 (5)	0.09 (6)
+5	0.05 (1)	0.06 (2)	0.84 (40)	0.05 (4)
+6	0.16 (5)	0.15 (9)	0.22 (5)	0.47 (28)
3' splice sites				
−6	0.06 (0)	0.39 (7)	0.06 (1)	0.47 (11)
−5	0.06 (5)	0.40 (8)	0.08 (0)	0.46 (6)
−4	0.23 (7)	0.29 (2)	0.23 (2)	0.23 (8)
−3	0.03 (1)	0.74 (11)	0.01 (0)	0.22 (7)
−2	1.00 (19)	0.00 (0)	0.00 (0)	0.00 (0)
−1	0.00 (0)	0.00 (0)	1.00 (19)	0.00 (0)
+1	0.28 (5)	0.13 (2)	0.49 (10)	0.10 (2)
+2*	0.17 (1)	0.22 (3)	0.24 (6)	0.37 (9)

APPENDIX 5

Analysis of functional domains within neurofibromin using the programme 'MOTIF' (GCG; Smith et al. 1991)

****
kinase II phosphorylation site *
****

inase II (CK-2) is a protein serine/threonine kinase that has activity
ent of cyclic nucleotides and of calcium. CK-2 phosphorylates many
t proteins. The substrate specificity [1] of this enzyme can be
ed as follows:

er comparable conditions Ser is favoured over Thr.
acidic residue (either Asp or Glu) must be present three residues to
C-terminal of the phosphate acceptor site.
itional acidic residues in positions +1, +2, +4, and +5 increase the
sphorylation rate. Most physiological substrates have at least one
idic residue in these positions.
is preferred to Glu as the provider of acidic determinants.
asic residue to the N-terminal of the acceptor site decreases the
sphorylation rate, while an acidic one will increase it.

sus pattern: [ST]-x(2)-[DE]
[S or T is the phosphorylation site]

his pattern is found most, but not all, of the known physiological
ates.

CK2_PHOSPHO_SITE (S,T)x2(D,E)

		(S)x{2}(D)	
15:	VQAVV	SRFD	EQLPI
		(T)x{2}(E)	
123:	CHFLH	TCRE	GNQHA
		(S)x{2}(D)	
421:	RIITN	SALD	WWPKI
		(T)x{2}(E)	
467:	MAPSL	TFKE	KVTSL
		(T)x{2}(E)	
482:	FKEKP	TDLE	TRSYK
		(S)x{2}(E)	
521:	PETQG	STAE	LITGL

		(T)x{2}(E)	
569:	DAPVE	TFWE	ISSQM
		(S)x{2}(E)	
592:	SHQML	SSTE	ILKWL
		(S)x{2}(D)	
665:	KGKGN	SSMD	SAAGC
		(T)x{2}(E)	
685:	CRQAQ	TKLE	VALYM
		(T)x{2}(E)	
737:	LPNYN	TFME	FASVS
		(S)x{2}(D)	
829:	GSIDL	SDTD	SLQEW
		(S)x{2}(E)	
833:	LSDTD	SLQE	WINMT
		(S)x{2}(E)	
1,118:	LLNDC	SEVE	DESAQ
		(T)x{2}(E)	
1,178:	LQTRA	TFME	VLTKI
		(T)x{2}(D)	
1,191:	ILQQG	TEFD	TLAET
		(T)x{2}(E)	
1,195:	GTEFD	TLAE	TVLAD
		(S)x{2}(D)	
1,234:	NVVPC	SQWD	ELARV
		(T)x{2}(D)	
1,245:	ARVLV	TLFD	SRHLL
		(T)x{2}(D)	
1,310:	LRIVI	TSSD	WQHVS
		(T)x{2}(E)	
1,324:	FEVDP	TRLE	PSESL
		(S)x{2}(E)	
1,331:	LEPSE	SLEE	NQRNL
		(S)x{2}(E)	
1,353:	HAIIS	SSSE	FPPQL
		(S)x{2}(E)	
1,399:	NPAIV	SPYE	AGILD
		(T)x{2}(E)	
1,435:	NHVLF	TKEE	HMRPF

		(S)x{2}(D)	
1,449:	NDFVK	SNFD	AARRF
		(S)x{2}(D)	
1,503:	GQYLS	SNRD	HKAVG
		(S)x{2}(E)	
1,546:	LNLTS	SKFE	EFMTR
		(T)x{2}(E)	
1,744:	TAVQV	TSAE	RTKVL
		(S)x{2}(E)	
1,765:	DIYYA	SEIE	EICLV
		(T)x{2}(E)	
1,808:	IIHIR	TRWE	LSQPD
		(S)x{2}(D)	
1,813:	TRWEL	SQPD	SIPQH
		(S)x{2}(E)	
1,916:	SGFSK	SSIE	LKHLC
		(T)x{2}(E)	
1,961:	KLITM	TINE	KQMYP
		(S)x{2}(E)	
2,131:	QVLRL	SLTE	FSLPK
		(S)x{2}(D)	
2,160:	IAFRS	SYRD	RSFSP
		(T)x{2}(E)	
2,179:	ETFAL	TSLE	TVTEA
		(T)x{2}(E)	
2,183:	LTSLE	TVTE	ALLEI
		(S)x{2}(E)	
2,334:	IFNDK	SPEE	VFMAI
		(T)x{2}(E)	
2,423:	LAALL	TVSE	EVRSR
		(T)x{2}(E)	
2,466:	DPSYR	TLKE	TQPWS
		(T)x{2}(E)	
2,554:	RRVAE	TDYE	METQR
		(S)x{2}(E)	
2,576:	HLRKV	SVSE	SNVLL
		(T)x{2}(E)	
2,609:	TLVKY	TTDE	FDQRI
		(T)x{2}(E)	
2,729:	GIDEE	TSEE	SLLTP

****
istoylation site *
****

reciable number of eukaryotic proteins are acylated by the covalent
n of myristate (a C14-saturated fatty acid) to their N-terminal residue
amide linkage [1,2]. The specificity of the enzyme responsible for
dification, myristoyl CoA:protein N-myristoyl transferase (NMT), has
rived from the sequence of known N-myristoylated proteins and from
using synthetic peptides. The sequence specificity seems to be the
ng:

N-terminal residue must be glycine.
osition 2, uncharged residues are allowed. Charged residues, proline
large hydrophobic residues are not allowed.
ositions 3 and 4, most, if not all, residues are allowed.
osition 5, small uncharged residues are all allowed (Ala, Ser, Thr,
Asn and Gly). Serine is favored.
osition 6, proline is not allowed.

sus pattern: G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}
[G is the N-myristoylation site]

we deliberately include as potential myristoylated glycine residues
which are internal to a sequence, for it could well be that the
ce under study represents a viral polyprotein precursor and that
uent proteolytic processing could expose an internal glycine as the N-
al of a mature protein.

L G~(E,D,R,K,H,P,F,Y,W)x2(S,T,A,G,C,N)~(P)

G~(E,D,R,K,H,P,F,Y,W)x(2)(T)~P

26: LPIKT GQQNTH TKVST

G~(E,D,R,K,H,P,F,Y,W)x(2)(A)~P

127: HTCRE GNQHAA ELRNS

G~(E,D,R,K,H,P,F,Y,W)x(2)(S)~P

140: RNSAS GVLFSL SCNNF

G~(E,D,R,K,H,P,F,Y,W)x(2)(A)~P

367: KPFSR GSQPAD VDLMI

		G~(E,D,R,K,H,P,F,Y,W)x{2}(A)~P	
455:	AVQGC	GAHPAI	RMAPS
		G~(E,D,R,K,H,P,F,Y,W)x{2}(A)~P	
772:	EHPTA	GNTEAW	EDTHA
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
875:	VSERK	GSMISV	MSSEG
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
1,160:	ANVDS	GLMHSI	GLGYH
		G~(E,D,R,K,H,P,F,Y,W)x{2}(A)~P	
1,277:	QTLFR	GNSLAS	KIMTF
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
1,382:	PQNSI	GAVGSA	MFLRF
		G~(E,D,R,K,H,P,F,Y,W)x{2}(A)~P	
1,481:	SFISD	GNVLAL	HRLLW
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
1,498:	NQEKI	GQYLSS	NRDHK
		G~(E,D,R,K,H,P,F,Y,W)x{2}(A)~P	
1,576:	IFYQA	GTSKAG	NPIFY
		G~(E,D,R,K,H,P,F,Y,W)x{2}(G)~P	
1,595:	RRFKT	GQINGD	LLIYH
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
1,678:	ERLLT	GLKGSK	RLVFI
		G~(E,D,R,K,H,P,F,Y,W)x{2}(T)~P	
1,786:	TIANQ	GTPLTF	MHQEC
		G~(E,D,R,K,H,P,F,Y,W)x{2}(N)~P	
1,831:	PKDVP	GTLLNI	ALLNL
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
2,000:	KTSAT	GGLGSI	KAEVM
		G~(E,D,R,K,H,P,F,Y,W)x{2}(A)~P	
2,003:	ATGGL	GSIKAE	VMADT
		G~(E,D,R,K,H,P,F,Y,W)x{2}(N)~P	
2,103:	RASTH	GLVINI	IHSLC
		G~(E,D,R,K,H,P,F,Y,W)x{2}(N)~P	
2,358:	MDHFV	GLNFNS	NFNFA
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
2,518:	TKKLL	GTRKSF	DHLIS
		G~(E,D,R,K,H,P,F,Y,W)x{2}(S)~P	
2,662:	LNPIH	GIVQSV	VYHEE

