Genetic analysis of hereditary sensory, motor and autonomic neuropathies, including a rat model

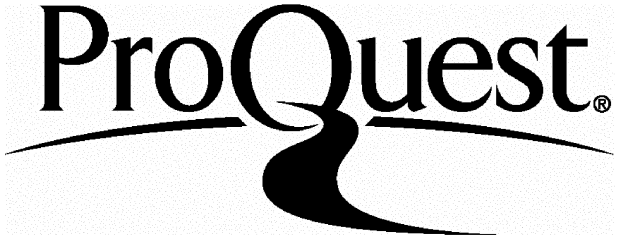
Ming-Jen Lee

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

August 2002

Institute of Neurology

University of London



To my wife and my parents

Who supported me throughout my studies.

Acknowledgements

I would like to thank Dr. Mike Groves and Professor Francesco Scaravilli, for their help in the animal breeding and phenotype characterization. I am also indebted to Dr. Dennis Stephenson, MacLaughlin Institute, Grate Fall, Montana, USA. for advice on positional cloning and on the construction of the rat/mouse/human comparative genetic map. His idea on BLAST searches between Celera mouse genome database and rat markers, made a breakthrough in selecting candidate genes, which lead me to identify the *mf* gene.

I am grateful to Dr. Pete Dixon and Dr. Mary Davis for their advice on genetic linkage and physical mapping. I would also like to thank Ms. Mary Sweeney, for her expert advice on genotyping. Thanks to Dr. Henry Houlden for his help in the initial stage of this project. I am thankful to both of my supervisors, Dr. Mary M. Reilly and Prof. Nicholas W. Wood. They initiated the entire project and I am truly grateful that I have the chance to be involved from the very beginning. I appreciated the support and guidance that both provided. I gratefully acknowledge the financial support from Brain Research Trust and from the Ministry of Education, Taipei, Taiwan. Lastly, I thank God for his blessing on my work and my family living in London.

Abstract

The subject of this thesis is genetic research in two inherited neuropathies, the hereditary sensory and autonomic neuropathies (HSAN) and Charcot-Marie-Tooth diseases (CMTs). Two sections are included; the first is dedicated to identifying the causative gene of a hereditary sensory neuropathy in a spontaneous mutant rat (*mf* rat). The second involves the sequencing of two known genes in patients with CMTs.

The locus for the *mf* disease was linked to the distal end of rat chromosome 14. Using the sequences of the closely linked rat microsatellite markers, database BLAST searches identified a three Megabase region on mouse chromosome 11 whose gene content was co-linear with that on human 2p15-p16. Sequencing the cDNA of genes in this interval revealed a c.1349G>A substitution in subunit 4 of the chaperonin containing TCP-1 protein (*Cct4*) in the mutant rat. This change replaced a highly conserved cysteine450 with tyrosine. This result demonstrates that a mutation in the Cct4 protein causes an inherited sensory neuropathy in rats.

The genes, myelin protein zero (*MPZ*) and gap junction protein beta1 (*GJB1*), were screened for mutations in patients with CMT. Most forms of CMT are either clearly demyelinating or axonal. Mutations in *MPZ* and *GJB1* can cause both types of CMT. Two novel *MPZ* mutations were identified in five Taiwanese patients with a demyelinating neuropathy. Forty-six British patients with axonal CMT were also screened but only one mutation was identified. The frequency of *MPZ* mutations in axonal CMT is low.

GJB1 is the gene responsible for X-linked CMT (CMTX). The clinical features of CMTX patients are variable and the neurophysiological findings can suggest a demyelinating or an axonal neuropathy. 24 mutations (6 novel ones) were identified in 139 patients with CMT. Detailed clinical and electrophysiological studies in these patients are presented.

In conclusion, this thesis identified the genetic defect of the *mf* rat, an animal model for HSAN, and identified mutations in *MPZ* and *GJB1* in CMT patients.

Contents

List of figures

List of tables

Abbreviations

AD	autosomal dominant
ADM	Abductor digiti minimi muscle
APB	Adductor pollucis brevis muscle
AR	autosomal recessive
ATP	Adenosine triphosphate
ATPase	Adenosine triphosphatase
Cct4	Chaperonin containing TCP-1 subunit 4
CHN	Congenital hypomyelination
CIPA	Congenital insensitivity to pain with anhidrosis
CMT	Charcot-Marie-Tooth disease
cM	Centi-Morgan
CMAP	Compound motor action potential
DEPC	Diethylpyrocarbonate
dNTP	Deoxynucleotide acid
DRG	Dorsal root ganglia
DSD	Dejerine-Sottas disease
EDTA	Ethylenediaminetetraacetic acid
EGR2	Early growth response 2 gene
EMG	Electromyography
GDAP1	Ganglioside-induced differentiation-associated protein 1
GJB1	Gap junction protein, beta 1
HNPP	Hereditary neuropathy with liability to pressure palsy
HSAN	Hereditary sensory and autonomic neuropathy
IKAP	IκB kinase complex-associated protein

KIF1Bβ	Isoform of kinesin motor protein 1B
LMNA	Lamin A/C
MTMR-2	myotubularin-related protein-2
MNCV	Motor nerve conduction velocity
MPZ	Myelin protein zero
MRI	Magnetic resonance image
NDRG1	N-myc downstream-regulated gene 1
NF-L	neurofilament light gene
NGF	Nerve growth factor
np-water	Nanopure H_2O
NTRK1	Nerve growth factor tyrosine kinase receptor-1
p75NTR	p75 nerve growth factor receptor
PCR	Polymerase chain reaction
PMP-22	Peripheral myelin protein 22
RF	Recombination frequency
RT-PCR	Reverse-transcriptase polymerase chain reaction
SAP	Sensory action potential
SCV	Sensory conduction velocity
SPTLC1	Serine palmitolyltransferase, long chain base subunit-1
SSCP	Single strand conformation polymorphism
TBE	Tris, Boric acid, EDTA buffer
TE	Tris-EDTA buffer
TEMED	Tetramethlyethylenediamine
TCP-1	Tailless complex polypeptide-1
UV	Ultra violet

Chapter 1

Introduction

Two of the most common types of inherited neuropathies in humans are the hereditary sensory and autonomic neuropathies (HSANs) and the hereditary motor and sensory neuropathies (Charcot-Marie-Tooth disease (CMT)). HSAN mainly affects the sensory peripheral nerves but also involves the motor peripheral nerves and the autonomic nervous system to a variable degree. CMT disease affects the motor and sensory nerves in the peripheral nervous system. This thesis is divided into two parts. The first part is devoted to identifying the causative gene for a mutant rat, which is an animal model for human hereditary sensory and autonomic neuropathy (HSAN). The second part is the study of mutations in two responsible genes and their genotype/phenotype correlation in CMT diseases.

1.1 Hereditary Sensory and Autonomic Neuropathies (HSANs)

The hereditary sensory and autonomic neuropathies (HSANs) are a rare group of disorders characterized by the developmental failure or degeneration of the primary sensory and autonomic neurons [Dyck 1993]. The predominant symptoms are of a sensory neuropathy and include distal sensory loss with complications including foot ulceration and neurotrophic changes and a mutilating and necrotic acropathy. Some patients have autonomic involvement

such as orthostatic hypotension, supine hypertension, arrhythmia, hyperhidrosis or anhidrosis with increased body temperature. Mild motor weakness and muscle atrophy can also be seen. As shown in table 1.1, hereditary sensory and autonomic neuropathy is classified into five types based on the clinical features, electrophysiological findings, mode of inheritance and pathological findings.

Table 1.1. Classification of the hereditary sensory and autonomic neuropathies (HSANs)

HSAN	**Inheritance**	**Genetic locus/ causative gene**
HSAN I	AD	SPTLC1
	AR	Unknown
HSAN II	AR	Unknown
HSAN III	AR	IKAP
HSAN IV	AR	NTRK1
HSAN V	AR	NTRK1
	AR	Unknown

Abbreviations

HSAN = Hereditary sensory and autonomic neuropathy

SPTLC1 = serine palmitolyltransferase, long chain base subunit-1

IKAP = IκB kinase complex-associated protein

NTRK1 = Nerve growth factor tyrosine kinase receptor-1

1.1.1 Classification and molecular genetics of the HSANs

1.1.1.1 HSAN I

In 1922, Hick first described a family with 10 affected members suffering from perforating ulcers of the feet, shooting pains throughout the body and deafness [Dyck 1993a]. The symptoms appeared between 15 and 36 years of age. A painless ulcer of the big toe was usually the first symptom. Bilateral hearing impairment occurred and progressed to total deafness in patients. On examination, the cranial nerves were normal except the auditory nerves. There was loss of pain, touch, and temperature sensation over the feet. In an autopsy of one of the affected family members, Denny-Brown described a marked loss of ganglion cells in the sacral and the lumbar dorsal root ganglia [Denny-Brown D. 1951]. The peripheral nerves and nerve roots showed a considerable loss of nerve fibers. The sciatic nerve had lost about one-third of its fibers at mid-thigh level. Denny-Brown defined this disease as hereditary sensory radicular neuropathy. Familial sensory neuropathy with foot ulcers and bony changes were also reported by Heller and Robb, Mandell and Smith, Campbell and Hoffman, DeLeon, Ervin and Sternback, and Silverman and Gilden [Heller and Robb 1955] [Mandell 1960] [Campbell 1964] [DeLeon 1969] [Ervin 1960] [Silverman 1959].

Taking all these case reports together, the characteristic clinical features of HSAN I include: (1) the mode of inheritance is usually autosomal dominant; however, sporadic cases are also described; (2) the onset is usually between the

second and fourth decades; (3) sensory involvement is more prominent than autonomic or motor involvement; (4) feet are generally much more affected than hands; (5) the neurophysiological and pathological features are those of chronic neuronal (axonal) degeneration; (6) the central nervous system is usually not affected. Foot complications, such as plantar ulcers, recurring paronychia of the toes, stress fractures of foot bones, recurring cellulitis, osteomyelitis, and resorption of bones, are a prominent feature in some kindreds [Giaccai 1952] [Granger 1960].

On examination, all modalities of sensation are affected, especially pain and thermal sensation. Sensation is decreased over the feet and distal legs and in some cases over the fingers and hands. Ankle reflexes are reduced or absent and knee reflexes are often reduced. In addition to a sensory neuropathy, pes cavus and distal muscle atrophy are found in some cases [England and Denny-Brown 1952] [Dyck *et al.* 1965]. In those kindreds with HSAN I and peroneal muscular atrophy, the classification into either HSAN I or CMT2 can be difficult. Neurophysiologically, sensory action potentials are usually absent with variable motor involvement. Nerve pathology shows markedly decreased numbers of unmyelinated, small myelinated, and large myelinated fibers in that order of decreasing severity.

In 1996, Nicholson *et al.* linked HSAN I in an Australian family to chromosome 9q22.1-q22.3 and the linkage was confirmed in a German family by Bejaoui *et al.* [Nicholson *et al.* 1996b] [Bejaoui *et al.* 1999]. Subsequently, mutations in the serine palmitoyltransferase long-chain base subunit 1 gene (*SPTLC1*) at this

region, were identified to cause the disease in 11 HSAN I families [Dawkins *et al.* 2001]. These mutations are associated with increased *de novo* glucosyl ceramide synthesis in lymphoblast cell lines in affected individuals. The ceramide triggers apoptosis which results in massive cell death during neural tube closure leading to cell loss in the dorsal root ganglia [Dawkins *et al.* 2001].

1.1.1.2 HSAN II

In 1952, Giaccai first reported four children from two consanguineous marriages with HSAN II. A sensory neuropathy with foot ulcers and acropathy was noted in these children [Giaccai 1952]. In 1964, Johnson and Spalding reported two unrelated boys with the same condition, each of whom had consanguineous parents [Johnson 1964]. The age of onset was in early childhood. The sensory disturbance involved all modalities with a slowly progressive course and without motor or autonomic dysfunction. In 1973, Ohta *et al* reported a Quebec kinship with congenital sensory neuropathy [Ohta 1973]. The disease began in the first decade. The sensory impairment involved all modalities. Loss of sensory action potentials was noted in nerve conduction studies. Nerve biopsy showed marked loss of large and small myelinated fibers with relatively less involvement of unmyelinated fibers. They suggested that this disease, which appeared to be autosomal recessive, should be called hereditary sensory and autonomic neuropathy type II to differentiate it from HSAN I where there is profound loss of unmyelinated fibers with autosmal dominant inheritance. Apart from this classification, HSAN type II had also been subclassified into two sub-groups, a

progressive and a non-progressive type [Bradley 1974] [Tamari *et al.* 1980]. Nukada *et al.* followed up four patients with HSAN type II where both clinical and pathological findings were progressive [Nukada *et al.* 1982]. However, other families with a recessive hereditary sensory neuropathy clinically and whose pathological findings were compatible with HSAN II, have been described to have a non-progressive course [Murray 1973] [Nordborg *et al.* 1981] [Ferrière 1992]. There is still no linkage for either type of HSAN II.

1.1.1.3 HSAN III

The first case report of HSAN III (also called familial dysautonomia) was over 50 years ago by Riley *et al.* [Riley 1949]. The clinical features of this disease are of an autonomic neuropathy with reduction of tears, emotional lability, paroxysmal hypertension, and cold peripheries with increased sweating. Absence of the fungiform papillae of the tongue is characteristic of this disease. Scoliosis can occur and be severe. Adults often have an unsteady gait and poor balance. Sympathovagal function is progressively involved with worsening of orthostatic hypotension, development of supine hypertension, and even bradyarrhythmias. Autopsy findings include demyelination in the medulla, pontine reticular formation and dorsal longitudinal tracts, as well as degeneration and loss of neurons in the autonomic ganglia [Brown 1964].

The disease is more common in Ashkenazi Jewish people. The incidence of familial dysautonomia is 1 in 3,703 among Ashkenazi Jews in Israel [Maayan *et*

al. 1987]. Patients with familial dysautonomia die in early adulthood [Axelrod 1982]. The diagnosis of this disease is suggested by lack of axon flare after intradermal injection of histamine, absence of fungiform papillae on the tongue, pupil miosis after conjunctival instillation of methacholine chloride, and diminished tear flow. These criteria are specific for familial dysautonomia (HSAN III) and are not seen in other hereditary sensory and autonomic neuropathies [Axelrod 1983]. A sensory neuropathy is common in familial dysautonomia with sural nerve biopsies showing a reduced transverse fascicular area, significantly reduced numbers of myelinated axons (particularly those of small diameter) and very few unmyelinated fibers. There is also reduced neuronal numbers in sympathetic and sensory ganglia [Pearson 1975].

The causative gene was linked to chromosome 9q31-q33 in 1993 [Blumenfeld *et al.* 1999]. In 2001, Slaugenhaupt *et al.* and Anderson *et al.* described the gene encoding the inhibitor of kappa light polypeptide gene enhancer in b cells, kinase complex (IKBK)-associated protein (*IKBKAP; IKAP*) to be the causative gene [Slaugenhaupt *et al.* 2001] [Anderson *et al.* 2001]. In Cohen's study showed that IKAP which can bind NF-kappaB- inducing kinase (NIK) and IkappaB kinase (IKKs) and assemble them into an active kinase complex, is a scaffold protein and a regulator for three different kinases involved in pro-inflammatory cytokine signaling [Cohen *et al.* 1998]. Krappmann *et al.* suggested that this protein might play a role in general gene-activation, since overexpression of IKAP interferes with the activity of both NF-κB-dependent and independent receptor genes [Krappmann *et al.* 2000]. Familial dysautonomia may be caused by aberrant expression of genes crucial to the development of the sensory and

autonomic nervous systems, secondary to the loss of a functional IKAP protein in specific tissues.

1.1.1.4 HSAN IV

In 1963, Swanson *et al.* first reported two male sibs with HSAN IV (also called congenital insensitivity to pain with anhidrosis (CIPA)) [Swanson 1963]. The children had never shown a painful response to injury. Their parents were normal and there was no consanguinity. Neither boy developed sweating when overheated during warm summer days. They showed no fear of injury and often hurt themselves with self-mutilation. X-ray examination demonstrated a fracture dislocation of the left elbow in both patients. The autonomic problem appeared to severely affect their ability to regulate their body temperature, and the younger bother died suddenly after a febrile illness during which his body temperature exceeded 43°C. An autopsy showed that the volume of Lissauer's tract in the spinal cord was decreased significantly [Swanson 1965]. A nearly uniform population of large fibers was seen with only a few scattered small fibers in the dorsal roots. In the dorsal root ganglia, an almost uniform population of large neurons replaced the normal pattern of large and small ganglion cells. The skin from both the abdomen and finger showed abundant sweat glands, which were anatomically normal. Pinsky and DiGeorge described the same disorder in 3 mentally retarded children with recurrent episodes of unexplained fever, repeated traumatic and thermal injuries, and self-mutilating behaviour [Pinksy and DiGeorge 1966]. Rafel *et al.* studied the pathology of the cutaneous branch of

the radial nerve in HSAN IV patients and found complete loss of small myelinated and unmyelinated fibers [Rafel *et al.* 1980]. They concluded that the disease is not a sensory neuropathy but a developmental defect of the neural crest and the peripheral nerve system. Ishii *et al.* described an affected girl from Japan who died at the age of 21 months [Ishii *et al.* 1988]. She suffered from recurrent episodes of unexplained high fever without sweating and was found to have lack of sensation to pain. She bit off the apical part of her tongue and began self-mutilating her lips and the tips of her fingers. More recently, Rosemberg *et al.* reviewed the literature which included 31 patients [Rosemberg *et al.* 1994]. This review demonstrated that 20% of patients died from hyperpyrexia, most of them before age 3. Most of the children are mentally retarded. Virtual absence of unmyelinated and marked reduction in small myelinated fibers of biopsied sural nerves has been found in most of the specimens from patients suffering from this disorder.

Ohta suggested that the congenital insensitivity to pain with anhidrosis should be called type IV HSAN. Autonomic involvement is significant in both HSAN type III and type IV, which differentiates them from type I and II [Ohta 1973].

In *trkA*-deficient ($trkA^{-/-}$) mice, lacking the tyrosine kinase receptor for nerve growth factor, all dorsal root ganglia neurons associated with nociceptive functions are lost and the mice share phenotypical features with CIPA [Smeyne *et al.* 1994]. Indo *et al.* screened the human *trkA* homologue gene, *NTRK1*, in three unrelated CIPA patients who had consanguineous parents. They detected a deletion, a missense and a splicing site mutation in these three patients. Their

findings strongly suggested that mutations in *NTRK1* cause CIPA and that the NGF-NTRK1 system has a crucial role in the development and function of the nociceptive sensation and thermoregulation in humans [Indo *et al.* 1996]. There have been more mutations of *NTRK1* reported in recent years associated with HSAN IV [Mardy *et al.* 1999]. This is an example of the identification of a causative gene in an animal model for a human hereditary disorder with subsequent confirmation that the same gene is responsible for the human disease.

1.1.1.5 HSAN V

Low *et al.* first reported a 6-year-old girl who suffered from recurrent ulceration of her extremities and a degenerative arthropathy of both ankles [Low *et al.* 1978]. Her pregnancy and birth history were normal. The disease began before 2 years of age when she sustained a fracture of her left ankle that remained undiagnosed for four days because it caused her no discomfort. Her hand was repeatedly injured without apparently causing pain leading to an ulcer, which did not heal for several months. Her cognitive function was normal. All tendon reflexes were present and normal. Touch-pressure, proprioceptive and vibration sense were unimpaired. However, pain perception was absent in both legs below the groin and in both arms below the mid-forearm. Temperature sensation was abnormal in the feet. Electrophysiological findings and the sweating pattern were within normal range. A sural nerve biopsy showed a marked loss of small myelinated fibers and the total myelinated fiber density was toward the lower limit of the controls. On electron microscopy, the fiber density and diameter

distribution of the unmyelinated fiber were normal. Subsequently, Dyck *et al.* reported a 26 month old girl with insensitivity to pain and anhidrosis [Dyck *et al.* 1983]. She was born after a normal pregnancy but she was hypotonic and had poor feeding from birth. She was noted not to react to pain. She appeared irritable during periods of hot weather and her body temperature was found to be elevated up to 39°C. At 15 months she had burned both hands without feeling pain. She frequently bit her tongue and lips causing them to bleed without pain. The motor nerve conduction velocity (MNCV) and sensory nerve action potentials were normal. Morphometric studies of a sural nerve biopsy demonstrated severe reduction of the small diameter myelinated fibers. The unmyelinated fibers appeared to be slightly reduced in number. These findings are similar to that found in Low *et al.*'s patient.

Donaghy *et al.* also published a family from Kashmir with a consanguineous marriage suffering from hereditary sensory neuropathy with neurotrophic keratitis [Donaghy *et al.* 1987]. Affected members suffered from ulceration of the tip of the tongue when they were 2 or 3 years old. Frequent whitlows led to loss of terminal finger pulp and deformities of the toes with autoamputation. At around 5 or 6 years of age, they could not feel pain or temperature in their hands and feet, anterior tongue or around the nose. Prior to the presentation of the sensory abnormalities, bilateral keratitis with corneal opacities was noted. Sensory and motor nerve conduction studies were normal. Sweating function was affected with only a few patches of skin in the abdomen, axilla, neck and forehead regions showing incomplete sweating. A sural nerve biopsy showed a lack of small myelinated fibers. A fiber size-frequency histogram confirmed a

striking selective reduction in the smaller diameter myelinated fiber peak. The unmyelinated fibers were morphologically normal on electron microscopy. The density of unmyelinated axons was increased. Although the above reports are very similar to HSAN IV, the loss of small myelinated fibers rather than unmyelinated fibers suggested that this may be a separate disease, and thus it was classified as HSAN V. Recently, a Pakistani boy with consanguineous parents who suffered from dry skin, hyperpyrexia, recurrent syncope attacks and anhidrosis was reported [Houlden *et al.* 2001]. The boy also had injured his right ankle without feeling pain. Pinprick and temperature senses were lost distally in the limbs. Deep pain sensation in his feet was absent bilaterally. Nerve conduction studies were normal. The clinical phenotype was similar to HSAN IV. However, the sural nerve biopsy showed a relative deficiency of small myelinated fibers similar to HSAN V. The unmyelinated axon density was slightly reduced compared with controls. A novel mutation of *NTRK1*, which is the responsible gene for HSAN IV, was identified in this patient [Houlden *et al.* 2001]. The two disorders are therefore likely to be allelic. However, this is still debatable since the autonomic involvement in this Pakistani boy is more prominent than those patients reported by Low *et al.* and Donaghy *et al.* There are no reports yet of *NTRK1* screening in the above-described families with a classical phenotype of HSAN V. It is therefore unknown, as to whether HSAN - IV and –V are due to abnormalities in the same gene.

1.2 Charcot-Marie-Tooth disease

Charcot-Marie-Tooth disease (CMT), also known as hereditary motor and sensory neuropathy (HMSN) is a relatively common disorder (CMT1, 1 in 5000 [Harding 1995]). The common clinical features include progressive distal muscle wasting and weakness, distal sensory loss, hyporeflexia and foot deformities. Electrophysiologically, there are two main types, demyelinating (type 1) and axonal (type 2). Inheritance can be autosomal dominant, autosomal recessive or X-linked. Knowledge of the molecular genetics in this disease has progressed dramatically in the last ten years.

1.2.1 Historical review

Charcot-Marie-Tooth disease is an inherited neuropathy with progressive muscle wasting, weakness, and sensory disturbances. Virchow, Eulenburg Friedreich and others had described peroneal musclular atrophy prior to the report of Charcot and Marie [Brody and Wilkins 1967]. In 1873, Eichhorst described a family with peroneal muscle wasting in six generations. Charcot and Marie described a specific form of progressive muscular atrophy in 1886 (review in [Brody and Wilkins 1967]). The disease was often familial and symptoms started in the feet and legs and later involved the hands. It was thought that the basis for the disorder was either a myelopathy or a neuropathy. Tooth's thesis on the peroneal type of progressive muscular atrophy was also published in 1886. He concluded that the disease affected peripheral nerves [Dyck *et al.* 1993b].

This type of disease was called Charcot-Marie-Tooth disease (CMT). Bell J. who reviewed patients with CMT, concluded that the disease could be inherited as a dominant, recessive, or X-linked trait and that some cases were sporadic [Bell 1935]. In 1968, using neurophysiological and pathological criteria, Dyck and Lambert differentiated this heterogeneous syndrome into type I (demyelinating) and type II (axonal) [Dyck and Lambert 1968a] [Dyck and Lambert 1968b]. Major advances have occurred in the understanding of the underlying genetics of CMT in the last decade, necessitating a more detailed classification (Table 1.2). In addition to the description of the causative genes for many of these disorders, progress is occurring in understanding the pathogenesis of CMT using knockout and transgenic mice and cell line or Xenopus öocyte expression systems in vitro.

1.2.2 Classification of Charcot-Marie-Tooth disease

The classification of CMT is constantly changing as the underlying genetic defects are being described. The term "CMT" is used widely in the genetic literature; however, "hereditary motor and sensory neuropathy" (HMSN) tends to be used in the clinical literature. CMT is a heterogeneous group of disorders. A genetic classification cannot cover all the clinical syndromes of HMSN yet, as the underlying genetic defect has not been described in all of these syndromes. The classification in table 1.2 combines a genetic and clinical classification.

Table 1.2 Classification of Charcot-Marie-Tooth disease (CMT)

CMT1 (HMSN I)	**Inheritance**	**Genetic locus/ causative gene**
CMT1A (HMSN IA)	AD	Duplication 17p11.2 PMP-22
CMT1B (HMSN IB)	AD	MPZ
CMT1C (HMSN IC)	AD	16p13.1- p12.3
CMT1D (HMSN ID)	AD	EGR2
CMTX (HMSN X)	X-linked, CMT X1	GJB1
HNPP		
HNPP A	AD	Deletion 17p11.2 PMP-22
DSD (HMSN III)		
DSD A	AD or AR	PMP-22
DSD B	AD or AR	MPZ
DSD	AD	EGR2
DSD	AD	8q23-q24
CHN		
CHN	AD	PMP-22
CHN	AD	MPZ
CHN	AD/AR	EGR2
INTERMEDIATE CMT (DI-CMT)		
DI-CMT1	AD	10q24.1-q25.1
DI-CMT2	AD	19p12-p13.2
CMT2 (HMSN II)		
CMT2A (HMSN IIA)	AD	KIF1Bβ
CMT2B (HMSN IIB)	AD	3q13-q33
CMT2C (HMSN IIC)	AD	Unknown
CMT2D (HMSN IID)	AD	7p14
CMT2E (HMSN IIE)	AD	NF-L
CMT2F (HMSN IIF)	AD	7q11-q21
CMT2 (HMSN II)	AD	MPZ
AR CMT, demyelinating		
CMT4A	AR	GDAP1
CMT4B1	AR	MTMR-2
CMT4B2	AR	11p15
CMT	AR	5q23-q33
HMSNL	AR	NDRG1
HMSNR	AR	10q23
CCFDN	AR	18qter
CMT 4F	AR	Periaxin
AR CMT, axonal		
CMT2	AR	LMNA
CMT2	AR	19q13.3
CMT4C1	AR	Unknown
CMT4C2	AR	8q21.3
CMT2	AR	GDAP1

Abbreviations:

CMT= Charcot-Marie-Tooth disease
DSD= Dejerine-Sottas disease
HNPP= Hereditary neuropathy with liability to pressure palsy
CHN= Congenital hypomyelination
AD= autosomal dominant
AR= autosomal recessive
PMP-22= peripheral myelin protein 22

MPZ= myelin protein zero
GDAP1 = ganglioside-induced differentiation-associated protein 1
GJB1= gap junction protein beta 1, connexin32
EGR2 = early growth response 2
NF-L= neurofilament light gene
MTMR-2= myotubularin-related protein-2
LMNA = Lamin A/C
KIF1Bβ = isoform of kinesin motor protein 1B
NDRG1= N-myc downstream-regulated gene 1

1.2.3 Clinical and pathological features of Charcot-Marie-Tooth disease

The classification of CMT was traditionally based on neurophysiology and nerve pathology [Dyck *et al.* 1993b]. Clinically it is still the most useful way to classify these disorders, but the combination of clinical features and genetic findings are being increasingly used for classification.

1.2.3.1 CMT1

CMT1 (HMSN type I) is the most common type of inherited neuropathy, with a prevalence of 1 in 5000 in the United Kingdom [Harding 1995]. The age of onset is usually between the 1st and 2nd decades. Clinically, patients usually have a slowly progressive distal muscle wasting and weakness, predominantly involving the anterior tibialis and peroneal muscles. Muscle weakness usually manifests as an abnormality of gait. Muscle weakness and atrophy begin insidiously in the feet and leg muscles, especially the intrinsic foot and peroneal muscles. Cramps, particularly after exercise, and fasciculations in affected muscles are common. Sometimes patients with CMT1 do not complain of sensory symptoms, although they may be aware that the feeling in their feet is diminished. Sensory loss usually affects all modalities and is present distally in the feet and to a lesser extent in the hands. Deformities such as pes cavus and hammer toes are frequently seen in patients with CMT1. Sometimes the peripheral nerves are hypertrophied and are palpable. The severity of this disorder varies in and between families.

CMT1 is characterized electrophysiologically by slow motor nerve conduction velocities (MNCV < 38 m/s in the median nerve), increased distal motor latencies and hence is defined as a demyelinating neuropathy. Pathological findings include demyelination, remyelination with onion bulb formation and secondary axonal changes. The onion bulbs consist of circumferentially directed Schwann cells and their processes. Most of the families are autosomal dominant; however, autosomal recessive and X-linked forms have also been described.

1.2.3.2 X-linked CMT

X-linked inheritance should be suspected if there is no male-to-male transmission in the family. In X-linked CMT (CMTX), male patients are more severely affected than females and the electrophysiological findings are compatible with a demyelinating neuropathy in males. The mean value of the ulnar MNCV of affected males was less than 38 m/s in a kindred described by Allan [Dyck *et al.* 1993b]. However, the MNCV is greater than 38 m/s in a few affected males in a large kindred described by Hahn [Hahn *et al.* 1990]. Females are less severely affected and may appear to have an axonal neuropathy [Timmerman *et al.* 1996]. The MNCV of median nerve is often greater than 38 m/s or normal.

1.2.3.3 HNPP

Hereditary neuropathy with liability to pressure palsy (HNPP) is an autosomal dominant neuropathy. Recurrent episodes of pressure palsy of individual nerves such as carpal tunnel syndrome, cubital ulnar nerve palsy and common peroneal nerve palsy are typical features. The diagnostic criteria for HNPP are: (1) an autosomal dominant inheritance, (2) the clinical presentation of a recurrent mononeuropathy simplex or multiplex, (3) a significant slowing of MNCVs and SCVs in affected and non-affected nerves, (4) tomaculous swelling of myelin sheath in nerve biopsy [Meier and Moll 1982]. Enlargement of peripheral nerves are sometimes palpable.

1.2.3.4 CMT2

The classical clinical features of CMT2 are much the same as CMT1, but the age of onset may be later. The main presenting symptoms are foot deformities and difficulty in walking, and the symptoms usually show a very slow progression, with only moderate disability. Weakness and wasting predominate in the distal lower limbs, occurring in the distal upper limbs in only half of cases, and tendon areflexia is usually confined to the lower limbs [Harding and Thomas 1980]. Sensory loss, mainly of vibration sense, is clinically detected in half the patients. In families described by Dyck *et al.*, there are subtle differences in the distribution of symptoms in patients with CMT2 as compared to CMT1: (1) onset of symptoms is usually later, (2) peripheral nerves are not clinically enlarged,

and (3) weakness of small hand muscles is less severe and, (4) weakness and atrophy of plantar flexor muscles of the ankles are more severe [Dyck *et al.* 1993]. Normal or nearly normal MNCVs with reduced amplitude is characteristic of CMT2 (median MNCV > 38 m/s [Harding and Thomas 1980]). Sural nerve biopsy shows a variable loss of myelinated fibers, mainly of large diameter and clusters of regenerating fibers. The disease is usually autosomal dominant, but there are recessive and X-linked families.

Some uncommon forms of CMT2 are described with atypical features or associated other neurological deficits. The atypical features include continuous muscle fiber activity, restless leg syndrome, severe sensory neuropathy and vocal cord paresis, etc [Gemignani and Marbini 2001]. In some CMT2 families, there are other additional neurological features including CMT2 with pyramidal features, CMT2 with optic atrophy and CMT2 with retinitis pigmentosa. Previously, these families have been designed as HMSN V, HMSN VI and HMSN VII [Dyck *et al.* 1993].

The subtype CMT2B is an axonal type motor and sensory neuropathy with a predominant sensory neuropathy. The patients usually suffer from progressive foot drop or foot weakness initially. Some patients also complain of diminished sensation in the feet and toe/foot ulceration. The sensory loss, which involves all modalities, is a prominent feature of CMT2B and it occurs predominantly in the lower extremities. However, there are no symptoms or signs referable to the autonomic nervous system such as abnormalities in sweating, tearing, orthostatic hypotension or sexual dysfunction. In the family reported by Elliott *et al.*, the

median MNCVs of the affected patients were greater than 38 m/s with mildly reduced amplitudes. The peroneal and tibial CMAPs were unobtainable. The sural nerve SAPs were absent as well. These findings indicate an axonal neuropathy.

Auer-Grumbach also described an Australian family with CMT2B [Auer-Grumbach *et al.* 2000a]. Patients developed signs of foot deformity, callus, sensory loss, ulcers and amputations at a young age. Sensory loss involved all modalities. Only a few patients in this family had distal weakness and atrophy in the lower limbs. Electrophysiologically, the MNCVs and SCVs were only moderately reduced or normal. The CMAPs and SAPs were reduced or were unobtainable. These findings suggest primary axonal degeneration. The clinical features of patients in these two families are similar to *mf* disease. However, there are no pathological findings published in these two families.

1.2.3.5 Dejerine-Sottas disease

Dejerine-Sottas disease (DSD) or HMSN III is defined as a severe demyelinating/hypomyelinating neuropathy of early onset. The criteria for a clinical diagnosis are an early onset (before 2 years of age) motor and sensory neuropathy usually with delayed motor developmental milestones and nerve conduction studies showing marked slowing of motor conduction velocities (usually less than 10 m/s in the median nerve) [Ouvrier *et al.* 1987]. Pathologically, severe demyelination is seen with evidence of hypomyelination,

basal lamina onion bulb or classic onion bulb formation [Gabreels-Festen and Gabreels 1993]. The mode of inheritance was originally considered to be autosomal recessive. However, molecular genetic studies have shown that *de novo* mutations in all of the genes involved in CMT1, i.e. *PMP22, MPZ* and *EGR2* are the most common cause of DSD [Parman *et al.* 1999] [Timmerman *et al.* 1999] [Warner *et al.* 1996] [Warner *et al.* 1997]. A few families with this disorder show autosomal dominant inheritance. There is an overlap clinically between DSD and CMT1 and DSD is now considered to be a more severe form of CMT1. However, not all families with DSD link to the loci for CMT1. There is one autosomal dominant DSD family with linkage to chromosome 8 [Ionasescu *et al.* 1996a].

1.2.3.6 CHN

Another rare disorder, called congenital hypomyelination (CHN), is characterized by very slow nerve conduction velocities (median MNCV below 6-7 m/sec) with onset at birth or soon after birth, and by hypo/amyelination or basal lamina onion bulbs on nerve biopsy. The clinical findings are similar to DSD; but pathologically, CHN shows no active demyelination and is considered to be a form of dysmyelination [Gabreels-Festen *et al.* 1994].

1.2.3.7 DI-CMT (dominant intermediate-CMT)

There are families with CMT, who have median MNCVs between 25 and 45 m/s [Davis *et al.* 1978]. These families show autosomal dominant inheritance. Recently it has been suggested that this form of CMT should be classified as dominant intermediate CMT (DI-CMT). This was first proposed by Davis et al. in 1978 and has recently become of interest with the description of two loci for this form of CMT [Davis *et al.* 1978] [Rossi *et al.* 1985] [Villanova *et al.* 1998].

1.2.3.8 Autosomal recessive CMT (ARCMT)

The autosomal recessive CMTs (ARCMT), either demyelinating or axonal, have been classified as CMT4 although as more loci and genes are described, the CMT4 classification has been found to be inadequate [Ben Othmane *et al.* 1993a]. By using clinical, electrophysiological and pathological criteria, three subtypes of CMT4 have been described from a large series of Tunisian families. CMT4A and CMT4B are characterized by slow MCVs but pathologically CMT4A has basal lamina onion bulbs whereas CMT4B has globular masses along the myelinated fibers called focally folded myelin sheaths. The MNCV is preserved and no myelin changes are seen in nerve biopsies from patients with CMT4C1, which is therefore an axonal form.

Many other forms of ARCMT have now been described that do not fit into the CMT4 classification. A form of demyelinating ARCMT has been described in

two Algerian families, characterized by a precocious and rapidly progressive scoliosis in combination with a relatively mild neuropathy. Pathology shows demyelination with onion bulb formation and loss of large myelinated fibers. Another rare group of families from the Balkan Gypsy population, first described in Lom, a small town in northwest Bulgaria, have been described with an autosomal recessive demyelinating neuropathy and deafness (labeled HMSNL) [Kalaydjieva *et al.* 1996]. Nerve biopsy demonstrated demyelination/remyelination with no indication of axonal atrophy. An unusual finding is the presence of intra-axonal accumulations of irregularly arranged curvilinear profiles resembling those findings seen in experimental vitamin E deficiency [King *et al.* 1999]. Tournev and his colleague described a distinct autosomal recessive neuropathy called congenital cataract facial dysmorphism neuropathy (CCFDN) [Tournev *et al.* 1999]. Most of the patients are Wallachian Gypsies. Congenital cataracts and microcorneas can be recognized at infancy. Patients suffered from a motor neuropathy from childhood and severe disability occurred in the third decade. Nerve conduction studies and pathology suggest a demyelinating/hypomyelinating neuropathy. Most of the affected patients have characteristic facial dysmorphism, short stature and some cognitive deficits. During the study of HMSNL, one large kindred of Romany Gypsies (called HMSNR), has been found to have another autosomal recessive demyelinating CMT, characterized by prominent sensory loss with moderately reduced MNCVs and a greatly increased threshold for electrical nerve stimulation [Rogers *et al.* 2000]. Nerve biopsy demonstrated hypomyelination with profuse regenerative activity. Another form of ARCMT has been described in a large inbred Lebanese family. Patients suffered from delayed motor development, muscle weakness

and atrophy at an early age with sensory ataxia, scoliosis and pes cavus deformities. Nerve conduction studies revealed total absence of sensory and motor evoked potentials. Nerve biopsy demonstrated a severe axonal loss following a chronic process of demyelination-remyelination [Delague *et al.* 2000].

1.2.4 Molecular genetics of CMT

1.2.4.1 CMT1, HNPP, DSD and CHN

Genetic studies of autosomal dominant CMT1 has shown that it is a heterogeneous condition involving at least 4 loci. In 1991, the first linkage was described in CMT1 to chromosome 1, which subsequently was defined as CMT1B or HMSN IB [Lebo *et al.* 1991]. Myelin protein zero (*MPZ*) gene encoding the main structural protein of myelin was mapped to chromosome 1q22-23 where is the same region as the CMT1B locus [Hayasaka *et al.* 1991]. Mutations have been found in the *MPZ* gene from patients with CMT1B, suggesting that *MPZ* is the causative gene for this neuropathy [Hayasaka *et al.* 1993a]. In addition, *de novo MPZ* point mutations have also been reported in patients with DSD [Hayasaka *et al.* 1993b] and CHN [Warner *et al.* 1996]. Currently, at least 83 mutations of *MPZ* have been described (Mutation database of Inherited Peripheral Neuropathies, available at http://molgen-www.uia.ac.be /CMTMutations/). Myelin protein zero is the major structural protein in myelin, accounting for at least 50% of peripheral myelin protein. Each MPZ protein is

composed of 248 amino acids and it contains intracellular, extracellular and transmembrane domains [Scherer and Chance 1995]. *MPZ* consists of six exons and each exon corresponds to a proposed functional domain. The extracellular domain, which is composed of exons 2 and 3, forms the immnoglobulin-like structure involved in myelin compaction by homophilic interactions [Lemke *et al.* 1988]. *MPZ* mutations, resulting in structural abnormalities in the MPZ protein itself or an unfavorable interface for the assembly of two adjacent proteins could interfere with myelin compaction and then, induce demyelination and onion bulb formation. Mutation screening of *MPZ* can be done by single strand conformation polymorphism analysis (SSCP) and/or direct sequencing of the coding region of the *MPZ* gene.