****
ein kinase C phosphorylation site *
****
-(81%)[Hit space to continue, Del to abort]

, protein kinase C exhibits a preference for the phosphorylation of
or threonine residues close to a C-terminal basic residue [1,2]. The
e of additional basic residues at the N- or C-terminal of the target
cid enhances the Vmax and Km of the phosphorylation reaction.

sus pattern: [ST]-x-[RK]
[S or T is the phosphorylation site]

SPHO_SITE (S,T)x(R,K)

		(T)x(R)	
101:	GQPKD	TMR	LDETM
		(T)x(R)	
123:	CHFLH	TCR	EGNQH
		(S)x(R)	
158:	VFSRI	STR	LQELT
		(S)x(K)	
259:	DGFAE	STK	RKAAV
		(T)x(R)	
260:	GFAES	TKR	KAAVW
		(S)x(R)	
302:	KLFLD	SLR	KALAG
		(T)x(K)	
467:	MAPSL	TFK	EKVTS
		(S)x(K)	
474:	KEKVT	SLK	FKEKP
		(S)x(K)	
488:	DLETR	SYK	YLLLS
		(S)x(R)	
657:	RTPGA	SLR	KGKGN
		(T)x(R)	
750:	SNMMS	TGR	AALQK
		(S)x(R)	
871:	PMGPV	SER	KGSMI
		(S)x(R)	
1,135:	RKRGM	SRR	LASLR
		(S)x(R)	

1,140:	SRRLA	SLR	HCTVL
		(T)x(K)	
1,343:	NLLQM	TEK	FFHAI
		(S)x(R)	
1,373:	LYQVV	SQR	FPQNS
		(S)x(R)	
1,503:	GQYLS	SNR	DHKAV
		(S)x(K)	
1,545:	SLNLT	SSK	FEEFM
		(T)x(K)	
1,577:	FYQAG	TSK	AGNPI
		(T)x(K)	
1,609:	YHVLL	TLK	PYYAK
		(S)x(R)	
1,630:	THTGP	SNR	FKTDF
		(S)x(R)	
1,682:	TGLKG	SKR	LVFID
		(S)x(K)	
1,733:	KDTKV	SIK	VGSTA
		(S)x(R)	
1,847:	GSSDP	SLR	SAAYN
		(S)x(K)	
2,004:	TGGLG	SIK	AEVMA
		(S)x(K)	
2,026:	NVKLV	SSK	VIGRM
		(S)x(R)	
2,096:	ATGPL	SLR	ASTHG
		(S)x(R)	
2,160:	IAFRS	SYR	DRSFS
		(T)x(K)	
2,201:	MRDIP	TCK	WLDQW
		(S)x(R)	
2,235:	VFGCI	SKR	VSHGQ
		(S)x(R)	
2,326:	LHTLD	SLR	IFNDK
		(T)x(R)	
2,388:	AIVAR	TVR	ILHTL
		(S)x(K)	
2,433:	VRSRC	SLK	HRKSL

		(S)x(R)	
2,463:	HHGDP	SYR	TLKET
		(T)x(K)	
2,466:	DPSYR	TLK	ETQPW
		(S)x(K)	
2,475:	TQPWS	SPK	GSEGY
		(S)x(R)	
2,494:	TVGQT	SPR	ARKSM
		(T)x(K)	
2,513:	PSQAN	TKK	LLGTR
		(T)x(K)	
2,519:	KKLLG	TRK	SFDHL
		(T)x(R)	
2,560:	DYEME	TQR	ISSSQ
		(S)x(R)	
2,789:	TGHCN	SGR	TRHGS
		(S)x(K)	
2,808:	QRSAG	SFK	RNSIK
		(S)x(K)	
2,813:	SFKRN	SIK	KIV*

****
lycosylation site *
****

s been known for a long time [1] that potential N-glycosylation sites are
fic to the consensus sequence Asn-Xaa-Ser/Thr. It must be noted that the
nce of the consensus tripeptide is not sufficient to conclude that an
agine residue is glycosylated, due to the fact that the folding of the
in plays an important role in the regulation of N-glycosylation [2]. It
een shown [3] that the presence of proline between Asn and Ser/Thr will
it N-glycosylation; this has been confirmed by a recent [4] statistical
sis of glycosylation sites, which also shows that about 50% of the sites
have a proline C-terminal to Ser/Thr are not glycosylated.

st also be noted that there are a few reported cases of glycosylation
with the pattern Asn-Xaa-Cys; an experimentally demonstrated occurrence
e--(12%)[Hit space to continue, Del to abort]
ch non-standard site is found in the plasma protein C [5].

ensus pattern: N-{P}-[ST]-{P}
[N is the glycosylation site]

LYCOSYLATION		N~(P)(S,T)~(P)		
		N~P(S)~P		
45:	KECLI	NISK	YKFSL	
		N~P(S)~P		
639:	IPSSG	NTSQ	MSMDH	
		N~P(S)~P		
664:	RKGKG	NSSM	DSAAG	
		N~P(T)~P		
839:	LQEWI	NMTG	FLCAL	
		N~P(T)~P		
974:	KNLLD	NHTE	GSSEH	
		N~P(S)~P		
1,472:	TSDAV	NHSL	SFISD	
		N~P(T)~P		
1,542:	HWSSL	NLTS	SKFEE	
		N~P(S)~P		
1,652:	GFAYD	NVSA	VYIYN	
		N~P(T)~P		
1,882:	LCIPA	NNTL	FIVSI	
		N~P(S)~P		
2,070:	LMLSF	NNSL	DVAAH	
		N~P(S)~P		
2,756:	ITANL	NLSN	SMTSL	

****
e zipper pattern *
****

ure, referred to as the 'leucine zipper' [1,2], has been proposed to how some eukaryotic gene regulatory proteins work. The leucine zipper of a periodic repetition of leucine residues at every seventh over a distance covering eight helical turns. The segments containing eriodic arrays of leucine residues seem to exist in an alpha-helical tion. The leucine side chains extending from one alpha-helix interact se from a similar alpha helix of a second polypeptide, facilitating tion; the structure formed by cooperation of these two regions forms a

oil [3]. The leucine zipper pattern is present in many gene regulatory , such as:

CATT-box and enhancer binding protein (C/EBP).
AMP response element (CRE) binding proteins (CREB, CRE-BP1, ATFs).
Jun/AP1 family of transcription factors.
yeast general control protein GCN4.
fos oncogene, and the fos-related proteins fra-1 and fos B.
c-myc, L-myc and N-myc oncogenes.
octamer-binding transcription factor 2 (Oct-2/OTF-2).

sus pattern: L-x(6)-L-x(6)-L-x(6)-L
ces known to belong to this class detected by the pattern: All those ned in the original paper, with the exception of L-myc which has a Met d of the second Leu.
sequence(s) detected in SWISS-PROT: some 600 other sequences from every ry of protein families.

as this is far from being a specific pattern you should be cautious in the presence of such pattern in a protein if it has not been shown to uclear DNA-binding protein.

_ZIPPER Lx6Lx6Lx6L
Lx(6)Lx(6)Lx(6)L
1,834: VPGTL LNIALLNLGSSDPSLRSAAYNL LCALT

****
saminoglycan attachment site *
****

lycans [1] are complex glycoconjugates consisting of a core protein to variable number of glycosaminoglycan chains (such as heparin sulfate, itin sulfate, etc.) are covalently attached. The glycosaminoglycans are d to the core proteins through a xyloside residue which is in turn is to a serine residue of the protein. A consensus sequence for the ent site seems to exist [2]. However, it must be noted that this us is only based on the sequence of three proteoglycans core proteins.

sus pattern: S-G-x-G
[S is the attachment site]
onal rule: There must be at least two acidic amino acids from -2 to -4 relative to the serine.