Most of the autosomal dominant CMT1 families have been shown to be linked to the short arm of chromosome 17, 17p11.2 and classified as CMT1A [Vance *et al.* 1991]. A large DNA duplication within 17p11.2 was found in individuals with CMT1A [Lupski *et al.* 1991] [Raeymaekers *et al.* 1991]. It is now known that most CMT1A families, and overall about 90% of patients with dominant CMT1, harbor the duplication [Pareyson 1999]. The duplication was shown to be mutated in 70.7% of 819 CMT1 patients, and the *de novo* duplication cases represented 6.7% of the total number of duplication patients in a recent European collaborative study of CMT1[Nelis *et al.* 1996]. It is not found in normal subjects. The duplication involves 1.5 megabases (Mb) of DNA, flanked by repeat sequences 17-29Kb in length (CMT1A-REP). The duplicated segment is constant in the majority of patients. The duplication probably results from intra- or inter-chromosomal exchange. Reciprocal products of the homologous

recombination are thought to be the cause of the deletion of this 1.5Mb segment found in patients with hereditary neuropathy with liability to pressure palsies (HNPP). *De novo* duplications, which are mainly from paternal origin, are assumed to be due to misalignment of the repeat sequence during meiosis. There is a 1.7Kb recombination hotspot within the CMT1A-REP found by Reiter et al. [Reiter *et al.* 1996]. Sequence analysis revealed a mariner-like element near the hotspot, which mediated strand exchange via cleavage by a transposase at or near the 3' end of the element [Reiter *et al.* 1998]. In a large series of *de novo* CMT1A patients, it was found that the 17p11.2 rearrangements that occur during spermatogenesis differ from those that occur during oogenesis. Duplication occurring in spermatogenesis results from the unequal meiotic crossing-over between two chromosome 17 homologues. However, the duplications and deletions from maternal origin are caused by an intra-chromosomal process, either unequal sister chromatid exchange or, in the case of deletions, excision of an intra-chromatidal loop [LeGuern *et al.* 1996a] [Lopes *et al.* 1997] [Lopes *et al.* 1998]. The duplication and deletion can be detected by a number of techniques; the scoring of three alleles or dosage differences on Southern blot hybridization, the quantification of alleles of fluorescent-labelled microsatellite markers using an automated sequencer and by fluorescence in situ hybridization (FISH) or pulsed-field gel electrophoresis.

The Trembler and Trembler-J mice are good animal models for CMT1 as they have an autosomal dominant demyelinating neuropathy. These mice were found to have point mutations in a new myelin gene (*Tr*), at the same time as the 17p duplication was described in CMT1A patients. The *Tr* locus maps to mouse

chromosome 11, which has syntenic homology to human chromosome 17p in the region of the CMT1A locus. The mouse point mutations were found in a gene coding for peripheral myelin protein 22 (*PMP22*). *PMP22* gene was then found to be within the CMT1A duplication segment in humans and therefore was thought to be the likely causative gene for CMT1A. Proof for this hypothesis was obtained when an autosomal dominant CMT1A family, who showed linkage to chromosome 17p but who did not have the duplication, were found to have a point mutation in *PMP22* [Valentijn *et al.* 1992]. It has been postulated that overexpression of the PMP22 protein in patients with CMT1A and the chromosome 17p duplication causes the neuropathy. Further findings supported the hypothesis as described below. One patient from two CMT1A parents both with a duplication of 17p11.2-12, has four copies of the *PMP22* gene and was severely affected. The overexpression of *PMP22* in transgenic mice or rats revealed gait abnormalities, slowing in nerve conduction velocities, severe hypomyelination and onion bulb formation [Sereda *et al.* 1996] [Huxley *et al.* 1996]. Huxley *et al.* in their study suggested that the ratio of human/mouse *PMP22* mRNA in the sciatic nerve of injected transgenic mice is an index for the severity of histological and electrophysiological findings [Huxley *et al.* 1998].

Currently, there are at least 43 *PMP22* mutations described. Some of these are *de novo* mutations and associated with the phenotype of DSD, CHN or HNPP and more rarely CMT1. The PMP22 protein accounts for only 2-5% of peripheral myelin proteins, suggesting that it does not play a major structural role. It may be involved in adhesions between axons and Schwann cells. By using co-immunoprecipitation experiments, D'Urso *et al.* showed that the MPZ

and PMP22 proteins form complexes in the myelin membrane [D'Urso *et al.* 1999]. In transfected cells under confocal microscopy, they found that MPZ and PMP22 are recruited and colocalize at the apposed plasma membranes of expressors [D'Urso *et al.* 1999]. Point mutations in *PMP22* gene can be detected by various methods, including SSCP and direct sequencing of the coding regions of the *PMP22* gene.

Recently, mutations in a new gene, the early growth response 2 gene (*EGR2*) have been found responsible for occasional cases of CMT1, DSD and CHN [Warner *et al.* 1998] [Warner *et al.* 1999]. The *EGR2* gene is located on human chromosome 10q21-22 and is homologous to the murine *Krox-20* gene. Topilko *et al.* found *Krox-20*, which encodes a Cys2/His2 zinc finger transcription factor, controls the myelination process of Schwann cells [Topilko *et al.* 1994]. It is speculated that EGR2, a zinc finger protein, partly controls myelination in human peripheral nerve. Warner *et al.* reported the first mutations in this gene causing both CMT1 and CHN [Warner *et al.* 1998]. Subsequently, mutations of this gene have also been found in patients with DSD [Timmerman *et al.* 1999]. Currently, there have been seven *EGR2* mutations described in patients with CMT, four of which were associated with CMT1, two with CHN and one with DSD. A microarray expression profile has identified that the neuropathy- associated mutated *EGR2* inhibits wild-type *Egr2*-mediated expression of the essential myelin genes (MPZ, PMP22 and myelin associated protein) to levels sufficiently low to result in the abnormal myelination [Nagarajan *et al.* 2001]. SSCP is usually used for screening mutations in the *EGR2* gene with subsequent sequencing to characterize the mutation.

Families have been described with dominant CMT1 that do not link to chromosome 17, 1, or 10. These families are labeled CMT1C. Linkage from two American CMT1C families (with Irish and English ancestry) has recently been described to chromosome 16p13.1-p12.3 [Street *et al.* 2002]. A family with autosomal dominant DSD has been described with linkage to chromosome 8q23-q24 [Ionasescu *et al.* 1996a].

1.2.4.2 HNPP

HNPP is due to a deletion of the 1.5Mb segment of chromosome 17p11.2, which is duplicated in CMT1A[Chance *et al.* 1993]. In 1996, a European study showed an overall deletion frequency of 84% in 162 European patients thought to have HNPP [Nelis *et al.* 1996]. The frequency of *de novo* deletions is 4.6% of the total number of deleted HNPP patients. The deletion has been found in patients with a disease suggesting HNPP, and also in more unusual clinical syndromes including progressive scapuloperoneal syndrome and progressive asymmetric proximal upper limb weakness. Only nine mutations in the *PMP22* gene have been found in HNPP patients without the deletion of 17p11.2 (http://molgen-www.uia.ac.be/CMT Mutations/). The mutations include two nonsense, two splice site, one missense and four frame-shift mutations. It is postulated that mutations of *PMP22* causing HNPP might be due to the loss of function of one copy of the gene. The deletion can be detected by the same methods as duplication detection, and mutations of *PMP22* are detected as described above.

1.2.4.3 X-linked CMT

Linkage studies in families with X-linked CMT mapped this disease to chromosome Xq13.1 in 1993 [Bergoffen *et al.* 1993a]. Since the gene for the gap junction protein beta 1 (*GJB1*; also called connexin32) also mapped to this region, it was considered to be a candidate gene for CMTX. Sequencing of this gene in patients with CMTX confirmed that mutations in *GJB1* gene cause this disease [Bergoffen *et al.* 1993b] [Ionasescu *et al.* 1994]. Currently, 230 mutations of *GJB1* have been described in patients with CMTX. In these families, some patients presented with classical features of CMT and with electrophysiological and pathological features compatible with a demyelinating neuropathy [Tabaraud *et al.* 1999]. In other patients, especially females, an axonal neuropathy was found, suggesting a diagnosis of CMT2 [Hahn, *et al.* 1990] [Senderek *et al.* 1998a]. GJB1 (connexin32) is a member of the family of gap junction proteins. The protein consists of four transmembrane domains, two extracellular loops, one intracellular loop, C- and N-terminals [Bennett *et al.* 1991]. Connexon, which is made up of six connexin32 monomers in a hexagonal structure with a central canal, assembles and is exported onto the cell surface. Two connexons are combined at two adjacent cells to construct a channel used for cell-cell trafficking. This gap junction channel admits passage of small molecules, including ions, second messengers and metabolites (<1000 Da). In the peripheral nervous system, connexin32 is present in non-compacted myelin, i.e. paranodal loop, periaxonal collar, and the Schmidt-Lanterman incisures [Bone *et al.* 1997]. It plays an important role in the communication between the axon-Schwann cells and inner-outer layers of myelin. Mutations,

causing abnormalities in the connexin32 structure or in the assembly of connexons can produce demyelination as well as axonal degeneration. Mutations can be detected by the same techniques described above, e.g. SSCP or direct sequencing.

1.2.4.4 DI-CMT

In 1998, Villanova et al. reported an Italian family with autosomal dominant CMT with the median MNCVs in affected patients ranging between 25 and 45 m/s [Villanova *et al.* 1998]. The localization of the gene responsible was found to be on chromosome 10q24.1-q25.1 [Auer-Grumbach *et al.* 2000b]. In the same year, an Australian family was described with similar clinical and pathological features, showing linkage to chromosome 19p12-p13.2 [Kennerson *et al.* 2001]. These two families with intermediate NCVs are linked to distinct loci different from CMT1 and CMT2. Therefore, DI-CMT seems to be a genuinely separate entity from CMT1 and CMT2, but further characterization of this disease with description of the responsible genes and the disease pathogenesis is necessary.

1.2.4.5 CMT2

The first linkage study for patients with CMT2 was in 1993 where linkage to chromosome 1p35-p36 in three CMT2 families and later one further family was described [Ben Othmane *et al.* 1993b] [Timmerman *et al.* 1996]. This type of

axonal CMT has been termed CMT2A. In the last ten years, as proteins in axonal transport were being studied, many members of the kinesin superfamily of molecular motor proteins (KIFs) were identified. Zhao *et al.* have identified and characterized an isoform of the conventional mitochondrial motor KIF1B protein, called KIF1Bβ [Zhao *et al.* 2001]. In addition to proving that the KIF1Bβ protein transports synaptic vesicle precursors in axons, they also characterized the functional significance of this protein in vivo, showing that knockout mice (*KIF1Bβ*$^{-/-}$) developed a peripheral neuropathy. Because the *KIF1Bβ* gene is located at the CMT2A locus, they sequenced this gene in CMT2A families and identified a missense mutation in affected patients [Zhao *et al.* 2001]. ATPase activity has been shown to be significantly reduced in the mutated protein and cells with the mutated protein have been shown to have lost microtubule motor activity.

A second form of CMT2, CMT2B was linked to chromosome 3q13-22, two years after the linkage of CMT2A [Kwon *et al.* 1995]. The patients with CMT2B have, in combination with weakness of the distal muscles and preserved ankle reflexes, very prominent sensory abnormalities with ulceration, leading to amputation of distal limbs [Elliott *et al.* 1997]. It was argued that the phenotype should be classified as hereditary sensory neuropathy. In an Austrian family with ulcero-mutilating neuropathy, linkage to the known loci for HSAN I and CMT2B has been excluded, suggesting that CMT2B may be genetically heterogeneous.

The gene responsible for CMT2C, an AD axonal neuropathy with laryngeal nerve palsy and/or diaphragmatic weakness, is unknown. In 1996, another family with CMT2 was linked to chromosome 7p14 and was labeled as CMT2D [Ionasescu *et al.* 1996b]. The disease onset was in the second or third decade with a progressive course. Muscle weakness in patients was more severe in the hands than in the feet and the deep tendon reflexes were absent in the upper limbs.

A large Russian family (from the Mordovian Republic) with a typical phenotype of CMT2 associated with hyperkeratosis has been linked to chromosome 8p21. The disease onset was between 2nd to 3rd decades. The patients had walking difficulties followed by weakness and wasting of the leg muscles and a variable degree of foot deformities. Electrophysiological studies in seven patients showed decreased CMAPs and there was EMG evidence of denervation of the distal limb muscles. This type of disease was designated as CMT2E. A p.Q333P (glutamine to proline) mutation in the neurofilament-light chain gene (*NF-L*) has been identified in this family [Mersiyanova *et al.* 2000]. More recently, a double missense mutation resulting in a P8R (proline to arginine) change in *NF-L*, was identified in a Belgian CMT2 family [De Jonghe *et al.* 2001]. The MNCVs of the patients in this family were sometimes severely slowed and could be classified as CMT1 if one only depended on the conduction velocities. However, the CMAPs are always severely reduced, suggesting the possibility of loss of fast conducting axons. Therefore, mutations in the *NF-L* gene should also be screened in CMT1 patients once the mutations in the CMT1 genes have been excluded. A further type of CMT2 has been described in a multi-generation

Russian family. This disease, which was designated as CMT2F, has been linked to a new locus at chromosome 7q11-q21 [Ismailov *et al.* 2001].

In recent years at least three groups found that three families with a CMT2 phenotype harbored *MPZ* mutations. Marrosu *et al.* first published a large Sardinian family with inherited motor and sensory neuropathy, which was then confirmed clinically and electrophysiologically as CMT2, with a missense mutation at Ser44Phe of *MPZ* [Marrosu *et al.* 1998]. De Jonghe and Chapon *et al.* also reported *MPZ* mutations in CMT2 families with a biopsy proven axonal neuropathy [De Jonghe *et al.* 1999] [Chapon *et al.* 1999].

1.2.4.6 ARCMT

ARCMT can be divided into ARCMT1 (demyelinating) and ARCMT2 (axonal). There are at least 8 types of ARCMT1. Demyelinating or axonal CMT with autosomal recessive transmission have been defined as CMT4 in Tunisian families. CMT4A, is described in families from Tunisia characterized by moderately slow motor conduction velocities and basal lamina onion bulbs, and has been linked to chromosome 8q13-21.1 [Ben Othmane *et al.* 1993a]. Recently, mutations in the gene encoding the ganglioside-induced differentiation-associated protein 1 (*GDAP1*), was found in three CMT4A Tunisian families with demyelinating neuropathy but also surprisingly in three Spanish families with an axonal CMT and vocal cord paresis [Baxter *et al.* 2002] [Cuesta *et al.* 2002]. The homologue *Gdap1* is highly expressed in mouse brain

and was identified after ganglioside-induced differentiation of the mouse neuroblastoma cells. This gene may be involved in a signal transduction pathway in neuronal development. Like *MPZ*, mutations in *GDAP1* can cause either a demyelinating or an axonal neuropathy.

CMT4B1, characterized by slow motor conduction velocities and focally folded myelin sheaths, has been linked to chromosome 11q23.1 [Bolino *et al.* 1996]. Bolino A. *et al* found the myotubularin-related protein-2 (*MTMR2*) gene at 11q22, to be the gene responsible for CMT4B1 [Bolino *et al.* 2000a] [Bolino *et al.* 2000b]. This gene belongs to a family of at least eight human genes encoding a dual specificity protein-tyrosine phosphatase (DSP). Phosphatidylinositol 3-phosphate (PI(3)P) is the highly specific substrate for the protein MTMR2 [Kim *et al.* 2002]. Mutations in this gene disrupt the phosphatase activity of MTMR2; it is probable that this disease results from improper regulation of PI(3)P and aberrant signal transduction. A linkage study on another big Tunisian family with demyelinating neuropathy and focally folded myelin showed linkage to chromosome 11p15, which is therefore classified as CMT4B2 [Ben Othmane *et al.* 1999].

LeGuern *et al* first identified a new locus on chromosome 5q23-33, in two Algerian families with ARCMT [LeGuern *et al.* 1996b]. This type of ARCMT is characterized by a precocious and rapidly progressive scoliosis in combination with a relatively mild neuropathy [Gabreels-Festen *et al.* 1999]. Subsequently, the locus has been refined to chromosome 5q31-q33 [Guilbot *et al.* 1999].

HMSNL, another form of ARCMT is described in Balkan Gypsies with an autosomal recessive demyelinating neuropathy and deafness and has been linked to chromosome 8q24. N-myc downstream-regulated gene 1 (*NDRG1*) has been found mutated from affected members in families with HMSNL [Kalaydjieva *et al.* 2000]. NDRG1 protein localizes in the Schwann-cell cytoplasm without expression in axons. The expression of the NDRG1 in different stages of the cell cycle shows that NDRG1 is involved in growth arrest and cell differentiation during development, and also in the maintenance of the differentiated state in the adult. Thus, it suggests that NDRG1 possibly works as a signaling protein shuttling between the cytoplasm and the nucleus [Kalaydjieva *et al.* 2000]. However, the pathogenesis of HMSNL is still not clear.

HMSNR, another form of ARCMT described in Balkan gypsies, is linked to chromosome 10q23 and EGR2 which is located centromerically and causes autosomal dominant CMT, DSD and CHN, has been excluded as the gene causing HMSNR [Rogers *et al.* 2000]. CCFDN, an AR neuropathy with congenital cataracts and facial dysmorphism also described in the Balkan gypsies, has been linked to chromosome 18qter [Angelicheva *et al.* 1999].

A large inbred Lebanese family suffering from AR demyelinating CMT has been linked to chromosome 19q13.1-13.3 and a mutation in *Periaxin,* which is located in this area has been shown to be responsible for this type of ARCMT, classified as CMT4F [Delague *et al.* 2000]. Periaxin is a protein expressed by myelinating Schwann cells of the developing mammalian peripheral nervous system. As myelin sheaths mature, periaxin becomes predominantly localized at the

abaxonal membrane, which suggests that it might participate in membrane-protein interactions that are required to stabilize the mature sheath. The homozygous periaxin knockout mice develop extensive PNS demyelination, which is identical to the morphology of the myelinated fibers in affected patients.

There are five types of ARCMT2. A Tunisian family with axonal degeneration and normal motor conduction velocities has been described. This has been defined as CMT4C1 but no locus been found. Another family with distal muscle wasting, weakness and pyramidal signs was described. The nerve biopsy showed a decrease of large myelinated fibers, increase in connective tissue, no onion bulb formation and rare axonal regeneration. Genetic analysis demonstrated that this disease linked to 8q21.3, which is close to the locus of CMT4A. This disease is designed as CMT4C2.

A Moroccan family with an autosomal recessive axonal type of neuropathy has been linked to chromosome 1q21.2-q21.3, flanking an interval of 1.7 cM [Bouhouche *et al.* 1999]. Two Algerian families also show linkage to the same locus. Motor deficits are distal in the upper limbs and both proximal and distal in the lower limbs. The mean MNCVs measured at the median nerve were 59 ± 10 m/s and the SAPs were undetectable at the sural, median and ulnar nerves. The nerve biopsy revealed a significant loss of large myelinated fibers. The *LMNA* gene encodes the nuclear-envelop protein lamin A/C. Sequencing analysis shows a "C" to "T" transition resulting in R298C change in Algerian families [Sandre-Giovannoli *et al.* 2002]. The knockout null mice (*LMNA* $^{-/-}$) show a significant reduction of axon density, an increase in axon diameter, and the

presence of non-myelinated axons, which is histologically similar to human patients with this disease [Sandre-Giovannoli *et al.* 2002].

Another big consanguineous family from Costa Rica is described to link to chromosome 19q13.3. All patients in this family presented with symmetrical weakness of plantar extensor and flexor muscles of the ankles. The MNCV of the median nerve is normal or slightly reduced (range 28.8 – 54.5 m/s; mean, 43.9 m/s). Sural nerve SAPs are not detectable in all patients and the CMAPs are reduced in the affected nerves [Leal *et al.* 2001]. EMG shows an increased number of polyphasic motor unit potentials and increased amplitudes of the motor-unit potentials.

The fifth responsible gene for the axonal ARCMT is *GDAP1,* which mutation can also cause demyelinating ARCMT (designed as CMT4A). Three ARCMT families with Spanish ancestry had slightly reduced MNCVs, reduced CMAPs and axonal degeneration on sural nerve biopsy. Three missense mutations and one frameshift mutation of *GDAP1* were identified and demonstrated segregation with the phenotype [Cuesta *et al.* 2002].

1.2.4.7 X-linked CMT2

In addition to recessive CMT2, there is an X-linked form of CMT2. A Pennsylvanian family of Italian ancestry is described with an unusual, slowly progressive form of X-linked recessive axonal neuropathy. The disease is

characterized by severe distal weakness, sensory loss, and areflexia at birth or within the first decade in male patients [Cowchock *et al.* 1985]. EMG studies reveal denervation, fasciculations, and large motor unit potentials. The median MNCVs range from 33.3 to 56 m/s in affected males and were normal in two carriers. The sural nerve biopsy shows axonal loss and sprouting with no segmental demyelination or onion bulb formation. Five of seven affected patients developed deafness and three of these five patients have mental retardation. There is no male-to-male transmission in the family and only the males are affected. Obligate heterozygous females are aysmptomatic. The disease has been found to be linked to Xq24-q26 [Priest *et al.* 1995].

1.2.5 Conclusion

CMT disease is very variable clinically and genetically. The clinical classification is based on electrophysiological and pathological criteria and the mode of inheritance. It provides useful clues for choosing which gene to screen. Molecular diagnosis for CMT has made rapid progress in recent years and accurate diagnosis using techniques, including southern blot hybridization, FISH, SSCP, high performance HPLC and sequencing can be done routinely in many laboratories. Nevertheless, careful characterization of patients, particularly the inheritance pattern (AD, AR or X-linked) and the neurophysiological type (demyelinating or axonal) remains the cornerstone of diagnosis particularly because screening for all responsible genes is laborious and expensive.

1.3 Rat genetic study

1.3.1 History

The rat (*Rattus norvegicus*) is one of the two most commonly used experimental animals in biomedical research. Further, it was the first mammal to be domesticated for scientific research 150 years ago [Lindsey 1979]. Crampe performed the first genetic study on rat coat color between 1877 and 1885 [Lindsey 1979]. Many genetically defined strains have been developed and maintained since then. However, because of its smaller size the mouse became the model of choice for mammalian geneticists, whereas the rat remained the model for physiologists and other biomedical researchers. Many individual laboratories developed specific rat strains for their studies. For example, between 1960 and 1970, rat genetics was dominated by research groups studying immunology. Until the 1980s, the two main drawbacks of the rat as a genetic model were the relatively poorly developed rat gene map and the paucity of useful polymorphic genetic markers. These difficulties were overcome between the mid-1980s and early 1990s. The most important finding in rat genetic mapping, was the discovery and development of microsatellite markers (SSLP, simple sequence repeat length polymorphism). Fluorescence in situ hybridization (FISH) developed recently also provided a new method for regional chromosome mapping. Nowadays, there are more than 5,000 markers in the rat genome [Watanabe *et al.* 1999]. These tools allow the mapping of single genes as well as quantitative trait loci (QTL) that cause or contribute to a disease phenotype. Rat disease models are usually developed either from phenotypic

selection and generating inbred strains or from isolating spontaneous mutants that model human diseases. Although these models do not always accurately represent the clinical human disease owing to species-specific differences, they are extremely useful, especially when little is known about the basis of a human disease. Not only are these animal models useful in identifying the disease-causing gene, they can also provide insight into disease pathogenesis.

1.3.2 Mutilated foot rat

In 1976, among the colony of Sprague-Dawley (SD) rats at the Clinical Research Centre, Harrow, Middlesex, UK, a rat was observed with an ataxic gait and progressive ulceration of the feet with distal sensory disturbance, eventually leading to toe and foot loss. The affected animal was described as the mutilated foot (*mf*) rat [Jacobs *et al.* 1981]. Genetic studies involving out-crossing of affected individuals to normal rats and intercrossing their offspring help to establish that the defect was inherited as an autosomal recessive mutation. The out-crossed offspring did not display the disease whereas only 25% of the intercross offspring did.

The onset of the disease is variable. At birth the clinical abnormalities seen in adults are not always evident. However, affected animals move less actively than the other littermates and occasionally, they may be smaller [Chimelli and Scaravilli 1986b]. Most frequently, the first signs appear at about two weeks of age. The hind limbs present first with an ataxic gait and are more severely

affected than the fore limbs. Affected animals have a wide based gait and stooping posture with legs pointing stiffly forwards or extended laterally (Fig. 1.1). As the disease progresses, the fore limbs are also affected and in the most severe cases, the animal moves with a seal-like motion. Although ataxia is usually the first sign observed, some affected animals develop mutilating lesions before limb ataxia. Occasionally ulceration has been seen on the feet in 1 or 2 day old animals. The foot ulcer usually starts with inflammation of the toes followed by the development of a bleeding ulcer; eventually the toes are lost leaving a swollen stump (Fig. 1.1). The rat is indifferent to pain, pinprick and temperature stimuli. The loss of sensation, including all sensory modalities, is obvious in affected rats, especially on the hind limbs.

1.3.3. Pathological findings of the *mf* rat

A peripheral sensory neuropathy is one of the major characteristics of the *mf* phenotype. Nerve pathology shows decreased numbers of axons in sciatic nerves and lumbar dorsal roots without any inflammatory changes or active processes of demyelination/ remyelination (Fig. 1.2) [Jacobs *et al.* 1981]. The involvement of small sensory fibers is more severe than the large fibers. The diameter of the myelinated fibers in sural nerves is bigger in the *mf* rat than in the normal SD rat. Loss of motor axons in the ventral root is also noted, but the severity is less marked than the loss of sensory axons in the dorsal root. The size of the dorsal root ganglia (DRG) in cervical and lumbar regions of affected animals is much

Fig. 1.1 The mf rat

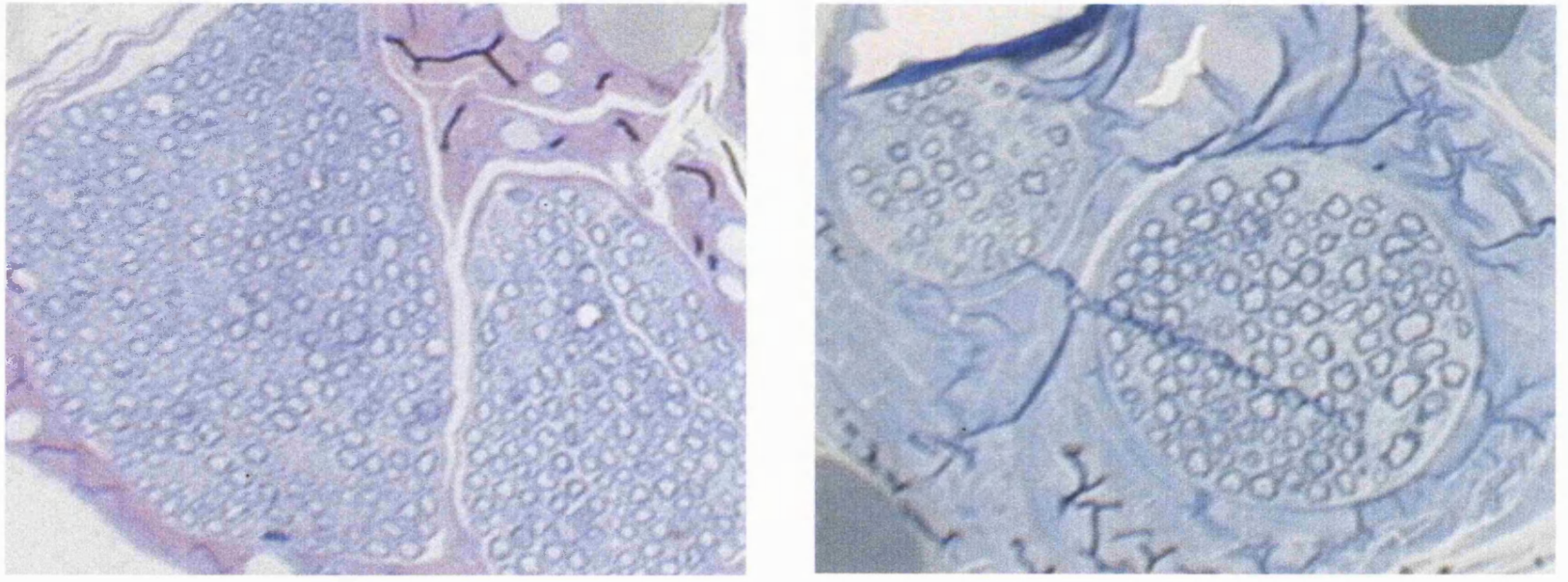


Fig. 1.2 The lumbar dorsal roots (L5) of *mf* (right) and normal SD (left) rats. The size and the number of axons in the L5 root of the *mf* rat is much smaller than the normal rat.

smaller than that of normal rats (Fig. 1.2). This phenomenon occurs in the fetal stage of development [Chimelli and Scaravilli 1986b].

The length of a normal pregnancy in rats is 21 days. In the normal development of DRG in rats, the loss of redundant neurons (neuronal apoptosis) can be detected at fetal day E15-E20 and under the electron microscope chromatolysis and aggregates of cell debris in cytoplasm can be seen in the neuron. Apoptosis is more conspicuous in the *mf* rats than in normal SD rats at E15-18 days [Chimelli and Scaravilli 1986b].

In the central nervous system, the size of the spinal cord is decreased in the *mf* rat, especially the fasciculus gracilis and cuneatus. In these fasciculi, more degenerating fibers are found in the *mf* rat than in the normal rat. There is significant degeneration in the white matter of the cerebellum and the spino-cerebellar tract in the spinal cord. In the granular cell layer of the cerebellum, there are many vacuoles and degenerating mossy fibers. Brains of affected animals show a variable degree of ventricular dilatation. The cortical ribbon is conserved with normal six-layered distribution of neurons without dysplasia. However, glial fibrillary acidic protein (GFAP) staining of the cortex shows discrete gliosis, predominantly in the fifth and sixth cortical layers and in the subjacent white matter. Moderate gliosis is also found in the hippocampus [Herberg *et al.* 1997].

The apoptosis of the DRG neurons that occurs in the embryonic stages of development is thought to be the crucial step causing the sensory neuropathy.

In vivo, nerve growth factor (NGF) can rescue the sensory neurons in the DRG from apoptosis [Geffen and Goldstein 1996]. There are two types of receptors for NGF, p75 nerve growth factor receptor (p75NTR) which is a low-affinity receptor, and neurotrophic tyrosine kinase receptor, type 1 (TRKA or NTRK1) which is a high-affinity receptor for NGF [Klein *et al.* 1991]. Mice carrying a disrupted *Ntrk 1* gene develop a severe sensory and autonomic neuropathy [Smeyne *et al.* 1994]. Similarly, mice with a targeted mutation of the *p75NTR* (*Ngfr*) in embryonic stem cells, develop a sensory neuropathy with mutilated feet [Lee *et al.* 1992]. These features are very similar to the *mf* rat. Therefore, *p75NTR* (*Ngfr*) and *Ntrk 1* are both good candidates as the disease-causing gene in the *mf* rat.

1.3.4. Animal model for HSAN

Animal models are useful in the study of human diseases. There are four categories of animal models [Gill 1980] [Davidson *et al.* 1987]:

1. **Induced or experimental models** that try to reproduce the diseases being studied.
2. **Spontaneous or natural models** that occur spontaneously but are recognized as similar to the disease of interest.
3. **Negative models** that are the normal counterparts of a diseased animal.
4. **Orphan models** these are animal diseases without counterparts in human or other animals.

The *mf* rat is a spontaneous mutant. The clinical and pathological features of the *mf* rat are very similar to the human disease, the hereditary sensory and autonomic neuropathies (HSANs), especially HSAN I and HSAN II (see section 1.2.1.1. and 1.2.1.2.). The *mf* mutant characteristics include autosomal recessive transmission, a non-progressive course, a congenital sensory neuropathy and conspicuous apoptosis of the DRG neurons. These features are most similar to patients with non-progressive congenital hereditary sensory neuropathy type II (HSAN II). There are some patients described as having HSAN I, without mutations in the *SPTLC1* gene who are also similar to the *mf* rat. The mild motor features, especially those described pathologically in *mf* rats also raise the possibility that this could also be an animal model for rare forms of CMT (especially CMT 2B (see section 1.2.3.4 & 1.2.4.5)). Therefore, the *mf* rat is a useful animal model to study HSAN and possibly CMT. This study aimed to identify the disease-causing gene for the *mf* rat. The identification of the causative gene for *mf* disease and the study of the pathogenesis of this disease may shed light on the pathogenesis of some forms of HSAN, especially if the causative gene for the *mf* disease in the rat is also found to cause a type of HSAN.

1.3.5. Genetic analysis of the *mf* rat

1.3.5.1 Strategies for disease gene identification

Several strategies have been devised to identify the gene responsible for a disease phenotype. These approaches include the following:

1. **Functional cloning**: this strategy is based on some information about the function of the gene, which is exploited to isolate a gene clone. For example, tyrosinase (the enzyme that is defective in certain types of albinism) was cloned by raising an antibody to the enzyme and using this to screen a bacterial expression library for colonies expressing the gene [Kwon *et al.* 1987]. The decreased activity of Hypoxanthine guanine phosphoribosyl transferase (*Hprt*) was found in patients with gout and Lesch-Nyhan syndrome. The mutation of this gene in a patient with gout was firstly reported in 1983 [Wilson *et al.* 1983].
2. **Positional-independent candidate gene approaches**: this approach is based upon similarities in disease phenotypes between species in which the gene has been identified in one of the two species. The brown (*b*) locus in the mouse was cloned because the antibody raised against tyrosinase identified this gene as well [Dakour *et al.* 1993].
3. **Positional cloning**: identification of gene based upon positional information without fore knowledge of pathogenesis or the biomedical function. Examples here are the identification of the responsible genes for cystic fibrosis [Tsui *et al.* 1985] and Huntington's disease in humans [Gusella 1989]

and for diabetes (*db*) = leptin receptor gene, and obese (*ob*) = leptin gene in the mouse [Chua *et al.* 1996]. Genetic linkage analysis established the chromosomal positional of the genetic defect, providing a potential opportunity for isolating the disease gene and characterizing its product in the absence of any knowledge of its biochemical function.

4. **Positional candidate gene approaches**: knowing something about the genetic position of a mutant gene can frequently help identify potential candidate genes. Since the completeness of the sequencing of human and mouse genome, the candidates that are located at the linkage region are easily retrieved from genome databases. This approach avoids the pain-taking and time-consuming steps of chromosome walking [Parrish and Nelson 1993] [Collins 1995]. For example, the HSAN I locus has been mapped to chromosome 9q22.1-22.3. *SPTLC1* gene, which has been mapped to this region was a candidate and mutations were identified in patients with HSAN I [Dawkins *et al.* 2001]. The same approach has also been employed in identification of new genes from knowledge of other species, such as the identification of the responsible gene (*KIF1Bβ*) for CMT2A. The gene mapped to the linkage interval of CMT2A. The heterozygous knockout mice ($KIF1B\beta^{+/-}$) have a defect in transporting synaptic vesicle precursors and suffer from progressive muscle weakness similar to Charcot-Marie-Tooth disease. The screening of *KIF1Bβ* identified mutations in CMT2A patients [Zhao *et al.* 2001].
5. **Mutational analysis approaches**: using approaches based upon information concerning the nature of the mutational event. There are some examples of mutagenesis, which provide clues for gene identification: 1. Retroviral

insertional mutagenesis, e.g., Dilute (d) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome [Jenkins *et al.* 1981]. DBA/2J mice homozygous for the reverse mutation to wild type at the dilute locus (d+2J) lack ecotropic virus-specific sequences, suggesting that the dilute mutation was caused by virus integration. 2. Deletional mutagenesis, e.g., the *Ph* mutated mouse carries a deletion in the genomic sequence associated with *Pdgfra* gene, which was identified by analyzing the restriction fragment length polymorphisms in interspecific F1 hybrids [Stephenson *et al.* 1991]. 3. Loss of heterozygosity, usually derived either from disabled caretaker genes, which protect the integrity of DNA, or from chromosome segregator genes, which mediate faithful chromosome disjunction, is a hallmark for genomic instability in cancer. Because of the high level of specificity, loss of heterozygosity has recently become invaluable as a marker for diagnosis and prognosis of cancer. Therefore, loss of heterozygousity can provide clues for identification of the chromosome caretaker or segregator genes [Thiagalingam *et al.* 2002].

Neither the function nor the position of the *mf* gene is known. However, the resemblance between the *mf* rat and, *Ntrk 1* and *p75NTR* transgenic 'knock-out' mice suggest that these genes might worth be sequencing to determining whether they are mutated in the rat (positional-independent candidate gene approach). Whilst the candidate gene approach is a valid way to proceed, the rationale for identifying likely genes relies too heavily on a limited amount of information about the genes themselves and the disease process itself. Thus, a genetic

approach involving positional cloning techniques will also be undertaken to identify the *mf* gene.

The process of identifying the molecular basis of a mutant gene by virtue of its position within a complex genome is a time consuming process involving simple Mendelian genetics. As mentioned above, the positional cloning approach (section 1.3.5.1) has been used successfully to clone a number of very prominent human genetic disease including Duchenne muscular dystrophy [Koenig *et al.* 1987], cystic fibrosis [Tsui *et al.* 1985] as well as several mouse disease genes. In principle, the approach involves a combination of genome analysis techniques: cytogenetic analysis, linkage analysis, physical mapping, chromosome walking, and candidate gene identification. However, as a result of the human genome initiative, resources are available to substantially minimize the amount of effort required to achieve objects that in the past have taken ten or more years of intensive activity.

1.3.5.2 Resources

The purposes of the human genome project (and other model organisms) are to understand the genomic organization and its relation to function; assembly of gene map; identification of disease genes; provision of DNA segments as overlapping contig; and then, eventually to determine the complete sequence of the genome. As the project has been completed, it provides detailed genetic and genomic information. The genetic analysis of the mouse genome, which has

been totally sequenced, is far more advanced than that of other rodents. Since the genomic segment is conserved among species, especially those closely related species in evolution, the human and mouse genomes are frequently used as a paradigm for other animal genomes. Many techniques and resources developed for the human and the mouse genomes have been transposed to the mapping of other animals' genome.

1.3.5.2.1 Marker loci

Positional cloning strategies for the human genome utilized unique markers called sequence-tagged sites (STSs) combined with cytogenetic mapping, genetic-linkage mapping, radiation-hybridization mapping and chromosome walking to map one or more putative genes or positional candidates. The markers used for the linkage analysis in our study are microstellite markers. The mammalian genome contains a large amount of highly repeated DNA sequence families which are transcriptional inactive. Microsatellite DNA families include small arrays of tandem repeats, which are simple in sequence (often 1-4 base pairs) and are interspersed throughout the genome. The dinucleotide repeats, CA/TG and CT/AG are most common in the human genome, which count for 0.5% and 0.2% of the genome, respectively. These markers are also common in mouse and rat genome [Miller 1997]. These highly interspersed small arrays of DNA are highly polymorphic and are specific individually. The size (in base pairs) of the marker is used as a genotype and the typing of serial linked markers is designated as a haplotype, practically.

1.3.5.2.2 Bioinformatic resources

Bioinformatics is the study of the inherent structure of biological information and biological systems. It brings together the avalanche of systematic biological data (e.g. genomes) with the analytic theory and practical tools of mathematics and computer science. It deals with the storage, manipulation and analysis of biological information via computer science, especially the worldwide website server. The first bioinformatic database was created by Dayhoff in 1981, which gathered all the nucleic acid sequence data available [Dayhoff *et al.* 1981]. Since then on, databases with free availability for scientific research, are created for specific purposes, e.g. genome database for the query of sequences of genes, markers, genetic linkage and radiation-hybrid maps; protein database for sequence of amino acids, protein domains and three-dimensional structures. These resources provide a virtual important tool for identifying an unknown gene. Table 1.3 lists the uniform resource locators (URLs) of these databases, which we used in this thesis.

Table 1.3 Genetic databases used in the study of the genetic analysis of the *mf* rat

Genome database	URL
Rat	http://ratmap.ims.u-tokyo.ac.jp/
	http://www.rgd.mcw.edu/
	http://www.well.ox.ac.uk/
	http://waldo.wi.mit.edu/rat/
	http://ratmap.gen.gu.se/
	http://www.informatics.jax.org/
	http://mapmgr.roswellpark.org/
Mouse	http://www.jax.org/
	http://www.mgc.har.mrc.ac.uk/
	http://mouse.info.nih.gov/
	http://mouse.ensembl.org/
Human	http://www.hgmp.mrc.ac.uk/
	http://www.ncbi.nlm.nih.gov/
	http://genome.ucsc.edu/
	http://www.ebi.ac.uk/
	http://www.ddbj.nig.ac.jp/
	http://www.gene.ucl.ac.uk/
	http://www.tigr.org/
	http://www.sanger.ac.uk/
	http://www.expasy.ch/

1.3.5.2.3 Homologus/ Syntenic mapping

The rat (*Rattus norvegicus*) is a vital animal model for a large number of human diseases, and it is the most widely studied experimental animal in biomedical research and physiology [Gill, III *et al.* 1989] [Szpirer *et al.* 1996]. However, the

mouse is the most frequently used animal for genetic study and traditionally, the mouse genetic study is far advanced than that of other rodents. A private company, Celera (Celera Genomics Rockvilli, Rockvilli, MD, USA), has completed the sequencing of the whole mouse genome. In addition to the study of mouse genome, the human genome project with international effort has finished the human genome sequencing as well [Lander *et al.* 2001]. The working draft sequence for individual chromosomes with 99% accuracy is available in the human genome database. The easily accessible and non-commercial genetic database from human and mouse genome nourished the biomedical studies worldwide.