AMINOGLYCAN SGxG
821: RMSHV SGGG SIDLS

****
sine kinase phosphorylation site *
****

ates of tyrosine protein kinases are generally characterized by a lysine
arginine seven residues to the N-terminal side of the phosphorylated
ne. An acidic residue (Asp or Glu) is often found at either three or
esidues to the N-terminal side of the tyrosine [1,2,3]. There are a
of exceptions to this rule such as the tyrosine phosphorylation sites
lase and lipocortin II.

nsus pattern: [RK]-x(2)-[DE]-x(3)-Y
or [RK]-x(3)-[DE]-x(2)-Y
[Y is the phosphorylation site]

OSPHO_SITE (R,K)x{2,3}(D,E)x{2,3}Y
(R)x{3}(E)x{2}Y

2,549: TPPKM RRVAETDY EMETQ

(R)x{2}(E)x{2}Y

2,550: PPKMR RVAETDY EMETQ

****
dation site *
****

-Consensus pattern: x-G-[RK]-[RK]
[x is the amidation site]

recursor of hormones and other active peptides which are C-terminally
ted is always directly followed [1,2] by a glycine residue which provides

nide group, and most often by at least two consecutive basic residues
or Lys) which generally function as an active peptide precursor cleavage
Although all amino acids can be amidated, neutral hydrophobic residues
s Val or Phe are good substrates, while charged residues such as Asp or
are much less reactive. C-terminal amidation has not yet been shown to
in unicellular organisms or in plants.

TION xG(R,K)(R,K)
xG(R)(K)

1,128: ESAQT GGRK RGMSR

xG(R)(R)

1,510: RDHKA VGRR PFDKM

****
GTPase-activating proteins signature *
****

oteins are membrane-associated molecular switches that bind GTP and GDP owly hydrolyze GTP to GDP [1]. This intrinsic GTPase activity of ras is ated by a family of proteins collectively known as 'GAP' or GTPase-ting proteins [2]. As it is the GTP bound form of ras which is active, proteins are said to be down-regulators of ras. Proteins known to s such activity are listed below:

malian GAP. GAP can down-regulate wild-type ras, but fails to do so with ogenic, mutated ras.

1 and IRA2, the functional equivalents of GAP in yeast. They regulate RAS-cyclic AMP pathway, controlling cell growth.
1, the fission yeast protein that regulates ras1 in that organism.
2, a yeast protein that activates BUD1/RSR1 and which participates in regulation of bud-site selection [3].
an neurofibromin (gene NF1) [4]. NF1 is associated with type 1 rofibromatosis, one of the most frequently inherited genetic diseases racterized, in part, by multiple neural tumors. NF1 has been shown etically and biochemically to interact with and stimulate the GTPase ivity of h-ras.
sophila Gap1 [5], which acts as a negative regulator of signalling by Sevenless receptor tyrosine kinase involved in eye development.

e above proteins are quite large (from 765 residues for sar1 to 3079 es for IRA2) but share only a limited (about 250 residues) region of ce similarity, referred to as the 'catalytic domain'. The most conserved within this domain contains a 12 residue motif which seems to be teristic of this family of proteins [2].

nsus pattern: F-L-R-x(3)-P-[AV]-x(3)-P
nces known to belong to this class detected by the pattern: ALL.
sequence(s) detected in SWISS-PROT: NONE.
there are distinctly different GAPs for the rap and rho/rac subfamilies s-like proteins (reviewed in reference [6]) that do not share sequence arity with ras GAPs.

PASE_ACTIV FLRx3P(A,V)x3P
FLRx{3}P(A)x{3}P
1,389: VGSAM FLRFINPAIVSP YEAGI

****
- and cGMP-dependent protein kinase phosphorylation site *
****

nas been a number of studies relative to the specificity of cAMP- and
ependent protein kinases [1,2,3]. Both types of kinases appear to share
erence for the phosphorylation of serine or threonine residues close to

st two consecutive N-terminal basic residues. It is important to note
nere are quite a number of exceptions to this rule.

nsus pattern: [RK](2)-x-[ST]
[S or T is the phosphorylation site]

PHOSPHO_SITE		(R,K)2x(S,T) (K)(2)x(T)	
583:	LFYIC	KKLT	SHQML
		(R)(2)x(S)	
815:	KTIVK	RRMS	HVSGG
		(R,K)(2)x(S)	
873:	GPVSE	RKGS	MISVM
		(R,K)(2)x(S)	
2,236:	FGCIS	KRVS	HGQIK
		(R,K)(2)x(S)	
2,573:	QHPHL	RKVS	VSESN
		(R,K)(2)x(S)	
2,810:	SAGSF	KRNS	IKKIV

REFERENCES

Ainsworth,P.J., Rodenheiser,D.I. and Costa,M.T. (1993) Identification and characterisation of inherited and sporadic mutations in exon 31 of the neurofibromatosis (NF-1) gene. *Hum. Genet.* **91**, 151-156.

Ahlgrenbeckendorf,J.A., Maggio,W.W., Chen,F. and Kent,T.A. (1993) Neurofibromatosis 1 mRNA expression in blood vessels. *Biochem. Biophys. Res. Comm.* **197,** 1019-1024.

Andersen,L.B., Ballester,R., Marchuk,D., Chang.,E. Gutmann,D.H., Saulino,A., Camonis,J., Wigler,M. and Collins,F. (1993a) A conserved alternative splice in the von Recklinghausen neurofibromatosis (NF-1) gene produces two neurofibromin isoforms, both of which have GTPase activating protein activity. *Molecular and Cellular Biology* **13**, 487-95.

Andersen,L.B., Fountain,J.W., Gutmann,D.H., Tarle,S.A., Glover,T.W., Dracopoli,N.C., Housmann,D.E. and Collins,F.S. (1993b) Mutations in the neurofibromatosis 1 gene in sporadic melanoma cell lines. *Nature Genet.* **3,** 118-121.

Anglani,F., Murgia,A., Bedin,S., Bresin,E., Bernardi,F., Clementi,M. and Tenconi,R. (1993) A new disease causing mutation in the GAP related domain of the NF-1 gene. *Hum. Mol. Genet.* **2,**1057-1059.

Antonarkis,S.E., Boehn,C.D., Serjeant,S.R., Theisen,C.E., Dover,G.J. and Kazazian,H.H. (1984) Origin of the beta-S gene in Blacks: the contribution of recurrent mutation or gene conversion or both. *Proc. Natl. Acad. Sci. USA* **81**, 853-856.

Bairoch,A. (1991) PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Research* **19**, 2241-2245.

Ballester,R., Marchuk,D., Boguski,M., Saulino,A., Letcher,R., Wigler,M. and Collins,F. (1990) The NF-1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. *Cell* **63**, 851-859.

Barbacid,M. (1987) ras genes *Ann. Rev. Biochem.* **56,** 779-828.

Barker,D., Wright,E., Nguyen,K., Cannon,L., Fain,P., Goldgar,D., Bishop,D.T., Carey,J., Baty,B., Kivlin,J., Willard,H., Waye,J.S., Greig,G., Leinwald,L., Nakamura,Y., O'Connell,P., Leppert,M., Lalouel,J.M., White,R. and Skolnick,M. (1987) Gene for Von recklinghausen neurofibromatosis is in the pericentromeric region on chromosome 17. *Science* **236**, 1100-1102.

Basu,T.N., Gutmann,D.H., Fletcher,J., Glover,T.W., Collins,F.S. and Downward,J. (1992) Aberrant regulation of ras proteins in malignant tumour cells from type-1 neurofibromatosis patients. *Nature* **356**, 713-715.

Bernards,A., Snijders,A.J., Hannigan,G., Murthy,A.E. and Gusella,J. (1993) Mouse neurofibromatosis type-1 sequence reveals high degree of conservation of both coding and non-coding mRNA segments *Hum. Mol. Genet.* **2**, 645-650.

Beutler,E. and Kuhl,W. (1990) The NT 1311 polymorphism of G6PD: G6PD Meditarranean mutation may have originated independently in Europe and Asia. *Am. J. Hum. Genet.* **47**,1008-1012.

Bird,A.P. (1986) CpG-rich islands and the function of DNA methylation *Nature* **321**, 209-213.

Bleyl,S., Ainsworth,P., Nelson,L., Viskochil,D. and Ward,K. (1994) An ancient Ta subclass L1 insertion results in an intragenic polymorphism in an intron of the NF-1 gene. *Hum. Mol. Genet.* **3**, 517-518.

Bollag,G. and McCormick,F. (1991) Differential regulation of rasGAP and neurofibromatosis gene product activities. *Nature* **351**, 576-579.

Bos J.L. (1989) Ras oncogenes in human cancer: A review. *Cancer research* **49**, 4682-89.

Boyer,M.J., Gutmann,D.H., Collins,F.S. and BarSagi,D. (1994) Crosslinking of the surface immunoglobulin receptor in B lymphocytes induces a redistribution of neurofibromin but not p120GAP. *Oncogene* **9**, 349-357.

Brannan,C.I., Perkins,A., Vogel,K., Ratner,N., Nordlund,M., Reid,S., Buchberg,A., Jenkins,N., Parada,L. and Copeland,N. (1994) Targeted disruption of the neurofibromatosis type-1 gene leads to developmental abnormalities in heart and various neural crest dreived tissues. *Genes and development* **8**, 1019-1029.

Buchberg,A.M., Bedigian,H.G., Taylor,B.A., Brownell,E., Ihle,J.N., Nagata,S., Jenkins,N.A. and Copeland,N.G. (1988). Localisation of evi-2 to chromosome 11: Linkage to other protoncogene and growth factor loci using interspecific backcross mice. *Oncogene Res.* **2,** 149-165.

Buchberg,A.M., Brownell,E., Nagata,S., Jenkins,N.A. and Copeland,N.G. (1989) A comprehensive genetic map of murine chromosome 11 reveals extensive linkage conservation between mouse and human. *Genetics* **122**, 153-161.

Buchberg,A.M., Cleveland,L.S., Jenkins,N.A. and Copeland,N.G. (1990) Sequence homology shared by neurofibromatosis type1 gene and IRA1 and IRA2 negative regulators of the RAS cAMP pathway. *Nature* **347**, 291-294.

Cawthon,R.M., Weiss,R., Xu,G., Viskochil,D., Culver,M., Stevens,J., Robertson,M., Dunn,D., Gesteland,R., O'Connell,P. and White,R. (1990a) A major segment of the neurofibromatosis type 1 gene: cDNA sequence, genomic structure and point mutations. *Cell* **62**, 193-201.

Cawthon,R.M., O'Connell,P., Buchberg,A.M., Viskochil,D., Weiss,R.B., Culver,M., Stevens,J., Jenkins,N.A., Copeland,N.G. and White,R. (1990b) Identification and characterisation of transcripts from the neurofibromatosis 1 region: The sequence and genomic structure of EV12 and mapping of other transcripts. *Genomics* **7**, 555-565.

Cawthon,R.M., Andersen,L.B., Buchberg,A.M., Xu,G., O'Connell,P., Viskochil,D., Weiss,R.B., Wallace,M.R., Marchuk,D.A., Culver,M., Stevens,J., Jenkins,N.A., Copeland,N.G., Collins,F.S. and White,R. (1991) cDNA sequence, and genomic structure of EV12B, a gene lying within an intron of the neurofibromatosis type-1 gene. *Genomics* **9**, 446-460.

Chomczynski,P. and Sacchi,N. (1987) Single step method of RNA isolation by acid guanidinium thiocyanate-phenol chloroform extraction. *Anal. Biochem.* **162**, 156-159.

Collins,F.S., Gutmann,D.H., Mitchell,A.L., Doran,S., Andersen,L.B., Gregory,P.E., Hajara,A., Swaroop,M. and Cole,J. (1994) Molecular genetics of neurofibromatosis type 1. *J. Cell. Biochem.* Supplement **18A**, 215.

Colman,S.D., Collins,F.S. and Wallace,M.R. (1993) Characterisation of a single base pair deletion in neurofibromatosis type-1. *Hum. Mol. Genet.* **2,** 1709-1711.

Cooper,D.N. and Krawczak,M. (1989) Cytosine methylation and the fate of CpG dinucleotides in disease. *Hum.Genet.* **83**,181-188.

Cooper,D.N. and Krawczak,M. (1991) Mechanisms of insertional mutagenesis in human genes causing genetic disease. *Hum. Genet.* **87**, 519-560.

Cooper,D.N. and Youssoufian,H. (1988) The CpG dinucleotide and human genetic disease *Hum. Genet.* **78**, 151-155.

Cooper,D.N. and Krawczak,M. (1993) Human Gene Mutation, Bios Scientific Publishers.

Cotton,R.G.H., Rodrigues,N.R. and Campbell,R.D. (1988) Reactivity of cytosine and thymine in single base pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc. Natl. Acad. Sci. USA* **85**, 4397- 4401.

Crowe,F.W., Schull,W.J. and Nell,J.V. (1956) A clinical, pathological, and genetic study of multiple neurofibromatosis. Charles C.Thomas, Springfield, Illinois.

Cummings,L.M., Glatefeller,A. and Marchuk,D.A. (1993) NF-1 related loci on chromosomes 2, 12, 14, 15, 20, 21, and 22: A potential role for gene conversion in the high spontaneous mutation rate of NF-1? *Am J. Hum. Genetics* **53,** Supp no 3.

Daston,M.M., Scrable,H., Nordlund,M., Sturbaum,A.K., Nissen,L.M. and Ratner,N. (1992) The protein product of the neurofibromatosis type-1 gene is expressed at highest abundance in neurons, Schwann cells and oligodendrocytes. *Neuron* **8**, 1-14.

Daston,M.M. and Ratner,N. (1993) Neurofibromin, a predominantly neuronal GTPase activating protein in the adult, is ubiquitiously expressed during development. *Dev Dyn* **19**, 216.

DeClue,J.E, Cohen,B. and Lowy,R.D. (1991) Identification and characterisation of the neurofibromatosis type-1 gene product. *Proc Natl Acad Sci. USA* **88**, 9914-18.

DeClue,J.E, Papageorge,A.G., Fletcher,J.A., Diehl,S.R., Ratner,N., Vass,W.C. and Lowy R.D. (1992) Abnormal regulation of mammalian p21ras contributes to malignant tumour growth in Von Recklinghausen (Type1) neurofibromatosis. *Cell* **69**, 265-273.

Easton,D.F., Ponder,M.A., Huson,S.M. and Ponder,B.A.J. (1993) An analysis of variation in expression of Neurofibromatosis (NF) type 1(NF-1): Evidence for modifying genes. *Am. J. Hum. Genet.* **53**, 305-313.

Erlich,H.A. (editor) (1989) PCR Technology Principles and applications for DNA amplification M.Stockton Press.

Estivill,X., Lazaro,C., Casals,T. and Ravella,A. (1991) Reccurence of a nonsense mutation in the NF-1 gene causing classical neurofibromatosis-1. *Hum. Genet.* **88**,185-188.

Fanning,T.G. and Singer,M.F. (1987) LINE-1: a mammalian transposable element *Biochimica et Biophysica Acta* **910,** 203-212.

Feinberg,A.P. and Vogelstein,B. (1983) A technique for radiolabelling DNA restriction endonucleases fragments to high specific activity. *Analytical Biochem.* **123**, 8-13.

Fischer,G. and Lerman,S. (1983) DNA fragments differing by single base substitutions are separated in denaturing gradient gels: correspondence with melting theory. *Proc. Natl. Acad. Sci. USA* **80**, 1579-1583.

Forrest,S.M., Dahl,H.H., Howells,D.W., Dianzani,I. and Cotton,R.G. (1991) Mutation detection in phenylketonuria by using chemical cleavage of mismatch: Importance of using probes from both normal and patient samples. *Am. J. Hum. Genet.* **49**, 175-183.

Fountain J.W., Wallace,M.R., Brereton,A.B., O'Connell,P., White,R.L., Rich,D.C., Ledbetter,D.H., Leach,R.J., Fournier,R.E.K., Menon,A.G., Gusella,J.F., Barker,D., Stephens,K. and Collins,F.S. (1989a). Physical mapping of the Von Recklinghausen region on chromosome 17. *Am.J.Hum.Genet.* **44**, 58-67.

Fountain J.W., Wallace,M.R., Bruce,M.A., Seizinger,B.R., Menon,A.G., Gusella,J.F., Michels,V.V., Schmidt,M.A., Dewald,G.W. and Collins,F.S. (1989b) Physical mapping of a translocation breakpoint in neurofibromatosis, *Science* **244**, 1085-1087.

Fu,Y.H., Kuhl,D.P.A., Pizzuti,A., Pierreti,M., Sutcliffe,J.S., Richards,S., Verkerk,A.J.M.H., Holden,J.J.A., Fenwick,R.G., Warren,S.T., Oostra,B., Nelson,D.L. and Caskey,C.T. (1991) Variation of CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* **67**, 1047-1058.