The comparative maps among different species are important in the genetic study of a species with relatively paucity of genomic information. Because of the highly conserved coding region of genes among species, the comparative mapping is usually based on the accumulated information of gene homology among different species. These data have shown that chromosomal segments of gene homology have been conserved in evolution ("evolutionarily conserved segments") [Gomez-Fabre *et al.* 2002]. Therefore, the potential for successful transferring of gene position data between rat and mouse (diverged 30-40 Myr ago) and even rat and human (diverged ~110 Myr ago) should be very high [Arnason *et al.* 2000]. According to the integrating radiation hybrid, FISH (fluorescent in situ hybridization), and zoo-FISH mapping data, Nilsson proposed a rat genetic map against mouse and human homology and estimated there was 83~100% of conserved regions between rat and mouse and 66~82% of the conserved regions between rat and human [Nilsson *et al.* 2001]. The rat-

mouse comparative map almost covers the complete genome. These comparative maps (rat-mouse and rat-human) can be used as a tool to predict the gene location and to bridge between the genomes of different species. These comparative maps or species homology data are also available in the human, mouse or rat genetic databases.

1.3.5.2.4 Clone resources

To identify the disease gene, moving from the linkage map to the physical map is an inevitable step. The linkage maps are based upon prediction from family pedigrees and crosses. The recombination frequency in the linkage map is assumed to be constant in all regions, which is not the case. The centi-Morgans (cM) are not accurate units. However, the physical maps are expressed in bases, which is constant and invariant. A marker may map 5 or 10 cM distant from another marker depending on various predictions and recombinations, but the marker is a DNA sequence that has an exact location in the physical map. How can we place markers from a linkage map into the physical map? There are two methods: 1. Fluorescent in situ hybridization (FISH): FISH can localize a marker to a chromosome band, but these bands are 4-8 mega-bases (Mb), which is the only rough location to the marker. 2. To create a high-resolution physical map: there are three steps. Firstly, cut the genomic DNA into very large overlapping pieces. This is usually been done by enzyme partial restriction and the DNA generated with this manner is large (1Mb) and overlapping. Secondly, clone the DNA fragments into yeast- or bacterial-artificial chromosome (YAC or BAC) or cosmid vectors [Burke *et al.* 1987] [Shizuya *et al.* 1992] [Collins and Hohn

1978]. These large insert genomic DNA libraries have served a pivotal role for the isolation and characterization of important genomic regions and genes. Finally, reconstruct the order of the cloned fragments in the genome, that is to use the restriction analysis or DNA hybridization of linkage markers and chromosome walking to construct the physical map with overlapping contigs. These clones in the physical map contain the sequence of the unknown gene to explore.

In this study, we used the positional candidate approach in the following steps, the genetic linkage study, refining the genetic map, physical mapping, identifying the mouse syntenic region in the comparative map, candidate gene screening and finally the identification of the mutation.

Chapter 2

Materials and Methods

This chapter is divided into two sections. Section 1 contains the materials and methods used to identify the molecular basis of the mutilated foot (*mf*) rat. The approach used is that for an unknown gene utilizing a positional dependent candidate gene approach. Section 2 describes the methods used to identify mutations in patients with Charcot-Marie-Tooth disease (CMT). This involved the sequencing of two genes (*MPZ* and *GJB1*) known to be involved in this disease. Most of the molecular genetic methods follow the protocols described in "A Laboratory Manual", by J. Sambrook, E. Fritsch and T. Maniatis, 2nd Edition, from the Cold Spring Harbour Press. Buffers and solutions are listed in the appendix I.

2.1 The rat project

2.1.1 Materials

2.1.1.1 Breeding protocol

The project and personal license for the breeding programme were approved by the Home Office. Animals were kept in the animal house at the Institute of

Neurology, Queen Square, London. All the procedures performed complied with the ANIMALS (Scientific Procedures) ACT 1986. Rats were housed under clean pathogen-free conditions. They were maintained at a constant temperature (20-26°C), humidity (55-70%) and had a 12 hours light-dark cycle. Food and water were available *ad Librium*.

The Lewis rat strain (LE) was chosen to mate with the *mf* rat because it was shown to be polymorphic at 48/51 micro-satellite repeat loci dispersed throughout the rat genome. Two male LE rats were purchased from B&K Universal Ltd, East Yorkshire, UK and imported into the Institute animal house. At 8-10 weeks of age, these males were mated with two *mf* female rats with a Sprague-Dawley (SD) background of the same age. The F1's from this cross were backcrossed with *mf* homozygotes. The resulting N2 rats were examined (see below, section 2.1.1.2) and sacrificed at seven days of age. A total 536 N2 rats were produced over a period of two years (254 in the first year and 282 in the second) in accordance with the conditions imposed by the Home Office on awarding the animal licence.

2.1.1.2 Phenotype characterization and tissue collection

Two neuropathologists examined the offspring at postnatal day 7 to ascertain whether they were affected. The examination consisted of clinical observation and physical sensation tests proposed by L. Chimelli and F. Scaravilli [Chimelli and Scaravilli 1986b]. When these animals were lifted by their tails, the affected

animals could not hold their hind limbs extended and the limbs moved in an abnormal fashion. Reaction to pinprick and pressure sensation on fore and hind limbs, tail and trunk was also examined. Pressure on the hind limbs in the affected N2 rats provoked no strong or prompt withdrawal reaction unlike that observed in normal animals. Mutilation of the feet was occasionally seen at birth.

After clinical examination, the rats were killed by a Schedule 1 method (overdose of anaesthetics) in the procedure room of the animal house. The soft tissues were collected and promptly frozen in liquid nitrogen. They were then kept at -70°C prior to extraction of either RNA or DNA. The corpses were kept in 95% ethanol in case of potential ambiguities in the genotype/phenotype data.

2.1.2 Methods

2.1.2.1 Genetic linkage analysis

2.1.2.1.1 Introduction

Use of inbred strains of laboratory animals ensures homozygosity for any polymorphisms that may exist within the population. Further, phase of linkage is also guaranteed. Neither of these conditions is met in random bred populations. The *mf* rat, which arose in 1978 on the SD strain background, has been maintained since then as an inbred strain. As an autosomal recessive mutation,

F1s produced by out-crossing *mf* homozygotes to another rat strain will not exhibit the phenotype. However, when these F1s are backcrossed with *mf* homozygotes, 50% of the N2 rats will be phenotypically normal and the other 50% will suffer from *mf* disease. To enhance the likely of producing backcross offspring that would be informative in a mapping study, the *mf* SD rat strain was genotyped using a group of primer pairs that specifically amplify well-defined micro-satellite repeat loci. The size of the PCR product for each locus was determined and compared with those reported for other strains of rat. The PCR primer pairs from the RatMap® panel were purchased from Research Genetics (Invitrogen Ltd., Paisley, UK). The LE strain was found to be polymorphic at 48/51 (94%) of the micro-satellite repeat loci typed in the *mf*-SD strain. Allele sizes differ by as much as 2 nt. to as many as 68 nt. (results in chapter 3). Thus, the LE strain was chosen to mate with the *mf* rat. The F1 rats were then backcrossed with *mf* rats to produce the N2 rats, which were used for linkage analysis. The methods for DNA extraction, polymerase chain reaction (PCR) and genotyping are presented in section 2.1.2.2, 2.1.2.3 and 2.1.2.4, respectively.

2.1.2.1.2 Genome-wide screening and linkage study

A "genome scan" was performed on a subset of the 536 N2 offspring in an attempt to assign linkage of *mf* disease to a particular chromosome within the rat genome. The genome scan means that by using a series of RatMap® pairs, the whole genome was genotyped at a density of one marker locus per 20 centiMorgans (cM). Because *mf* disease is autosomal recessive, the clinically

affected N2 rats should be homozygous at the *mf* locus. The normal N2 rats should be heterozygous at this locus (Fig. 2.1 & 2.2). If a marker is tightly linked to the disease gene, the alleles from the two chromosomes (one from F1, the other from *mf*) in the affected N2 rats should be homozygous. The same marker should be heterozygous in the unaffected N2 rats. Therefore, a heterozygous marker in the affected N2 rats indicates that a recombination has occurred between the disease gene and the marker. The recombination frequency is proportionate to the distance between the disease gene and the marker. According to the Kosambi function, the higher the recombination frequency, the lower the probability of linkage between two loci [Ott 1991]. Therefore, the calculation of the recombination fraction will give the genetic distance between the marker and the disease locus. The LOD score is a measurement of the likelihood of genetic linkage between two loci (LOD score: The log of the odds that the loci are linked rather than unlinked). All genotyping data were imported into the MapManager programme (MapManager QTXb06, http://mcbio.med.buffalo.edu/ mapmgr.html) which is designed to analyze linkage data and calculate the LOD score and genetic distances between the linked loci.

Firstly, DNA from thirty-four affected N2 rats along with their grandparents (2 LE males and 3 *mf* females) were genotyped using primers from the RatMap® panel, Research Genetics (Invitrogen Ltd., Paisley, UK). We chose markers approximately 20 cM apart on each chromosome. The recombination frequency was calculated depending on the number of heterozygous rats obtained. After finding a marker with no recombinants in the 34 affected N2 rats, 34 unaffected

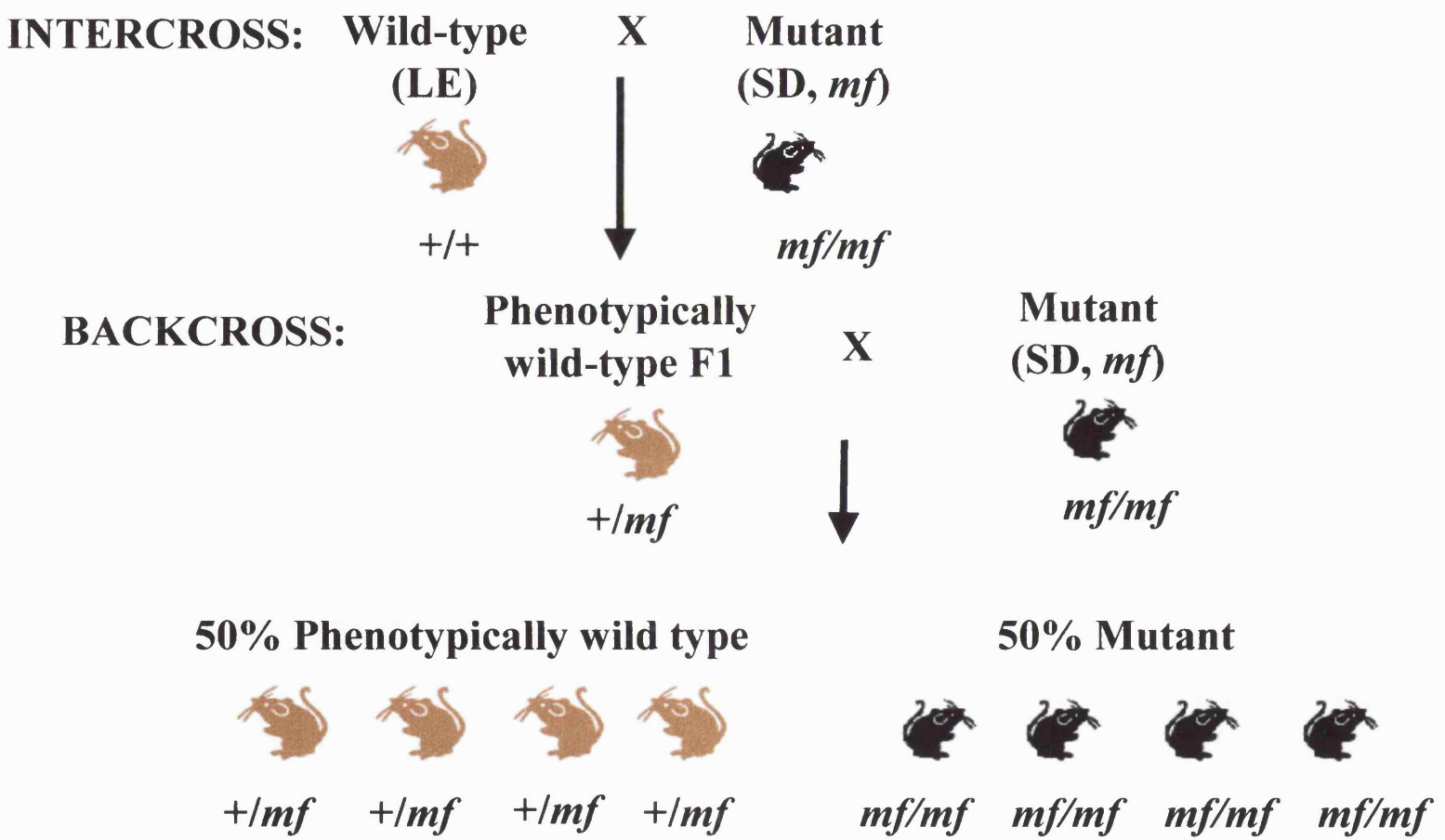


Fig. 2.1. Segregation of *mf* disease using a backcross breeding programme.

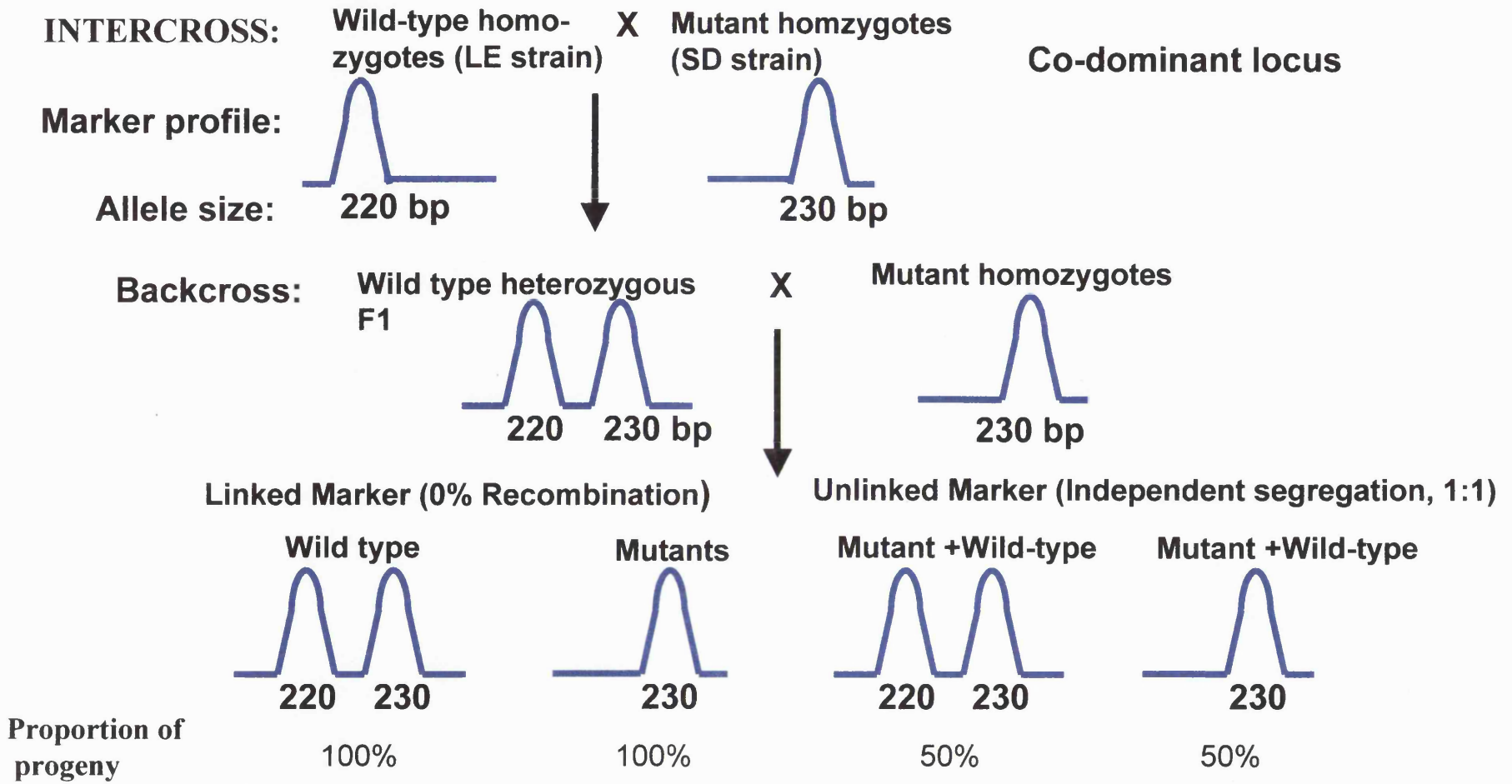


Fig. 2.2. Example of microsatellite marker profile polymorphic loci segregating in the backcross programme. bp = base pairs.

N2 rats were analysed followed by analysis of all 254 N2 rats. The 282 N2 rats from the second year were also genotyped to refine the area of linkage.

2.1.2.1.3 Selecting the BAC clones

Markers that were either non-recombinant with or lay within a cM of the mutant locus (based upon the analysis of the larger sample set) were used to screen a genomic library. In 1998, Woon *et al.* constructed a rat genomic library by P1 bacterial artificial chromosomes (BACs) with 10-fold genomic coverage [Woon *et al.* 1998]. DNA pools of BACs were obtained from the German Genome Resource Centre (RZPD) (http://www.rzpd.de/). The pooled DNA was screened to select the BAC clones that contained the relevant marker sequences. The primer pairs for *D14Got72, D14Got73, D14Got76* and *D14Rat32* were used to screen the primary pool DNA by PCR. The plate, column and row number on the secondary pool defined the positive clones. These clones were also ordered from RZPD. The host of these BACs is E-coli, which was grown on Agar plates under chloramphenicol drug selection. Extraction of the BAC DNA was performed using a Qiagen large-construct DNA extraction kit. The DNA was then end-sequenced. The method for the BAC DNA extraction and the sequencing reaction are presented in section 2.1.2.2.4 and section 2.1.2.3.6, respectively.

2.1.2.1.4 Anchoring the *mf* locus on the mouse genome

Sequence drafts of the mouse genome were accessible commercially (Celera Genomics Rockvilli, Rockvilli, MD, USA). The company also provided BLAST software to search for sequence homology from individual mouse chromosomes. In general, conservation of linkage and similarities in gene sequence are greater in more closely related species than in those that are evolutionary more divergent. This conservation can be used to extrapolate genetic information between species. The unique sequences associated with the closely linked microsatellite repeat markers were used to BLAST the human and mouse sequence databases available through NCBI (http://www.ncbi.nlm.nih.gov/), Ensemble (http://www.ensembl.org/) and the commercial site provided by Celera. The search of the latter database was performed in collaboration with Dr. D.A. Stephenson at the MacLaughlin Research Institute. Similar searches were also performed with the BAC end sequence. Figure 2.3 is the search result obtained from the Celera database using approximately 100-150 nucleotides (nt.) of unique sequences associated with *D14Got73* (351 nts). Of the 88 nts (nt 209~296) BLASTed in this instance, a contiguous stretch of 78 nt displayed sequence similarities found in a segment of DNA (Celera scaffold identification number: GA_x5J8B7W498Y, nucleotide (nt) 1000001..1500000) on mouse chromosome 11. The probability (Expected) of this event occurring by chance was calculated to be $1x10^{-18}$, an extremely rare event suggesting significant sequence similarities. Another two segments with length 32 and 40 nts (nt 12-43

and 92~130, respectively) had shown similarity with probability by chance 4x 10^{-6} and $10^{0.26}$, respectively.

D14Got73

```
1.  GA_x5J8B7W498Y:1000001..1500000 /organism=Mus musculus
/order=3/ga_uid=196000025290015 /len=3916428
          Length = 500000

 Score = 97.6 bits (49), Expect = 1e-18
 Identities = 78/88 (88%)
 Strand = Plus / Plus

Query: 209..268
tgcaagcactggagagctggaaagccaacagtgcaattcagaaccagccaacgtcctgan
|||||||||||||||||||||||  | || || ||||||||||||| || | |||||
tgcaagcactggagagctggaaagttagcaatgaaattcagaaccagcaaatggcctgag
Sbjct: 466473..466532

Query: 269    aagtactgaggctaatggtgatactctc 296
               |||||||||||||||| ||||||||||
Sbjct: 466533 aagtactgaggctaatgatgatactctc 466560


1.  GA_x5J8B7W498Y:1000001..1500000 /organism=Mus musculus
/order=3
              /ga_uid=196000025290015 /len=3916428
          Length = 500000

 Score = 56.0 bits (28), Expect = 4e-06
 Identities = 31/32 (96%)
 Strand = Plus / Plus
Query: 12     taatggaaataggagagctggctttaaatctg 43
               |||||||||||||| ||||||||||||||||
Sbjct: 466262 taatggaaataggaaagctggctttaaatctg 466293

 Score = 40.1 bits (20), Expect = 0.26
 Identities = 36/40 (90%), Gaps = 1/40 (2%)
 Strand = Plus / Plus

Query: 92     gccaaagtatatgtagccctaaagt-aatacctagcatgt 130
               |||||||||| |||||||||| | |||||||| |||||
Sbjct: 466362 gccaaagtatacgtagccctaaattcaatacctaccatgt 466401
```

Fig. 2.3 Report from the BLAST software, which confirmed the sequence homology between the sequence of D14Got73 and the nucleotide 10000001 to 1500000 in scaffold GA_x5J8B7W498Y of mouse chromosome 11. The upper "Query" sequence was from *D14Got73* and the bottom "Sbjct" sequence was the subject scaffold sequence.

Using the same strategy, the end-sequences of the BAC clones containing the micro-satellite repeat loci *D14Rat72* were also BLASTed against the Celera database. In the example presented in Fig. 2.4, the sequence of Sp6 end (Sp6 is a primer of the BAC vector) of clone RPCIB657O13359 was found to be homologous to the sequence in between nt 2000001and nt 2500000 of scaffold GA_x5J8B7W498Y of mouse chromosome 11. Two segments of this end sequences (nt. 106~218 and 272~435) have nucleotide identity 93% (Expect, $2x10^{-39}$) and 81% (Expect, $4x10^{-12}$), respectively.

RPCIB657O13359-Sp6

Sequence Report: GA_x5J8B7W498Y:2000001..2500000
sequence location - chr.11 17836454 - 21752882

```
1.  GA_x5J8B7W498Y:2000001..2500000 /organism=Mus musculus
/order=5/ga_uid=196000025290015 /len=3916428
          Length = 500000

 Score =  168 bits (85), Expect = 2e-39
 Identities = 106/113 (93%)
 Strand = Plus / Plus

Query: 106..165
ttgttgttttaaaacaaaatctgttcttccatatgtgtttcagaaactaatagcaggacc
|||||||||||||||||||||||||||| ||  ||||||||||||||||||||||||||
ttgttgttttaaaacaaaatctgttctttcaggtgtgtttcagaaactaatagcaggacc
Sbjct: 138997..139056


Query: 166..218
agcctctgaggatagttggcttttatgagctgccaccaggaagcaggctagga
|| |||||||||||| ||| ||||||| |||||||||||||||||||||||||
agtctctgaggatagtgggcctttatgaactgccaccaggaagcaggctagga
Sbjct: 139057..139109

 Score = 77.8 bits (39), Expect = 4e-12
 Identities = 135/166 (81%), Gaps = 2/166 (1%)
 Strand = Plus / Plus


Query: 272..331
cctggagctcactgtgtagaccagaacagccttaaacccgcagagggttccctcttccta
|||||| |||||||||||||||||   |||||||| | ||||||| | | | ||  |
cctggaactcactgtgtagaccagaatgaccttaaacacacagagggctacttgtttttg
Sbjct: 139172..139231


Query: 332..389
cctcctaag--ctggggttaaaggactgcaccgtcatgcccagctgacgagagactccat
|||||  |||  |||||||||||||| | |||| |||||  ||||||| |||| ||| |||
cctcccaagtgctggggttaaaggagtacaccatcatgtacagctgatgagaaacttcat
Sbjct: 139232..139291


Query: 390    atgtgagtcatgcagttcatttcgggtcttctatagatgttcttca 435
              |||||||||||||||  |||| ||| || ||||| ||||| ||||
Sbjct: 139292 atgtgagtcatgcagtatattttggggctcctatatatgtttttca
139337
```

Fig. 2.4 Sequence homology between the BAC end-sequence, which contains *D14Got72* and the scaffold GA_x5J8B7W498Y:1000001..1500000 of mouse chromosome 11. The upper "Query" sequence was from the BAC DNA on the Sp6 end and the bottom "Sbjct" sequence was the subject scaffold sequence.

2.1.2.1.5 Candidate gene screening

The BLAST searches identified mouse and human syntenic regions that were likely to contain the homologue of the rat gene for *mf*. Genes located close to the linked region were chosen as candidates for screening. The cDNA sequences were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db =Nucleotide). The primer sequences were designed using Generunner software (Generunner version 3.05, http://www.generunner.com/). Degenerate primers from mouse sequences were designed to screen the coding regions of candidate genes when rat cDNA sequences were not available. The primer sequences used to screen *Otx1*, *Kiaa0127*, *Bcl11a*, and *Cct4* are shown in appendix II.

2.1.2.2 Nucleic acid extraction

2.1.2.2.1 Extraction of genomic DNA from rat soft tissue

The rat soft tissue was chopped up and immersed in 500µl of nuclei lysis buffer (appendix I). The proteins were digested by adding 100µL of proteinase K (13 unit) to this suspension and incubating it overnight at 55°C. The DNA was then purified using organic extraction techniques. Basically, an equal volume of phenol was added to the aqueous solution and the tube inverted several times to obtain an emulsion. The aqueous phase was separated form the organic phase by centrifugation and transferred to a fresh tube. This step was repeated several times until the inter-phase cleared; however, a mixture of phenol/chloroform (in

a 1:1 ratio) was substituted for pure phenol. The phenol was removed using a pure chloroform step. Following this, the DNA was precipitated with pure ethanol (containing 0.5% 3N Sodium Acetate). The precipitate was pelleted by centrifugation. After washing with 70% ethanol, the sample was centrifuged again and then, dried and resuspended in 1x TE (Appendix I).

2.1.2.2.2 Quantification of DNA

The concentration of DNA following the organic extraction procedure was estimated by its optical density (O.D.) at a wave length 260 nm and 280 nm. These readings were performed on GeneQuant *pro* spectrophotometer (Amersham Pharmacia Biotech Ltd., Buckinghamshire, UK). 10µl of the DNA solution was diluted to 1000µl with 1x TE and placed in a quartz cuvette. The O.D. was measured against the blank sample containing 1x TE. 50 µg/ml of double stranded DNA gives an O.D. reading of 1.0 at 260nm. Thus, the sample DNA concentration was calculated according to the following equation:

$$OD260\ nm \times 50 \times 100^{*} = DNA\ concentration\ (\mu g/ml)$$

* the dilution factor

The 260/280nm OD ratio was used to assess the purity of the DNA sample. A ratio greater than 1.8 is indicative of fairly pure DNA. Thus, samples giving a ratio of 1.5 or less were re-extracted with phenol/chloroform because of probable protein contamination.

2.1.2.2.3 Extraction of total RNA from fresh rat tissue

All the solutions for RNA work were treated with 0.1% v/v diethylpyrocarbonate (DEPC, Sigma-Aldrich Company Ltd. Dorset, England) to inactivate ribonucleases. All the plastic ware was sterile and individually wrapped. A small piece of soft tissue was chopped and put in an eppendorf tube containing 200μl of Trizol (Invitrogen Ltd, Paisley, UK). An automatic sterile pestle was used to grind the tissue in the tube until all the tissue was homogenised. 200μl of chloroform and 800μl of Trizol were added to the homogenised mixture and the tubes were vortexed for 1 minute. The sample was spun at 12,000g for 15 minutes in a JA18-1 rotor in a Beckman J2-21 centrifuge (JA18-1 rotor, 10,000rpm). The aqueous phase was pipetted into microcentrifuge tubes and 500μl of isopropanol was added to precipitate RNA. The RNA was pelleted at centrifugation 12,000g for 10 minutes and the supernatant was discarded. The pellet was washed with 75% ethanol and spun at 7,500g for 5 minutes (7,500rpm). Before being dissolved in 100μl of DEPC treated water, the ethanol was discarded and the pellet air-dried for 5 minutes. The total RNA was electrophoresized on a 1% agarose gel to check the quality and the O.D. in 260 nm was used to evaluate the concentration of RNA.

2.1.2.2.4 Bacterial culture and DNA extraction from BAC clones

Rat bacterial artificial chromosome (BACs) resources were purchased from a Germany genetic database (RZPD, Berlin-Charlottenburg, Germany). Individual

clones were plated out on small Lactose broth (LB) agar plates, containing 20 μg/ml chloramphenicol selection. The vector in the E-coli is pBACe3.6, which has a capacity for larger inserts. One colony from each clone was picked and inoculated into 500ml of LB broth. After incubation in 37°C for 20~22 hours, the bacterial culture reached the exponential growth phase and was subjected to extraction of the BAC DNA.

The Qiagen Large-construct DNA extraction kit (Qiagen Ltd, Crawley, West Sussex, UK) was used to isolate the BAC DNA following the manufacturers alkali/SDS protocol. Bacterial cells were harvested by centrifugation at 6000*g* for 15 minutes in a Beckman J2-21 centrifuge (JA-14 rotor, 10,000rpm) at 4°C. The supernatant was discarded and the cells resuspended in 20ml of buffer P1. A further 20ml of buffer P2 was added to the suspension. Following a five-minute incubation at room temperature 20ml of buffer P3 was added and mixed gently. The mixture was then incubated for a further 10 minutes on ice before 20,000g centrifugation for 30 minutes at 4°C. The supernatant, containing the BAC DNA, was removed and filtered through a flat filter paper pre-wetted with distilled water. The genomic DNA in the clear lysate was precipitated by adding 0.6 volumes (≅ 36ml) isopropanol and was pelleted centrifugation at ≥ 15,000*g* for 30 minutes. The supernatant was decanted carefully. Five ml of 70% ethanol was used to wash the DNA pellet and the sample was centrifuged at 15,000*g* for 15 minutes. The DNA pellet was placed upside down to air-dry for 2-3 minutes. Once dry, the DNA was re-dissolved in 9.5ml EX buffer (Qiagen) and 200μl of ATP-dependent exonuclease and 300μl of ATP solution were added to the dissolved DNA. The mixture was incubated in a heating block at 37°C for one

hour before adding 10mL of QS buffer. In the meantime, a Qiagen-tip column was equilibrated with 10ml of QTB buffer. The solution containing the DNA was applied into the Qiagen-Tip 500 and allowed to flow through the resin by gravity. The columns were washed two times with 30mL QC buffer and then the DNA was eluted from the column with 15ml pre-warmed (65°C) QF buffer. The DNA sample was precipitated by 0.7 volumes (10.5ml) isopropanol. Once precipitated, the DNA was pelleted, washed with 70% ethanol before drying and resuspended in TE buffer. DNA prepared in this way was then used for sequencing.

2.1.2.3 Polymerase chain reaction

The polymerase chain reaction (PCR) is a well-established technique for amplification of segments of DNA identified by sequence specific oligo-nucleotides employing a heat stable DNA polymerase extracted from *Thermus aquaticus*. In essence, the process involves successive rounds of DNA denaturation, annealing and synthesis from site specific oligo-nucleotide primers resulting in an exponential growth of specific DNA products.

2.1.2.3.1 Amplification of rat genomic DNA

The reaction volume for all PCR reactions was 50µl. The reaction mix contained 50-100ng of template DNA, 20-40 pMol of both a forward and reverse primer, 200mM dNTPs (dATP, dCTP, dGTP and dTTP), 1unit TaqGold polymerase (Applied Biosystem, Foster City, California, USA), 1.5mM $MgCl_2$, 75mM Tris-HCl pH 9.0, 20mM $(NH_4)_2SO_4$, and 0.01% Tween 20. All PCRs were performed on a Perkin-Elmer 9700 Thermal Cycler (Applied Biosystems). The following standard protocol was used to amplify most DNA specific sequences:

94°C for 15 minutes	Denaturation step (5 minutes when using Taq polymerase from Promega, Promega, Southampton, UK)	
94°C for 30 seconds	Denaturation step	30-35 cycles
50-70°C for 30 seconds	Annealing step	
72°C for 45 seconds	Extension step	
72°C for 7 minutes	Final extension	

The annealing temperature was sometimes adjusted to match the specific requirements of individual primer pairs. For DNA products larger than 1Kb the extension time was increased by one minute for each kilobase of additional sequence. When degenerate primers were use, the number of cycles was increased to 50-60 and the annealing temperature was reduced.

2.1.2.3.2 Reverse transcription PCR (RT-PCR)

Reverse transcription (RT) PCR consists of two steps; an initial first strand cDNA synthesis followed by the amplification steps. The SuperScript™ first-strand synthesis system from Invitrogen (Invitrogen Ltd, Paisley, UK) was used to synthesize the first strand cDNA. The RNA/primer mixture containing 1μl of 10mM dNTPs, 1μl of Oligo$(dT)_{12-18}$ (0.5μg/μl), 7μl of DEPC water and 1μl of total RNA at a concentration of 1μg/μl, was mixed in a 0.2mL sterile eppendorf tube and placed at 65°C for 5 minutes followed by 5 minutes at 4°C. After this step, 9μl of a solution containing 2μl of 10x RT buffer, 4μl of 25mM $MgCl_2$, 2μl of 0.1M dithiothreiotol (DTT) and 1μl ribonuclease inhibitor was added. The combined solution was incubated at 25°C for 2 minutes before the addition of 1μl (50 units) of SuperScript II reverse transcriptase and the incubation step continued for a further 10 minutes. After this period, the incubating temperature was then raised to 42°C for a further 50 minutes. The RT activity was terminated by heating to 70°C for 15 minutes and then, placing the tube on ice. Two μl of cDNA was used as template for the 50μl PCR reaction.

2.1.2.3.3 Degenerate PCR

Some degenerate PCR primers were designed from the mouse DNA sequence to amplify the same gene from rat genomic DNA or cDNA. The concentration of each of the primers was increased to 50pMol. The protocol for the degenerate PCR reaction was as follows:

94°C for 15 minutes	Denaturation step	
94°C for 45 seconds	Denaturation step	50~60 cycles
30~50°C for 60 seconds	Annealing step	
72°C for 2 minutes	Extension step	
72°C for 5 minutes	Final extension	

The annealing temperature and running cycles were adjusted according to the specificity and the amount of PCR product, respectively.

2.1.2.3.4 Fluorescent incorporation PCR

For genotyping on the automatic sequencer, a fluorescently tagged nucleotide ([F]dUTPs, Applied Biosystem) was added to the PCR reaction mix for incorporation into newly synthesised DNA. The [F]dNTP kit comes in three colours: blue ([R110]), green ([R6G]) and yellow ([Tamra]). The reaction concentration of [R110] and [R6G] is 1μM (1μl of fluorescent dye in 99μl of sterile water) and of [Tamra] is 4μM (4μl of [Tamra] in 96μl of sterile water). The ratio of the final concentration of non-labelled dNTPs to [F]dUTPs is 200:1 in the 30 cycles PCR reaction. The mixture was put in a 96-well PCR plate and the reaction was performed on a PE-9700 PCR machine. After amplification the product was precipitated by the addition of 4 volumes of ethanol (with 5% 3M sodium acetate) added into each well for precipitation. Precipitation was

encouraged/enhanced by placing the sample at –70 °C for at least 30 minutes. The PCR plate was spun at 4500rpm for 45 minutes in Rotina 46, Hettich zentrifugen (Tuttlingen, Germany). After centrifugation, the plate was placed upside down to allow all the ethanol to drain away. To remove residual ethanol the plate was spun upside down for one minute at less than 150rpm. Samples were stored at -20°C prior to electrophoresis using the GeneScan acrylamide gel system.

2.1.2.3.5 Purification of PCR products

After the amplification, the quality of PCR product was checked by running a 5µl aliquot on a 1% agarose. If discrete product was detected after staining with ethidium bromide, the remaining PCR product was purified for sequencing. Purification was necessary to remove excess primers, salts and unincorporated dNTPs to prevent mis-priming of the sequencing reaction. This was achieved using Qiagen PCR purification kits. Five times excess of the Qiagen PB buffer was added to the PCR product. After mixing, the solution was transferred to a Qiagen spin column. The column was spun at 13000rpm for 1 minute in a Micromax centrifuge (IEC). The flow through was discarded and 750µl of PE buffer was used to wash the column. After washing, the column was spun twice to ensure that the entire PE buffer was removed. The column was placed into a new eppendorf tube and 50µl of water was pipetted onto the spin column membrane to elute the DNA. The column was left to stand for at least 1 minute

before the final spin (13000rpm for 1 minute). This allowed the maximum yield of DNA. After centrifugation, the DNA was stored at -20°C until used.

2.1.2.3.6 Sequencing reaction

The purified PCR product was sequenced using Bigdye™ Terminator Cycle Sequencing Ready Reaction Kit version 2 (Applied Biosystems). The same PCR primers (both forward and reverse) were used to generate sequence data but at a lower concentration (3.2pMol). The Bigdye Terminator Ready Reaction Mix contains four different colours of labelled terminator nucleotides, dNTPs, AmpliTaq DNA polymerase, FS, $MgCl_2$, and Tris-HCl buffer. The total 16µl reaction volume consisted of 4µl of Terminator Mix, 1µl of forward or reverse primer and 11µl of purified PCR product. The reaction was performed in a 96-well plate on a PE-9700 thermal cycler using following protocol:

96°C for 1 minute

96°C for 10 seconds
50°C for 5 seconds
60°C for 4 minutes
} Repeat for 25 cycles

Rapid ramp to 4°C and then place in a -20°C freezer until used

To purify the extension products and to remove excess Bigdye terminators, 80µl of 70% ethanol was added into each well. After covering with adhesive tape, the plate was placed on a shaker and shaken for an hour to precipitate the product.

The plate was then spun in a Beckman centrifuge (model GS-6R) at 3000rpm for 45 minutes. The supernatant was removed by quickly inverting the plate and blotting on filter paper. Residual supernatant was removed by spinning the plate upside down for one minute at 150rpm. The samples were washed with 80µl of 70% ethanol again. This was also removed by centrifugation. The pelleted product was then dried by placing the plate at 55°C. The samples were then stored at -20°C until they could be run on a sequencing gel.

In order to sequence the BAC ends using primers for the vector's Sp6 (5'-gttttttgcgatctgccgtttc-3') and T7 (5'-gccgctaatacgactcactatagggagag-3') promoter sites, the standard protocol needed modification to accommodate the difference in molarity associated with a large genomic clone. Approximately 0.8-1.0µg of BAC template DNA was used in these sequencing reactions. The reaction volume of 40µl was made up as follows: 16µl of Bigdye polymerase mixture, 10 pMol of primer and 0.8-1.0 µg of template. Prior to amplification, the DNA was denatured by heating to 65°C for 10 minutes. The reaction was performed in a 96-well plate on a PE-9700 thermal cycler using the following protocol:

96°C for 5 minute

96°C for 30 seconds
53°C for 20 seconds } 57 cycles
60°C for 4 minutes

Rapid ramp to 4°C and then put into -20°C freezer until use

The sequencing product was purified using the Qiagen DyeEx spin columns following the manufacturer's recommendations. The protocol in the DyeEx spin kit was followed. Prior to use, the columns were vortexed and spun at 750*g* for 3 minutes in order to redistribute the column resin. The sample was carefully applied onto the column resin and, after another spin, collected in a micro-centrifuge tube. It was then dried in a lyophiliser before being resuspended and run on an ABI sequencer.

2.1.2.4 Electrophoresis

2.1.2.4.1 Agarose gel electrophoresis

Agarose gels were used to check the quality of the nucleic acids (DNA and RNA) and to resolve fragments according to molecular weight size. The concentration of agarose gel varied from 0.8% to 3.5% depending on the size of the molecules to be resolved. Generally, small fragments (100-300bp) were separated in 3.2% agarose gel. Larger molecular weight DNA was usually resolved on 0.8% gel. All gels contained 0.6μg/ml ethidium bromide so that the nucleic acids could be seen under long wave ultra-violet light. Gels were cast containing 1x TBE buffer (Appendix I). Prior to loading, samples were mixed with 5μl of loading buffer (Appendix I). This buffer provided density to the sample and contained a tracking dye (Orange G) so that the migration of the sample through the gel could be monitored. The gel was run at a constant voltage, of 50V. Molecular weight size of fragments and PCR products was

estimated by running a molecular weight marker sample (100bp or 1kb ladder DNA, Promega) along with the samples.