Fu,Y.H., Pizzuti,A., Fenwick,R.G., King,J., Rajanarayan,S., Dunne,P.W., Dubel,J., Nasser,G., Ashiwaza,T., DeJong,P., Wieringa,B., Korneluk,R., Perryman,M.B., Epstein,H. and Caskey,T. (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy *Science* **255**, 1256-58.

Gasparini,P., Grifa,A., Origone,P., Coviello,D., Antonacci,R. and Rocchi,M. (1993) Detection of a neurofibromatosis type 1 (NF-1) sequence by PCR: Implications for the diagnosis and screening of genetic diseases. *Mol. Cell. Probes* **7**, 415-418.

Gibbs,R.A., Nguyen,P.N., Edwards,A., Civitello,A.B. and Caskey,C.T. (1990) Multiplex DNA deletion detection and and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families. *Genomics* **7**, 235-244.

Glover,T.W., Stein,C.K., Legius,E., Andersen,L.B., Brereton,A. and Johnson,S. (1991) Molecular and cytogenetic analysis of tumours in Von Recklinghausens neurofibromatosis *Genes, Chromosomes, Cancer* **3**, 62-70.

Goldgar,D.E., Green,P., Parry,D.M. and Mulvihill,J.J. (1989) Multipoint linkage analysis in neurofibromatosis type1: An international collaboration. *Am.J.Hum.Genet.* **44**, 6-12.

Golubic,M., Tanaka,K., Dobrowolski,S., Wood,D., Tsai,M., Marshall,M., Tamanoi,F. and Stacey,D.W. (1991) The GTPase stimulatory activities of the neurofibromatosis type1 and the yeast IRA2 proteins are inhibited by arachidonic acid. *EMBO J.* **10**, 2897-2093.

Gregory,P.E., Gutmann,D.H., Mitchell,A., Park,S., Boguski,M., Jacks,T., Wood,D.L., Jove,R. and Collins,F.S. (1993) Neurofibromatosis type 1 gene product (neurofibromin) associates with microtubules. *Som. Cell. Mol. Genet.* **19**, 265-274.

Grompe,M. (1993) The rapid detection of unknown mutations in nucleic acids. *Nature Genet.* **5**, 111-117.

Gutmann,D., Wood,D. and Collins,F. (1991) Identification of the neurofibromatosis type 1 gene product. *Proc. Natl. Acad. Sci. USA* **88**, 9658-62.

Gutmann,D., Andersen,L.B., Cole,J., Swaroop,M. and Collins,F.S. (1993) An alternatively spliced mRNA in the carboxy-terminus of the neurofibromatosis type-1 is expressed in muscle. *Hum. Mol. Genet.* **2**, 989.

Gyllensten,U.B. and Erlich,H.A. (1988) Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing at the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA* **85,** 7652-7656.

Hajra,A., Martin-Gallardo,A., Tarle,S., Freedman,M., Wilson-Gunn,S., Bernards,A. and Collins,F. (1994) DNA sequences in the promoter region of the NF-1 gene are highly conserved between human and mouse. *Genomics* **21**, 649-652.

Hall,A. (1990) ras and GAP-Who's controlling whom? *Cell* **61**, 921-923.

Hall,J.G. (1990) Genomic imprinting: Review and relevance of genetic diseases *Am. J. Hum. Genet.* **46**, 857.

Han,J.W., McCormick,F. and Macara,I. (1991) Regulation of Ras-GAP and the neurofibromatosis gene product by eicosanoids. *Science* **252,** 576-579.

Hinrichs,S.H, Nerenberg,M., Reynolds,R.K., Khoury,G. and Jay,G. (1987) A transgenic mouse model for human neurofibromatosis. *Science* **237,** 1340.

Hochberg,A., Gonik,B., Goshen,R. and DeGroot,N. (1994) A growing relationship between genomic imprinting and tumourigenesis. *Cancer Genet. cytogenet.* **73,** 82-83.

Hoffmeyer,S., Assum,G., Kaufmann,D. and Krone,W. (1994) Unequal expression of NF-1 alleles. *Nature Genetics* **6**, 331.

Horiuchi,T., Hatta,N., Matsumoto,M., Ohtsuka,H., Collins,F.S., Kobayashi,Y. and Fujita,S. (1994) Nonsense mutations at Arg-1947 in two cases of familial neurofibromatosis type 1 in Japanese. *Hum. Genet.* **93**, 81-83.

Hofman,K.J., HarrisE.L., Bryan,R.N. and Denckla,M.B (1994) Neurofibromatosis-1: the cognitive phenotype. *Journal of pediatrics*, **124,** 1-8.

Huntingtons Disease Collaborative Research Group (1993) A novel gene containing a trinucleotide repeat that is expanded and stable on Huntington's disease chromosomes. *Cell* **72**, 971-983.

Huson, S.M., Clark,D., Compston,D.A. and Harper,P.S. (1989) A genetic study of von Recklinghausens neurofibromatosis in South east wales I: Prevalence, fitness, mutation rate, and effect of paternal transmission on severity. *J. Med. Genet.* **26**, 704.

Hutter,P., Antonarkis,S.E., Delozier-Blanchet,C. and Morris,M.A., (1994) Exon skipping associated with a A to G transition at +4 of the IVS 33 splice donor site of the neurofibromatosis type 1 (NF-1) gene. *Hum. Mol. Genet.*, **3**, 663-665.

Hu,X. and Worton,R.G (1992) Partial gene duplication as a cause of human disease. *Human Mutation* **1**, 3-12.

Innis,M.A., Gelfland,D.H., Sninsky,J.J., White,T.J (1990) PCR protocols-A guide to methods and applications. Academic press Inc.

Jacks,T., Shih,S., Schmitt,E.M., Bronson,R., Bernards,A. and Weinberg,R. (1994) Tumour predisposition in mice hetrozygous for a targeted mutation in NF-1. *Nature Genetics* **7,** 353-361.

Jadayel,D., Fain,P., Upadhyaya,M., Ponder,M.A., Huson,S.M., Carey,J., Fryer,A., Mathew,C.G.P., Barker,D.F. and Ponder, B.A.J. (1990), Paternal origin of new mutations in Von Recklinghausens neurofibromatosis. *Nature* **343**, 558-559.

Johnson, M.R., Look,A.T., DeClue,J.E., Valentine,M.B. and Lowy,D.R. (1993) Inactivation of the NF-1 gene in human melanoma and neuroblastoma cell lines without impaired regulation of GTP-ras. *Proc. Natl. Acad. Sci. USA* **90**, 5539-5542.

Jones,C.T., McIntosh,I., Keston,M., Ferguson,A. and Brock,D.J. (1992) Three novel mutations in the cystic fibrosis gene detected by chemical cleavage: analysis of variant splicing and a nonsense mutation. *Hum. Mol. Genet.* **1**, 11-17.

Jorde,L.B., Watkins,W.S., Viskochil,D., O'Connell,P.O. and Ward,K. (1993) Linkage disequilibrium in the neurofibromatosis 1 (NF-1) region: Implications for gene mapping. *Am. J. Hum. Genet.*. **53**, 1038-1050.

Kayes,L.M., Riccardi,V.M., Burke,W., Bennet,R.L. and Stephens,K. (1992) Large de novo DNA deletion in a patient with sporadic neurofibromatosis 1, mental retardation and dysmorphism. *J. Med. Genet.* **29**, 686-690.

Kayes,L.M., Burke,W., Riccardi,V.M., Bennet,R.L., Ehrlich,P., Rubenstein,A. and Stephens,K. (1994) Deletions spanning the neurofibromatosis-1 gene: Identification and phenotype of five patients. *Am. J. Hum. Genet.* **54**, 424-436.

Kazazian,H.H.Jr, Waber,P.G., Boehm,C.D., Lee,J.I., Antonarkis,S.E. and Fairbanks,V.F.(1984) Hemoglobin E in Europeans: further evidence for multiple origins of the beta-E gene. *Am. J. Hum. Genet.* **36**, 212-217.

Knudson,A.G. (1991) Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. USA* **68**, 820-823.

Koide,R., Ikeuchi,T., Onodera,O., Tanaka,H., Igarshi,S., Endo,K., Takahashi,H., Kondo,R., Ishikawa,A., Hayashi,T., Saito,M., Tomoda,A., Miike,T., Naito,H., Ikuta,F. and Tsuji,S. (1994) Unstable expansion of CAG repeat in hereditary dentataorubral-pallidoluysian atrophy (DRPLA) *Nature Genetics*, **6**, 9-13.