2.1.2.4.2 Single-strand conformation polymorphism (SSCP) gel electrophoresis

SSCP gel electrophoresis was used for mutation and polymorphism screening. 10μl of PCR product was mixed with 6μl of SSCP loading blue dye (see appendix I) and denatured by heating at 96°C for 3 minutes. The sample was chilled on ice after denaturation and then loaded into the well of a pre-cast 10% TBE gel (Novex 10% TBE gel, Invitrogen). The gel was immersed in a tank filled with 1x TBE buffer and electrophoresis was carried out vertically under a constant voltage of 200V, at 10°C and 20°C, for 2 to 3 hours, depending on the size of PCR product.

The gel was removed from the Manufacturer's casting cassette and placed in a bath of 1x TBE containing 100ng/ml ethidium bromide. After staining for 20 minutes on a shaker, excess ethidium bromide was washed away and the gel examined on a long-wave ultra-violet transilluminator.

2.1.2.4.3 Photography of electrophoresis gel under UV light transilluminator

An UV transilluminator (Syngene, Bioimaging system, A Division of Synoptics Ltd, Cambridge, UK) was used to check the ethidium bromide stained gel. The photograph was viewed on a computer screen and stored as an image file or printed with a Sony video graphic printer (model UP-860 CE).

2.1.2.4.4 Polyacrylamide gel electrophoresis

2.1.2.4.4.1 Polyacrylamide gel preparation

The ingredients for the gel mixture for the ABI Prism 377 and 373 automatic sequencers are the same; but the volume varies (see appendix I). Before cleaning the gel plates, the gel mixture was prepared and stirred until all the urea had dissolved. For the 377 sequencer, the 36cm well-to-read glass plate was used and the 24cm well-to-read plate was used for the 373 sequencer. Both bottom and cover plates were scrubbed by 1% Decon after removing the old gel. Large amounts of warm tap water was used to wash the plates and then rinsed with nano-pure water. The gel and well spacers were also washed using the same procedure. The air-dried bottom plate was set in a 377 cassette with two spacers on each side. The cover plate was covered on top of the bottom plate with the screw turned to tighten both plates. The plates for the 373 sequencer were held together by bulldog clips. In the preparation of the gel for the 377 sequencer (50ml), 250µl of 10% ammonium persulfate (APS, Sigma) and 35µl of N, N, N',

N'-tetramethylethylenediamide (TEMED, Sigma) were added into the filtered gel mixture, whereas 400µl of 10% APS and 45µl of TEMED were used for the 80ml 373 sequencing gel mixture. A 50ml syringe was used to suck the gel mixture after gently mixing and then, the mixture was injected in between the plates. When the gel went through, the cover plate was lightly tapped to prevent bubble formation. The gel was left to polymerise for two hours before running. The spacer strip was removed and the plate cleaned again so that no gel was left on the top of the gel. A 36 or 48 well shark-tooth comb was inserted into the gel with the tips dipped into the gel for 2mm. The plates were then placed into the 373 or 377 sequencer chamber. The Genescan collection software (version 2.5) was used on the ABI 377 sequencer and Sequencing collection software (version 2.0) was used for sequencing. The laser checked the background of the plates. The heat plate and buffer chamber were assembled and 1.3 liter 1x TBE running buffer was poured into the upper and lower chambers after the plate was checked. The machine pre-runs the gel until the temperature was high enough (usually 10 minutes) for sample loading.

2.1.2.4.4.2 Gel loading and conditions for electrophoresis

The 1x TBE buffer was used to wash the wells carefully until no bubbles were left. The Genescan sample was dissolved in 5µl of mixture, containing 2.5µl deionized formamide, 1.5µl nanopure water, 0.5µl size standard (GS-400HD [ROX], Applied Biosystems), and 0.5µl gel loading dye. For the sequencing gel, 5µl of loading buffer (Formamide: gel loading dye = 5:1) was added into the

sequence pellets. The sample was denatured at 96°C for 4 minutes and chilled on ice immediately. The 0.2mm duckbilled flat tips (Anachem) was used to suck 1.5-2µl of the samples, which was then loaded into alternate lanes (odd number). Before loading the even number wells, the gel electrophoresis was run for 3 minutes at 3000V under the prerun model. The machine was paused and the samples were loaded into the even wells. The parameters for Genescan run on a 377 sequencer were as follows,

Plate check module	Plate check A
Pre run module	GS PR 36A-2400
Run module	GS Run 36A-2400
Matrix file	GS Matrix Set –A
Lanes	96 or 48 lanes
Run time	2.5 hours

The parameters for sequencing gel electrophoresis were as follows:

DNA sequencing on the ABI automated 377 Sequencer

Plate check module	Plate check E
Pre run module	GS PR 36E-1200
Run module	GS Run 36E-1200
Matrix file	dRhod
Lanes	48 lanes
Run time	7 hours

DNA sequencing on the ABI automated 373 Sequencer

Plate check module	Plate check
Pre run module	Pre Run
Run module	Run
Matrix file	dRhod
Lanes	48 lanes
Run time	12 hours

2.1.2.4.5 Data analysis

2.1.2.4.5.1 Genescan data analysis

The 377 sequencer generated the genotyping data and was stored as a gel file with image pictures. The software GeneScan® 3.1.2 was employed to analyse the gel file. The software opened the file and auto-tracked the lanes. The data of each lane was extracted including the standard and the samples. Before determining the size of the product, we defined the standard and the software analysed each lane comparing the sample to the standard automatically.

2.1.2.4.5.2 DNA sequencing data analysis

The software for sequencing analysis was Sequencing Anaysis 3.4.1, which converted the four different wavelengths of fluorescence into four equivalent

bases. After auto-tracking the lane, the data was extracted from each lane. The parameters for the analysis are as follows: ABI100 for Basecaller, 500 bp for Basecaller Settings, 0 for Peak 1 Location, 0 for Start Point, 12000 for Stop point, DT{BD Set Any-Primer} for Dye Set/Primer file and dRhod_Bigdye for Instrument file. The software analysed the sequence automatically after setting all parameters. The data was then exported into the software Sequence Navigator™ 1.0.1, which contained the software Factura™ 1.2.0 (Applied Biosystems). Both the tested and normal control sequences were aligned. The software highlighted any base changes among the alignments.

2.1.2.5 Southern Blot hybridisation

The method for DNA hybridisation followed the transfer and detection protocols for GeneScreen™ & GeneScreen Plus ® hybridisation transfer membranes (NEN® Research Products, Boston, Massachusetts, USA). The genomic DNA from *mf* rat, normal SD rat, Guinea pig and human was extracted and quantified. 5μg of DNA was digested with the following restriction endonucleases: *Bam*HI, *Hind*III and *Bgl*I (Promega) using the protocols recommended by the manufacturer. Once digested, the DNA samples were loaded into a 0.6% agarose gel and the electrophoresis was performed under 70Volts for 14 hours. Quality of the restricted DNA was assessed by staining the gel with ethidium bromide and examining it with a long-wave length ultra-violet transilluminator. The gel was then ready for salt transferring. The membrane was equilibrated in 2X SSC

(see appendix I) for 15 minutes. The gel was subjected to a depurination step by treatment with 0.25N HCl for 10 minutes. This ensures better transfer of the higher molecular weight DNA. After rinsing in distilled water for 5 minutes the DNA was denatured by treatment with 0.4N NaOH / 0.6M NaCl for 30 minutes. Once denatured, the gel was neutralised by treatment with 1.5M NaCl / 0.5M Tris-HCl pH7.5 for 30 minutes. The gel was then placed on a blotting apparatus, with wick and reserve tank containing 2X SSC. A pre-soaked and sized membrane filter was placed on top of the gel. Additional blotting material (paper towels) was placed on top and left over night to draw the 2X SSC through the gel and membrane by capillary action. Thus, moving the DNA through the gel onto the blotting membrane. After the transfer, the membrane was removed and rinsed with 0.4N NaOH for 1 minute and then neutralised by rinsing in 0.2M Tris-HCl pH 7.5 / 1X SSC for 1 minute.

The DNA bound membrane was pre-hybridised for two hours at 65°C in the same solution as the hybridisation solution (appendix I) minus the labelled probe DNA. During the prehybridisation step, the *p75NTR* PCR amplicon was radioactively end-labelled using the bacteriophage T4 kinase reaction. Unincorporated label was separated by spinning (1500rpm for five minutes) the sample through a Sephadex G60 column. 300µl of salmon sperm DNA (10 mg/ml) was added to the end labelled probe, and then the DNA was denatured by heating to 95°C for 5 minutes. The labelling probe and salmon sperm DNA were added to the pre-hybridisation solution in a sealed bag and the hybridisation continued over-night at 65°C. After the over-night hybridisation, the blot was washed two times in a solution containing 2X SSC 1% SDS at room temperature

for five minutes each. An additional wash in the same solution was performed at 65°C for 5 minutes. The membrane was removed from the wash solution, blotted dry with Whatman filter paper, wrapped in saran wrap and placed on Kodak X-ray film in a film cassette. The cassette was then placed at -70°C for 4 hours before the film was developed.

2.1.2.6 Northern Blot hybridisation

A Northern slot blot was performed to examine the expression of the *Cct4* gene in different rat tissues. Probes were designed from rat *Cct4* cDNA sequence (Table 2.1). The expression of the rat glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) gene was chosen as a positive control and two probes were designed from the *Gapdh* cDNA sequence (Table 2.1). To distinguish the possibility of cross-hybridisation with genomic sequences, primers were designed that bridged exon sequences.

Table 2.1 Probe sequences for rat *Cct4* and *Gapdh* genes

	Probe	Sequences
Cct4	1	163-taggtccaaggcttgttctaatggcatcggcaaccgctttggc-121
	2	1131-tgtaatcttgaacagttttccagaaccattc-1101
GAPDH	1	30-gccaaatccgttgacaccgaccttcaccatcttgtctatgagac-(-14)
	2	970-ccaccaccctgttgctgtagccatattcattgtcataccaggaa-927

(-14) = the 14th nucleotide in the 5'untranslated region of *GAPDH*

2.1.2.6.1 Probe tailing and determination of labelling efficiency

Using the DIG oligonucleotide tailing kit (Roche, Diagnostics GmbH, Mannheim Germany), the probes were labelled with Dig-dUTP following the Manufacturer's protocol. The concentration of the probes was diluted to 100pMol/µl. In a total reaction volume of 20 µl, there were 4 µl of reaction buffer, 4µl of $CoCl_2$-solution, 1 µl of DIG-dUTP solution, 1µl of dATP solution and 1µl of 50 U Terminal transferase). The mixture was centrifuged briefly and then stood in ice. After a 15 minute- incubation period at 37°C the reaction was terminated by the addition of EDTA (pH 8.0). The labelling efficacy was determined by binding a serially diluted aliquot of the reaction mix to a nylon membrane. After drying at 80°C for 30 minutes, the unincorporated label was washed away by rinsing the membrane in a solution (Buffer 1) containing 0.1 M maleic acid and 0.15M NaCl at pH 7.5 (appendix I). The membrane was then immersed in blocking buffer (1% solution diluted 1:10 with buffer 1) for 30 minutes and further incubated in solution containing the anti-digoxigenin antibody-peroxidase (anti-DIG-POD), which was diluted to 1:500~ 1:1000 in blocking buffer for 30 minutes. The membrane was rinsed twice with washing buffer, 0.3% Tween 20 in buffer 1. After briefly rinsing with PBS, the membrane was stained for peroxidase activity. The stain, which contained Vector® NovaRED™ substrate (Vector Laboratories Ltd., Peterborough, UK), was freshly prepared and used under reduced light conditions for 10 minutes. The reaction was terminated by washing with buffer 4, containing 10mM Tris-HCl and 1mM EDTA (pH 8.0) and the activity was detected by dark red coloration.

2.1.2.6.2 Transferring and fixing the RNA onto a nylon membrane

RNA extracted from soft tissue was quantified and approximately 40μg denatured by treatment with 10 mM NaOH and1mM EDTA solution and slot-blotted onto a nylon membrane. A Bio-Rad (Bio-Rad Laboratories Ltd., Hertfordshire, UK) vacuum slot-blot apparatus was used to draw the RNA sample on to pre-soaked membrane. Once the sample had been pulled down on the membrane, the wells were washed with 500μl of the denaturing solution. The RNA was fixed onto the membrane in an oven at 80°C for 30 minutes.

2.1.2.6.3 Pre-hybridisation and hybridisation

Using 20ml of solution for each 100cm^2 of membrane, the Northern slot-blot was pre-hybridised with hybridisation solution (appendix I) in a Hybaid (Hybaid Ltd., Middlesex, UK) hybridisation oven at 68°C for an hour. After this period, the pre-hybridisation solution was replaced with 3.5ml of hybridisation solution (containing either *Cct4* or *Gapdh* labelled probe, 10 pMol/ml) for every 100 cm^2 of membrane. Hybridisation was continued at 68°C overnight. The membrane was then washed three times at room temperature in a solution containing 0.2X SSC and 0.1% SDS for a duration of 15 minutes each.

2.1.2.6.4 Signal detection

Hybridisation was detected using the anit-digoxigenin-POD peroxidase detection system. In essence, the blot was exposed to a peroxidase/antibody conjugate that recognises the Dig-dUTP tail of the probe DNA. After washing, peroxidase activity was detected by the production of coloured precipitate from a colourless substrate involving the turn over of hydrogen peroxide. The Vector® NovaRED™ substrate kit (Vector Laboratories Ltd., Peterborough, UK) was specifically designed for the detection of the peroxidase enzyme and it was used according to the Manufacturer's protocol.

2. 2 Human CMT disease project

2.2.1 Patients

2.2.1.1 Patients for *MPZ* mutation screening

2.2.1.1.1 Taiwanese patients

Five Taiwanese index patients with Charcot-Marie-Tooth (CMT) disease were selected by Professor Sung-Tsang Hsieh from National Taiwan University Hospital, Taipei, Taiwan. The clinical summaries are presented in Table 2.2. There were three female and two male index patients. The disease began in early childhood and progressed slowly. These five patients presented with distal muscle wasting and weakness of the feet and lower limbs, walking difficulties and sensory disturbances. Three patients had reduced tendon reflexes. Two index patients had a positive family history. The electrophysiological data in Table 2.3 showed that the motor conduction velocities of the median nerves were below 38 m/s in all patients, which is compatible with demyelinating CMT. Genomic DNA from these five patients was extracted and sent from Taiwan.

Table 2.2. Summary of the clinical features of the five Taiwanese patients

Pt.	Age (y/o) / Sex	age of onset	Symptoms and signs	Family history
#051	65/ female	Unknown	Walking difficulty, pes cavus, generalised hyporeflexia	Unavailable
#056	15/ male	7 y/o	Unsteadiness, walking difficulty, positive Rombergism, areflexia	No
#092	32/ female	6-10y/o	Distal muscle weakness and atrophy, unsteady gait and sensory impairment	Yes
#094	63/ female	Unknown	Aching pain with sensory impairment on the lower limbs	Unavailable
#099	20/ male	10 y/o	Slowly progressive distal muscle wasting and weakness in lower limbs, areflexia	Yes

Pt. = Patient
y/o = years of age

Table 2.3. Nerve conduction studies of the Taiwanese patients and their families.

	Motor conduction velocities (m/sec)		Sensory action potentials (μV)
Patient	Median	Ulnar	
#051	30 (R)	27 (R)	All absent
	23 (L)	30 (L)	
#056	15 (R)	13 (R)	All absent
#092	Absent	Absent	All absent
Son of #092	Absent	5 (R)	All absent
		5 (L)	
Daughter of #092	7 (R)	4 (R)	All absent
#094	28 (R)	24 (R)	All absent
	31 (L)	27 (L)	
#099	13 (R)	13 (R)	All absent
Brother of #099	11 (R)	14 (R)	

m/sec = meter per second
R = right side
L = left side
μV = micro-voltage

2.2.1.1.2 British CMT2 patients

Forty-six patients mostly from England, presenting with progressive distal muscle wasting, weakness and sensory disturbance were collected. Difficulties in walking, pes cavus and foot deformities were found in most of the patients. There was no involvement of the cranial nerves, cerebellum or autonomic nervous system in any patient except two patients had pigmentary retinopathy. Thirty-three patients had a positive family history of neuropathy. Mutations in *GJB1* (gap junction protein beta 1) were excluded in patients without male-to-male transmission in their family history. The average age of onset was 30.9 ± 22.0 years of age (y/o) (the range is 2-68 y/o). The nerve conduction studies showed decreased sensory and motor action potentials with minimal slowing of nerve conduction velocities. The range of the conduction velocities in the right median nerves were between 33.3 and 59 m/s (average is 47.2 ± 5.7 m/s). Axonal degeneration with minimal secondary demyelination was noted in the biopsied sural nerve from eight patients. DNA was extracted from these patients for mutation screening.

2.2.1.2 Patients for *GJB1* mutation screening

One hundred and thirty nine patients with a diagnosis of CMT were collected. Eighty patients had been seen in the peripheral neuropathy or neurogenetic clinics at the National Hospital for Neurology and Neurosurgery. Fifty-nine patients were referred from Neurologists elsewhere in the United Kingdom.

Most of the patients came from the south of England. All patients showed distal muscle weakness and wasting, sensory disturbance and some had foot deformities. The clinical features and electrophysiological studies of these patients were used to classify them into 44 with CMT1, 38 with CMT2, 6 with CMT3 (Dejerine-Sottas disease), 21 with CMTX and 30 with CMT (without adequate details to classify further). Eighty patients had a family history without male-to-male transmission. Family history or clinical details were not available in thirty patients. However, long disease duration and the presence of foot deformity, such as pes cavus suggested that these patients suffered from an inherited neuropathy. The chromosome 17p duplication had been excluded in all CMT1 patients. The mean age of the patients was 43.5± 17.8 y/o (range from 2-77 y/o) and there were 64 females and 75 males. Ethical approval for genetic studies in neuropathy had been obtained from the National Hospital Ethics Committee.

2.2.2 Methods

2.2.2.1 Extraction of genomic DNA from human whole blood

10 ml of venous blood in EDTA coated bottles was collected from the index patients and their families. The samples were divided into two falcon tubes and lysed with 40ml Reagent A (see appendix I). The sample was spun in a Beckman centrifuge (model GS-6R) at 2600rpm for 20 minutes to pellet the white blood cells. The pellet was washed with 5ml PBS (see appendix I) after

discarding the supernatant. The resuspended samples were centrifuged again in 2600rpm for 8 minutes and this step was repeated. The supernatant was discarded and then, 500µl of a solution containing 25µl 10% SDS, 500µl CVS buffer and 100µl proteinase K solution was added (see appendix I). The samples were incubated at 55°C overnight. 500µl of phenol was added to each sample after cooling and inverted a few times to mix the sample. The mixture was spun at 2600rpm for 15 minutes. The top aqueous phase was transferred to a new tube. The same volume of phenol was added to the aqueous phase and mixed well and spun at 2600rpm again. 500µl of chloroform was added to the aqueous phase and then, the tube was inverted and mixed thoroughly. The sample was spun at 2600rpm for 15 minutes to remove the phenol. The DNA was precipitated with 2 times volume of absolute ethanol with 0.5% 3M sodium acetate. After centrifuging, the DNA was washed with 70% ethanol. Then, the sample was left to dry in the air for 10 minutes and the DNA was resuspended in 300µl 1x TE (see appendix I). The method for assessing the yield is described in section 2.2.1.2.

2.2.2.2 PCR

The protocol, the method and the machines used for PCR were the same as described in section 2.1.2.2.. A few reactions needed a touch-down protocol. In the initial 25 cycles, the annealing temperature was decreased by each cycle and then in the final 12 cycles a constant temperature was used as in a standard reaction. The stepwise decrease of the annealing temperature increased the

chances of annealing the primers to the DNA. Published primers were used to amplify *MPZ* and *GJB1* [Roa *et al.* 1996] [Bergoffen *et al.* 1993b].

2.2.2.3 Restriction fragment length polymorphism (RFLP) analysis

In order to screen a large number of normal control chromosomes for mutations, restriction polymorphism is efficient and simple. The software Generunner 3.05 (freely available from http://www.generunner.com/) was used to detect the cutting enzyme. In general, 7-10μl of PCR product of exons with a mutation was used for the digestion reaction in a mixture containing 1μl of restriction enzyme, 2μl of enzyme buffer and sterile water to make a total volume of 20μl. A few enzymes needed bovine serum albumin to facilitate the reaction. The mixture was incubated at the optimal temperature for the enzyme for 2 hours. Then, the product was electrophoresed through a 3.2% agarose gel to allow separation of bands.

2.2.2.4 Genotyping and Sequencing

The genotyping and sequencing were performed using the same procedures as described in section 2.1.2.3.4. For human genotyping, some of the primers were labelled with fluorescence. Following the screening of *GJB1* gene in patients with CMT, two families sharing the same mutation was found with four mutations. In order to look for a founder effect, five markers (DXS990, DXS986, DXS8052, DXS8107, DXS991) close to the *GJB1* (in distance less than 1 cM) were genotyped.

Chapter 3

Results of genetic analysis in the mutilated foot (*mf*) rat

This chapter describes the results of the genetic analysis and identification of the disease-causing gene for the mutilated foot (*mf*) rat. Section 3.1 details the results of the screening of two candidate genes for *mf* disease, *NTRK1* and *p75NTR*. A detailed linkage study by genome wide screening with genetic markers and the refinement of the linkage map are presented in section 3.2. Section 3.3 describes the process of identifying the responsible gene for *mf* disease.

3.1 Positional independent candidate gene study

Before embarking on genome-wide screening in the rat, two potential candidate genes for *mf* disease, *Ntrk1* and *p75NTR*, were investigated because the clinical features of transgenic and knockout mice of these two genes are similar to *mf* rats [Lee *et al.* 1992] [Smeyne *et al.* 1994]. RT-PCR and direct cDNA sequencing showed that no sequence differences were detected in the coding region of the *Ntrk1* gene in the *mf* rat as compared to a normal Sprague-Dawley (SD) rat and the published rat sequence. While sequencing the coding region of *p75NTR*, three nucleotide changes (c. 183G>A, c.325A>G and c.1379A>G) were identified as compared to the published rat sequence. The c.1379A>G did not

cause an aminoacid change (Thr422Thr). However, c.183G>A and c.325A>G result in missense changes Val24Met and Asn71Met, respectively. It is unlikely that these changes are significant because both normal SD and *mf* rats have the same sequence changes (Fig. 3.1 a, b). To rule out possible large-scale genomic rearrangement in *p75NTR*, Southern blot hybridisation were employed. Genomic DNA from *mf*, normal SD and Wistar rats was cut with 3 six base pair recognition site restriction enzymes. Using the coding sequence of the *p75NTR* as a probe, no large-scale genetic rearrangements such as deletions or duplications of the gene were detected by Southern blot hybridisation.

3.2 Genetic linkage analysis

3.2.1 Choosing mating strain

Although the *mf* mutation arose on the SD strain background, it has been maintained as an isolated colony for some years. Thus, it was necessary to ascertain how close it was genetically to the SD strain and how different it was from other strains. This was achieved by PCR amplification of genomic DNA from the *mf* stock using primer pairs to micro-satellite repeat loci that spanned the rat genome. By comparing the product size with those of the published data, we were able to determine that the LE strain was sufficiently different (48/51 allelic size variations were detected between the two strains) (Table 3.1).

Fig. 3.1a

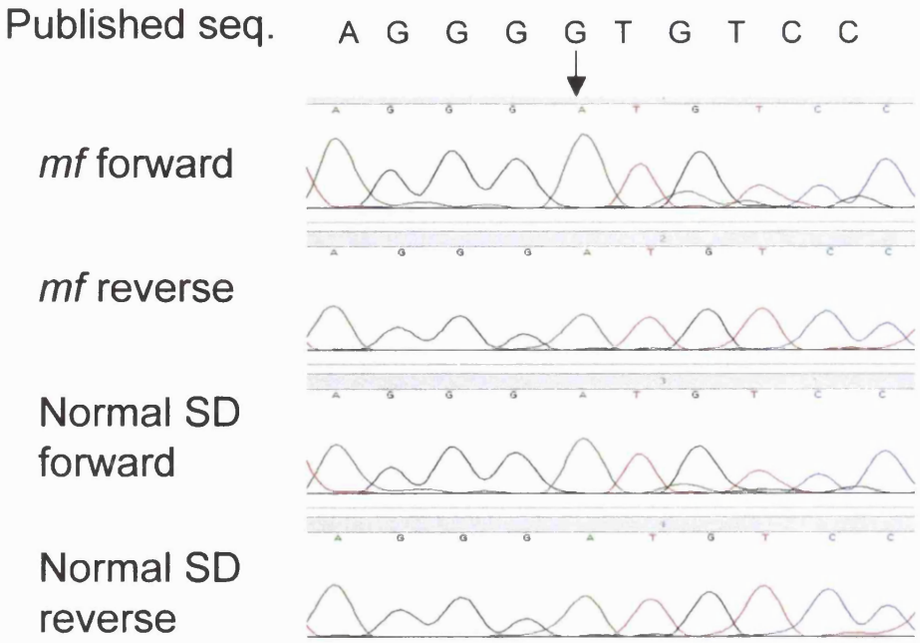


Fig. 3.1b

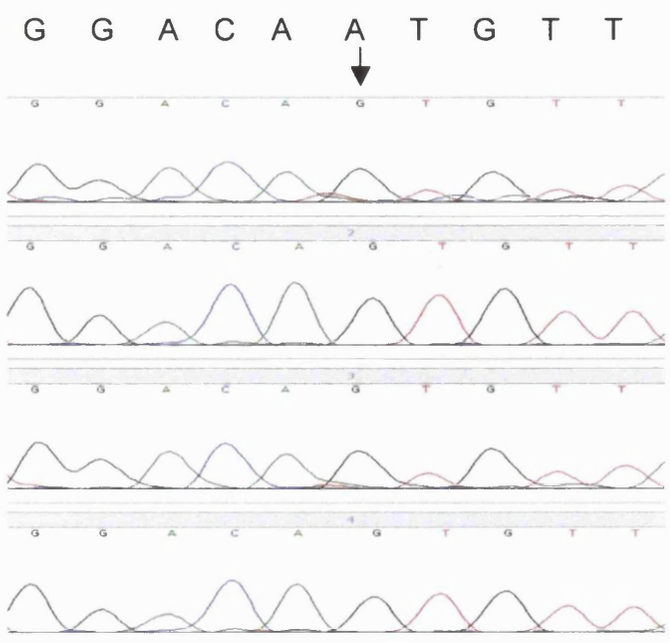


Fig. 3.1. Sequence of *p75NTR* in *mf* and normal SD rats. As compared to the sequence published in the database a G to A change at nt 183 causing Val24Met (arrow in Fig. 3.1a) and a A to G change at nt 325 resulting in Asn71Met (arrow in Fig. 3.1b) are shown both in the *mf* and normal SD rats (Published seq. = published sequence from Genbank).

The alleles of *mf* and Lewis rats were variable in size; the smallest difference was only 2 base pairs (bp) (*D13Mit5* and *D19Mit3*), whereas the largest difference was 68 bp (*D5Mit4*). No differences were detected with *D2Mit15, D9Mgh2* and *D14Mgh2*, which were not informative for the linkage study. These data justified using the Lewis (*LE*) rat as a mating partner of the *mf* rat in our mapping study.

Table 3.1 Allele sizes in base pairs of the *mf* rat and Lewis rat

Markers	*mf* rat	Lewis rat	Size difference (bp)	Markers	*mf* rat	Lewis rat	Size difference (bp)
D1Mgh1	248	238	10	*D9Mit3*	149	178	29
D1Mgh5	148	141	7	*D9Mgh2*	106	106	0
D1Mgh19	218	200	18	*D10Mit8*	172	184	12
D1Mit7	172	178	6	*D10Mit10*	198	210	12
D2Mgh6	113	117	4	*PPY*	171	165	6
D2Mit6	187	197	10	*D11Mgh1*	154	104	50
D2Mit13	230	242	12	*D11Mgh5*	229	234	5
D2Mit15	166	166	0	*D12Mgh3*	115	120	5
D3Mgh6	104	110	6	*D12Mit6*	142	146	4
D3Mgh10	120	131	11	*D13Mit5*	248	246	2
D3Mit2	214	225	11	*D13Uw1*	217	198	19
D3Mit10	219	170	49	*D14Mgh2*	133	133	0
D3Mit15	138	149	11	*D14Mit7*	238	300	62
D4Mgh8	121	124	3	*D15Mgh1*	149	130	19
D4Mit2	253	240	13	*D15Mgh3*	153	147	6
D4Mit17	134	147	13	*D16Mgh4*	148	155	7
D5Mit4	241	309	68	*D16Mit2*	168	179	11
D5Mit10	104	101	3	*D17Mgh4*	120	111	9
D6Mit6	146	116	30	*D17Mit7*	216	177	39
D6Mit8	210	203	7	*D18Mgh2*	195	217	22
D6Mit9	150	157	7	*D18Mit1*	290	309	19
D7Mit4	223	235	12	*D19Mit3*	116	118	2
D7Mit7	249	245	4	*TAT*	327	307	20
D7Mit9	271	268	3	*D20Mgh1*	199	236	37
D8Mgh7	214	225	11	*D20Uw1*	144	160	16
D8Mgh11	157	162	5				

3.2.2 Genome wide screening

34 affected N2 rats were analysed for segregation of alleles at 93 microsatellite markers spanning the entire rat genome (i.e., approximately 20cM between loci). Despite our best attempts to ensure maximum genetic differences in the founder crosses, approximately 50% (48/93) of markers were invariant between the two strains. However, using the remaining 45 markers, we were able to establish clear linkage between the *mf* locus and microsatellite markers on chromosome 14 (Table 3.2). For those microsatellite repeat markers that were clearly polymorphic between the two parental strains, most displayed a recombination frequency with the *mf* locus that was not significantly different from independent segregation with LOD scores well below the level needed to establish linkage (Table 3.2). However, *D14Rat22* displayed significant linkage to the *mf* locus (with a LOD score of 9.3). No recombinants between the two loci were detected in the initial set of 34 affected N2 rats. Another 34 unaffected N2 rats were typed and they all showed heterozygosity at the *D14Rat22* locus. The strong linkage with *D14Rat22* ruled out the possibility that the *mf* locus might be on one of the other chromosomes for which we had no linkage data. The markers close to rat *p75NTR* gene, are *D10Rat98, D10Rat27, D10Rat124* and *D10Rat99*, none of which showed significant linkage to the *mf* locus.

Table 3.2. Screened markers, the frequency of recombination and the LOD scores for the affected 34 N2 rats.

Chromosomes (Chr)	Markers	Approximate position (cM)	Recombination frequency (%) in 34 affected N2 rats	LOD Score
Chr 1	*D1Rat173*	75.5	50.0	0
	D1Rat286	94.4	44.4	0.1
	D1Rat197	109.8	50.0	0
Chr 2	*D2Rat202*	26.7	53.1	-
	D2Rat161	46.7	48.5	0
	D2Rat70	117.7	61.3	-
Chr 3	*D3Rat82*	13.4	44.8	0.1
	D3Rat45	24.6	43.8	0
	D3Rat132	92.3	48.5	0.1
Chr 4	*D4Rat115*	24.9	36.4	0.5
	D4Rat24	40.2	34.1	0.1
	D4Rat233	56.7	53.3	-
	D4Rat66	92.1	47.1	0
Chr 5	*D5Rat82*	26.5	39.4	0.3
	D5Rat154	60.6	45.2	0.1
	D5Rat41	87.1	50	0
Chr 6	*D6Rat105*	23.2	45.2	0.1
	D6Rat121	62.0	64.5	-
	D6Rat79	87.5	61.8	-
Chr 7	*D7Rat32*	11.1	47.1	0
	D7Rat100	56.5	70.6	-
	D7Rat4	81.9	29.4	1.3
Chr 8	*D8Rat55*	9.6	43.8	0.1
	D8Rat156	40.6	66.7	-
	D8Rat205	50.1	72.2	-
	D8Rat123	68.7	63.6	-
Chr 9	*D9Rat79*	15.6	48.5	0
	D9Rat13	46.7	47.1	0
	D9Rat109	73.3	50	0
Chr 10	*D10Rat43*	32.7	42.4	0.2
	D10Rat98	53.1	51.6	-
	D10Rat27	57.0	40.0	0.3
	D10Rat124	63.1	38.9	0.2
	D10Rat99	68.6	43.3	0.1
	D10Rat88	87.6	34.6	0.5
Chr 11	*D11Rat18*	11.1	47.1	0
	D11Rat37	39.9	69.7	-
Chr 14	*D14Rat55*	17.9	29.4	0.9
	D14Rat12	28.8	32.4	0.6
	D14Rat39	41.0	17.6	3.4
	D14Rat22	**68.4**	**0**	**9.3**
Chr 16	*D16Rat1*	16.4	47.1	0
	D16Rat53	33.3	45.2	0.1
Chr 20	*D20Rat5*	17.6	50	0
	D20Rat39	33.9	55.9	-

3.2.3 Refinement of the linked region

To refine the linked region, both more loci and more backcross animals were included. The following markers flanking *D14Rat22* were chosen for typing: *D14Rat21, D14Rat38, D14Rat49, D14Rat130, D14Rat 138*, which were located proximally to *D14Rat22* and *D14Rat95, D14Got73, D14Got72, D14Got76, D14Rat32, D14Rat131, D14Rat 109, D14Rat132, D14Rat139*, which were located distally. Although the consensus mapping data identified these 14 loci in the region of *D14Rat22*, we were only able to successfully integrate 10 of them. We used all 254 N2 rats in our expanded linkage study. As shown in Figures 3.2, expansion of the sample size resulted in the identification of 2/254 recombinants between the *mf* locus and *D14Rat22*. However, no recombinants were detected between marker *D14Rat72* and the *mf* locus in this larger sample set. Based upon this data, it can be deduced that the two loci (disease locus and *D14Got72*) are within 1cM of one another at the 95% confidence interval. Furthermore, both loci are flanked proximally by *D14Got73* (1/254 recombinant) and distally by both *D14Got76* and *D14Rat32* (1/254 recombinant each). Thus, the *mf* locus lies within a 0.2-1.5cM region defined by the flanking loci (Fig. 3.2). In order to further refining the linked region, we bred another 282 N2 rats in the second year. We typed markers *D14Got73, D14Got76* and *D14Rat32* and *D14Got72*. In the total 536 N2 rats, we found one additional recombinant at each of the flanking markers *D14Got73, D14Got76* and *D14Rat32* but there were still no recombinants found at *D14Got72*. This suggests that the disease gene is very tightly linked to this marker.

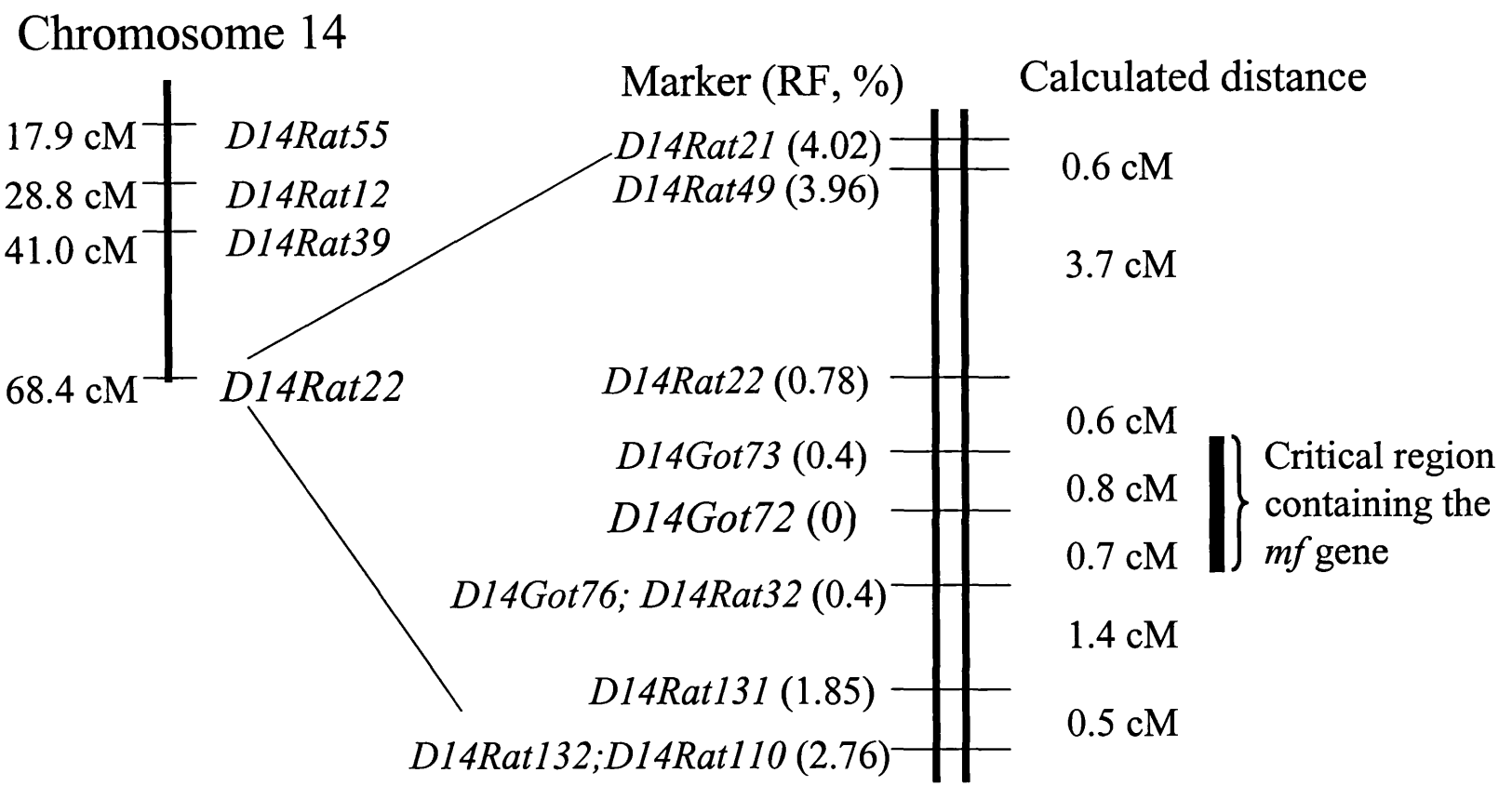


Fig. 3.2. Genetic localization of the *mf* locus and its associated flanking markers. (RF = recombination frequency)

These results indicated that marker *D14Got72* lies within 0.6 cM of the *mf* locus at the 95% confidence interval. This would place the loci within a megabase or so of one another. We were unable to further refine the linked region due to a lack of available markers between *D14Got72* and both flanking markers.

3.3 Identification of the responsible gene

3.3.1 Selecting BAC clones

The entire rat genomic library (RPCIB657) at ten-fold coverage is contained in 78 tertiary BAC DNA pools (from PP1 to PP78). The strategy for pooling the BAC DNA was as follows: 1) the individual clones were identified in the primary plate by a specific row and column number, 2) there were 6 columns in a secondary pool plate; the 8 wells in the first column represented the plate number, the 2nd and 3rd columns (16 wells) represented the row number and the rest of the columns (24 wells) represented the column number of the clone, 3) one well in the tertiary pool represented the pooling of all wells of a secondary pools. We screened the tertiary BAC genomic library with primers from the four markers *D14Got72, D14Got73, D14Got76* and *D14Rat32*. The positive tertiary pools are listed in Table 3.3. We identified eight tertiary BAC pools for *D14Got73*, seven for *D14Got72*, nine for *D14Got76* and six for *D14Rat32*. The frequency of positive pools is well within expectation for a ten-fold genome coverage.

Table 3.3. Positive signals in tertiary pool DNA screened by the linked markers. The tertiary pool DNA consisted of the rat genomic DNA library.

Markers	*D14Got73*	*D14Got72*	*D14Got76*	*D14Rat32*
	PP-7	PP-6	PP-6	PP-6
	PP-10	PP-8	PP-8	PP-8
	PP-12	PP-21	PP-20	PP-16
	PP-13	PP-41	PP-27	PP-35
	PP-29	PP-45	PP-33	PP-39
	PP-34	PP-52	PP-38	PP-68
	PP-48	PP-71	PP-45	
	PP-78		PP-48	
			PP-71	

Three of the positive tertiary pools were chosen from each marker. The positive tertiary pools were subjected to a second round of screening to confirm the initial result. The appropriate secondary pools were ordered from the German Resource Centre and screened to identify the primary plates containing the positive clone(s). The positive signal amplified from each secondary pooled plate is shown in Table 3.4. For example, three positive signals were picked up from the secondary pool plate (PP12) for marker *D14Got73*. The first positive signal (P89) represents the plate number containing the specific clone. The second (PP12-4) and third (PP12-19) positive signals represent the row and column number of this clone, respectively. The plate, row and column numbers determine the location of the BAC clones, which are shown in Table 3.5. Because of overlapping coverage on the rat genome, we only picked up two clones for *D14Got73* (RCPIB657L1298, RPCIB657D02381) and *D14Rat32* (RPCIB657C03312, RPCIB657K13541) and three clones for *D14Got72*

(RPCIB657O1132, RPCIB657O13359, RPCIB657B21324) and *D14Got76* (RPCIB657H21157, RPCIB657O1062, RPCIB657F21359).