Kozak,M. (1986) Point mutations define a sequence flanking the AUG initiatior codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 284-292.

Krawczak,M. and Cooper,D.N. (1991) Gene deletions causing human genetic disease: Mechanisms of mutagenesis and the role of the local DNA sequence environment. *Human Genetics* **86**, 425-441.

KulozicA.E., Wainscoat,J.S., Sergeant,G.R., Kar,B.C., Al-Awamy,B., Essan,G.J.F., Falusi,A.G., Haque,S.K., Hilali,A.M., Kate,S.,

Ranasinghe,W.A.E.P. and Weatherall,D.J. (1986) Geographical survey of beta-S-globin haplotypes: Evidence for an independent Asian origin of the sickle cell mutation. *Am.J.Hum.Genet.* **39**, 239-244.

Kunkel,L.M., Smith,K.D., Boyer,S.H. and Borgaonkar,D.S. (1977) Analysis of human Y chromosome specific reiterated DNA in chromosome variants. *Proc. Natl Acad Sci USA* **68**, 120-124.

Laird,P.W, Zijderveld,A., Linders,K., Rudnicki,A., Jaenisch,R. and Berns,A. (1991) Simplified mammalian DNA isolation procedure. *Nucl. Acids Res.* **19,** 4293.

Lakich,D., Kazazian,H., Antonarkis,S.E. and Gitschier,J. (1993) Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nature Genetics* **5,** 236-241.

LaSpada,A.R., Wilson,E.M., Lubahn,D.B., Harding,A.E. and Fishbeck,K.H. (1991) Androgen gene receptor mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77-79.

Lazaro,C. and Estivill,X. (1992) Mutation analysis of genetic diseases by asymmetric PCR-SSCP and ethidium bromide staining: application to neurofibromatosis and cystic fibrosis. *Molecular and cellular probes* **6**, 357-359.

Lazaro,C., Gaona,A., Ravella,A., Volpini,V., Casals,T., Fuentes,J.J. and Estivill,X. (1993) Novel alleles, hemizygosity and deletions at an Alu repeat within the neurofibromatosis type-1 gene. *Hum. Mol. Genet.* **6**, 725-730.

Leach,R.J., Thayer,M.J., Schafer,A.J. and Fournier,R.E.K. (1989). Physical mapping of human chromosome 17 using fragment containing microcell hybrids. *Genomics* **5**,167-176.

Ledbetter,D.H., Rich,D.C., O'Connell,P., Leppert,M. and Carey,J.C. (1989) Precise localisation of NF-1 to 17q11.2 by balanced translocation. *Am. J. Hum. Genet.* **44**, 20-24.

Legius,E., Marchuk,D., Hall,B.K., Andersen,L.B., Wallace,M.R., Collins,F.S. and Glover,T.W. (1992) NF-1 related locus on chromosome 15. *Genomics* **13,** 1316-1318.

Legius,E., Marchuk,D.A., Collins,F.S. and Glover,T.W. (1993) Somatic deletion of the neurofibromatosis type 1 gene in a neurofibrosarcoma supports a tumour suppressor gene hypothesis. *Nature Genet.* **3**, 122-126.

Legius,E., Descheemaeker,M.J., Spaepen,A., Caeser,P. and Fryns,J.P. (1994a) Neurofibromatosis type 1 in childhood: A study of the neuropsychological profile in 45 children. *Genetic counselling* **5,** 51-60.

Legius,E., Hall,B.K., Wallace,M.R., Collins,F. and Glover,T.S (1994b) Ten base pair duplication in exon 38 of the NF-1 gene. *Hum. Mol. Genet.* **3,** 829-830.

Levinson,B., Kenwrick,S., Lakich,D., Hammonds,G. and Gitschier,J. (1990). A transcribed gene in an intron of the factor VIII gene. *Genomics* **7**, 1-11.

Lewin,B. (1994) Genes V, Oxford University press.

Li,Y., Bollag,G., Clark,R., Stevens,J., Conroy,L., Fults,D., Ward,K., Freidman,E., Samowitz,W., Robertson,M., Bradley,P., McCormick,F., White,R. and Cawthon,R. (1992) Somatic mutations in the neurofibromatosis-1 gene in human tumours. *Cell* **69**, 275-281.

Marchuk,D.A., SaulinoA.M., Tavakkol,R.L., Swaroop,M., Wallace,M.R., Andersen,L.B., Mitchell,A.M., Gutmann,D,H. and Collins,F.S. (1991). cDNA cloning of the type-1 neurofibromatosis gene: complete sequence of the NF-1 gene product. *Genomics* **11**, 931-940.

Marchuk,D., Tavakkol,R., Wallace,M., Brownstein,B., Taillon-Miller,P., Fong,C.T., Legius,E., Andersen,L.B., Glover,T.W. and Collins,F.S. (1992) A yeast artificial chromosome contig encompassing the type-1 neurofibromatosis gene. *Genomics* **13**, 672-680.

Martin,G.A., Viskochil,D., Bollag,G., McCabe,P.C., Crosier,W.J., Haubruck,H., Conroy,L., Clark,R., O'Connell,P., Cawthon,R.M., Innis,M.A. and McCormick,F. (1990) The GAP related domain of the neurofibromatosis type 1 gene product interacts with ras p21. *Cell* **63**, 843-849.

Martuza,R.L., Gusella,J.F. and Seizinger,B.R. (1990) Chromosome 17p deletions and p53 mutations are associated with the formation of malignant neurofibrosarcomas in von Recklinghausen's neurofibromatosis. *Proc. Natl. Acad. Sci. USA* **87,** 5434-5439.

McCormick,F. (1989) rasGTPase activating protein:signal transmitter and signal terminator. *Cell* **56,** 5-8.

McPherson,M.J., Quirke,P., Taylor,G. (editors) (1992) PCR: A practical approach. Oxford University Press.

Menon,A.G, Ledbetter,D.H., Rich,D.C., Seizinger,B.R., Rouleau,G.A., Michels,V.V., Schmidt,M.A., Dewald,G., DallaTorre,C.M., Haines,J.L. and

Gusella,J.F. (1989) Characterisation of a translocation within the von Recklinghausen neurofibromatosis region of chromosome 17. *Genomics* **5**, 245-249.

Menon,A.G, Anderson,K.M., Riccardi,V.M., Chung,R.Y., Whaley,J.M., Yandell,D.W., Farmer,G.E., Freiman,R.N., Lee,J.K., Li,F.P., Barker,D.F., Ledbetter,D., Kleider,A., Martuza,R.L., Gusella,J.F. and Seizinger,B.R. (1990) Chromosome 17p deletions and p53 gene mutations associated with the formation of malignant neurofibrosarcomas in von Recklinghausen's neurofibromatosis *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5435-5439.

Mikol,D.D., Gulcher,J.R. and Stefansson,K. (1990). The oligodendrocyte-myelin glycoprotein belongs to a distinct family of proteins and contains the HNK-1 carbohydrate. *J.Cell.Biol.* **110**, 471-480.

Mgone,C.S., Lanyon,W.G, Moore,M.R. and Connor,J.M. (1992) Detection of seven point mutations in the porphobilinogen deaminase gene in patients with acute intermittent porphyria by direct sequencing of in vitro amplified cDNA. *Hum. Genet.* **90**, 12-16.

Myers,R.M., Fischer,S.G., Lerman,L.S. and Maniatis,T (1985) Nearly all single base substitutions in DNA fragments joined to a GC clamp can be detected by denaturing gradient gel electrophoresis *Nucl. Acid res.* **13**, 3131-3145.

Myers,R.M., Larin,Z. and Maniatis,T. (1989) Detection of single base substitutions by ribonuclease A cleavage of mismatches in RNA:DNA duplexes. *Science* **230,** 1242-1246.

Nagamine,C.M., Chan,K. and Lau,Y.F.C. (1989) A PCR artifact: Generation of heteroduplexes. *Am. J. Hum. Genet.* **45**, 337-339.

Nakamura,T., Hara,M. and Kasuga,T. (1989) Transplacental induction of peripheral nervous tumour in the Syrian golden hamster by N-nitrosos-N-ethylurea. *Am. J. Pathology* **135,** 251.

Naylor,J., Brinke,A., Hasscock,S., Green,P.M. and Gianelli,F. (1993) Characteristic mRNA abnormality found in half the patients with severe Haemophilia A is due to large DNA inversions. *Hum. Mol. Genetics* **2**, 1773-1778.