Table 3.4. Positive signals shown in plates of secondary DNA pool screened by primers of the listed markers. (No. = number)

Makers	Secondary pool plates	Plate No.	Row No.	Column No.
D14Got73	PP12	P89	PP12-4	PP12-19
	PP13	P98	PP13-12	PP13-28
	PP48	P381	PP48-4	PP48-18
D14Got72	PP8	P62	PP8-15	PP8-27
	PP41	P324	PP41-2	PP41-37
	PP45	P359	PP45-15	PP45-29
D14Got76	PP8	P62	PP8-15	PP8-26
	PP20	P157	PP20-8	PP20-27
	PP45	P359	PP45-6	PP45-37
D14Rat32	PP16	P128	PP16-9	PP16-35
	PP39	P312	PP39-3	PP39-19
	PP68	P441	PP68-11	PP68-29

Table 3.5. BAC clones chosen from the linked markers

Markers			
D14Got73	RPCIB657L1298	RPCIB657D02381	
D14Got72	RPCIB657O1162	RPCIB657O13359	RPCIB657B21324
D14Got76	RPCIB657H21157	RPCIB657O1062	RPCIB657F21359
D14Rat32	RPCIB657C03312	RPCIB657K13541	

3.3.2 Anchoring linked rat segment to the mouse genome

We used the unique sequences of the closely linked rat markers (*D14Rat138, D14Rat139, D14Rat130, D14Rat95, D14Rat131, D14Rat132, D14Rat21, D14Rat34, D14Rat22, and D14Got73*) and the gene, *Otx1*, to BLAST the mouse and human genome sequence databases to establish whether the rat sequences had sequence identity with these two genomes for which there is complete sequence data available. Although the unique sequences associated with the closely linked microsatellite repeat loci failed to identify significant homology with sequences in the human genome, this was not true of the mouse. Significantly, those loci with sequence homology in the mouse genome all mapped to the same region of mouse chromosome 11 (Table 3.6). The BLAST results shown in Table 3.6 demonstrate the conservation of sequence between rat and mouse's genomic sequence of the markers. The number and percentage of identity between nucleotides have been presented. For example, with *D14Rat34*, there were two fragments of sequence used from the rat; one had 66 and the other 81 nucleotides. The BLAST results showed that the first fragment had 56 identical nucleotides (56/66, 84%) with mouse scaffold GA_x5J8B7W4T6T, which was located on mouse chromosome 11. The expected probability of this happening by random was only 0.072 (Probability Expect). According to the sequence homology, the *mf* locus was located at a 3.5Mb region on mouse chromosome 11 defined by markers *D14Rat22* and *D14Rat 131*. This region is also syntenic with human chromosome 2p15 (Fig. 3.3).

Table 3.6 BLAST result summary identifying regions of significant sequence homology between the microsatellite sequence on rat chromosome 14 and mouse chromosome 11 (nt. = nucleotide; X..Y = from X to Y nucleotide)

Rat marker (nt.)	BLAST Query fragment (nt.)	Nucleotide identity (%)	Event probability	Scaffold in Mouse chromosome 11 from Celera Database
D14Rat34 *(505)*	1. 47..112 (66)	56/66 (84)	0.072	GA_x5J8B7W4T6T:
	2. 239..319 (81)	71/81 (87)	2e-14	1..500000
D14Rat130 (670)	3. 8..161 (154)	135/154 (87)	1e-32	GA_x5J8B7W4498:
	4. 314..372 (59)	52/59 (88)	1e-04	500001..1000000
D14Rat21 (622)	1. 11..302 (292)	268/300 (89)	4e-39	GA_x5J8B7W4498: 1000001..1500000
D14Rat95 (551)	1. 85..218 (134)	118/139 (84)	2e-18	GA_x5J8B7W3EF1:
	2. 302..416 (115)	100/115 (86)	4e-19	1..500000
D14Rat139 (593)	1. 2..46 (45)	41/45 (91)	6e-06	GA_x5J8B7W3EF1:
	2. 74..174 (101)	83/101 (82)	9e-05	1500001..1989279
	3. 232..252 (21)	21/21 (100)	0.33	
	4. 250..311 (62)	54/62 (87)	1e-06	
D14Rat138 (840)	1. 6..26 (21)	21/21 (100)	0.48	GA_x5J8B7W3EF1:
	2. 33..202 (170)	145/170 (85)	2e-25	1..500000
	3. 537..575 (39)	35/39 (89)	0.008	
D14Rat22 (621)	1. 26..66 (41)	41/41 (100)	4e-13	GA_x5J8B7W498Y:
	2. 64..154 (91)	85/91 (93)	1e-28	500001..1000000
	3. 250..341 (92)	80/93 (86)	1e-06	
Otx1 (1068)	1. 2..139 (138)	136/138 (98)	6e-66	GA_x5J8B7W498Y:
	2. 139..293 (155)	153/155 (98)	5e-76	1000001..1500000
	3. 290..863 (574)	549/574 (95)	0	
	4. 951..1168 (218)	210/218 (96)	2e-99	
D14Got73 (351)	1. 9..96 (88)	78/88 (88)	1e-18	GA_x5J8B7W498Y:
	2. 12..43 (32)	31/32 (96)	4e-06	1000001..1500000
	3. 92..130 (39)	36/40 (90)	0.26	
RPCIB657B21324 -T7 (381)	1. 155..203 (49)	43/49 (87)	9e-04	GA_x5J8B7W498Y:
	2. 88..131 (44)	39/44 (88)	0.003	2000001..2500000
	3. 32..56 (25)	24/25 (96)	0.21	
RPCIB657O13359 -Sp6 (439)	1. 106..218 (113)	106/113 (93)	2e-39	GA_x5J8B7W498Y:
	2. 272..435 (164)	135/166 (81)	4e-12	2000001..2500000
RPCIB657B21324 -Sp6 (331)	1. 140..256 (117)	102/120 (85)	2e-16	GA_x5J8B7W498Y: 2000001..2500000
RPCIB657O13359 -T7 (439)	1. 40..332 (293)	255/300 (85)	3e-50	GA_x5J8B7W498Y:
	2. 403..439 (37)	33/37 (89)	0.25	2000001..2500000
D14Rat131 (906)	1. 36..75 (40)	37/40 (92)	3e-05	GA_x5J8B7W4TBK: 1000001..1500000
D14Rat132 (579)	1. 193..318 (126)	112/126 (88)	2e-30	GA_x5J8B7W4TBK: 1000001..1500000

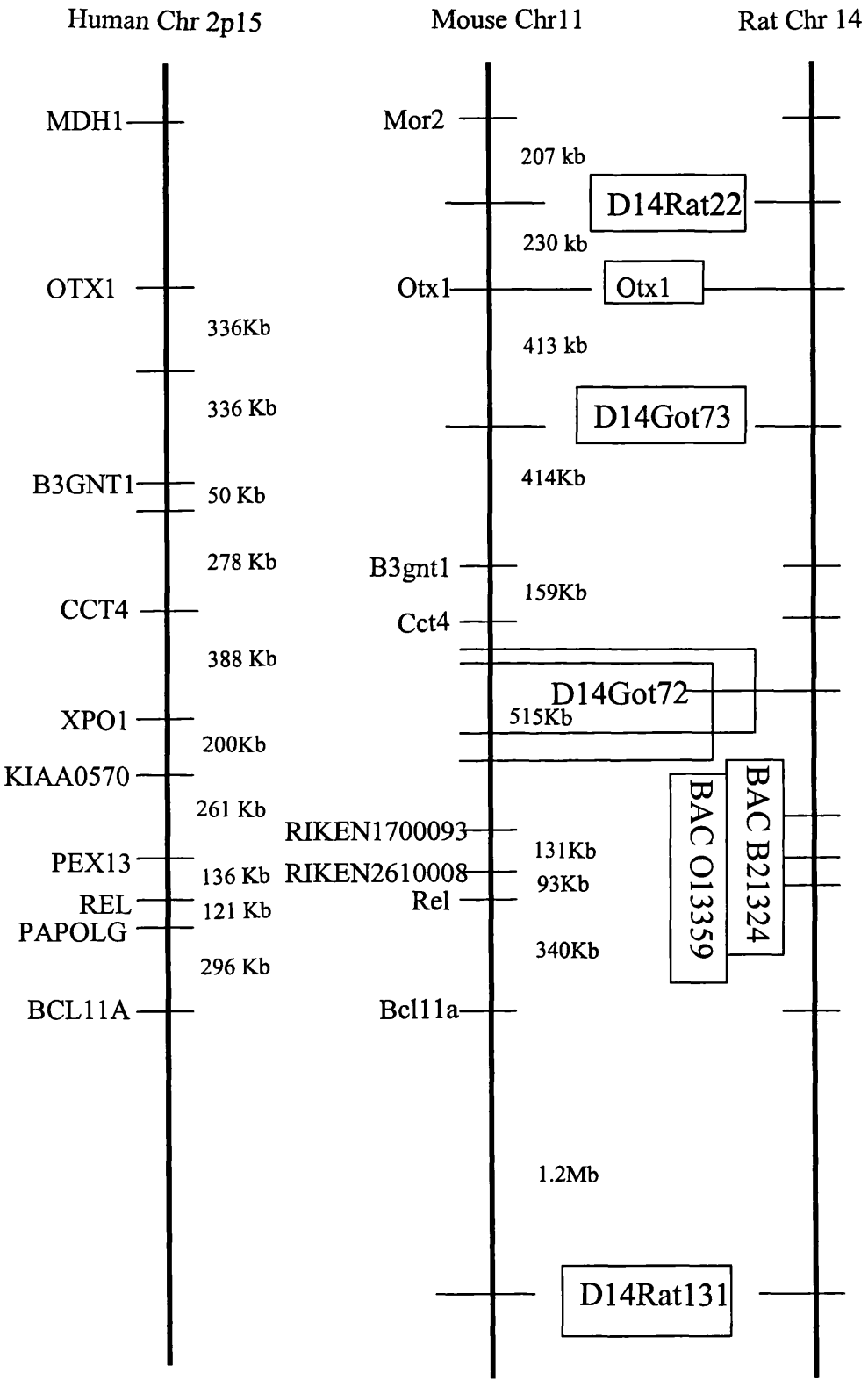


Fig. 3.3. Comparative genetic map for rat/ mouse / human at the *mf* locus based upon Celera database information. (Mb = Megabases; Kb = Kilobases)

As *D14Got73* and *D14Rat131* flank the *mf* locus and, if we assume that there is no disruption of synteny, the gene should fall within this 2.9Mb of mouse sequence.

Whilst sequence homology was detected with some of the Rat markers, the most important locus (*D14Got72*) (because of its tight linkage with the *mf* locus) failed to identify any significant homology in the mouse genome. However, a BLAST search with the rat BAC end sequences containing the *D14Got72* locus, did find significant homology in the mouse genome. Furthermore, the homologous sequence fell within the interval defined by *D14Got73* and *D14Rat131*. Similar results were obtained using sequence from both ends of the BAC (sequences in appendix IV) as well as sequence from two independent clones (RPCIB657O13359 and RPCIB657B21324) (Table 3.6, 3.7 and Fig. 3.3). The detailed BLAST reports for the rat markers and BAC end sequences are presented in Appendix III and IV, respectively. The position of the BAC ends helped define the length of the two BACs, which were 139.5kb and 152.8kb for RPCIB657O13359 and RPCIB657B21324, respectively. These findings were compatible with the rat BAC length suggested by RZPD. As shown in Table 3.7, the distance of the interval between *D14Got73* and *D14Rat131* is ~3Mb in the mouse genome, which is a distance highly consistent with the calculated distance in the rat linkage map. Therefore, we postulated that cross species synteny was probably maintained at this region. Gene identification in this region, presented in Table 3.7, was based upon gene prediction programmes such as GeneScan (http://genes.mit.edu/GENESCAN. html) and confirmed by the demonstration of homology to known genes or partial cDNA trascripts.

Table 3.7. Relative position of the rat microsatellite marker and the end of BAC sequences on the mouse gene map of chromosome 11.

Gene transcript	Gene Name	Symbol	Sequence Begins	Sequence Ends	Protein family assignment
		D14Rat34	498,405	498,040	
		D14Rat130	13,848,115	13,847,804	
		D14Rat21	14,728,352	15,158,126	
		D14Rat95	15,863,741	15,949,270	
mCT20790			15,791,905	15,902,376	TALE/MEIS, TALE/KNOX HOMEOBOX PROTEIN
		D14Rat139	15,932,535	16,087,063	
		D14Rat138	16,236,470	16,236,486	
mCT20791			16,568,723	16,569,909	GLUTAMINE SYNTHETASE
mCT20789			16,637,114	16,638,956	POL POLYPROTEIN
mCT11230			16,910,421	16,911,239	ACTIN
mCT11231			16,925,836	16,932,539	ACTIN
mCT11229	RAB1, member RAS oncogene family	*Rab1*	17,036,262	17,060,060	RAS SUPERFAMILY
mCT49159			17,124,363	17,125,638	RIBOSOMAL PROTEIN L11P
mCT11239			17,136,515	17,166,419	AMINO ACID TRANSPORTER-RELATED
mCT49160			17,168,955	17,179,743	GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE
mCT11233			17,285,571	17,286,005	RIBOSOMAL PROTEIN S17E
mCT49869			17,409,243	17,409,732	RIBOSOMAL PROTEIN L15E
mCT11237			17,658,406	17,663,048	GALECTIN
mCT2361	RIKEN cDNA 2810468L03 gene	*2810468L03 Rik*	18,035,295	18,049,820	
mCT56220			18,049,966	18,089,822	RETROVIRUS-RELATED GAG POLYPROTEIN-RELATED
mCT2360			18,219,408	18,268,649	UTP--GLUCOSE-1-PHOSPHATE URIDYLYLTRANSFERASE-RELATED
mCT58614			18,366,453	18,433,819	RETROVIRUS-RELATED POLYPROTEIN
mCT61608			18,402,822	18,403,038	SER/THR FAMILY OF PROTEIN KINASES-RELATED
mCT2366	Malate dehydrogenase, soluble	*Mor2*	18,452,113	18,466,687	LACTATE DEHYDROGENASE-RELATED
mCT2370			18,577,841	18,579,551	PHOSPHOGLYCERATE MUTASE
		D14Rat22	18,659,064	18,659,343	
mCT54273			18,768,915	18,771,027	POL POLYPROTEIN

mCT54212			18,845,662	18,846,013	RIBOSOMAL PROTEIN L15E
mCT8679			18,888,246	18,893,802	ANTP/PBX HOMEOBOX PROTEIN
		OTX1	18,889,671	18,891,992	
mCT8684			18,899,176	18,988,373	MYOSIN HEAVY CHAIN-RELATED
mCT8687			18,992,779	19,178,139	ACTIN-CROSSLINKING PROTEINS
mCT8680			19,230,293	19,231,973	RETROVIRUS-RELATED GAG POLYPROTEIN-RELATED
mCT8685			19,278,172	19,278,671	RIBOSOMAL PROTEIN S2P
		D14Got73	19,302,928	19,302,862	
mCT8686			19,632,163	19,659,929	METHIONINE AMINOPEPTIDASE
mCT8682	beta-1,3-N-acetylglucosaminyltransferase 1	*Beta3gnt*	19,717,383	19,746,052	BETA-1,3-N-ACETYLGLUCOSAMINYLTRANSFERASE-RELATED
mCT8875	Chaperonin subunit 4 (delta)	*Cct4*	19,876,740	19,889,485	HSP60 OR TCP-1 CHAPERONIN FAMILY MEMBER
		B21324-T7	19,953,912	19,953,982	
		O13359-Sp6	19,975,452	19,975,792	
		B21324-Sp6	20,106,602	20,106,721	
		O13359-T7	20,114,537	20,114,968	
mCT8876			20,137,483	20,172,390	CHROMOSOME REGION MAINTENANCE PROTEIN 1-RELATED
mCT8877			20,196,793	20,281,444	LAMININ-RELATED
mCT8884			20,314,364	20,319,945	UBIQUITIN CARBOXYL-TERMINAL HYDROLASE
mCT8873	RIKEN cDNA 1700093K21 gene	*1700093K21 Rik*	20,392,031	20,394,885	
mCT8878	RIKEN cDNA 2610008O20 gene	*2610008O20 Rik*	20,523,710	20,541,047	TYROSINE KINASE RECEPTOR-RELATED
mCT8880	Reticuloendotheliosis oncogene	*Rel*	20,616,770	20,645,765	REL/DORSAL-RELATED
mCT8887			20,736,976	20,768,316	POLY(A) POLYMERASE
mCT11858	B-cell CLL/lymphoma 11A (zinc finger proten)	*Bcl11a*	20,956,876	21,043,527	KRUEPPEL-SUBFAMILY C2H2-TYPE ZINC-FINGER PROTEIN
		D14Rat131	22,181,747	22,441,386	
		D14Rat132	23,118,309	23,414,124	

B21324 = Rat BAC, RPCIB657B21324

O13359 = Rat BAC, RPCIB657O13359

3.3.3 Candidate gene screening

The rat-mouse and rat-human comparative genetic maps are available from the website http://ratmap.ims.u-tokyo.ac.jp/cgi-bin/Mapview_rat.pl?RNO14. This comprehensive comparative map from rat, mouse, and human, is based on the extensive radiation hybrid-mapping data of rat genes. The comparative map shows that the region with markers *D14Rat22, D14Got72* and *D14Rat32* is syntenic to mouse chromosome 11 and human chromosome 2p13-16. A BLAST search of the species genome sequence databases clearly substantiates this linkage conservation between the three species. Furthermore, the co-linearity of real or predicted genes between mouse and human (Table 3.7 and Fig. 3.3), suggests that the order is likely to be maintained in the rat gemone also. Thus, the genome data from both species can be used to predict the likely genes in the homologous region of rat chromosome 14. Of particular interest are those genes predicted or known to exist within the sequences flanked by *D14Got73* and *D14Rat131*. In addition to the gene location, the expression pattern and the function of the genes were also been used as criteria for selecting a candidate gene.

3.3.3.1 *OTX1*

The gene *Otx1*, which is a homeobox-containing gene similar to orthodenticle (*otd*) in Drosophila, is located at the linked region. This gene plays an important

role in early patterning of the cephalic development (rostral brain) and the mechanism is conserved from Drosophila to mouse [Finkelstein and Boncinelli 1994]. The whole coding region (exons 1~3) was sequenced. There were no nucleotide changes detected among the Lewis, normal SD and *mf* rats.

3.3.3.2 *Kiaa0127*

On the 4th June 2001, more detailed comparative maps were released from the rat genome database (http://rgd.mcw.edu/VCMAP/). These maps show the chromosome of the backbone species (e.g. the rat) plus the syntenic regions from the two non-backbone species (e.g. the mouse and the human). The maps allow us to obtain the information of the mapped markers (microsatellite as well as radiation hybrid markers) and the expression sequence tag (EST) clusters in a specific syntenic chromosomal regions among different species. The tightly linked marker *D14Got72* is located at a distance of 834.90 centi-Ray (cR) of rat chromosome 14 in the radiation hybrid map. An EST cluster at 836.90 cR called Rn.15977 had a syntenic human gene, *KIAA0127*. The EST syntenic to KIAA0127 is the closest gene to marker *D14Got72*. The coding region of *Kiaa0127* of the SD rat was sequenced, but we did not find any nucleotide changes between normal SD and *mf* rats.

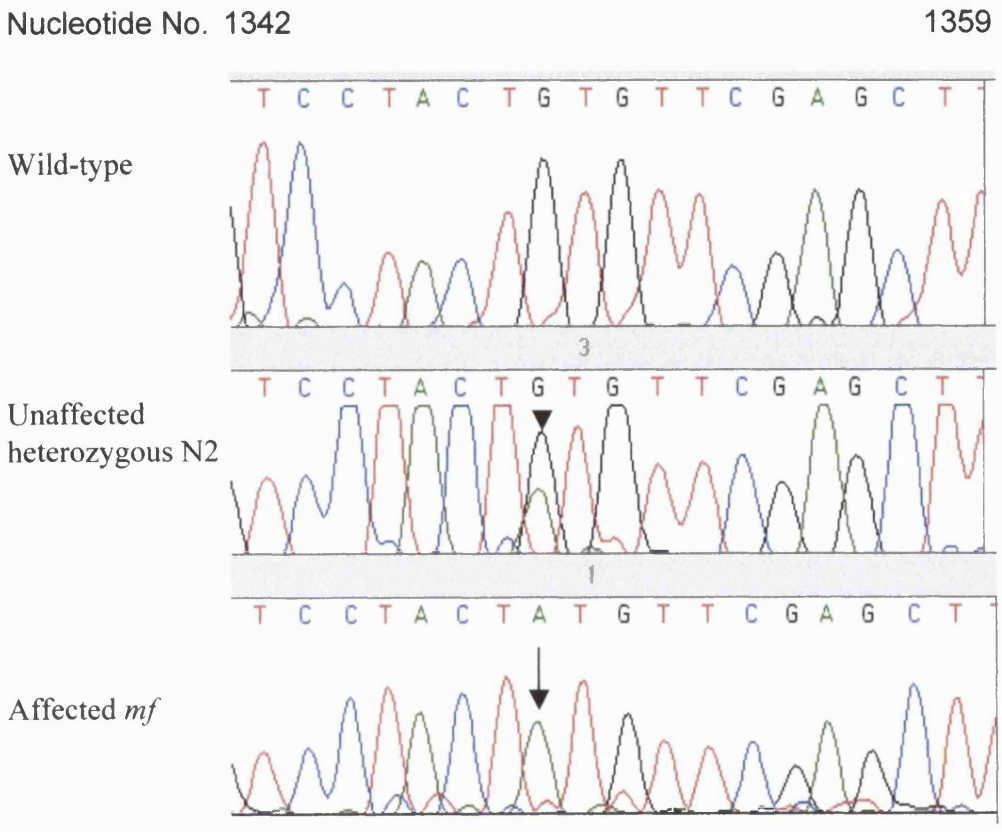
3.3.3.3 *Bcl11a*

The gene *Bcl11A*, which is located at the mouse syntenic region, is known to encode a zinc-finger protein which plays an important role in the differentiation of myeloid cells during hematopoiesis [Martin-Subero *et al.* 2002]. Bcl11A is one member of the Bcl11 protein family. It has been suggested that REL (reticuloendotheliosis protein), a NF-kappaB gene family member, mapping within the same region on mouse chromosome 11, is the pathologic target for the Bcl11 protein [Satterwhite *et al.* 2001]. The REL protein is important in the process of neuronal apoptosis during development. The gene has been screened by direct sequencing of the cDNA and the untranslated region. There was no nucleotide change in the *mf* rat as compared to the normal SD rat.

3.3.4. Identifying the gene

In the anchored comparative map (Fig. 3.3), the gene *Cct4* (Chaperonin containing TCP-1, delta subunit) is one of the closest genes to marker *D14Got72*. This gene had previously been cloned and mapped to mouse chromosome 11 [Nabetani *et al.* 1996]. Both the genetic proximity of *Cct4* to the *D14Got72* locus and the function of this gene justify the screening of *Cct4* as a candidate gene. Tcp-1 (tailless complex polypeptide-1) and related proteins facilitate the folding of cytosolic proteins, especially actin and tubulin. Both these proteins are main components of the cytoskeleton and play important role in axonal transport [Llorca *et al.* 1999] [Hynes and Willison 2000].

The coding region of *Cct4* was screened using mRNA extracted from a 16-day *mf* fetus. The primers were designed from the mouse *Cct4* cDNA sequence, which was divided into five fragments for sequencing the rat *Cct4*. The sequence from fragment 5 of the *Cct4* cDNA from the *mf* rat revealed a "c.1349G>A" transition (the nucleotide sequence of rat *Cct4* is presented in Appendix V). The transition causes a Cysteine450 to Tyrosine change (p.C450R) in the protein sequence (Fig. 3.4). The amino acid sequences of the rat and mouse Cct4 proteins are 98% identical. There is 96% homology between the rat and the human CCT4 protein (Fig. 3.5). Furthermore, the amino acid Cysteine450 is highly conserved among different species, from Neurospora, Leshimanina, fish, birds, mouse, monkey to human (Fig. 3.6). Sequence data obtained from 25 affected and 25 unaffected rats confirmed the mutation: all affected animals were homozygous "A" at nucleotide 1349 whereas the normal N2 animals were A/G heterozygotes. Four N2 rats had shown recombination at two flanking markers (Rat204 and Rat483 showed recombination at *D14Got73*, and Rat 206 and Rat431 at *D14Got76*). Clinically, Rat431 was normal but the other three are affected. *Cct4* sequencing in these four rats showed that Rat 431 is heterozygous "G/A" and the others are homozygous "A/A" at nucleotide position 1349 (Fig. 3.7).



Wild-type	Ser	Tyr	Cys	Val	Arg
	TCC	TAC	TGT	GTT	CGA
			↓		
mf	TCC	TAC	TAT	GTT	CGA
	Ser	Tyr	Tyr	Val	Arg

Fig. 3.4. The sequence of *Cct4* for wild-type (+/+), heterozygotes N2 (+/*mf*) and mutant homozygous (*mf/mf*) rat. No. = number

```
CLUSTAL W (1.81) multiple sequence alignment

RatCct4         MPENVASRSGPPAAGPGNRGKGAYQDRDKPAQIRFSNISAAKAVADAIRTSLGPKGMDKM 60
MouseCct4       MPENVASRSGAPTAGPGSRGKSAYQDRDKPAQIRFSNISAAKAVADAIRTSLGPKGMDKM 60
HumanCCT4       MPENVAPRSGATAGAAGGRGKGAYQDRDKPAQIRFSNISAAKAVADAIRTSLGPKGMDKM 60
                ******.***..:...*.***.****************************************

RatCct4         IQDGKGDVTITNDGATILKQMQVLHPAARMLVELSKAQDIEAGDGTTSVVIIAGSLLDSC 120
MouseCct4       IQDGKGDVTITNDGATILKQMQVLHPAARMLVELSKAQDIEAGDGTTSVVIIAGSLLDSC 120
HumanCCT4       IQDGKGDVTITNDGATILKQMQVLHPAARMLVELSKAQDIEAGDGTTSVVIIAGSLLDSC 120
                ************************************************************

RatCct4         TKLLQKGIHPTIISESFQKALEKGLEILTDMSRPVQLSDRETLLNSATTSLNSKVVSQYS 180
MouseCct4       TKLLQKGIHPTIISESFQKALEKGLEILTDMSRPVQLSDRETLLNSATTSLNSKVVSQYS 180
HumanCCT4       TKLLQKGIHPTIISESFQKALEKGIEILTDMSRPVELSDRETLLNSATTSLNSKVVSQYS 180
                ***********************:**********:***********************

RatCct4         SLLSPMSVNAVMKVIDPATATSVDLRDIKIVKKLGGTIDDCELVEGLVLTQKVANSGITR 240
MouseCct4       SLLSPMSVNAVMKVIDPATATSVDLRDIKIVKKLGGTIDDCELVEGLVLTQKVANSGITR 240
HumanCCT4       SLLSPMSVNAVMKVIDPATATSVDLRDIKIVKKLGGTIDDCELVEGLVLTQKVSNSGITR 240
                ************************************************:***********

RatCct4         VEKAKIGLIQFCLSAPKTDMDNQIVVSDYAQMDRVLREERAYILNLVKQIKKTGCNVLLI 300
MouseCct4       VEKAKIGLIQFCLSAPKTDMDNQIVVSDYAQMDRVLREERAYILNLVKQIKKTGCNVLLI 300
HumanCCT4       VEKAKIGLIQFCLSAPKTDMDNQIVVSDYAQMDRVLREERAYILNLVKQIKKTGCNVLLI 300
                ************************************************************

RatCct4         QKSILRDALSDLALHFLNKMKIMVVKDIEREDIEFICKTIGTKPVAHIDQFTPDMLGSAE 360
MouseCct4       QKSILRDALSDLALHFLNKMKIMVVKDVEREDIEFICKTIGTKPVAHIDQFTADMLGSAE 360
HumanCCT4       QKSILRDALSDLALHFLNKMKIMVIKDIEREDIEFICKTIGTKPVAHIDQFTADMLGSAE 360
                ************************:**:************************.*******

RatCct4         LAEEVSLNGSGKLFKITGCTSPGKTVTIVVRGSNKLVIEEAERSIHDALCVIRCLVKKRA 420
MouseCct4       LAEEVSLNGSGKLFKITGCTSPGKTVTIVVRGSNKLVIEEAERSVHDALCVIRCLVKKRA 420
HumanCCT4       LAEEVNLNGSGKLLKITGCASPGKTVTIVVRGSNKLVIEEAERSIHDALCVIRCLVKKRA 420
                *****.*******:*****:************************:***************

RatCct4         LIAGGGAPEIELALRLTEYSRTLSGMESYCVRAFADAMEVIPSTLAENAGLNPISTVTEL 480
MouseCct4       LIAGGGAPEIELALRLTEYSRTPSGMESYCVRRFADAMEVIPSTLAENAGLNPISTVTEL 480
HumanCCT4       LIAGGGAPEIELALALTEYSRTLSGMESYCVRAFADAMEVIPSTLAENAGLNPISTVTEL 480
                ************** ******* ********* ****************************

RatCct4         RNRHAQGEKTTGINVRKGGISNILEEMVVQPLLVSVSALTLATETVRSILKIDDVVNTR 539
MouseCct4       RNRHAQGEKTTGINVRKGGISNILEEMVVQPLLVSVSALTLATETVRSILKIDDVVNTR 539
HumanCCT4       RNRHAQGEKTAGINVRKGGISNILEELVVQPLLVSVSALTLATETVRSILKIDDVVNTR 539
                **********:***************:**********************************
```

Fig. 3.5. Protein sequence of Cct4 from rat, mouse and human.

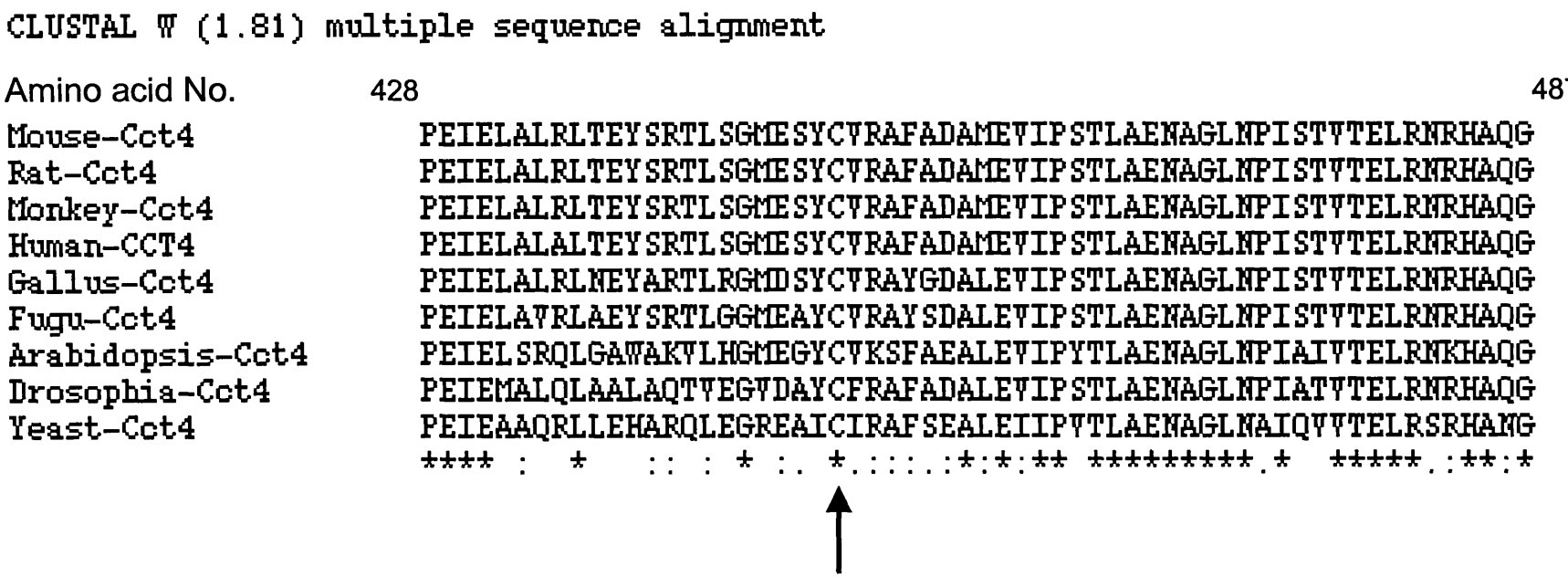


Fig. 3.6. Predicted Cct4 peptide sequence displaying the highly conserved cysteine at amino acid 450. (* = conserved amino acid in evolutionary divergent species.)

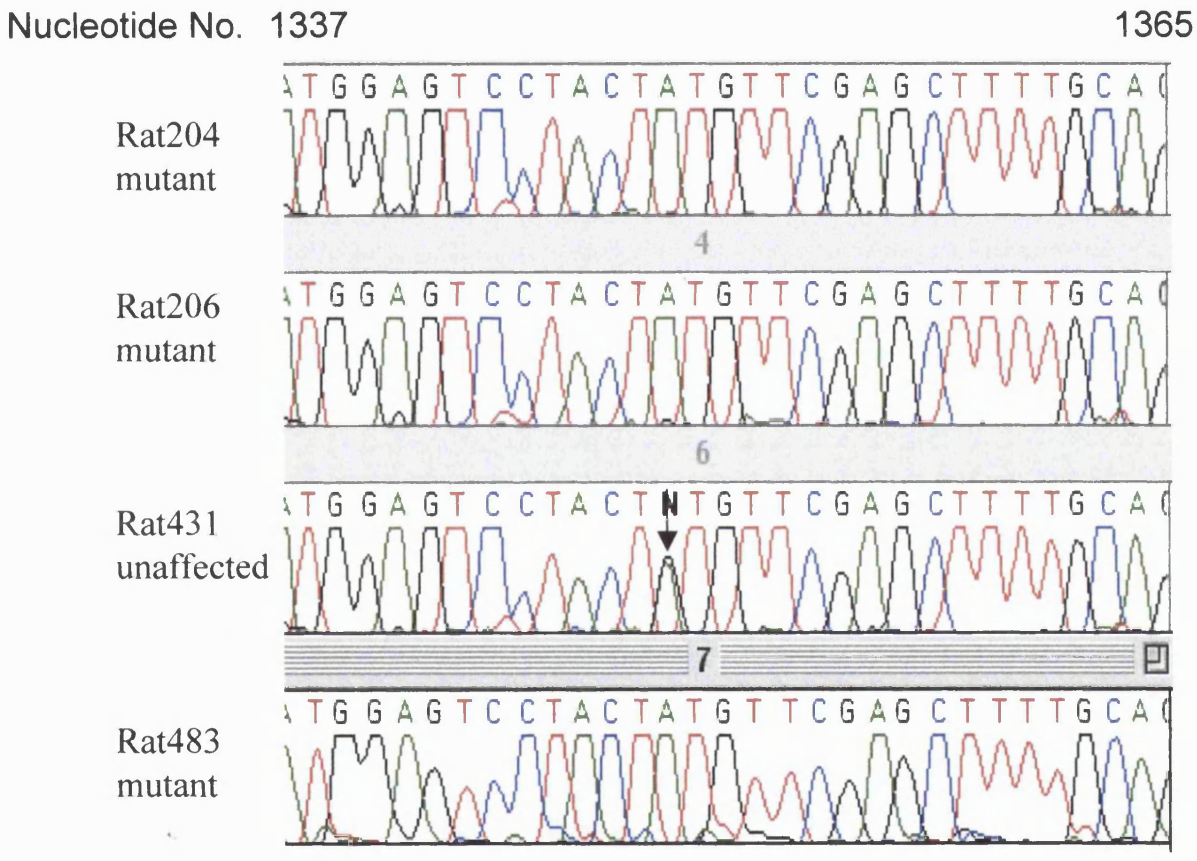


Fig. 3.7. *Cct4* sequence from N2 rats showing recombination at closely linked loci flanking the *mf* locus. No. = number

To provide further evidence that this mutation is indeed pathogenic and not a polymorphism, we screened *Cct4* in eight different rat strains. All eight strains were homozygous "G/G" at nucleotide position 1349 (Fig. 3.8). Thus there is overwhelming evidence that *Cct4* is the disease-causing gene for *mf* disease.

3.3.5 The expression of *Cct4*

In order to investigate the expression of this gene in rat tissue, a Northern slot blot analysis and RT-PCR were performed. The results from both experiments showed that the expression of *Cct4* mRNA is ubiquitous in several rat tissues (Fig. 3.9 & Fig. 3.10).

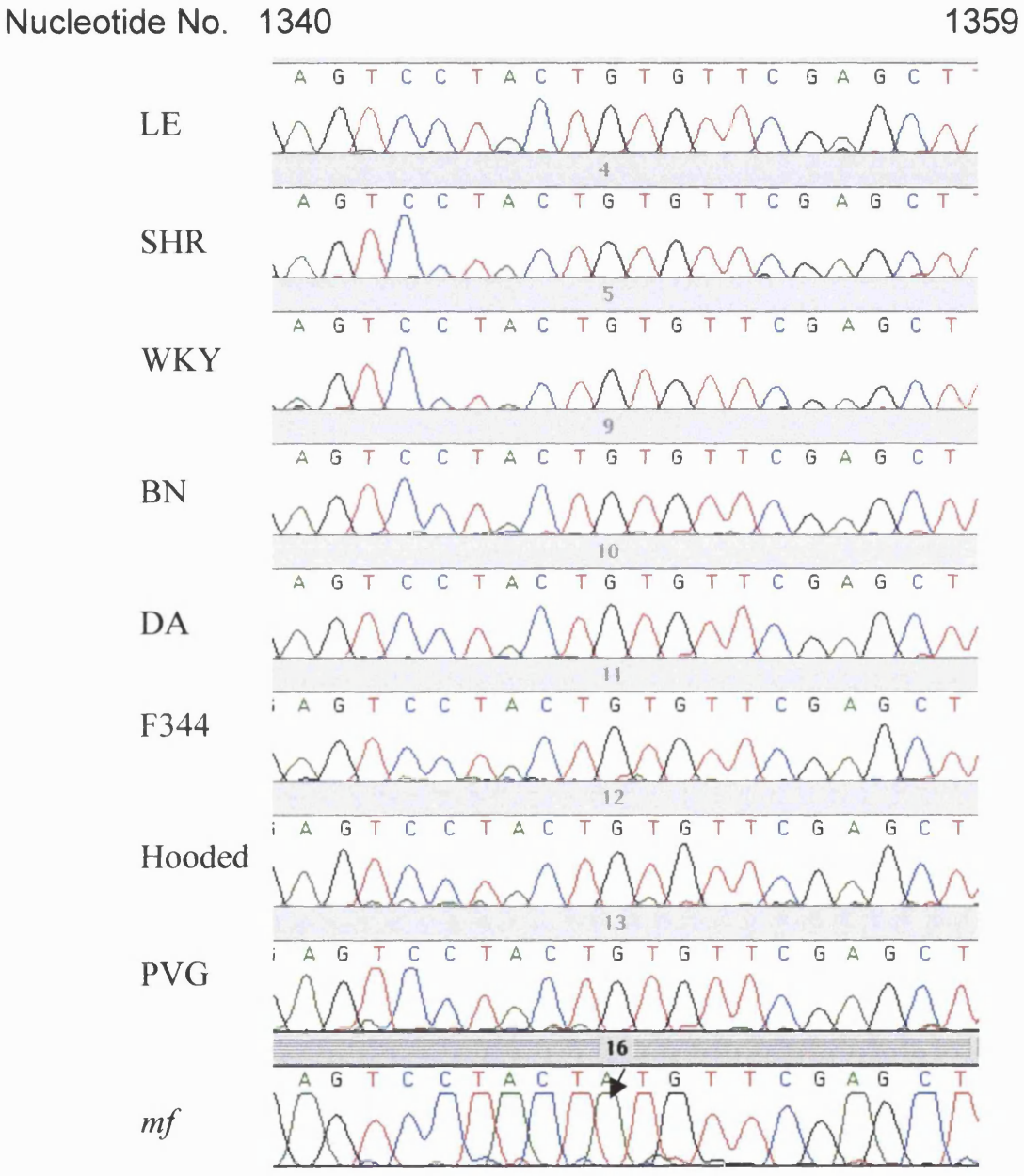


Fig. 3.8. Nucleotide sequence of *Cct4* from 8 different rat strains. Only the *mf* rat shows a homozygous "A/A" transition from "G/G". No. = number

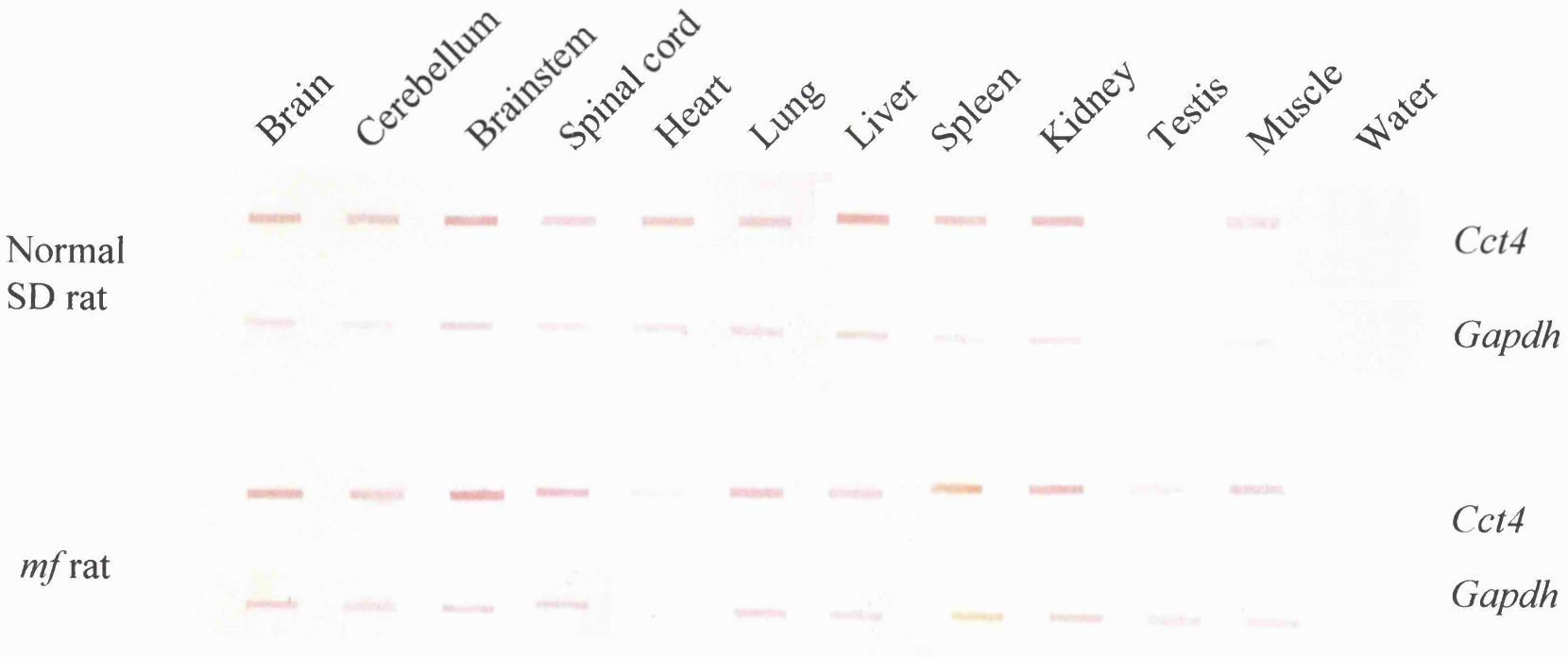


Fig. 3.9 Tissue expression of *Cct4* and *Gapdh* in normal and *mf* rats.