Nishi,T., Lee,P.S., Oka,K., Levi,V.A. and Tanase,S. (1991) Differential expression of two types of the neurofibromatosis type I (NF-1) gene transcripts related to neuronal differentiation. *Oncogene* **6,** 1555-1559.

Nordlund,M., Gu,X., Shipley,M.T. and Ratner,N. (1993) Neurofibromin is enriched in the endoplasmic reticulum of CNS neurons. *Neuroscience* **13**, 1588-1600.

North,K., Joy,P., Yullie,D., Cocks,N., Mobbs,E., Hutchins,P, McHugh,K. and DeSilva,M. (1994) Specifc learning disability in children with neurofibromatosis type-1: Significance of MRI abnormalities. *Neurology* **44**, 878-883.

O'Connell,P., Leach,R.J., Ledbetter,D.H., Cawthon,R.M., Culver,M., Eldridge,J.R., Frej,A.K., Holm,T.R., Wolff,E., Thayer,M.J., Schafer,A.J., Fountain,J.W., Wallace,M.R., Collins,F.S., Skolnick,M.H., Rich,D.C., Fournier,R.E.K., Baty,B.J., Carey,J.C., Leppert,M.F., Lathrop,G.M., Lalouel,J.M. and White,R. (1989) Fine structure DNA mapping studies of the chromosomal region harbouring the genetic defect in neurofibromatosis type-1. *Am.J. Hum.Genet.* **44**, 51-57.

O'Connell,P., Leach,R.J., Cawthon,R.M., Culver,M., Stevens,J., Viskochil,D., Fournier,R.E.K., Rich,D.C., Ledbetter,D.H. and White,R. (1990a) Two NF-1 translocations map within a 600 kilobase fragment of 17q11.2. *Science* **244**, 1087-1088.

O'Connell,P., Viskochil,D., Buchberg,A.M., Fountain,J., Cawthon,R.M., Culver,M., Stevens,J., Rich,D.C., Ledbetter,D.H., Wallace,M., Carey,J.C., Jenkins,N.A., Copeland,N.G., Collins,F.S. and White,R. (1990b) The human homologue of murine evi2 lies between two von Recklinghausen neurofibromatosis translocations. *Genomics* **7**, 547-554.

Orita,M., Suzuki,Y., Sekiya,T. and Hayashi,K. (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**, 874-879.

Orr,H.T., Chung,M.Y., Banfi,S., Kwiatkowski,T.J., Servadio,A., Beaudet,A., McCall,A.E., Duvick,L.A., Ranum,L. and Zoghbi,H. (1993) Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type-1. *Nature Genetics*, **4**, 221-226.

Padgett,R.A., Grabowski,P.J., Konarska,M.M., Seiler,S. and Sharp,P.A. (1986) Splicing of messenger RNA precursors. *Ann. Rev. Biochem.* **55**, 1119-11150.

Pagnier,J., Mears,J.G., Dunda-Behlkhodja,O., Schaefer-Rego,K.E., Beldjord,C., Nagel,R.L. and Labie,D. (1984) Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. *Proc. Natl. Acad Sci USA* **81**,1771-1773.

Pakula,A.A. and Sauer,R.T. (1989) Genetic analysis of protein stability and function. *Ann. Rev. Genet.* **23**, 289-310.

Poullet,P., Lin,B., Esson,K. and Tamanoi,F. (1994) Functional significance of lysine 1423 of neurofibromin and characterisation of a second site suppressor which rescues mutations at this residue and suppresses RAS2 (Val-19) activated phenotypes. *Molecular and cellular biology* **14**, 815-821.

Purandare,S.M., Davidson,H.R., Lanyon,W.G. and Connor,J.M. (1994a) A novel insertional mutation of a single base in exon 34 of the neurofibromatosis -1 gene. *Hum. Mut.* **3,** 76-78.

Purandare,S.M., Lanyon,W.G. and Connor,J.M. (1994b) Characterisation of inherited and sporadic mutations in neurofibromatosis type-1. *Hum. Mol. Genet.* **3**, 1109-1115.

Riccardi,V.M. (1994) Genotype, malleotype, phenotype, and randomness: lessons from neurofibromatosis (NF-1). *Am J Hum. Genet* **53,** 301-304.

Robberson,B.L., Cote,G.J. and Berget,S.M. (1990) Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**, 84-94.

Rodenheiser,D.I., Coulter-Mackie,M.B. and Singh,S.M. (1993) Evidence of DNA methylation in the neurofibromatosis type-1 (NF-1) gene. *Hum. Mol. Genet.* **2**, 439-444.

Roest,A.M., Roberts,R., Sugino,S., Van-Ommen,G.B. and DenDunnen,J. (1993) Protein truncation test for rapid detection of translation terminating mutations. *Hum Mol Genet,* **2**, 1719-1721.

Rolfs,A., Schuller,I., Finckh,U. and Weber-Rolfs,I. (1992) PCR:Clinical diagnostics and research. Springer Verlag.

Schmale,M.C., and Hensley,G.T. (1988) Transmissibility of a neurofibromatosis like disease in bi-color damselfish. *Cancer Research* **48**, 3828.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour.

Sanger,F., Nicklens,S. and Coulson,A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Sapienza,C. (1991) Genomic imprinting and carcinogenesis. *Biochimica et biophysica acta* **1072**, 51-61.

Sarkar,G., Yoon,H.S. and Sommer,S.S. (1992a) Screening formutations by RNA single strand conformational polymorphism (rSSCP): comparison with DNA SSCP. *Nucleic acids research* **20**, 871-878.

Sarkar,G., Yoon,H.S. and Sommer,S.S. (1992b) Dideoxyfingerprinting (ddE) a rapid and efficient screen for the presence of mutations. *Genomics* **13**, 441-443.

Sawadago,M., VanDyke,M.W. (1991) A rapid method for purification of deprotected oligonucleotides. *Nucleic acids research* **19**, 674.

Schafer,G.L, Ciment,G., Stocker,K.M. and Baizer,L. (1993) Analysis of the sequence and embryonic expression of chicken neurofibromin mRNA. *Molecular and chemical neuropathology* **183**, 267-278.

Schmidt,M.A., Michels,V.V. and Dewald,G.W. (1987) Cases of neurofibromatosis with rearrangements of chromosome 17 involving band 17q11.2. *Am.J.Med.Genet.* **28**, 771-777.

Seizinger,B.R., Rouleau,G.A, Ozelius,L.J., Lane,A.H., Faryniarz,A.G., Chao,M.V., Huson,S., Korf,B.R., Parry,D.M., Periack-VAnce,M.A., Collins,F.S., Hobbs,W., Falcone,B.G., Iannazzi,J.A., Roy,J.C., St.George-Hyslop,P.H., Tanzi,R.E., Bothwell,M.A., Upadhyaya,M., Harper,P., Goldstein,A.E., Hoover,D.L., Bader,J.L., Spence,M.A., Mulvihill,J.J., Aylaworth,A.S., Vance,J.M., Rossenwasser,G., Gaskell,P.C., Roses,A.D., Martuza,R.L., Breakfield,X.O. and Gusella,J.F. (1987) Genetic Linkage of vonRecklinghausen neurofibromatosis to the nerve growth factor receptor gene. *Cell* **49,** 589-594.

Seizinger,B.R. (1993) NF-1: a prevalent cause of tumourigenesis in human cancers? *Nature Genetics* **3,** 97-99.

Sellner,L.N. Coelen,M. and McKenzie,I. (1992) Reverse transcriptase inhibits Taq polymersase activity. *Nucleic Acids Research* **20**, 1487-1490.

Senapathy,P., Shapiro,M.B. and Harris, N.L. (1990) Splice junctions, branchpoint sites, and exons: sequence statistics, identification and applications to the genome project. *Methods in enzymology* **183**, 252-278.

Sharland,M., Taylor,R., Patton,M.A. and Jeffrey,S. (1992) Absence of linkage to Noonan syndrome to the neurofibromatosis type-1 locus. *J. Med. Genet.* **29,** 188-190.

Sheffield,V.C., Cox,D.R., Lerman,L.S. and Myers,R.M. (1989) Attachment of a 40 base pair G+C rich sequence (GC-clamp) to genomic DNA fragments

by the polymerase chain reaction results in improved detection of single base changes. *Proc. Natl. Acad. Sci. USA* **86**,232-236.

Shen,M.H. and Upadhyaya,M. (1993a) A de novo nonsense mutation in exon 28 of the neurofibromatosis type 1 (NF1) gene. *Hum. Genet.* **92,** 410-412.

Shen,M.H., Harper,P.S. and Upadhyaya,M. (1993b) Neurofibromatosis type1 (NF-1): The search for mutations by PCR heteroduplex analysis on Hydrolink gels. *Hum. Mol. Genet.* **2**, 1861-1864.

Skuse,G.R., Kosciolek,B. and Rowley,P.T. (1989) Molecular genetic analysis of tumours in von Recklinghausen neurofibromatosis: Loss of heterozygosity for chromosome 17. *Genes chromosomes Cancer* **1**,36-41.

Skuse,G.R., Kosciolek,B. and Rowley,P.T. (1991) The neurofibroma in von Recklinghausen neurofibromatosis has a unicellular origin. *Am. J. Hum. Genet.* **49,** 600-607.

Smith,H.O., Annau,T.M. and Chandrasegaran,S. (1990) Finding sequence motifs in groups of functionally related proteins. *Proc. Natl. Acad. Sci. USA* **87**, 826-830.

Southern,E.M. (1975) Detection of specifc sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503-517.

Stark,M., Assum,G. and Krone,W. (1991) A small deletion and an adjacent base exchange in a potential stem loop region of the NF-1 gene. *Hum. Genet.* **87**, 685-687.

Stephens,K., Kayes,L., Riccardi,V.M., Rising,M., Sybert,V.P. and Pagon, R.A. (1992) Preferential mutation of the neurofibromatosis type I gene in paternally derived chromosmes *Hum Genet* **88**, 279.

Stern,H.J., Saal,H.M., Lee,J.S., Fain,P.R., Goldgar,D.E., Rosenbaum,K.N. and Barker,D.F (1992) Clinical variability of type-1 neurofibromatosis: Is there a neurofibromatosis-Noonan syndrome? *J. Med. Genet.* **9,** 184-187.

Stoflet,E.S., Koeberl,D.D., Sarkar,G. and Sommer,S.S. (1988) Genomic amplification with transcript sequencing. *Science* **239,** 491-494.

Suzuki,Y., Suzuki,H., Kayama,T., Yoshimoto,T. and Shibahara,S. (1991) Brain tumours predominantly express the neurofibromatosis type-1 gene transcripts containing the 63-base insert in the region coding for the GTPase activating protein related domain. *Biochimica Biophysica Acta Res. Comm.* **181,** 955-961.

Step by step protocols for DNA sequencing with sequenase version 2.0. 5th edition U.S.B

Tassabehji,M., Read,A.P., Newton,V.E, Harris,R., Balling,R., Gruss,P. and Strachan,T. (1992) Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature*, **355**, 635-636.

Tassabehji,M., Strachan,T., Sharland,M., Colley,A., Donnai,D., Harriss,R. and Thakker,N. (1993) Tandem duplication within a NF-1 gene exon in a family with features of Watson syndrome and Noonan syndrome. *Am. J. Hum. Genet.* **53**, 90-95.

The,I., Murthy,A.E., Hannigan,G.E., Jacoby,L.B., Menon,A.G., Gusella,J.F. and Bernards,A. (1993) Neurofibromatosis type 1 gene mutations in neuroblastoma. *Nature Genet.* **3**, 362-366.

Trahey,M. and McCormick,F. (1987) A cytoplasmic protein stimulates normal n-ras p21 GTPase, but does not affect oncogenic mutants. *Science* **238**, 542-545.

Upadhyaya,M., Cherryson,A., Broadhead,W., Fryer,A., Shaw,D.J., Huson,S., Wallace,M.R., Andersen,L.B., Marchuk,D.A., Viskochil,D., Black,D., OConnell,P., Collins,F.S. and Harper,P.S. (1990) A 90kb deletion associated with neurofibromatosis type-1. *J. Med. Genet.* **27**, 738-741.

Upadhyaya,M., Shen,M., Cherryson,A., Farnham,J., Maynard,J., Huson,S.M. and Harper,P.S. (1992) Analysis of mutations at the neurofibromatosis-1 (NF-1) locus. *Hum. Mol. Genet.* **1**, 735-740.

Valero,M.C., Velasco,E., Moreno,F. and Hernandez-Chico,C. (1994) Characterisation of four mutations in the neurofibromatosis type 1 gene by denaturing gradient electrophoresis (DGGE). *Hum. Mol. Genet.* **3**, 639-641.

VanTuinen,P., Rich,D.C., Summer,K.M. and Ledbetter,D.H. (1987) Regional mapping panel for human chromosome 17: Application to neurofibromatosis type1. *Genomics* **1**, 371-381.

Viskochil,D., Buchberg,A.M, Xu,G., Cawthon,R.M., Stevens,J., Wolff,R.K., Culver,M., Carey,J.C., Copeland,N.G., Jenkins,N.A., White,R. and O'Connell,P. (1990) Deletions and translocations interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell* **62**, 187-192.

Viskochil,D., Cawthon,R.M., O'Connell,P., Xu,G., Stevens,J., Culver,M., Carey,J.C. and White,R. (1991). The gene encoding the oligodendrocyte-

myelin protein is embedded within the neurofibromatosis type-1 gene. *Mol.Cell. Biol.* **11**, 906-912.

Viskochil,D., White,R. and Cawthon,R.M. (1993) The neurofibromatosis type-1 gene. *Annual review of neuroscience 1993* **16,** 183-205.

Wallace,M.R., Marchuk,D.A., Andersen,L.B., Letcher,R., Odeh,H.M., Saulino,A.M., Fountain,J.W., Breeton,A., Nicholson,J., Mitchell,A.L., Brownstein,B.H. and Collins,F.S. (1990a) Type 1 neurofibromatosis gene: Identification of a large transcript disrupted in 3 NF-1 patients. *Science* **249**, 181-186.

Wallace,M.R., Andersen,L.B., Fountain,J.W., Odeh,H.M., Viskochil,D., Marchuk,D., O'Connell,P., White,R. and Collins,F.S. (1990b) A chromosome jump crosses a translocation breakpoint in the von Recklinghausen neurofibromatosis region. *Genes Chromosomes Cancer* **2**, 271-277.

Wallace,M.R., Andersen,L.B., Saulino,A.M., Gregory,P.E., Glover,T.W. and Collins,F.S. (1991) A de novo *Alu* insertion results in neurofibromatosis type 1. *Nature* **353**, 864-866.

Watson,J.D., Gilman,M., Witkowski,J., Zoller,M. (editors) (1992) Recombinant DNA, Scientific American books.

Weatherall, D.J. (1991) The New Genetics and clinical practice. Oxford University Press.

Xu, W., Yu,Q., Lihzi,L., Ponder,M., Wallace,M., Gangfeng,X. and Ponder,B. (1992) Molecular analysis of neurofibromatosis type 1 mutations. *Hum. Mut.* **1**, 474-477.

Weismuller,L. and Wittinghoffer,A. (1992) Expression of the GTPase activating domain of the neurofibromatosis type I gene in Eschericia coli and the role of the conserved lysine residue. *J. Biol. Chem.* **267**, 10207.

White,M.B., Carvalho,M., Derse,D., O'Brien,S.J. and Dean,M. (1992) Detecting single base substitutions as heteroduplex polymorphisms. *Genomics* **12**, 301-306.

Xu,G., O'Connell,P., Viskochil,D., Cawthon,R., Robertson,M., Culver,M., Dunn,D., Stevens,J., Gasteland,R., White,R. and Weiss,R. (1990a) The neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* **62**,599-608.

Xu,G., Lin,B., Tanaka,K., Dunn,D., Wood,D., Gesteland,R., White,R., Weiss,R. and Tamanoi,F. (1990b) The catalytic domain of the

neurofibromatosis type I gene product stimulates rasGTPase and complements ira mutants of S.Cerevisae. *Cell* **63**, 835-8411.

Xu,G., Nelson,L., O'Connell,P., White,R. (1991) An Alu polymorphism intragenic to the neurofibromatosis type 1 gene (NF1). *Nucleic Acids Research* **19**, 3764.

Xu,W., Mulligan,L.M., Ponder,M.A., Liu,L., Smith,B.A., Mathew,C.P.G., Ponder,B.A.J. (1992) Loss of NF-1 alleles in phaeochromocytomas from patients with type-1 neurofibromatosis. *Genes chromosomes Cancer* **4**, 337-342.

Yu,C.L., Tsai,M.H. and Stacey,D.W. (1988) Cellular ras activity and phospholipid metabolism. *Cell* **52**, 63-71.

Zhong,J., Speigel,R., Boltshauser,E. and Schmid,W. (1993) Two novel mutations 5108delAG and 5816insG in the NF-1 gene detected by SSCP analysis. *Hum. Mol. Genet.* **2,** 1491-1492.