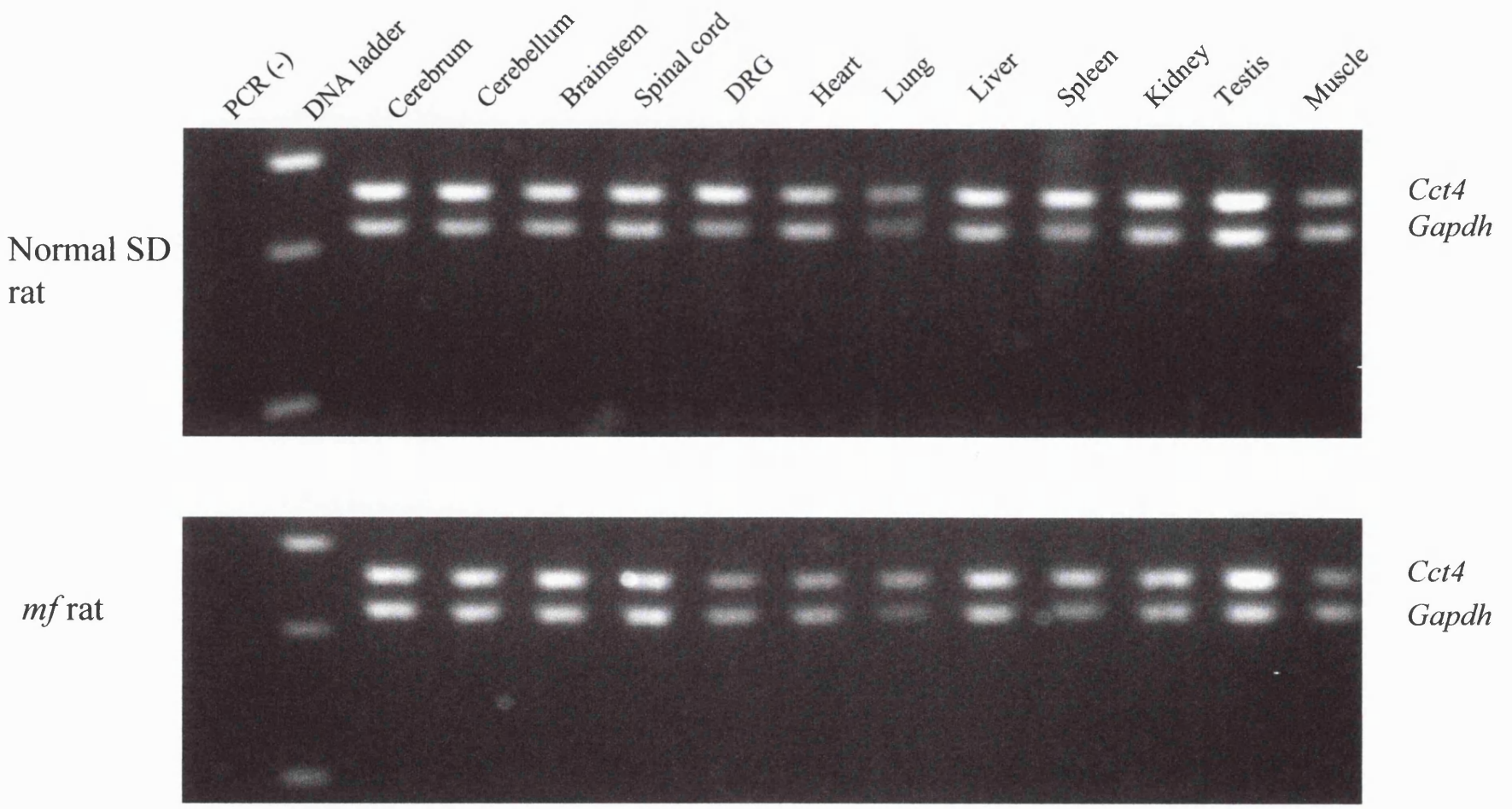


Fig. 3.10. The amplified fragments of *Cct4* (362 bp) and *Gapdh* (321bp) from different tissues of normal SD and *mf* rat by RT-PCR.

Chapter 4

Discussion: Rat genetic study

4.1 Summary of the rat project

Mutilated foot (*mf*) disesase in the rat is an autosomal recessive hereditary sensory neuropathy. Clinically symptoms of the disease include loss of all sensory modalities [Jacobs *et al.* 1981]. Pathological changes include loss of sensory neurons and axons in the dorsal root ganglia (DRG). In order to gain a better understanding of this animal model of human hereditary sensory autonomic neuropathy (HSAN), a genetic study was undertaken to identify the genetic defect associated with the mutation.

Two different approaches were used to identify the molecular defect associated with this disease (positional independent and dependent strategies). The positional independent approach involved the sequencing of candidate genes (*p75NTR* and *NTRK1*) based upon phenotypic similarities with known mutations in the mouse [Smeyne *et al.* 1994] [Lee *et al.* 1992]. This approach failed to provide any useful information as the sequence obtained showed no variation between wild type and mutant rat. The approach that ultimately proved successful was the positional dependent approach. Using backcross linkage analysis, the mutant locus was shown to map to a 1.5cM region at the distal end of rat chromosome 14. BLAST search sequence analysis involving linked and

flanking markers identified a specific syntenic region on mouse chromosome 11 (Table 3.6). The gene content and physical properties of this interval in the mouse were also conserved with that of human chromosome 2p15 (Fig. 3.4). Identification of significant BAC end sequence (containing the tightly linked marker) similarities with those on mouse chromosome 11, supports the assumption that gene order might be preserved in the rat as well. Based on this assumption, the rat genes predicted to lie in the interval based upon the conservation data from both mouse and human, were sequenced. This lead to the identification of specific nucleotide substitution in the sequence of the *Cct4* gene associated with the mutant genotype (Fig. 3.5). As a result of this substitution, a highly conserved cysteine was replaced by tyrosine, "p.C450R" of peptide sequence. Authenticity of this mutation was established by sequencing the same gene in 50 N2 rats, several different rat strains and the recombinant individuals involving the two flanking loci (*D14Got73, D14Got76* and *D14Rat32*) identified during the mapping studies. The expression studies by Northern slot blot and RT-PCR show that *Cct4* protein is ubiquitous expressed in rat which is different from which reported in the mouse [Kubota *et al.* 1999].

4.2 The function of Cct4 and its role in *mf* disease

4.2.1 Cct4 function

Cct4 gene encodes the delta subunit of the cytosolic chaperonin (CCT, chaperonin containing TCP-1 protein) complex. This complex, comprised of

eight subunits encoded by different genes, is responsible for ensuring that newly synthesised proteins assume their correct three-dimensional structure [Rommelaere *et al.* 1993]. A vitally important process, as mis-folded proteins will frequently form aggregates with significant pathological consequences to the cell [Hartl and Hayer-Hartl 2002].

Synthesis of chaperonins and chaperones are greatly increased when cells are placed under physiological stress conditions that result in the denaturation of native proteins. As heat is frequently used to induce the synthesis of these proteins, they are known collectively as heat shock proteins (Hsp) [Gething and Sambrook 1992]. Chaperones (Hsp70, Hsp40) and chaperonins (CCT) participate in *de novo* protein folding. This process involves cycles of substrate binding and release regulated by ATPase activity and other cofactors [Bukau and Horwich 1998] [Hartl and Hayer-Hartl 2002].

The chaperonins are conserved large double ring complexes enclosing a central cavity. They include two subgroups with similarities in architecture but distinct peptide sequence. Group I chaperonins are found in eubacteria and endosymbiotic organelles (such as mitochondria and chloroplasts) of which the GroEL complex from *E. coli* is the most extensively studied. Group II or CCT chaperonins are found in archaea and in the cytosol of eukaryotic cells [Willison 1999].

The CCT (Chaperonin containing tailless complex polypeptide 1, TCP-1, or TCP-1 ring complex, TriC) complex contains of eight orthologous subunits (α,

β, γ, δ, ε, ζ, θ and η, CCT1 through 8 respectively) per ring that differ primarily in their apical domains. Two rings stack upon one another to form a functional unit [Llorca *et al.* 1999] [Willison 1999].

The most abundant substrates of CCT are the cytoskeletal proteins, actin and tubulin. The CCT proteins from mouse, rabbit, and cow tissue can facilitate the folding of the newly synthesised actin and tubulin and their related proteins [Gao *et al.* 1992] [Yaffe *et al.* 1992] [Melki *et al.* 1993]. Actin interacts with two specific subunits (CCT4, CCT-delta and CCT2, CCT-beta) of the CCT complex before activation of the folding machinery [Llorca *et al.* 1999]. The mechanism for protein binding, folding and release of both chaperonin groups is ATP dependent. ATP binding induces encapsulation of the protein by the protruding apical-domains of the complex [Llorca *et al.* 2001]. However, this mechanism is not well understood. Tubulins interact with the subunits β and η suggesting that heterogeneity in the subunit content of the CCT complex may be an important factor in determining the diversity of protein that can be accommodated by the complex [Llorca *et al.* 2001].

4.2.2. Pathogenesis of *mf* disease

Pathology suggests that the main reason for sensory impairment in the rat is due to excessive neuronal loss in the DRG. This phenomenon occurs at late fetal to early post-natal stage. However, during development of the nervous system, more neurons are produced than are actually required for normal sensory/motor function. Thus, there is natural wastage of neuronal cells [Oppenheim 1991]. The time of natural neuronal loss in DRG occurs at the same developmental

stages as the excessive loss observed in the mutant (i.e., from E14~16 days to postnatal). Evidence suggest that cell survival is dependent upon target tissue contact [Johnson and Deckwerth 1993] [Gordon 1995]. If they fail to establish suitable contact, the neurons will die through the mechanism of programmed cell death (apoptosis).

Nerve growth factor was the first chemotrophic agent to be identified with specific effect on axonal growth and neuronal survival [Levi-Montalcini and Angeletti 1966]. It is secreted by tissues requiring neurological innovation that are themselves receptive to the growth factor (i.e., have a suitable receptor). They include sympathetic neurons and some neural crest-derived sensory neurons such as those of the DRG. *In vivo* treatment of the cervical DRG in developing chick embryos has been shown to sustain neuronal survival and reduce apoptosis [Geffen and Goldstein 1996].

Programmed cell death (or apoptosis) is an extremely complex process involving the interplay of a large number of effector/suppressor molecules [Hunot and Flavell 2001]. Some of the key elements, such as Caspases, in this process were original identified in *Caenorhabditis elegans* (*C. elegans*) because of the well defined cellular map that exist for this organism. The developmental cell death in this organism is controlled by a genetic pathway involving the sequential and coordinated action of EGL-1, CED-9, CED-4 and CED-3 [Horvitz 1999]. Ellis and co-workers have identified several genes in *C. elegans* that affect this process in specific cell linkages. Cell death was found in those that express a functional *ced-3* gene but not the non-functional *ced-3* gene. Mutations in *ced-3*

and *ced-4* cause survival of almost all cells that would normally die [Ellis and Horvitz 1986]. These results suggest that *ced-3* and *ced-4* genes encode proteins that act within the dying cells. Caspases, a family of 14 mammalian cysteine-containing, aspartate-specific protease, show structural and functional similarity with the CED protein families in *C. elegans*.

4.2.2.1 Actin and apoptosis

The mammalian homologue of the *C. elegans ced-3* gene is CPP32 (Caspase-3) [Xue *et al.* 1996]. This gene is specifically expressed in the trigeminal ganglia, facio-acoustic ganglion complex, and DRGs of mouse as early as E10.5 days. Caspase-3 like proteases are activated during apoptosis of DRG neurons in the absence of NGF and serum [Mukasa *et al.* 1997]. Mashima *et al.* showed that caspase-3 specifically cleaves the actin protein into two (15Kand 30K) fragments *in vitro* as well as *in vivo* [Mashima *et al.* 1995] [Mashima *et al.* 1997]. Actin is the major component of cytoskeleton. Cytoskeletal degradation has indeed been identified as an early event of apoptosis [Guenal *et al.* 1997] [van Engeland *et al.* 1997]. Disruption of micro-filaments by cytochalasins [Rubtsova *et al.* 1998] or depolymerization of actin by thymosine β10 [Hall 1995] will induce apoptotic events. This would suggest that disruption of the cytoskeleton could be a key factor in the initiation of apoptosis.

In addition to being one of the basic elements of cytoskeleton, actin is also an inhibitor of the endonucleolysis which is mediated by DNaseI [Kayalar *et al.* 1996]. The overexpression of Dnase I exhibited progressive nuclear destruction,

i.e., pynknosis, cytoplasmic shrinkage and internucleosomal DNA degradation which is typical for apoptosis [Polzar *et al.* 1993]. Thus, disruption in actin biogenesis and function could, in addition to regulating apoptotic events, mediate its effect through its interaction with DNaseI. This, in turn, could also impact on cell viability.

Nascent actin has been shown to bind with a specific heat shock protein (or chaperon) hsp27. This association is important during gastrulation; neural tube development and the migration of neural tube derived cell lineage's including those that give rise to the DRG [Walsh *et al.* 1997]. Although these early developmental events might not be wholly consistent with that of the phenotype associated with the rat mutation, other studies have shown that disruption of this association can lead to apoptotic events in the neural retina [Tezel and Wax 2000].

The missense mutation in *Cct4* of *mf* rat resulted in cysteine to tyrosine change. The free energy of the disulfide bond (S-S) bridging two cysteines, is much stronger than the hydrogen bond (S-OH) between cysteine and tyrosine. The distance between the two cysteines is less than the distance between a cysteine and a tyrosine. This may further destabilise the molecule in the absence of enhanced stability afforded by the formation of a disulphide bridge that could form between adjacent cysteines. Therefore, we can speculate that the mutation in *Cct4* will cause the distortion of the normal 3-D structure of this subunit as well as the whole CCT protein. As the actins are the major substrate of CCT and the apical domain of Cct4 is one of the binding sites for nascent actin protein

prior to folding, this change is predicted to reduce the affinity of actin to the CCT ring complex resulting in either degradation of actin or diminishing in its polymerisation. The aberration might also influence the binding of ATPase to CCT protein, which cleaves the ATP to provide the energy for protein folding. Therefore, the abnormalities in actin may eventually cause the destruction of the cytoskeleton and start the process of neuronal apoptosis.

4.2.2.2 Independent role in neuronal apoptosis

In addition to the role of the TCP-1 complex in protein folding, individual subunits can also act independently. As shown in the chick embryo sympathetic neurons, over-expression of TCP-1δ (CCT4) accelerated the neuronal apoptosis induced by dopamine treatment [Zilkha-Falb *et al.* 2000]. Other subunits did not produce this effect. The authors suggest that over-expression of CCT4 in sympathetic neurons accelerated dopamine induced neuronal apoptosis. The use of antisense oligonucleotide treatment to inhibit CCT4 expression significantly reduced dopamine induced neuronal death in sympathetic neurons. These results suggest that the CCT4 protein is probably a positive mediator in the process of neuronal apoptosis. This study therefore demonstrated that in the process of inducing neuronal death, CCT4 played a role unrelated to the whole CCT protein. Also this study supports other research demonstrating that Cct subunits can act independently. Wu-Baer *et al.* purified a cofactor SRB, also known as human CCT4, which worked with an elongation factor 1α and polypyrimidine tract binding protein to facilitate the binding of RNA polymerase II and TRP-185

to TAR RNA (in HIV-1) to increase gene expression [Wu-Baer *et al.* 1996]. Another study showed that the α, γ, ζ, and θ subunits can form smaller complexes than those of the CCT complex and act as a microtubule-associated protein [Roobol *et al.* 1999]. These studies suggest that either as an independent subunit or as a component of the TCP-1, the CCT4 protein may play an important role in the process of neuronal apoptosis, which is the pathological hallmark of *mf* disease.

4. 3 Conclusion

The pathological defect associated with the *mf* mutation suggests that the sensory neuropathy is due to a loss of neurons in DRG. At this stage, it is unclear whether this apoptosis is due to a developmental defect involving the 'wiring' of the peripheral nervous system mediated through a defect in actin. It may be that actin is either having a direct impact on apoptosis or an indirect effect because the axons are failing to make "survival contact". It is also possible that the apoptosis is due to a general problem with protein folding that is unrelated to actin. Alternatively, the excess apoptosis may be due to some independent function of the CCT4 protein itself. Whatever the mechanism, it is clear that a mutation in *Cct4* can have serious pathological/clinical consequences in rat and possibly humans as well. Clearly more work needs to be done in order to understand the molecular mechanism that underlies that pathological/clinical defect.

Chapter 5

Results of screening of *MPZ* in CMT patients

This chapter describes the mutation screening of the myelin protein zero (*MPZ*) gene in Taiwanese CMT1 patients and British CMT2 patients. Section 5.1 presents the mutation screening of the *MPZ* gene from five Taiwanese patients with clinical features of demyelinating type CMT disease. Two mutations in *MPZ* were found in two index patients. In addition to causing CMT1, *MPZ* mutations can also cause CMT2 [Marrosu *et al.* 1998]. In section 5.2, the screening of 46 British patients with CMT2 is described; one patient was found to harbour a *MPZ* mutation.

5.1 *MPZ* mutations in CMT1 patients

5.1.1. Molecular genetic findings

The coding region of the myelin protein zero gene (*MPZ*) in five Taiwanese patients was sequenced. There were two mutations in two index patients (patients #056 and #092). The sequence revealed a heterozygous "T" to "C" change at nucleotide 151 (c.151T>C) in exon 2 of patient #056 (Fig. 5.1a). This transition changes amino acid serine51 to proline.

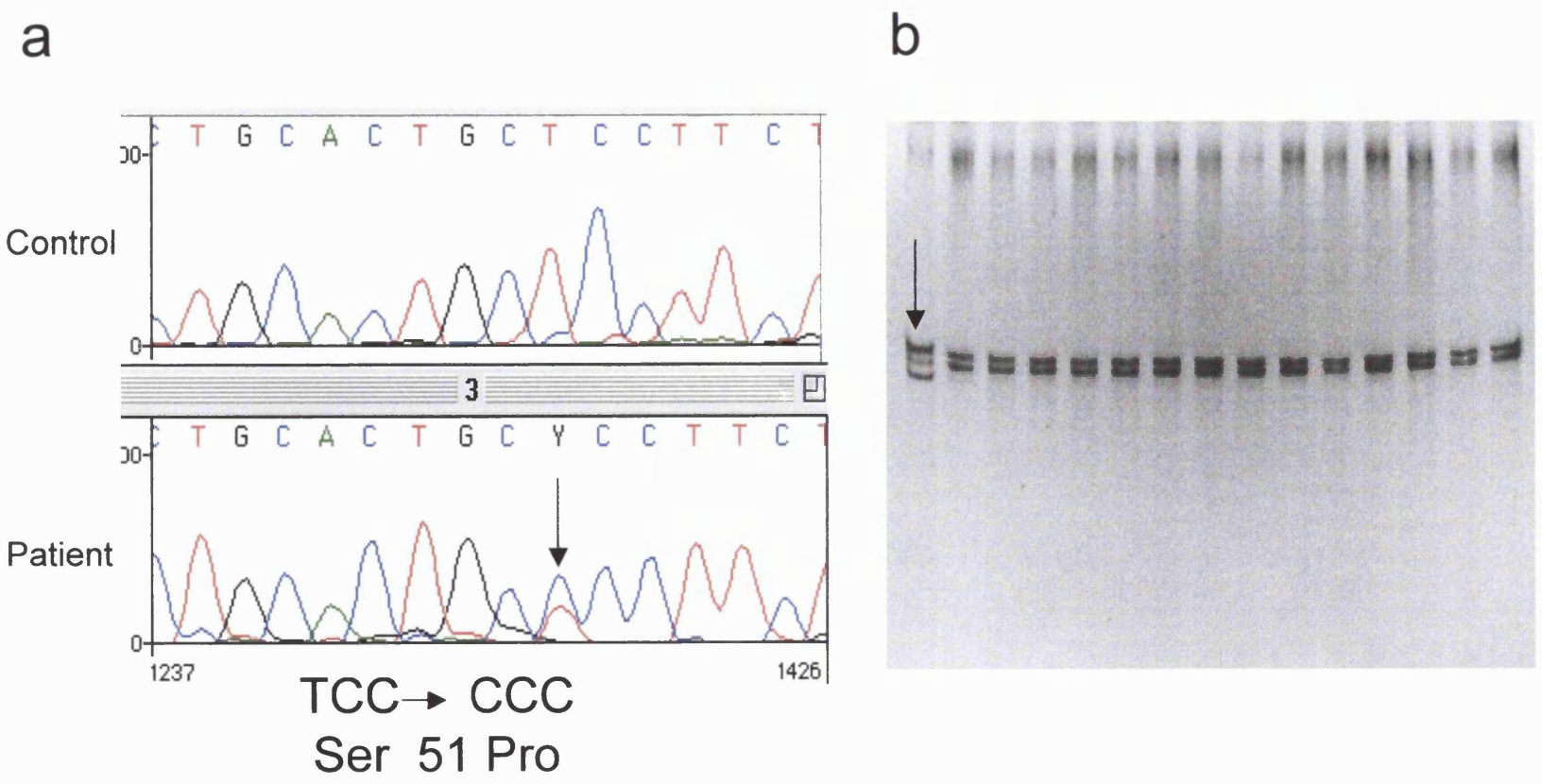


Fig. 5.1 *MPZ* mutation in patient #056. A 'c.151T>C' transition was shown in sequence (Fig. 5.1a) and this transition results in a band shift on SSCP gel (arrow in Fig. 5.1b).

This mutation does not create or lose a restriction enzyme cutting site. However, the PCR fragment of exon 2 from this patient showed a significant band shift on an SSCP gel, which was not detected in fifty Taiwanese controls (i.e. 100 chromosomes) (Fig. 5.1b).

A heterozygous "c.173T>A" transversion was found in exon 2 of the *MPZ* gene, causing a Valine to Aspartic acid change at codon 58 in patient #092 (Fig. 5.2a). This mutation created a new cutting site for the restriction enzyme Hinf I, which was used to show the absence of this mutation in 50 Taiwanese controls (Fig. 5.2b). In exon 5, a G600A silent polymorphism was also identified. Connexin32 (gap junction protein beta-1, *GJB1*) and its promotor, *PMP22* and *EGR2* were screened in the other three patients, but no mutations were found.

5.1.2 Clinical features of patients with *MPZ* mutations

Patient #056 is a 15-year-old boy. His birth and early development were normal. He had a 10-year history of unsteadiness, clumsiness and calf cramps. He had no sensory symptoms. On examination, the cranial nerves were normal. Muscle tone was normal and there was no wasting or weakness. Tendon reflexes in both upper and lower limbs were absent. There were no sensory signs except a positive Romberg's sign. His gait was ataxic.

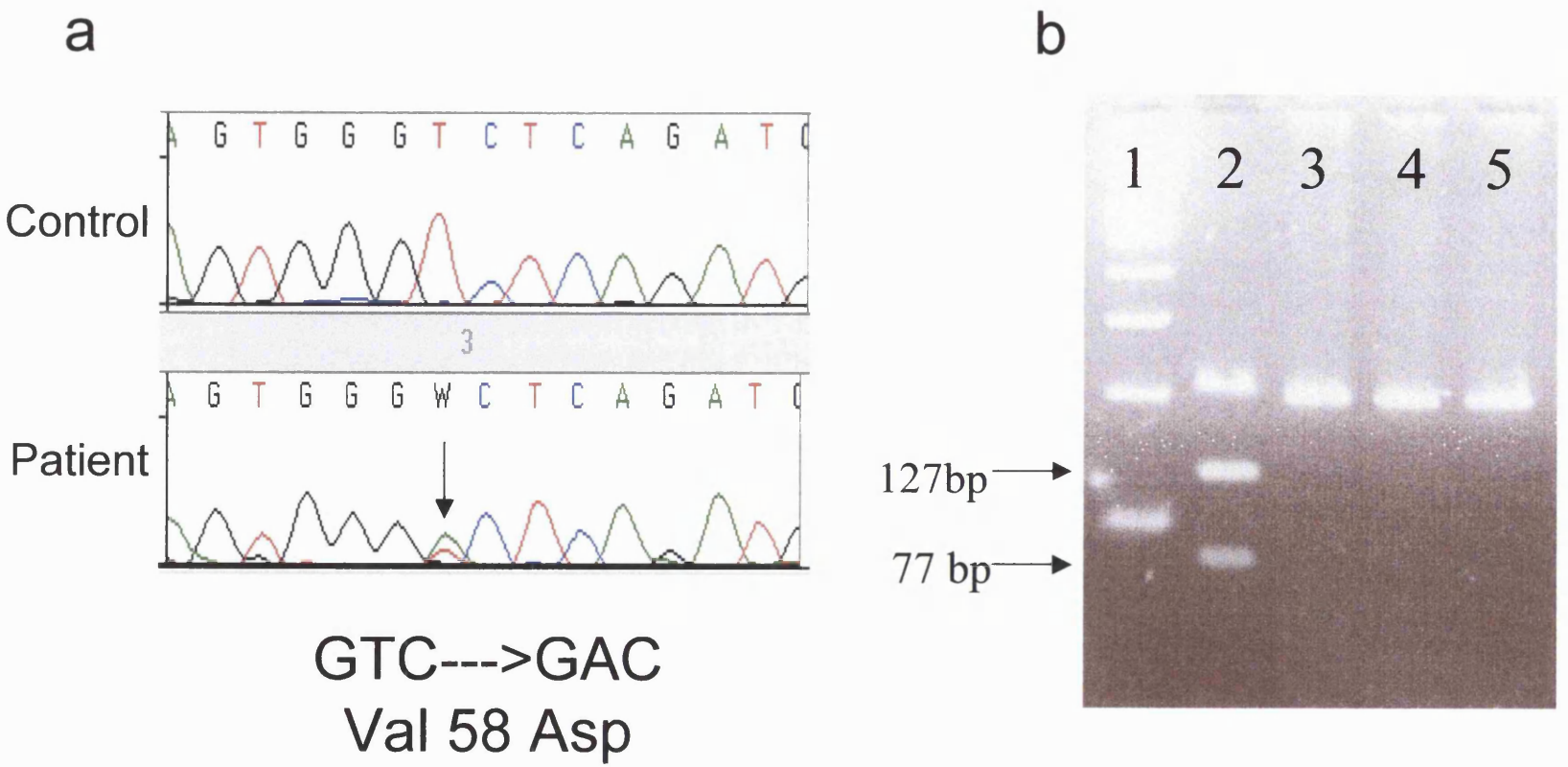


Fig.5.2. *MPZ* mutation in patient #092. Fig. 5.2a shows the sequence of a heterozygous 'c.173T>A' change in patient #092. This change creating a Hinf I cutting site. In fig. 5.2b, lane 1 is a DNA ladder, lane 2 is from the index patient, lane 3~5 are normal controls.

Nerve conduction studies showed markedly slow motor nerve conduction velocities (MCV) in the right median (15 m/sec) and the right ulnar (13 m/sec) nerves. Compound motor action potentials (CMAP) were 5.5 mV and 2.5 mV in the median and ulnar nerves, respectively. The F wave latency in the right median nerve was 104 milliseconds. Lower limb nerves (peroneal and tibial) were inexcitable. No H-reflex was recorded from the tibial nerve. Sensory action potentials (SAP) were absent from the median, ulnar and sural nerves. An EMG of the anterior tibialis muscle showed no active denervation changes.

The patient has one sister, one half brother and two half sisters. Clinical and electrophysiological examination of both parents and his sister were normal. There is no information available from his other siblings.

Patient #092 is a 32-year-old female with gait disturbance since childhood. She complained of progressive weakness and wasting of her distal lower limbs for 10 years. She had been reasonably stable, walking independently and able to ride a bicycle until 5 years before presentation. At that time following a pregnancy, she progressively deteriorated over two years with marked wasting and weakness of all four limbs. This was complicated by the development of diabetes mellitus during the pregnancy, which persisted after the pregnancy, and was difficult to control, despite the addition of insulin. By four years after the pregnancy, she had progressed to the point where she was paraplegic and wheelchair bound. Her deterioration over four years was thought to be partially due to diabetes mellitus but the question of a superimposed inflammatory neuropathy arose. CSF study was performed and was normal except for a mildly raised protein (61 mg/dl). A

sural nerve biopsy was done which confirmed demyelination with onion bulb formation. No evidence of inflammation was seen. A trial of plasma exchange resulted in some improvement and azathioprine was added. Over the next 2 years she had a further episode of deterioration, responding again partially to plasma exchange.

On examination, the cranial nerves were unremarkable. Muscle power in the upper limbs was MRC grade 3 proximally and grade 2 distally. In the lower limbs, knee flexion and extension was MRC grade 2 and there was no movement below the ankle. She was areflexic. Sensory examination revealed decreased pinprick and pain sensation in a glove-and-stocking distribution with severe loss of vibration and proprioception.

Motor and sensory action potentials were unobtainable in all tested nerves, including median, ulnar, peroneal, tibial and sural nerves. The distal motor latencies of the right musculocutaneous and axillary nerves were 20.7 and 15.3 milli-seconds, respectively.

There is a family history in that both her son and daughter have walking difficulties. All upper and lower limbs SAPs in her son and daughter were absent. The MCV of both ulnar nerves were 5 m/sec in her son. The median MCV was 7 m/sec and ulnar was 4 m/sec in her daughter. These findings clearly indicated that the family had an autosomal dominant demyelinating neuropathy. The final diagnosis was of CMT1 and diabetes mellitus with possible superimposed chronic inflammatory demyelinating polyneuropathy.

5.2 *MPZ* mutation in CMT2 patients

5.2.1 Molecular genetic findings

MPZ gene was screened in 46 British patients with CMT2. There was one patient (patient II:8 in family pedigree, Fig. 5.3a) who was shown to have a missense mutation. As shown in Fig.5.3b, a "C" to "T" transition (c.227C>T) causes a coding amino acid change from Alanine76 to Valine in index patient (patient II:8). The nucleotide change was not found in 50 Caucasian control sequences.

5.2.2 Clinical features of patient with *MPZ* mutation

The index patient is a 57-year-old woman. She suffered from insidious onset of muscle wasting in her lower limbs for 14 years. She complained of numbness and thickness on both feet and felt unsteady for many years. Thirteen years ago, she fell and sustained a back injury because of her unsteadiness. Since then, the sensory symptoms have persisted and progressed from the lower limbs to the upper limbs in two years. Apart from lower limb weakness, she had difficulties in grip and frequently dropped things.

On examination, the cranial nerves were unremarkable. Muscle wasting was found in both upper and lower limbs. Mild weakness (MRC grade 4) was noted in both APB and ADM muscles.

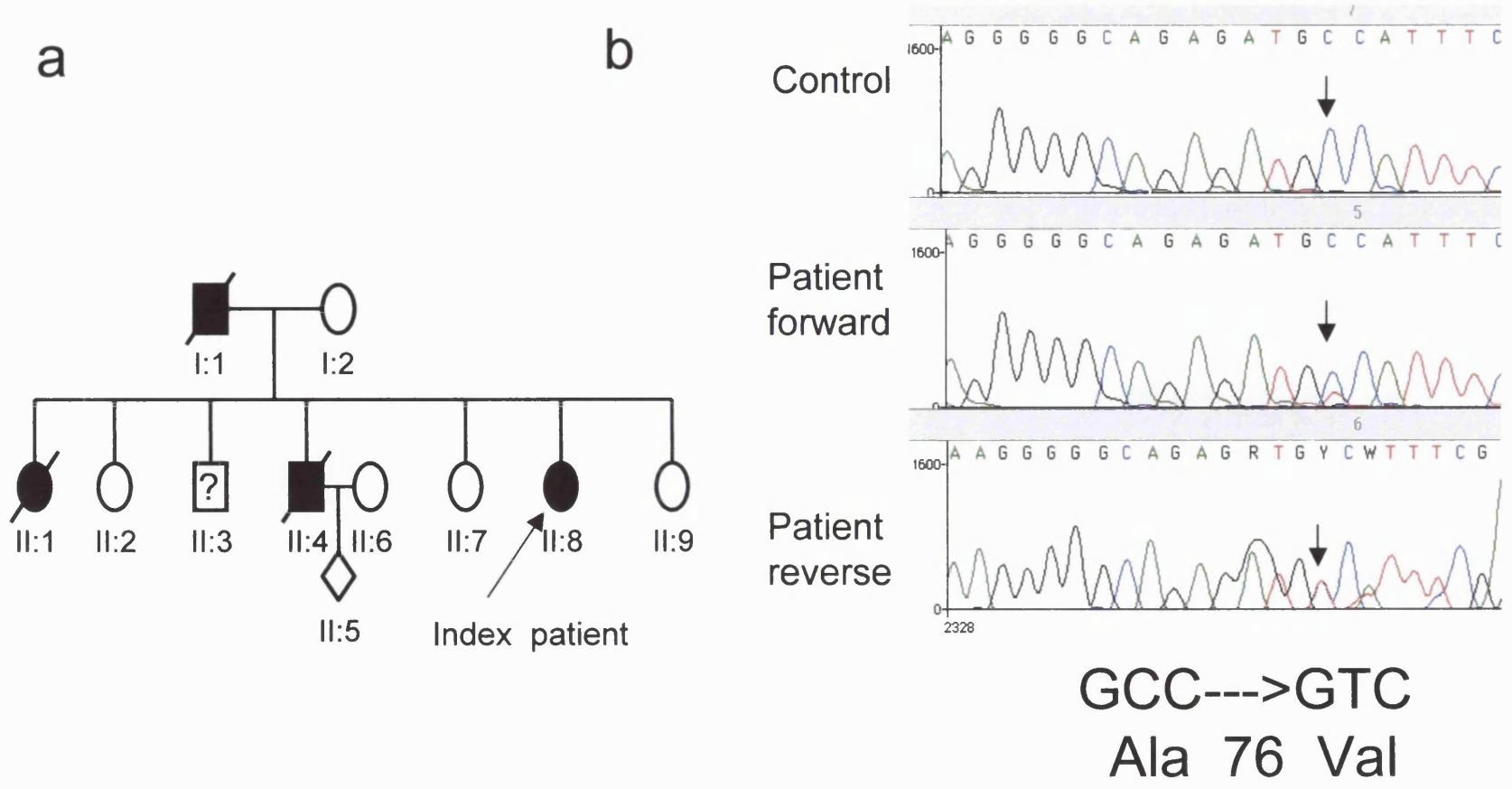


Fig. 5.3 Family pedigree of the CMT2 patient (fig. 5.3a) and *MPZ* sequence of the index patient and controls (fig. 5.3b). The 'c.227C>T' missense mutation causes a A76V change.

Muscle power for ankle and toe dorsiflexion was MRC grade 3+ and the proximal muscles power were grade 4+. On sensory examination, disturbance of all sensory modalities was noted from both feet to mid calf. She had bilateral pes cavus. The tendon reflexes were preserved on her first visit. However, reduced reflexes were noted three years later. Fig. 5.3a shows the pedigree of this patient. Her father had sensory loss in his feet and difficulties in walking. Her elder sister (II:1) and brother (II:4) both have pes cavus and a slapping gait.

Table 5.1 & 5.2 show the electrophysiological findings of the index patient. Sensory nerve conduction shows reduced SAPs (3.8μV for the median; 5μV for the ulnar nerve) with only mild slowing of sensory conduction velocities (46m/s for the median; 43m/s for the ulnar nerve) in the upper limbs. No responses were obtainable from either sural nerves. The CMAPs were decreased in the right median with only mild slowing of MNCV (43 m/s). The CMAPs from the right tibial nerve were not obtainable. These findings were compatible with an axonal sensorimotor polyneuropathy. No nerve pathology was available.

Table 5.1 Sensory nerve conduction studies in the index patient (patient II:8)

SAP	Amplitude (μV)	SCV (m/s)
Rt Median	3.8	46
Rt Ulnar	5.0	43
Rt Radial	5.5	52
Rt Sural	Absent	
Lt Sural	Absent	

SAP = sensory action potentials
SCV = sensory conduction velocities
Rt = right
Lt = left

Table 5.2 Motor nerve conduction velocities in the index patient

Motor	DML (ms)	MNCV (m/s)	CMAP (mV)	F wave (ms)
Rt Median	3.8	43	3.5 (wrist) 3.1 (elbow)	absent
Rt Ulnar	3.5	43	7 (wrist) 6 (below elbow) 6 (above elbow)	37.2 ms
Rt Tibial	5.5	-	0.4 (ankle) no response (fibular neck)	

Rt = right
DML = distal motor latency
MNCV = motor nerve conduction velocities
CMAP = compound motor action potentials
- = unobtainable

Chapter 6

Discussion: *MPZ* mutations in CMT patients

6.1 Genetic findings

In the first part of the myelin protein zero (*MPZ*) study, we screened five Taiwanese patients with CMT1. Two novel missense mutations were identified from two index patients (patient #056 and #092). The *MPZ* mutation in patient #056 (c.151C>T) caused a Ser51Pro change. There was no history of neuropathy in his parents. DNA was not available from the parents to check the segregation of the mutation. A "c.173T>A" transversion resulting in Val58Asp in exon 2 of *MPZ* was identified in patient #092. These two mutations are probably pathogenic because neither of them was detected in fifty Taiwanese controls by SSCP or restriction polymorphism. A nucleotide polymorphism in the exon 5 of *MPZ* was also noted in patient #092.

MPZ mutation screening in 46 British patients with CMT2 identified a missense mutation "c.227C>T" which causes an amino acid Ala76Val change. This mutation was not found in 50 Caucasian controls.

6.2 Phenotypes in patients with *MPZ* mutations

Clinically, patient #056 had no muscle wasting, weakness and no sensory symptoms. However, an ataxic gait, positive Robergism and tendon areflexia were suggestive of a peripheral neuropathy. The disease onset was in early childhood (~ 5 years of age) and progressed slowly. The electrophysiology was clearly in the demyelinating range. No detailed family history was available. The phenotype in this patient was compatible with previously described CMT1 patients with mutations in *MPZ* (CMT1B).

In patient #092, although the early history was compatible with an inherited neuropathy, the relapsing and remitting nature of the clinical features, the increased protein in the CSF and the partial response to plasma exchange and azathioprine suggest a superimposed inflammatory neuropathy, despite the absence of inflammation in the nerve biopsy. Rare cases of CMT disease associated with chronic inflammatory demyelinating neuropathy have been reported [Almsaddi *et al.* 1998] [Bird and Sladky 1991]. Recently, Donaghy *et al.* reported a CMT family with a *MPZ* mutation and a neuropathy that responded to steroid treatment [Donaghy *et al.* 2000]. Assessing the sera of 21 patients with chronic inflammatory demyelinating neuropathy (CIDP), revealed the presence of anti-myelin protein zero immunoglobulin G in the sera of six patients by Yan *et al.* [Yan *et al.* 2001]. All of the above studies suggests that myelin protein zero may be antigenic. Our patient may also have CMT1B and a superimposed inflammatory neuropathy, although the diagnosis of an inflammatory neuropathy could not be proven with nerve conduction studies (inexcitable nerves) and a

sural nerve biopsy, and the possibility remains that some of the deterioration was due to the diabetes mellitus.

In 1998, Marrosu *et al.* reported a missense mutation (Ser44Phe) of *MPZ* in a large Sardinian family with clinical features of axonal CMT [Marrosu *et al.* 1998]. The MNCVs of the median nerves in the patients ranged from 42.8 to 57.3 m/s and in some patients the MNCVs were not available. De Jonghe *et al.* reported a missense mutation (Thr124Met) in *MPZ* in seven Belgian CMT families, and raised the possibility of a common ancestor. The clinical features were characterised by late onset (later than 30 years of age), marked sensory abnormalities and in some patients, deafness and pupillary reflex abnomalities. The MNCVs of the median nerve ranged from 24 to 59 m/s. Sural nerve biopsy in an affected patient showed loss of myelinated axons and mutiple small and large clusters of regenerating axons. In the same year, Chapon *et al.* reported a French family with late onset motor and sensory neuropathy associated with deafness and Argyll-Robertson-like pupils. The MNCVs in the upper limbs ranged from 34 to 57 m/s with very low compound motor action potentials (CMAPs of all patients < 3.5 mV). Nerve biopsy from two patients showed loss of myelinated fibers and rare clusters of small myelinated fibers. Axonal degeneration was demonstrated in a teased fiber examination. Genetic studies also showed a Thr124Met missense mutation identical to the mutation in the Belgian families.

We screened 46 patients with a diagnosis of axonal CMT. A missense mutation (Ala76Val) was identified in one patient (patient II:8 in Fig. 5.3). Clinically the

first symptom was noticed at a late age (second decade or older) with a slowly progressive course. Sensory involvement of both small and large sensory fibers was prominent. The MNCVs of the right median and ulnar nerves were mildly slow with reduced CMAPs. The SAPs were significantly reduced in both upper and lower limbs. These data were compatible with a diagnosis of CMT2. There was no evidence of cranial nerve involvement (deafness or pupil response abnormalities) in this patient. Only one in 46 CMT2 patients had a *MPZ* mutation, suggesting that *MPZ* mutations are a rare cause of CMT2.

6.3 Functional implications for *MPZ* mutations

Myelin protein zero (MPZ) is the major myelin protein expressed in Schwann cells. The *MPZ* gene consists of six exons and each exon corresponds to a proposed functional domain. The extracellular domain, which is composed of exons 2 and 3, forms the immnoglobulin-like structure, which is involved in myelin compaction by homophilic interactions [Lemke *et al.*1988]. The structure of the extracellular domain consists of ten antiparallel β strains organised into two β sheets, which has been confirmed by X-ray crystallography data [Lemke *et al.* 1988]. Similar to the structure of immnoglobulin, the D-, E-, B-, A-βstrains make up one β sheet and the A'-, G-, F-, C-, C'- and C"-β strains make up the other sheet. MPZ crystal structure analysis suggests that individual MPZ molecules segregate and emanate from the membrane as tetramers surrounding a large central hole, like a doughnut whose base is in the myelin lipid bilayer [Shapiro *et al.* 1996]. Shapiro *et al* reported the fractional solvent accessibility

of each individual amino acid in the extracellular domain of MPZ protein. Serine51, valine58 and alanine76 correspond to amino acids 22, 29 and 47 in the extracellular domain, respectively. Serine-22 is one of the amino acids in the B strain, valine-29 is in the proximal end of the B-C loop and alanine-47 is in the C strain.

The B-C loop is important for forming interfaces between two adjacent MPZ molecules. There are three interfaces between the sheets and loops. The first (fourfold) interface consists of interactions between side chains from the B-C loop of one molecule with side chains of the C"-D and E-F loops on the next molecule. The second interface called a putative adhesive interface, is a stacking interaction between the His-52 ring of one MPZ and the guanidino group of Arg-45 of another MPZ. The third interface is a twofold symmetric head-to-head contact at the apex of the large B-C loop. The exposed residual tryptophan 28 makes Van der Waals contact with the main chain atoms of the opposing B-C loop. All of these interfaces among four MPZ molecules form a tetramer [Spiryda 1998]. When serine-22 changes to proline as in our patient #056, the large benzine ring might either destroy the structure of the B β strain or change the solvent accessibility of this amino acid. Therefore, the shape of two β sheets may not be maintained. The amino acid valine-29 is a small and neutral molecule. The replacement of valine-29 by aspartic acid as occurred in patient #092, is predicted to introduce an acidic hydrophilic side chain in the B-C loop and may insert new hydrogen bonds between two B-C loops. In terms of fractional solvent accessibility, the alanine-47 is important for the putative adhesive interface. Replacing alanine with valine at this position is predicted to

increase the volume of the side chain and change the context of the interface. Any structural changes could interfere with any of the interfaces among MPZ molecules. The oligomerization to form the tetramer and then, the homophilic contact between two tetramers in intralipid lines might be distorted by these mutated proteins. The exact pathogenic mechanism of each individual *MPZ* mutation warrants further studies.

6.4 Conclusion

Most of the people in Taiwan came from the provinces of China, Fu-Jien and Canton. Only rare original inhabitants are from south Asia and now most of them live in the high mountain areas. Hereditary motor and sensory neuropathy accounts for only 4.23% of neuropathies in 520 patients with generalised neuropathy surveying from 5 neurological centres in this small island [Lin *et al.* 1993]. Most of the genetic defects underlying CMT in Caucasian people have been well studied; however, only rare reports of the genetic defects underlying CMT in Chinese people exist. The two *MPZ* mutations reported here are novel; but the polymorphism in the exon5 of patient #92 has been reported in Caucasian people (http://molgen-www.uia.ac.be/ CMTMutations/). It will be interesting to compare the genetic defects underlying CMT in people from different ethnic origin.

In this study, mutations of *MPZ* have been demonstrated to cause either a demyelinating or axonal neuropathy i.e. CMT1 or CMT2. This reinforces the

concept that a modern classification of CMT needs to be based on clinical, electrophysiological, pathological and genetic data. The pathogenic mechanisms how *MPZ* mutations cause an axonal neuropathy remain unknown. It is now becoming clear that the *MPZ* gene is activated, albeit at relatively low levels, prior to myelination not only in all immature Schwann cells but also in Schwann cell precursors [Lee *et al.* 1997]. MPZ had also been seen even earlier in a subpopulation of cells in the trunk neural crest, the tissue that gives rise to Schwann cell precusors [Lee *et al.* 1997 & 2001]. The early expression of *MPZ* in the neural crest cells might suggest a direct involvement of *MPZ* mutation in axonal development.

Chapter 7

Results of screening *GJB1* in CMT patients

This chapter describes the mutations in the gap junction protein beta 1 (*GJB1*) in patients with CMT diseases (CMT1, CMT2 and CMTX). The clinical as well as the electrophysiological features of the index patients and their families are also described.

7.1 Molecular genetic findings

Sequencing of the *GJB1* gene in 139 CMT patients revealed twenty-four mutations, as shown in Table 7.1. These mutations include nineteen missense, one nonsense and four frameshift mutations. Two individuals from different families with the same mutation were found with mutations Arg107Trp, Val139Met, Arg164Trp and Arg220Stop. All of the amino acids found with missense mutations are conserved in all published sequences of mammalian species. The amino acids Asn14, Trp24, Ala40, Thr55, Val63, Pro70 and Val95 of GJB1 (connexin32) protein are conserved among all vertebrates and also conserved among the protein members of the connexin family, e.g. connexin-26, -30, -41, -43, etc.

Table 7.1. *GJB1* mutations found in patients with Charcot-Marie-Tooth disease

Missense and nonsense mutations			Frame shift mutations			
Codon	AA change	Nt change	Codon	AA change	Nt changes	Stop codon at
14	Asn—Lys	AAC—AAA	42	Ser—FS	GAGT—GGT, del A	120
21	Gly—Asp	GGT—GAT	73	His—FS	CATG—ATG, del C	120
22	Arg—Gln	CGA—CAA	109*	Glu—FS	GAG—TGAG, insT	109
24*	Trp—Cys	TGG—TGT	191*	Thr—FS	ACC—AACC, insA	242
26	Ser—Leu	TCG—TTG				
40	Ala—Thr	GCA—ACA				
55*	Thr—Arg	ACA—AGA				
55	Thr—Ile	ACA—ATA				
63	Val—Ile	GTT—ATT				
70	Pro—Ser	CCC—TCC				
83	Leu—Arg	CTA—CGA				
95	Val—Met	GTG—ATG				
107	Arg—Trp	CGG—TGG				
125*	Val---Asp	GTC---GAC				
139	Val—Met	GTG—ATG				
142	Arg—Trp	CGG—TGG				
153*	Phe—Ser	TTT—TCT				
164	Arg—Trp	CGG—TGG				
219	Arg—Cys	CGC—TGC				
220	Arg—Stop	CGA—TGA				

AA= amino acid
Nt= nucleotide
FS= frame shift
*= novel mutation

The four frame shift mutations, i.e. deletion of nucleotide "A" at codon 42, deletion of a "C" at codon 73, insertion of a "T" at codon 109 and insertion of a "C" at codon 191, induce a premature stop codon at 120, 120, 109 and 242, respectively. Mutations of Trp24Cys, Ala40Thr, Val125Asp, Phe153Ser, Glu109fs and Thr191fs are novel and the sequence pictures are shown in Fig. 7.1. The distribution of the amino acids with missense mutations was as follows: one in the N-terminal domain, seven in the four transmembrane domains, six in the two extracellular domains, two in the intracellular loop, and one in the C-terminal domain (Fig. 7.2). The nonsense mutation, Arg220Stop located at the C-terminal results in a truncated protein. Disease segregation with mutations was confirmed in ten families, in which DNA from other family members was available (Fig. 7.3 shows two examples).

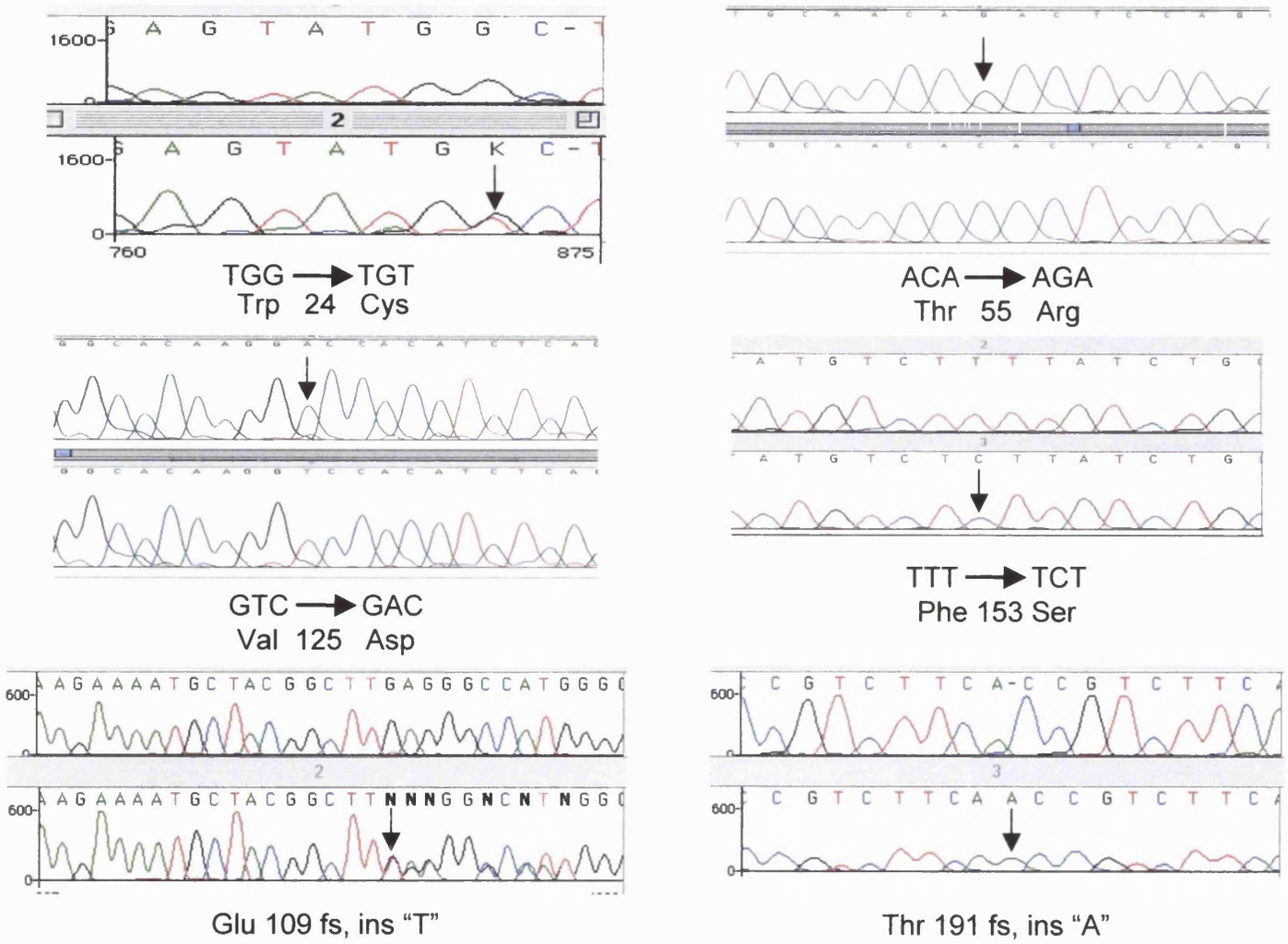


Fig. 7.1 Sequencing pictures for the six novel *GJB1* mutations (arrows). fs = frame shift, ins = insertion.

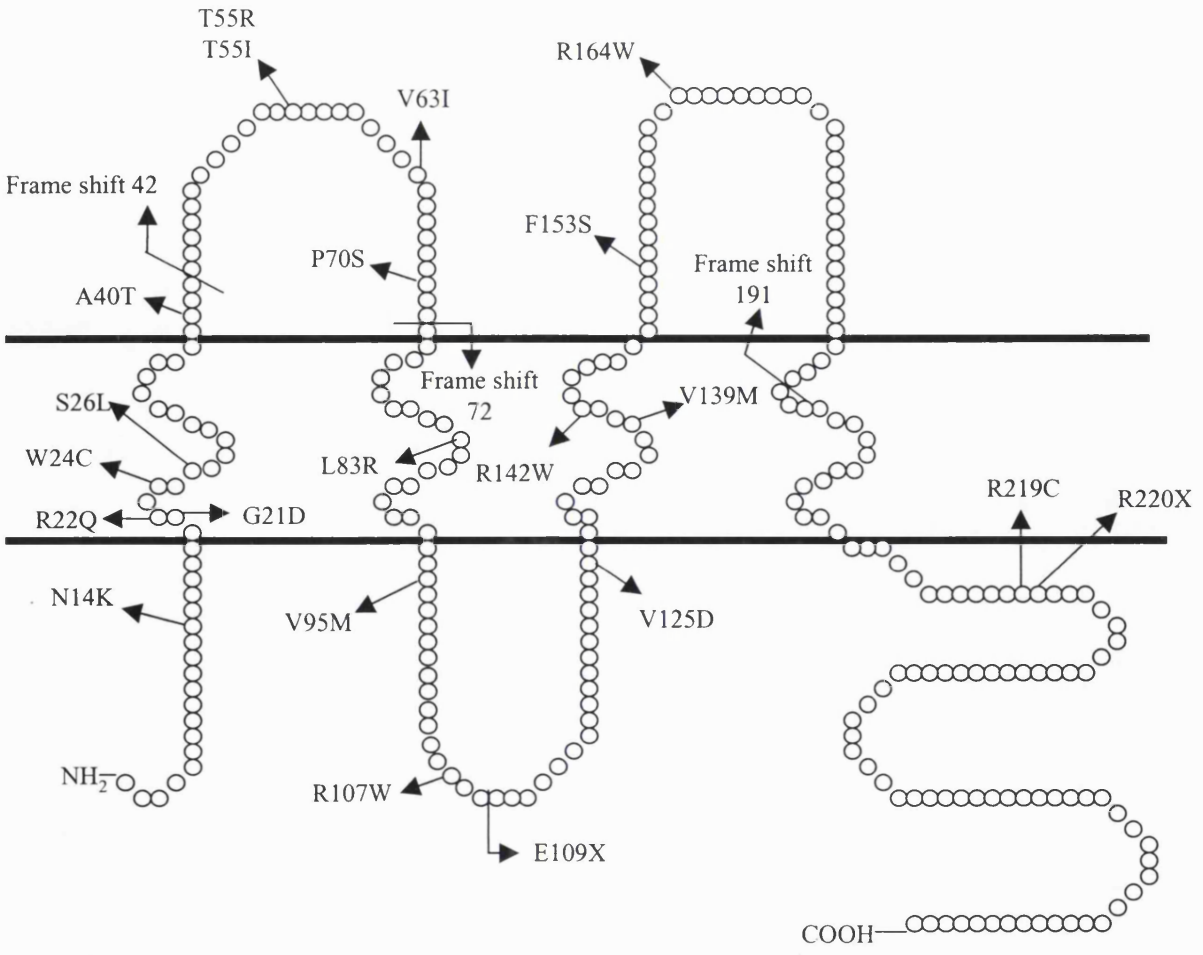


Fig. 7.2 Schematic presentation of the membrane topology of GJB1 protein and the distribution of mutations.

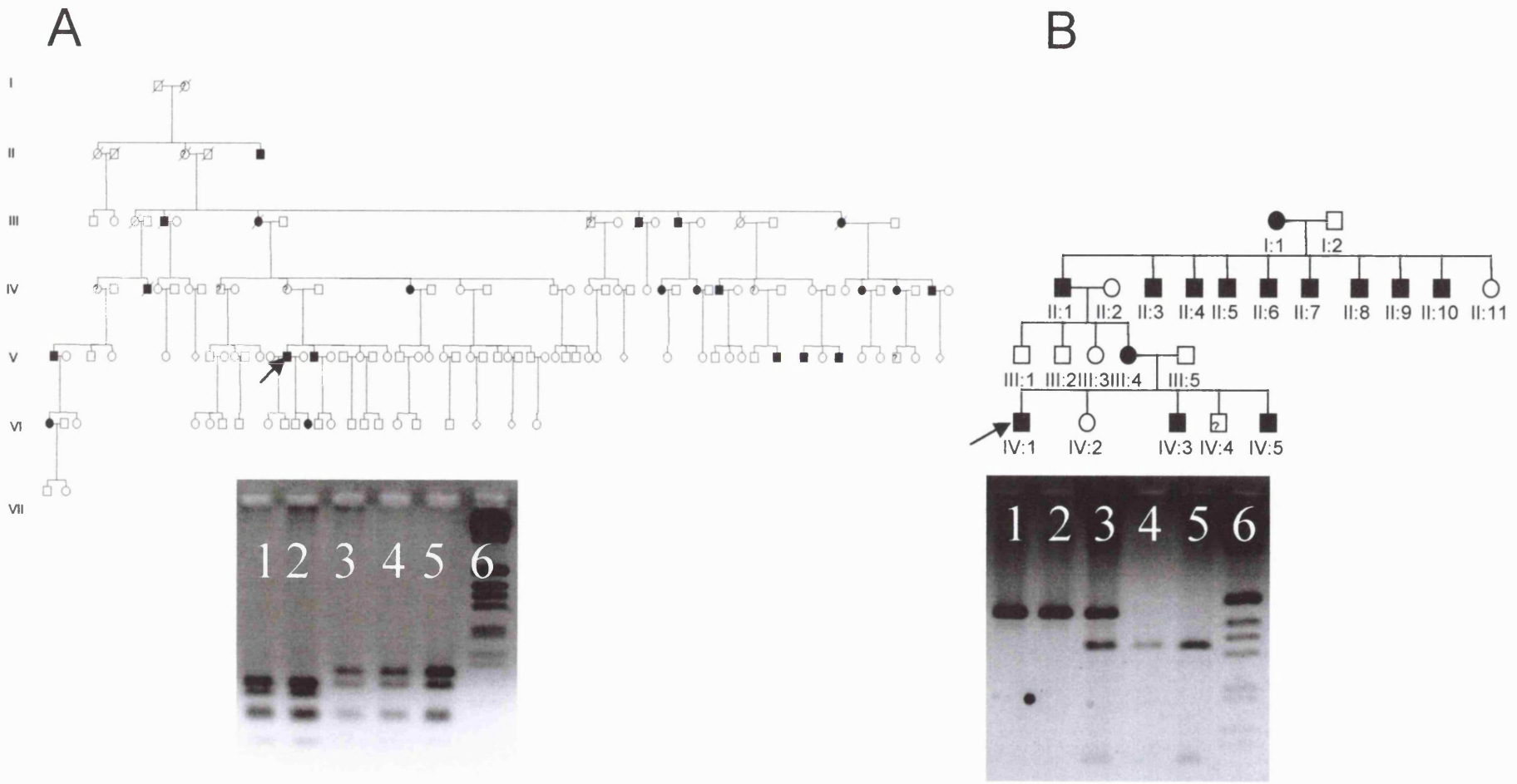


Fig. 7.3 Segregation of the *GJB1* mutations in two CMT families. 7.3A & 7.3B show the family pedigrees and the restriction polymorphisms with mutations Val139Met and Glu109fs, respectively. 7.3A Lanes 1 & 2 index patient (arrow) and his affected brother. Lanes 3~5 are from normal siblings. Lane 6 is a DNA ladder. 7.3B Lane 1, IV:1 (index patient); Lane 2, II:1; Lane 3, III:4 Lane 4 & 5, normal controls; Lane 6, a DNA ladder.

7.2 Clinical features of patients with *GJB1* mutation

The table 7.2 and 7.3 show the clinical features and electrophysiological data of the index patients and their families. The mean age of onset for the males was 16.8 ± 10.0 years of age (y/o) (range from 15 months - 32 y/o) and for the females was 24.4 ± 13.3 y/o (range from 5-52 y/o). Among family members, the age of onset of affected males was always younger than the affected females. In addition, the motor and sensory involvement of the affected males was more severe than that of the females in the same family. The degree of muscle wasting was variable from minimal to severe in both upper and lower limbs and was more severe distally. The involvement of large sensory fibers (vibration and joint position sense) was found earlier and more commonly than that of the small fibers (pinprick and temperature sensation). The severity of the sensory disturbance was not proportionate to the motor weakness or muscle wasting. Patchy distribution or asymmetrical involvement is not uncommon in these patients as shown in Table 7.2. Based only on the clinical and electrophysiological criteria, these twenty-eight families had been diagnosed as CMT1 (six families), CMT2 (twelve families) and CMTX (ten families). Brisk tendon reflexes and/or extensor plantar responses were noted in five families (families with mutations Gly21Asp, Trp24Cys, Pro70Ser, Phe153Ser and Thr191fs). Two index patients (mutations Thr55Arg and Thr191fs) had deafness and abnormal brainstem auditory response was noted in one patient (Thr55Arg). Further suggestions of associated central nervous system involvement included diplopia (Arg22Gln), medullary astrocytoma (Val139Met), generalised tonic-clonic seizure (Arg164Trp) and abnormal signal on brain MRI (Glu109fs).

Table 7.2. Clinical features of the index patients with *GJB1* mutations

Location	AA change	Clinical manifestation						
		Age of onset	Sex	Muscle wasting (Upper/Lower)	Sensory disturbance	Symmetry	Clinical Dx	Others
N-terminus	Asn14Lys	13#	F	D>P, ±/+	Large, mild	Yes	CMT2	
Transmembrane domains	Gly21Asp	9	M	D>>P, ++/+++	No	Yes	CMTX	
		11#	M	D>>P, ±/++	No	Yes		
		30	F	D>P, -/++	No	Yes		Brisk reflexes
	Arg22Gln	30#	F	D>P, +/++	Large fiber, moderate	No, L>R	CMT1	Diplopia
	Trp24Cys	15	F	D>P, +/++	Large fiber, mild	No, R>L	CMTX	
		11	F	D>P, +/++	Both, severe	No, L>R		Ext. plantar response
		12#	M	D>>P, +++/+++	Both, severe	Yes		
	Ser26Leu	15#	F	D>P, +/++	Both, moderate	Yes	CMT2	
	Leu83Arg	N.A.					CMT1	
	Val139Met*	22#	M	D>P, +/++	Both, mild	Yes	CMT1	Medulla astrocytoma
		13	M	D>P, ++/++	Both mild	Yes		
		36	F	D>P, ±/+	Minimal	Yes		
		N.A.					CMTX	
	Arg142Trp	10	M	D>>P, ++/+++	Large fiber, mild	No, R>L	CMT2	
		35#	F	D>P, ±/+	Large fiber, mild	No, R>L		
Extracellular loop	Ala40Thr	N.A.					CMT2	
	Thr55Ile	N.A.					CMTX	
	Thr55Arg	16#	M	D>>P, ++/+++	Both, moderate	Yes	CMTX	Deafness and abnormal ABR
	Val63Ile	32#	M	D>>P, ++/±	Minimal	Yes	CMT2	
	Pro70Ser	52#	F	D>P, +/+	Small fiber, mild	Yes	CMTX	Brisk reflexes except ankle jerk; ext. plantar response fasciculation on back
		25	M	D>P, ++/++	Large, moderate	Yes		
		26	F	D>P, ±/+	Small fiber, mild	Yes		

	Phe153Ser	8#	M	D>P, ±/+	No	Yes	CMT2	Brisk reflexes
	Arg164Trp*	30#	M	D>P, +/+++	Large fiber, mild	No, L>R	CMT2	
		30#	M	D>P, ++/+++	Both, mild	Yes	CMTX	GTC seizure
Intracellular loop	Val95Met	20# 40	M M	D>P, +/+ D>P, +/++	Both, mild Both, mild	No, L>R Yes	CMT1	
	Arg107Trp*	17#	M	D>P, +/++	No	No, patchy	CMT2	
		26#	M	D>P, +/++	Both, moderate	Yes	CMT1	
	Val125Asp	9#	M	D>>P, ++/+++	Large fiber, mild	Yes	CMTX	
C-terminus	Arg219Cys	18#	M	D>P, +/++	No	Yes	CMT2	IDDM
	Arg220Stop*	13#	M	D>P, +/++	Large fiber, moderate	No, R>L	CMT2	
		11#	M	D>P, +/++	No	Yes	CMT2	
Frame shift mutations	Ser42	6#	M	D>P, ±/+	No	Yes	CMT1	
	His73	5#	F	D>P, +/++	Large fiber, mild	Yes	CMT2	
	Glu109Stop	25@ 4# 15 mon	F M M	D>P, +/++ D>P, +/++ D.P, ±/+	No No No	Yes Yes Yes	CMTX	@MRI brain: high signal in the cerebral hemisphere and right mid brain in T2WI
	Thr191	10#	M	D>>P, ++/+++	Large fiber, moderate	Yes	CMTX	Deafness, ext. plantar response

ABR = auditory brainstem response Both = large and small sensory fiber involvement D= distal GTC = generalised tonic-clonic
IDDM = insulin dependent diabetes mellitus L = left N.A.= clinical detail is not available except the diagnosis P= proximal R = right
*= two families with the same mutation ± = minimal + = mild ++ = moderate +++ = severe
= index patient T2WI = T2 weighted image

Table 7.3. Electrophysiological findings in CMT patients with *GJB1* mutations

Location	AA change	Nerve conduction study						
			Median nerve				Peroneal nerve	
		Age/Sex	CMAP	C.V. (m)	SAP	C.V. (s)	CMAP	C.V.
N-terminus	Asn14Lys	21/F#	1.1	45	-	-	-	-
		50/M	0.1	27	-	-	-	-
Transmembrane domains	Gly21Asp	18/M	5	32	-	-	-	-
		31/M#	0.2	31	-	-	-	-
		50/F	9.1	37	3	37	-	-
	Arg22Gln	31/F#	0.6	33	-	-	-	-
	Trp24Cys	28/F	2.8	41.5	3.5	-	0.15	40
		65/M#	0.1	-	-	-	0.1	-
	Ser26Leu	45/F#	3.4	41	3	38		
		30/F	5.5	46	3	36		
	Leu83Arg	N.A.						
	Val139Met	44/M#	0.3	24	-	-	-	-
		61/F	2.0	40	-	-	-	-
		17/M#		31	4.0	30		
		37/F		40	3.0	40		
	Arg142Trp	26/M	1.0	32	-	-	-	-
		45/F#	4.0	52	6	37	1.5	39
Extracellular loop	Ala40Thr	N.A.						
	Thr55Ile	N.A.						
	Thr55Arg	71/M#	-	-	-	-		
			2.5 (ulnar)	33 (ulnar)				
	Val63Ile	54/M#	0.3	44	-	-	0.6	39
	Pro70Ser	53/F#	4.8	36	6	36	2.1	37
	Phe153Ser	12/M#	12.3	51	11.3	47	-	-
		42/F	11.6	55	-	-		
		52/M	5.4	44.7	-	-	-	-

	Arg164Trp	57/M#	2.3	34	-	-	-	-
		35/M#	3.0	24	-	-	-	-
Intracellular loop	Val95Met	49/M#	0.03	27	-	-	-	-
		65/M	0.02	32	-	-	0.1	21
	Arg107Trp	40/M#	0.4	37	-	-	2.9	36
		48/M#	2.5	28	-	-	-	-
	Val125Asp	16/M#	6.5	32	-	-	0.2	31
C-terminus	Arg219Cys	47/M#	0.5	35	-	-	-	-
	Arg220Stop	35/F#	0.25	45	-	-	-	-
		33/M#	-	-	-	-	-	-
Frame shift	Ser42	12/M#					0.02	22*
	His73	N.A.						
	Glu109Stop	31/F	0.72	49				
		12/M#	1.2	37				
		5/M	2.4	49				
	Thr191	54/M#	-	-	-	-	-	-

* = Sural nerve SAP, 4μV, mother has demyelinating neuropathy
CMAP= compound motor action potential, mV SAP = sensory action potential, μV
C.V. = conduction velocity, meter per second m = motor s = sensory
N.A. = not available - = not elicitable # = index patient

Clinical features including age of onset, disease severity, and CNS involvement did not correlate with mutation type or mutation position.

Electrophysiological findings showed variable features among the index patients and their families (Table 7.3). Among family members with the same mutation, CMAPs of the median nerve of females were usually greater than that of their relative males. The median CMAPs were obtainable in all female patients, but the CMAPs were not obtainable in four of the male index patients. Motor nerve conduction velocities (MNCVs) of the median nerve were slower in males as compared to females in the same family (Table 7.3). In males, the average median MNCV was 34.2 ± 7.7 m/sec (range from 24-51 m/sec), while in females, it was 43.1 ± 6.4 m/sec (range from 33-55 m/sec). 38.5% (5/13) of female median MNCVs were in the intermediate range (25-40 m/sec). The median MNCVs of all other female patients (8/13) were higher than 40 m/sec. The MNCVs in male patients were slower and 70% (14/20) of median MNCVs of male patients fell in the intermediate range. Two male patients had median MNCVs of 24 m/sec, which were slow and clearly in the demyelinating range. Only four males had median MNCVs higher than 40 m/sec. SAPs were usually more severely reduced than CMAPs. Median SAPs were obtainable in only 27% (9/33) of patients.

Motor conduction studies showed that the CMAPs and MNCVs of the median nerve (upper limb) were usually greater than that of the peroneal nerve (lower limb). However, the CMAPs of the peronal nerve were greater than that of the median nerve in patients with Val63Ile, Val95Met, and Arg107Trp mutations.

The same phenomenon has been reported by Senderek *et al.* [Senderek *et al.* 1999]. These findings may imply that the mutations cause a patchy involvement rather than a length dependent involvement. The CMAPs, MNCVs and patchy involvement were not correlated with either the type or the location of the *GJB1* mutation.

As shown in table 7.4, *GJB1* has been screened in 139 patients with variable types of CMT. Mutations were identified in six out of 44 (13.6%) CMT1 patients, 12 out of 38 (31.6%) CMT2 patients and 10 out of 21 (47.6%) CMTX patients. The percentage of the *GJB1* mutation in CMT2 was higher than that of CMT1.

Table 7.4 Numbers and percentage of *GJB1* mutations in patients with variable type of Charcot-Marie-Tooth disease

CMT Type	No. of sequenced Patients	No. of mutation	% of mutation in the relative type of CMT
CMT1	44	6	13.6
CMT2	38	12	31.6
CMTX	21	10	47.6
DSD	6	0	0
Other CMT (AR, no family history or history unavailable)	30	0	0
Total	139	26	

CMT = Charcot-Marie-Tooth disease
DSD = Dejerine-Sottas Disease
No. = number
% = percentage

7.3 Founder effect

Four mutations (Arg107Trp, Val139Met, Arg164Trp, and Arg220Stop) were each identified in two independent families. These families were from England. The families with mutations of Arg107Trp and Val139Met, did not share haplotypes (Table 7.5), but the four families with mutations Arg164Trp and Arg220Stop did, suggesting a common ancestor.

Table 7.5. Allele size in patients with the same *GJB1* mutations from different families (Families divided by vertical lines.)

Family	**Arg107Trp**			**Val139Met**					**Arg164Trp**		**Arg220Stop**				
Patient	1	2	3*	4	5	6	7	8	9	10	11	12	13	14	15
DXS990	5|-	4|-	3|3	5|-	5|-	4|5	5|-	3|4	5|-	5|-	5|-	3|5	5|5	2|5	3|-
DXS986	7|-	4|-	5|7	5|-	5|-	5|5	4|-	4|5	4|-	5|-	4|-	4|4	4|4	4|4	5|-
DXS8052	4|-	5|-	4|5	4|-	4|-	4|4	4|-	4|5	5|-	5|-	3|-	3|6	NA	3|2	3|-
DXS8107	3|-	4|-	3|11	12|-	12|-	12|9	3|-	3|3	3|-	3|-	12|-	12|8	12|12	12|12	12|-
DXS991	5|-	5|-	8|9	4|-	4|-	4|4	4|-	4|4	6|-	N.A.	6|-	4|6	6|8	6|7	6|-

* = unaffected sister
NA = not available

Chapter 8

Discussion of the mutation screening of *GJB1* in patients with CMT diseases

8.1 Summary of genetic findings

X-linked Charcot-Marie-Tooth disease (CMTX) is the second most common cause of CMT [Pareyson 1999]. The clinical features are similar to CMT1 and CMT2. In this study, we sequenced the gap junction protein β1 (*GJB1*) gene in 139 patients with CMT and without male-to-male transmission in the family history. The chromosome 17p duplication had been excluded in all CMT1 patients. We have identified 24 *GJB1* mutations in these families, of which 6 are novel mutations. Two unrelated families harboring the same mutation are described with four mutations. The mutations of *GJB1* include missense, nonsense and frameshift mutations. The mutations are found throughout all the domains of the GJB1 protein.

8.2 Functional implications of the mutations

Gap junction protein β1 (GJB1, connexin32) is one member of the family of gap junction proteins. This protein consists of four transmembrane (TM) domains, two extracellular (EC) loops, one intracellular (IC) loop, C- and N-terminals (N-

and C-ter) [Bennett *et al.*1991]. GJB1 protein is expressed in many organs including liver, pancreas as well as the nervous system. A connexon, which is made up of six GJB1 proteins in a hexagonal structure, is assembled and exported to the cell surface. Two connexons of two adjacent cells form a channel for cell-cell trafficking. This gap junction channel admits passage of small molecules, including ions, second messengers and metabolites (<1000 Da). In the peripheral nervous system, GJB1 protein is present on non-compacted myelin, i.e. paranodal loop, periaxonal collar, and the Schmidt-Lanterman incisures [Bergoffen *et al.* 1993b] [Bone *et al.* 1997]. It plays an important role in the communication between the axon-Schwann cells and inner-outer layers of myelin. By applying intracellular dye injection, Balice-Gordon *et al.* showed that functional gap junctions are present within the myelin sheath and allow small molecules to diffuse between the adaxonal and perinuclear Schwann cell cytoplasm [Balice-Gordon *et al.* 1998]. Some mutations have been expressed in Xenopus öocytes and in other *in vitro* systems and it is clear that certain mutations interfere with specific physiological functions of the GJB1 protein [Castro *et al.* 1999] [Peracchia *et al.* 1999]. Nevertheless, the exact mechanism of how channel dysfunction causes a neuropathy is still unknown.

8.3 Phenotypes of patients with *GJB1* mutation

In our study, males are more severely affected than females, with an earlier age of onset, more severe muscle wasting, weakness, and sensory loss, a findings which is similar to that in a large series of previously reported patients with

GJB1 mutations [Dubourg *et al.* 2001]. In 1998, Birouk *et al* reported a patient with a frameshift *GJB1* mutation, Arg22Stop, which had a severe neuropathy and an early age of onset [Birouk *et al.* 1998]. This premature stop codon results in a truncated protein with structure instability, which may explain the severity of the phenotype. The authors suggested that the shorter the GJB1 protein synthesised, the more severe the neuropathy. In our study, some of the nonsense and frameshift mutations with premature stop codons (Ser42fs, His73fs, Glu109fs, and Thr191fs) were associated with an earlier disease onset and a more severe phenotype. However, the severity of the phenotype seems not only to depend on the size of the protein, as patients (Ser42fs and His73fs) with a stop codon at 120, were not more severely affected than patients (Thr191fs) with a stop codon at 242. In addition, the disease severity, including the age of onset, motor and sensory involvement, neurophysiological abnormalities did not correlate with either the mutation type or the position of the mutation in this study (Tables 2 and 3).

In a European collaborative study for non-duplicated CMT1 families, the frequency of *PMP22*, *MPZ* and *GJB1* mutations was estimated as being 4%, 13%, and 28%, respectively [Nelis *et al.* 1996]. Another CMT1 study in Europe showed that 20% of non-duplicated non-dominant families had *GJB1* mutations [Janssen *et al.* 1997]. In 1998, Nicholson *et al.* reported that *GJB1* mutations were found in 50% of non-dominant, non-duplicated, CMT1 families [Nicholson *et al.* 1998]. In this study only 13.6% of the non-duplicated CMT1 family had a *GJB1* mutation (Table 4). This discrepancy may reflect the lack of availability of

the information about family pedigrees and in some cases the lack of detailed clinical description.

In a large previously reported kindred (six generations) with a *GJB1* mutation, absent or severely reduced CMAPs in the peroneal nerves (42%) and reduced sensory action potentials (SAPs) in sural nerves (75%) indicate a severe axon neuropathy [Hahn 1990]. Nerve biopsy showed a severe loss of myelinated and unmyelinated fibers. In our study, all but two (Pro70Ser and Arg107Trp) patients showed absent or severely reduced CMAPs in the peroneal nerve. All patients also have reduced SAPs in the median nerve (Table 3). These findings suggest early axonal loss in sensory nerves and in a length dependent fashion in motor nerves resulting from mutations of the gap junction protein GJB1. In addition, twelve out of thirty-eight (31.6%) CMT2 patients had *GJB1* mutations. In these families, the MNCVs of the median nerves of the affected females are greater than 38 m/sec with significantly reduced CMAPs, which suggested an axonal neuropathy [Harding and Thomas 1980]. Although the MNCVs of the median nerve of the male patients in these families are reduced and often fell in the intermediate or demyelinating ranges (25-40 m/sec), the CMAPs were disproportionally reduced. Silander *et al.* reported from a Finnish population and Latour *et al.* from a French population that some patients with *GJB1* mutations might be probably classified incorrectly as having an axonal neuropathy [Silander *et al.* 1997] [Latour *et al.* 1997]. By extracting DNA from embedded nerve samples with a diagnosis of axonal neuropathy, Senderek *et al.* also identified two out of 30 with (6.7%) *GJB1* mutations [Senderek *et al.* 1998b]. Our findings are in line with these reports and the frequency (31.6%) is higher

than hitherto reported. It has been suggested that *GJB1* mutations could result in early axonal involvement by means of interfering with the communication between Schwann cells and the axons.

Central nervous system involvement in patients with *GJB1* mutations has been reported previously [Nicholson and Corbett 1996a] [Bahr *et al.* 1999]. Hearing impairment with slowing of brainstem evoked potentials is characteristic in these patients. A CMTX family with identical twins from our laboratory was previously described with a missense mutation Ala39Val [Marques *et al.* 1999]. The proband as well as his twin brother have typical clinical features with intermediate MNCVs in the median nerve; however, both of them have extensor plantar responses. In the present study, at least five families in the 24 mutated families have been found to have pyramidal signs with either very brisk tendon reflexes or extensor plantar responses. In addition, cranial nerve involvement is also noted in two patients suffering from hearing loss and vertigo and one patient had diplopia. The brainstem auditory evoked responses are slow in the index patient from the family with a Thr55Arg mutation. A brain MRI of a patient from the family with the Glu109fs mutation shows high signal in the cerebral hemisphere and these findings suggest that central nervous system involvement is not uncommon in patients with *GJB1* mutation.

8.4 Mutation hotspots

Four mutations (Arg107Trp, Val139Met, Arg164Trp and Arg220Stop) occurred twice in our population. We looked for a founder effect in these families. The two families with the Arg107Trp and the two with Val139Met mutations, each have different haplotypes, which suggests that the *GJB1* mutation has occurred independently. At least seven groups have reported these two mutations (Arg107Trp and Val139Met) from different ethnic groups (http://www.uia.ac.be/CMTMutations/). These findings imply that these two amino acids might be hotspots for spontaneous mutations.

8.5 Conclusion

In conclusion, CMTX is a clinically heterogeneous neuropathy with variable clinical and electrophysiological features. *GJB1* mutations are responsible for CMTX. The pathogenic mechanisms are still unclear, and it is controversial whether demyelination or axonal degeneration is the primary pathological process. Mutations of the *GJB1* gene occurred in families with a clinical diagnosis of CMT1, CMT2 and CMTX and neither the mutation position nor the type of mutation predicted the clinical severity or electrophysiological features. Therefore, sequencing of the *GJB1* gene should always be undertaken in patients with an inherited neuropathy and without male-to-male transmission in the family history.

Chapter 9

Conclusions

9.1 Thesis Summary

In summary, this thesis is divided into two parts. The first part was a project designed to identify the molecular defect associated with the mutilated foot (*mf*) mutation in the rat. This particular mutant serves as an animal model for human hereditary sensory and autonomic neuropathy (HSAN). The second part involved screening patients with Charcot-Marie-Tooth (CMT) disease for mutations in two genes associated with this disease: myelin protein zero (*MPZ*) and gap junction protein β 1 (*GJB1*). Novel mutations were identified and detailed phenotypes for these patients are presented.

9.2. Identification of the causative gene for *mf* disease

9.2.1 Genetic analysis of the *mf* rat

Using a positional candidate gene approach, a substitutional mutation was identified in the δ subunit of the chaperonin containing tailless complex polypeptide-1 (*Cct4*). This approach combined conventional backcross linkage analysis, interrogation of sequence databases using BLAST search protocols and

extrapolation of conserved linkage between species (mouse, human and rat) to identify likely genes. These genes were sequenced in wild-type and mutant rats to identify modification in the sequence that could be attributed to the mutant phenotype. Of the four identified candidate genes, only one gene (*Cct4*) displayed sequence variation between the mutant and wild-type. This variation was authenticated by: 1) sequencing the same gene in several different strains of rat; 2) demonstrating that the wild-type sequence was conserved across 50 chromosomes; and, 3) the demonstration that the substitutional mutation replaces a cysteine for a tyrosine that was highly conserved across species.

9.2.2 Genetic confirmation that the mutation in *Cct4* is pathogenic.

Although the substitution of a highly conserved cysteine for a tyrosine in the *Cct4* gene of the *mf* rat provides strong evidence that this is the causative mutation for this phenotype, it will be useful to confirm this using transgenic techniques in the future. Protocols for producing transgenic rats, using a number of different rat strains (including SD), are now well established [Venkatesh *et al.* 1997]. Thus, it should be a fairly routine procedure to produced transgenic SD-*mf* rats expressing a wild-type cDNA *Cct4* transgene. Correction of the mutant phenotype by the transgene would provide definitive proof that the mutation in *Cct4* is the cause of *mf* disease [Kobayashi *et al.* 1997]. It will also be of interest to screen humans with HSAN for mutations in this gene. Another approach will be to generate the same mutation in mouse embryonic stem cells and study the resulting phenotype.

9.2.3 Function of *Cct4*

Having identified the causative gene for *mf* disease, the next step is to try to understand how a defect in this gene gives rise to a profound sensory neuropathy. As indicated in section 4.2.2 of chapter 4, apoptosis is a major aspect of the pathological process in *mf* disease and is almost certainly responsible for the reduced number of axons observed in the DRG. The relationship of mutations in *Cct4* to apoptosis needs to be explored. Apoptosis may be triggered by a disruption in the normal operation of the chaperonin complex leading to folding anomalies in a range of proteins or in some very specific cytoskeletal proteins such as actin? Alternatively, apoptosis might be triggered by the failure of the axon to make 'survival' contact with its target site. There also may be as yet unidentified independent functions of *Cct4*, the disruption of which could lead to apoptosis.

9.2.4 Future functional studies

9.2.4.1 Translation and transcription studies

RT-PCR results and the Northern slot-blot data demonstrate that the mutant gene is ubiquitously transcribed. It is unknown whether the mutant gene produces a mature protein or whether that protein is stable. To address these issues, it will be necessary to perform Western blot analysis using antibodies against the CCT4 protein. Pulse-chase experiments will also be necessary to determine whether

there is a change in protein turnover. If Northern and Western blot analyses establish no unusual changes in transcription or translation, it will be important to determine what effect the amino acid substitution has on CCT4 function/structure.

9.2.4.2 Structural abnormalities

As a highly conserved cysteine is replaced by a tyrosine, it is likely that this cysteine is functionally significant. The most probable role of cysteine is to form a disulphide bridge with another cysteine within the CCT4 peptide sequence or with a cysteine from another protein with which the subunit forms a complex. Three-dimensional modelling of the CCT4 subunit identifies a likely internal cysteine with which the affected cysteine might form a disulphide bridge. By employing some basic biochemical procedures (tryptic digestion and mass spectophotometric analysis), it could be established whether a disulphide bridge is formed that involves either an internal or an external cysteine.

Even in the absence of a detectable disulphide bridge, the substitution of the cysteine for a tyrosine could also have structural/functional stability implications for the CCT4 protein. Thus, it would be sensible to ascertain whether the mutation affects the structure of the delta subunit using electron microscopic techniques developed for this type of analysis [Llorca *et al.* 1999 & 2001]. The same approach can be used to analyse any impact the mutant CTT4 protein might have on the structure of the CCT complex.

9.2.4.3 Interaction with other proteins

Whilst the amino acid substitution might have no overall impact on structure, it may affect its ability to interact with specific substrates such as actin which is know to interact in a specific way with the CCT4 subunit of the CCT complex [McCormack *et al.* 2001]. The interaction deserves further study.

9.2.4.4 Developmental aspects of the disease phenotype and the effect of the *Cct4* mutation

Although the *Cct4* gene is ubiquitously expressed in adult tissues, little is known about the developmental regulation of this gene other than some over expression studies in the developing chick embryo [Zilkha-Falb *et al.* 2000]. Insights into this disease phenotype might be obtained by ascertaining when and where this gene is expressed during foetal development. These studies should be performed in conjunction with studies of the other CCT complex genes to determine whether CCT4 is acting independently of the other members of the complex.

The CCT4 subunit is known to interact with actin [Llorca *et al.* 1999]. Actin plays an important role in nerve growth cone motility and axonal path finding [Letourneau 1996]. It will be important to study the behaviour of actin both in terms of its distribution in the cytoskeleton and its turnover in the cell. It also might prove worthwhile studying the behaviour of the growth cone form the DRG of mutant rats. If they exhibit abnormal behaviour compared with wild-

type axons, it may suggest that the loss of neurones through apoptosis might be due to the failure of the axons to make survival contact with their target tissues.

9.2.4.5 Identification of other genes that might be affected by the *Cct4* mutation

The defect in *Cct4* has significant pathological consequences for the development of the peripheral nervous system, and as a result there are likely to be a great number of other genes whose expression is affected by this mutation. Given the pathology, it is possible to speculate on some of these genes (e.g., caspases) however; the involvement of other genes may be less predictable. The use of cDNA microarrays may provide some insights as to which genes are affected by the mutation.

9.3 Molecular genetic studies in CMT diseases

9.3.1 *MPZ* mutations in CMT patients

In this section, the myelin protein zero (*MPZ*) gene was screened in five Taiwanese CMT1 patients and forty-six British CMT2 patients. Two novel mutations (Ser51Pro, Val58Asp) were identified from the CMT1 patients. The clinical features and electrophysiological data of Patient #056 are compatible with a hereditary demyelinating neuropathy. There was no family history for this

patient. The second mutation was found in patient #092 who has a family history of peripheral demyelinating neuropathy. In addition to the CMT, this patient had a fluctuating neuropathy raising the suspicion of a superimposed inflammatory neuropathy.

One mutation, Ala76Val, in *MPZ,* was identified in forty-six British CMT2 patients. Clinically, this patient had a family history of chronic neuropathy. The electrophysiological study was compatible with an axonal neuropathy. Thus, this confirms that mutations in *MPZ* can be seen both in demyelinating and axonal CMT diseases.

9.3.2 *GJB1* mutations in CMT patients

The mode of disease transmission is sometimes difficult to determine in families with CMT. It is often difficult to differentiate autosomal dominant or X-linked inheritance in families without male-to-male transmission. Gap junction protein β1 (*GJB1*) gene is one of the responsible genes for X-linked CMT (CMTX). The clinical features and electrophysiological data from patients with mutations are reported to be quite variable. Disease severity and the type of neuropathy (demyelinating or axonal) are often different in males and females. In this study, the *GJB1* gene was sequenced in 139 British patients with variable types of CMT. Twenty-four mutations including six novel mutations were identified. The disease phenotypes including the clinical and electrophysiological features of patients with mutations are presented. We did not find any

genotype/phenotype correlations. The incidence of *GJB1* mutations was high in patients with a clinical diagnosis of CMT2.

9.3.3 Pathogenesis of *MPZ* and *GJB1* mutations

We screened two genes, *MPZ* and *GJB1* in patients with CMT. Clinically, both demyelinating and axonal neuropathies can be seen in patients with mutations in either of these genes. This observation raises many questions about Schwann cell/axon interactions and the role of MPZ and GJB1 in these interactions.

9.3.4 Future studies

CMT disease is a common disorder in neurology clinics. Future studies in CMT will need to focus in at least two directions. Firstly, as more and more causative genes are identified, the genetic screening of CMT is becoming more expensive and time consuming. Less expensive methods providing rapid and mass screening need to be introduced for service laboratories. Screening by SSCP and HPLC are alternatives for mutation detection in big populations and are used for screening both *MPZ* and *GJB1* but the sensitivity of these techniques needs to be improved.

Secondly, from a basic research angle, studies of tissue expression, cellular function and transgenic animals have already been performed for *MPZ* and *GJB1*

[Previtali *et al.* 2000] [Abel *et al.* 1999]. Functional studies of MPZ suggested that the disruption of compaction of myelin caused by *MPZ* mutations, may result in demyelination or dysmyelination [Previtali *et al.* 2000]. The mechanism by which *MPZ* mutations cause an axonal neuropathy is still unknown.

The functional implications of the *GJB1* mutations have been studied for many years. As this protein is important in forming channels between axons and Schwann cells, *in vitro* studies of mutated GJB1 proteins on channel conductance and gate function have been performed [Balice-Gordon *et al.* 1998] [Peracchia *et al.* 1999] [Ri *et al.* 1999]. The role of the GJB1 protein in stabilising myelin and the axon-Schwann cell unit has been established using a transgenic mouse model [Abel *et al.* 1999] [Scherer *et al.* 1998]. Histopathological studies of the sural nerve from CMTX patients lacking a functioning GJB1 protein showed demyelination, notably in a paranodal distribution, and distal accentuated axonal degeneration [Hahn *et al.* 2000]. These findings are similar to that found in *GJB1*-null mice. In summary, the GJB1 protein plays an important role in the stabilisation of myelin and in the rapid communication between the Schwann cell's nuclear domain and its distant periaxonal cytoplasm or the axon itself. Future studies in CMT secondary to *MPZ* and *GJB1* mutations will not only need to focus on pathogenesis, but eventually on developing treatment strategies for these complex diseases.

References

Abel A, Bone LJ, Messing A *et al.* Studies in transgenic mice indicate a loss of connexin32 function in X- linked Charcot-Marie-Tooth disease. J.Neuropathol.Exp.Neurol 1999; 58: 702-710.

Almsaddi M, Bertorini TE, Seltzer WK. Demyelinating neuropathy in a patient with multiple sclerosis and genotypical HMSN-1. Neuromuscul.Disord. 1998; 8: 87-89.

Anderson SL, Coli R, Daly IW *et al.* Familial dysautonomia is caused by mutations of the IKAP gene. Am.J.Hum.Genet. 2001; 68: 753-758.

Angelicheva D, Turnev I, Dye D *et al.* Congenital cataracts facial dysmorphism neuropathy (CCFDN) syndrome: a novel developmental disorder in Gypsies maps to 18qter. Eur.J.Hum.Genet. 1999; 7: 560-566.

Arnason U, Gullberg A, Gretarsdottir S *et al.* The mitochondrial genome of the sperm whale and a new molecular reference for estimating eutherian divergence dates. J.Mol.Evol. 2000; 50: 569-578.

Auer-Grumbach M, Wagner K, Timmerman V *et al.* Ulcero-mutilating neuropathy in an Austrian kinship without linkage to hereditary motor and sensory neuropathy IIB and hereditary sensory neuropathy I loci. Neurology 2000a; 54: 45-52.

Auer-Grumbach M, De Jonghe P, Wagner K *et al.* Phenotype-genotype correlations in a CMT2B family with refined 3q13-q22 locus. Neurology 2000b; 55: 1552-1557.

Axelrod FB, Abularrage JJ. Familial dysautonomia: a perspective study of survival . J.Pediat. 101, 234-236. 1982.

Axelrod FB, Pearson J, Tepperberg J *et al.* Congenital sensory neuropathy with skeletal dysplasia. J.Pediat. 102, 727-730. 1983.

Bahr M, Andres F, Timmerman V *et al.* Central visual, acoustic, and motor pathway involvement in a Charcot- Marie-Tooth family with an Asn205Ser mutation in the connexin 32 gene. J.Neurol Neurosurg.Psychiatry 1999; 66: 202-206.

Balice-Gordon RJ, Bone LJ, Scherer SS. Functional gap junctions in the schwann cell myelin sheath. J.Cell Biol. 1998; 142: 1095-1104.

Baxter RV, Ben Othmane, K, Rochelle JM *et al.* Ganglioside-induced differentiation-associated protein-1 is mutant in Charcot-Marie-Tooth disease type 4A/8q21. Nat.Genet. 30, 21-22. 2002.

Bejaoui K, McKenna-Yasek D, Hosler BA *et al.* Confirmation of linkage of type 1 hereditary sensory neuropathy to human chromosome 9q22. Neurology 1999; 52: 510-515.

Bell J. "Peroneal Type of Progressive Muscular Atrophy" Fisher, R. A. ed. Treasury of Human Inheritance: Nervous Diseases and Muscular Dystrophies. 4, pp 69-pp 139. 1935. London, Cambridge University Press.

Ben Othmane K, Hentati F, Lennon F *et al.* Linkage of a locus (CMT4A) for autosomal recessive Charcot-Marie-Tooth disease to chromosome 8q. Hum.Mol.Genet. 1993a; 2: 1625-1628.

Ben Othmane K, Middleton LT, Loprest LJ *et al.* Localization of a gene (CMT2A) for autosomal dominant Charcot-Marie- Tooth disease type 2 to chromosome 1p and evidence of genetic heterogeneity. Genomics 1993b; 17: 370-375.

Bennett MV, Barrio LC, Bargiello TA *et al.* Gap junctions: new tools, new answers, new questions. Neuron 1991; 6: 305-320.

Bergoffen J, Trofatter J, Pericak-Vance MA *et al.* Linkage localization of X-linked Charcot-Marie-Tooth disease. Am.J.Hum.Genet. 1993a; 52: 312-318.

Bergoffen J, Scherer SS, Wang S *et al.* Connexin mutations in X-linked Charcot-Marie-Tooth disease. Science 1993b; 262: 2039-2042.

Bird SJ, Sladky JT. Corticosteroid-responsive dominantly inherited neuropathy in childhood. Neurology 1991; 41: 437-439.

Birouk N, LeGuern E, Maisonobe T *et al.* X-linked Charcot-Marie-Tooth disease with connexin 32 mutations: clinical and electrophysiologic study. Neurology 1998; 50: 1074-1082.

Blumenfeld A, Slaugenhaupt SA, Liebert CB *et al.* Precise genetic mapping and haplotype analysis of the familial dysautonomia gene on human chromosome 9q31. Am.J.Hum.Genet. 1999; 64: 1110-1118.

Bolino A, Brancolini V, Bono F *et al.* Localization of a gene responsible for autosomal recessive demyelinating neuropathy with focally folded myelin sheaths to chromosome 11q23 by homozygosity mapping and haplotype sharing. Hum.Mol.Genet. 1996; 5: 1051-1054.

Bolino A, Levy ER, Muglia M *et al.* Genetic refinement and physical mapping of the CMT4B gene on chromosome 11q22. Genomics 2000a; 63: 271-278.

Bolino A, Muglia M, Conforti FL *et al.* Charcot-Marie-Tooth type 4B is caused by mutations in the gene encoding myotubularin-related protein-2. Nat.Genet. 2000b; 25: 17-19.

Bone LJ, Deschenes SM, Balice-Gordon RJ *et al.* Connexin32 and X-linked Charcot-Marie-Tooth disease. Neurobiol.Dis. 1997; 4: 221-230.

Bouhouche A, Benomar A, Birouk N *et al.* A locus for an axonal form of autosomal recessive Charcot-Marie-Tooth disease maps to chromosome 1q21.2-q21.3. Am.J.Hum.Genet. 1999; 65: 722-727.

Bradley WG. Disorders of Peripheral Nerves. pp 184-pp 188. 1974. Oxford, Blackwell.

Brody IA, Wilkins RH. Charcot-Marie-Tooth disease. Arch Neurol 1967; 17: 552-557.

Brown WJ, Beauchemin JA Linde LM. A neuropathological study of familial dysautonmia (Reily-Day syndrome) in siblings. J.Neurol.Neurosurg.Psychiatry 27, 131-139. 1964.

Bukau B, Horwich AL. The Hsp70 and Hsp60 chaperone machines. Cell 1998; 92: 351-366.

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

Campbell AMG, Hoffman HL. Sensory radicular neuropathy associated with muscle wasting in two cases. Brain 87, 67-74. 1964.

Castro C, Gomez-Hernandez JM, Silander K *et al.* Altered formation of hemichannels and gap junction channels caused by C- terminal connexin-32 mutations. J.Neurosci. 1999; 19: 3752-3760.

Chance PF, Alderson MK, Leppig KA *et al.* DNA deletion associated with hereditary neuropathy with liability to pressure palsies. Cell 1993; 72: 143-151.

Chapon F, Latour P, Diraison P *et al.* Axonal phenotype of Charcot-Marie-Tooth disease associated with a mutation in the myelin protein zero gene. J.Neurol.Neurosurg.Psychiatry 1999; 66: 779-782.

Chimelli L, Scaravilli F. The abnormal development of the gracile nucleus in the neurological mutant rat mf. Brain Res. 1986a; 394: 193-205.

Chimelli L, Scaravilli F. The development and pathogenesis of the sensory neuropathy in the mutant rat mf. Brain 1986b; 109 (Pt 4): 629-647.

Chua SC, Jr., Chung WK, Wu-Peng XS *et al.* Phenotypes of mouse diabetes and rat fatty due to mutations in the OB (leptin) receptor. Science 1996; 271: 994-996.

Cohen L, Henzel WJ, Baeuerle PA. IKAP is a scaffold protein of the IkappaB kinase complex. Nature 1998; 395: 292-296.

Collins FS. Positional cloning moves from perditional to traditional. Nat.Genet. 1995; 9: 347-350.

Collins J, Hohn B. Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. Proc.Natl.Acad.Sci.U.S.A 1978; 75: 4242-4246.

Cowchock FS, Duckett SW, Streletz LJ *et al.* X-linked motor-sensory neuropathy type-II with deafness and mental retardation: a new disorder. Am.J.Med Genet. 1985; 20: 307-315.

Cuesta A, Pedrola L, Sevilla T *et al.* The gene encoding ganglioside-induced differentiation-associated protein 1 is mutated in axonal Charcot-Marie-Tooth type 4A disease. Nat.Genet. 2002; 30: 22-25.

D'Urso D, Ehrhardt P, Muller HW. Peripheral myelin protein 22 and protein zero: a novel association in peripheral nervous system myelin. J.Neurosci. 1999; 19: 3396-3403.

Dakour J, Jimbow K, Vinayagamoorthy T *et al.* Characterization of melanosome-associated proteins by establishment of monoclonal antibodies and immunoscreening of a melanoma cDNA library through an anti-melanosome antibody. Melanoma Res. 1993; 3: 331-336.

Davidson MK, Lindsey JR, Davis JK. Requirements and selection of an animal model. Isr.J.Med.Sci. 1987; 23: 551-555.

Davis CJ, Bradley WG, Madrid R. The peroneal muscular atrophy syndrome: clinical, genetic, electrophysiological and nerve biopsy studies. I. Clinical, genetic and electrophysiological findings and classification. J.Genet.Hum. 1978; 26: 311-349.

Dawkins JL, Hulme DJ, Brahmbhatt SB *et al.* Mutations in SPTLC1, encoding serine palmitoyltransferase, long chain base subunit-1, cause hereditary sensory neuropathy type I. Nat.Genet. 2001; 27: 309-312.

Dayhoff MO, Schwartz RM, Chen HR *et al.* Nucleic acid sequence database. DNA 1981; 1: 51-58.

De Jonghe P, Timmerman V, Ceuterick C *et al.* The Thr124Met mutation in the peripheral myelin protein zero (MPZ) gene is associated with a clinically distinct Charcot-Marie-Tooth phenotype. Brain 1999; 122: 281-290.

De Jonghe P, Mersivanova I, Nelis E *et al.* Further evidence that neurofilament light chain gene mutations can cause Charcot-Marie-Tooth disease type 2E. Ann.Neurol 2001; 49: 245-249.

Delague V, Bareil C, Tuffery S *et al.* Mapping of a new locus for autosomal recessive demyelinating Charcot- Marie-Tooth disease to 19q13.1-13.3 in a large consanguineous Lebanese family: exclusion of MAG as a candidate gene. Am.J.Hum.Genet. 2000; 67: 236-243.

DeLeon GA. Progressive ventral sensory loss in sensory radicular neuropathy and hypertrophic neuritis. Johns Hopkins Med.J. 125, 53-61. 1969.

Denny-Brown D. Hereditary sensory radicular neuropathy. J.Neurol.Neurosurg.Psychiatry 14, 237-252. 1951.

Donaghy M, Hakin RN, Bamford JM *et al.* Hereditary sensory neuropathy with neurotrophic keratitis. Description of an autosomal recessive disorder with a selective reduction of small myelinated nerve fibres and a discussion of the classification of the hereditary sensory neuropathies. Brain 1987; 110: 563-583.

Donaghy M, Sisodiya SM, Kennett R *et al.* Steroid responsive polyneuropathy in a family with a novel myelin protein zero mutation. J.Neurol Neurosurg. Psychiatry 2000; 69: 799-805.

Dubourg O, Tardieu S, Birouk N *et al.* Clinical, electrophysiological and molecular genetic characteristics of 93 patients with X-linked Charcot-Marie-Tooth disease. Brain 2001; 124: 1958-1967.

Dyck PJ, Kennel AJ, Magal IV *et al.* A Virginia kinship with hereditary sensory neuropathy: peroneal muscular atrophy and pes cavus. Mayo Clin.Proc. 1965,40: 685-690.

Dyck PJ, Lambert EH. Lower motor and primary sensory neuron diseases with peroneal muscular atrophy. II. Neurologic, genetic, and electrophysiologic findings in various neuronal degenerations. Arch Neurol. 1968a; 18: 619-625.

Dyck PJ, Lambert EH. Lower motor and primary sensory neuron diseases with peroneal muscular atrophy. I. Neurologic, genetic, and electrophysiologic findings in hereditary polyneuropathies. Arch Neurol. 1968b; 18: 603-618.

Dyck PJ, Mellinger JF, Reagan TJ *et al.* Not 'indifference to pain' but varieties of hereditary sensory and autonomic neuropathy. Brain 1983; 106: 373-390.

Dyck PJ. Neuronal atrophy and degernation predominantly affecting sensory and autonomic neurons. Dyck PJ, Thomas PK, Griffin JW, Low PA, and Poduslo JF. Peripheral neuropathy. 2 (3rd ed.), pp 1065-pp 1093. 1993a. Philadelphia, W.B. Saunders Company.

Dyck PJ, Chance P, Carney JA. Hereditary motor and sensory neuropathies. Dyck PJ, Thomas PK Griffin JW Low PA Poduslo JF. Peripheral Neuropathy. 2 (3rd ed.), pp1094-pp1136. 1993b. Philadelphia, Saunders.

Elliott JL, Kwon JM, Goodfellow PJ *et al.* Hereditary motor and sensory neuropathy IIB: clinical and electrodiagnostic characteristics. Neurology 1997; 48: 23-28.

Ellis HM, Horvitz HR. Genetic control of programmed cell death in the nematode C. elegans. Cell 1986; 44: 817-829.

England A.C. and Denny-Brown D. Severe sensory changes, and trophic disorder, in peroneal muscular atrophy (Charcot-Marie-Tooth type). Arch Neurol Psychiatry 67, 1-14. 1952.

Ervin FR, Sterback RA. Hereditary insensitivity to pain. Trans.Am.Neurol. Assoc. 86, 70-74. 1960.

Ferrière G, Guzzetta F, Kulakowski S *et al.* Nonprogressive type II hereditary sensory autonomic neuropathy: a homogeneous clinicopathologic entity. J.Child Neurol. 7, 364-370. 1992.

Finkelstein R, Boncinelli E. From fly head to mammalian forebrain: the story of otd and Otx. Trends Genet. 1994; 10: 310-315.

Gabreels-Festen A, Gabreels F. Hereditary demyelinating motor and sensory neuropathy. Brain Pathol. 1993; 3: 135-146.

Gabreels-Festen AA, Gabreels FJ, Jennekens FG *et al.* The status of HMSN type III. Neuromuscul.Disord. 1994; 4: 63-69.

Gabreels-Festen A, van Beersum S, Eshuis L *et al.* Study on the gene and phenotypic characterisation of autosomal recessive demyelinating motor and sensory neuropathy (Charcot-Marie- Tooth disease) with a gene locus on chromosome 5q23-q33. J.Neurol.Neurosurg.Psychiatry 1999; 66: 569-574.

Gao Y, Thomas JO, Chow RL *et al.* A cytoplasmic chaperonin that catalyzes beta-actin folding. Cell 1992; 69: 1043-1050.

Geffen R, Goldstein RS. Rescue of sensory ganglia that are programmed to degenerate in normal development: evidence that NGF modulates proliferation of DRG cells in vivo. Dev.Biol. 1996; 178: 51-62.

Gemignani F, Marbini A. Charcot-Marie-Tooth disease (CMT): distinctive phenotypic and genotypic features in CMT type 2. J.Neurol Sci. 2001; 184: 1-9.

Gething MJ, Sambrook J. Protein folding in the cell. Nature 1992; 355: 33-45.

Giaccai L. Familial and sporadic neurogenic acro-osteolysis. Acta Radiol. 1952; 38, 17-29.

Gill TJ. The use of randomly bred and genetically defined animals in biomedical research. Am.J.Pathol. 1980; 101: S21-S32.

Gill TJ, Smith GJ, Wissler RW *et al.* The rat as an experimental animal. Science 1989; 245: 269-276.

Gomez-Fabre PM, Helou K, Stahl F. Predictions based on the rat-mouse comparative map provide mapping information on over 6000 new rat genes. Mamm.Genome 2002; 13: 189-193.

Gordon N. Apoptosis (programmed cell death) and other reasons for elimination of neurons and axons. Brain Dev. 1995; 17: 73-77.

Granger ME. Sensory neuropathy with ulcerative mutilating acropathy. Neurology , 725. 1960.

Guenal I, Risler Y, Mignotte B. Down-regulation of actin genes precedes microfilament network disruption and actin cleavage during p53-mediated apoptosis. J.Cell Sci. 1997; 110 (Pt 4): 489-495.

Guilbot A, Kessali M, Ravise N *et al.* The autosomal recessive form of CMT disease linked to 5q31-q33. Ann.N.Y.Acad.Sci. 1999; 883: 453-456.

Gusella JF. Location cloning strategy for characterizing genetic defects in Huntington's disease and Alzheimer's disease. FASEB J. 1989; 3: 2036-2041.

Hahn AF, Brown WF, Koopman WJ *et al.* X-linked dominant hereditary motor and sensory neuropathy. Brain 1990; 113: 1511-1525.

Hahn AF, Ainsworth PJ, Naus CC *et al.* Clinical and pathological observations in men lacking the gap junction protein connexin 32. Muscle Nerve 2000; 999: S39-S48.

Hall AK. Thymosin beta-10 accelerates apoptosis. Cell Mol.Biol.Res. 1995; 41: 167-180.

Harding AE, Thomas PK. The clinical features of hereditary motor and sensory neuropathy types I and II. Brain 1980; 103: 259-280.

Harding AE, Reilly M. Molecular genetics of inherited neuropathies. Ashbury AK, Thomas PK. Peripheral Nerve Disorder. 2, pp 118-pp 139. 1995. London, Butterworth-Heinmann.

Hartl FU, Hayer-Hartl M. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 2002; 295: 1852-1858.

Hayasaka K, Nanao K, Tahara M *et al.* Isolation and sequence determination of cDNA encoding the major structural protein of human peripheral myelin. Biochem.Biophys.Res.Commun. 1991; 180: 515-518.

Hayasaka K, Himoro M, Sato W *et al.* Charcot-Marie-Tooth neuropathy type 1B is associated with mutations of the myelin P0 gene. Nat.Genet. 1993a; 5: 31-34.

Hayasaka K, Himoro M, Sawaishi Y *et al.* De novo mutation of the myelin P0 gene in Dejerine-Sottas disease (hereditary motor and sensory neuropathy type III). Nat.Genet. 1993b; 5: 266-268.

Heller I.H. and Robb P. Hereditary sensory neuropathy. Neurology 1995;5, 15-29.

Herberg LJ, Rose IC, Davison F *et al.* Spontaneous epileptiform seizures but increased resistance to kindled seizures in a mutant Sprague-Dawley rat (mf/mf). Pharmacol.Biochem.Behav. 1997; 58: 993-1001.

Horvitz HR. Genetic control of programmed cell death in the nematode Caenorhabditis elegans. Cancer Res. 1999; 59: 1701s-1706s.

Houlden H, King RH, Hashemi-Nejad A *et al.* A novel TRK A (NTRK1) mutation associated with hereditary sensory and autonomic neuropathy type V. Ann.Neurol. 2001; 49: 521-525.

Hunot S, Flavell RA. Apoptosis. Death of a monopoly? Science 2001; 292: 865-866.

Huxley C, Passage E, Manson A *et al.* Construction of a mouse model of Charcot-Marie-Tooth disease type 1A by pronuclear injection of human YAC DNA. Hum.Mol.Genet. 1996; 5: 563-569.

Huxley C, Passage E, Robertson AM *et al.* Correlation between varying levels of PMP22 expression and the degree of demyelination and reduction in nerve conduction velocity in transgenic mice. Hum.Mol.Genet. 1998; 7: 449-458.

Hynes GM, Willison KR. Individual subunits of the eukaryotic cytosolic chaperonin mediate interactions with binding sites located on subdomains of beta-actin. J.Biol.Chem. 2000; 275: 18985-18994.

Indo Y, Tsuruta M, Hayashida Y *et al.* Mutations in the TRKA/NGF receptor gene in patients with congenital insensitivity to pain with anhidrosis. Nat.Genet. 1996; 13: 485-488.

Ionasescu V, Searby C, Ionasescu R. Point mutations of the connexin32 (GJB1) gene in X-linked dominant Charcot-Marie-Tooth neuropathy. Hum.Mol.Genet. 1994; 3: 355-358.

Ionasescu VV, Kimura J, Searby CC *et al.* Dejerine-Sottas neuropathy family with a gene mapped on chromosome 8. Muscle Nerve 1996a; 19: 319-323.

Ionasescu V, Searby C, Sheffield VC *et al.* Autosomal dominant Charcot-Marie-Tooth axonal neuropathy mapped on chromosome 7p (CMT2D). Hum.Mol.Genet. 1996b; 5: 1373-1375.

Ishii N, Kawaguchi H, Miyakawa K *et al.* Congenital sensory neuropathy with anhidrosis. Arch Dermatol. 1988; 124: 564-566.

Ismailov SM, Fedotov VP, Dadali EL *et al.* A new locus for autosomal dominant Charcot-Marie-Tooth disease type 2 (CMT2F) maps to chromosome 7q11-q21. Eur.J.Hum.Genet. 2001; 9: 646-650.

Jacobs JM, Scaravilli F, Duchen LW, Mertin J. A new neurological rat mutant "mutilated foot". J.Anat. 1981; 132: 525-543.

Janssen EA, Kemp S, Hensels GW *et al.* Connexin32 gene mutations in X-linked dominant Charcot-Marie-Tooth disease (CMTX1). Hum.Genet. 1997; 99: 501-505.

Jenkins NA, Copeland NG, Taylor BA, Lee BK. Dilute (d) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. Nature 1981; 293: 370-374.

Johnson EM, Jr., Deckwerth TL. Molecular mechanisms of developmental neuronal death. Annu.Rev Neurosci. 1993; 16: 31-46.

Johnson RH, Spalding JMK. Progressive sensory neuropathy in children. J.Neurol.Neurosurg.Psychiatry 1964; 27, 125-130.

Kalaydjieva L, Hallmayer J, Chandler D *et al* . Gene mapping in Gypsies identifies a novel demyelinating neuropathy on chromosome 8q24. Nat.Genet. 1996; 14: 214-217.

Kalaydjieva L, Gresham D, Gooding R *et al.* N-myc downstream-regulated gene 1 is mutated in hereditary motor and sensory neuropathy-Lom. Am.J.Hum.Genet. 2000; 67: 47-58.

Kayalar C, Ord T, Testa MP *et al.* Cleavage of actin by interleukin 1 beta-converting enzyme to reverse DNase I inhibition. Proc.Natl.Acad.Sci.U.S.A 1996; 93: 2234-2238.

Kennerson ML, Zhu D, Gardner RJ *et al.* Dominant intermediate Charcot-Marie-Tooth neuropathy maps to chromosome 19p12-p13.2. Am.J.Hum.Genet. 2001; 69: 883-888.

Kim SA, Taylor GS, Torgersen KM *et al.* Myotubularin and MTMR2, phosphatidylinositol 3-phosphatases mutated in myotubular myopathy and type 4B Charcot-Marie-Tooth disease. J.Biol.Chem. 2002; 277: 4526-4531.

King RH, Tournev I, Colomer J *et al.* Ultrastructural changes in peripheral nerve in hereditary motor and sensory neuropathy-Lom. Neuropathol.Appl.Neurobiol. 1999; 25: 306-312.

Klein R, Jing SQ, Nanduri V, O'Rourke E, Barbacid M. The trk proto-oncogene encodes a receptor for nerve growth factor. Cell 1991; 65: 189-197.

Kobayashi T, Mitani H, Takahashi R *et al.* Transgenic rescue from embryonic lethality and renal carcinogenesis in the Eker rat model by introduction of a wild-type Tsc2 gene. Proc.Natl.Acad.Sci.U.S.A 1997; 94: 3990-3993.

Koenig M, Hoffman EP, Bertelson CJ *et al.* Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell 1987; 50: 509-517.

Krappmann D, Hatada EN, Tegethoff S *et al.* The I kappa B kinase (IKK) complex is tripartite and contains IKK gamma but not IKAP as a regular component. J.Biol.Chem. 2000; 275: 29779-29787.

Kubota H, Yokota S, Yanagi H *et al.* Structures and co-regulated expression of the genes encoding mouse cytosolic chaperonin CCT subunits. Eur.J.Biochem. 1999; 262: 492-500.

Kwon BS, Haq AK, Pomerantz SH *et al.* Isolation and sequence of a cDNA clone for human tyrosinase that maps at the mouse c-albino locus. Proc.Natl.Acad.Sci.U.S.A 1987; 84: 7473-7477.

Kwon JM, Elliott JL, Yee WC *et al.* Assignment of a second Charcot-Marie-Tooth type II locus to chromosome 3q. Am.J.Hum.Genet. 1995; 57: 853-858.

Lander ES, Linton LM, Birren B *et al.* Initial sequencing and analysis of the human genome. Nature 2001; 409: 860-921.

Latour P, Levy N, Paret M *et al.* Mutations in the X-linked form of Charcot-Marie-Tooth disease in the French population. Neurogenetics. 1997; 1: 117-123.

Leal A, Morera B, Del Valle G *et al.* A second locus for an axonal form of autosomal recessive Charcot-Marie- Tooth disease maps to chromosome 19q13.3. Am.J.Hum.Genet. 2001; 68: 269-274.

Lebo RV, Chance PF, Dyck PJ *et al.* Chromosome 1 Charcot-Marie-Tooth disease (CMT1B) locus in the Fc gamma receptor gene region. Hum.Genet. 1991; 88: 1-12.

Lee KF, Li E, Huber LJ *et al.* Targeted mutation of the gene encoding the low affinity NGF receptor p75 leads to deficits in the peripheral sensory nervous system. Cell 1992; 69: 737-749.

Lee M, Brennan A, Blanchard A *et al.* P0 is constitutively expressed in the rat neural crest and embryonic nerves and is negatively and positively regulated by axons to generate non-myelin-forming and myelin-forming Schwann cells, respectively. Mol.Cell Neurosci. 1997; 8: 336-350.

Lee MJ, Calle E, Brennan A *et al.* In early development of the rat mRNA for the major myelin protein P(0) is expressed in nonsensory areas of the embryonic inner ear, notochord, enteric nervous system, and olfactory ensheathing cells. Dev.Dyn. 2001; 222: 40-51.

LeGuern E, Gouider R, Ravise N *et al.* A de novo case of hereditary neuropathy with liability to pressure palsies (HNPP) of maternal origin: a new mechanism for deletion in 17p11.2? Hum.Mol.Genet. 1996a; 5: 103-106.

LeGuern E, Guilbot A, Kessali M *et al.* Homozygosity mapping of an autosomal recessive form of demyelinating Charcot-Marie-Tooth disease to chromosome 5q23-q33. Hum.Mol.Genet. 1996b; 5: 1685-1688.

Lemke G, Lamar E, Patterson J. Isolation and analysis of the gene encoding peripheral myelin protein zero. Neuron 1988; 1: 73-83.

Letourneau PC. The cytoskeleton in nerve growth cone motility and axonal pathfinding. Perspect. Dev.Neurobiol. 1996; 4: 111-123.

Levi-Montalcini R, Angeletti PU. Second symposium on catecholamines. Modification of sympathetic function. Immunosympathectomy. Pharmacol.Rev 1966; 18: 619-628.

Lin KP, Kwan SY, Chen SY *et al.* Generalized neuropathy in Taiwan: an etiologic survey. Neuroepidemiology 1993; 12: 257-261.

Lindsey, J. R. The Laboratory Rat. Baker, H. J. Lindsey J. R. & Weisbroth S. H. 1-36. 1979. New York, Academic.

Llorca O, McCormack EA, Hynes G *et al.* Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. Nature 1999; 402: 693-696.

Llorca O, Martin-Benito J, Grantham J *et al.* The 'sequential allosteric ring' mechanism in the eukaryotic chaperonin- assisted folding of actin and tubulin. EMBO J. 2001; 20: 4065-4075.

Lopes J, Vandenberghe A, Tardieu S *et al.* Sex-dependent rearrangements resulting in CMT1A and HNPP. Nat.Genet. 1997; 17: 136-137.

Lopes J, Ravise N, Vandenberghe A *et al.* Fine mapping of de novo CMT1A and HNPP rearrangements within CMT1A-REPs evidences two distinct sex-dependent mechanisms and candidate sequences involved in recombination. Hum.Mol.Genet. 1998; 7: 141-148.

Low PA, Burke WJ, McLeod JG. Congenital sensory neuropathy with selective loss of small myelinated fibers. Ann.Neurol. 1978; 3: 179-182.

Lupski JR, Oca-Luna RM, Slaugenhaupt S *et al* . DNA duplication associated with Charcot-Marie-Tooth disease type 1A. Cell 1991; 66: 219-232.

Maayan C, Kaplan E, Shachar S *et al.* Incidence of familial dysautonomia in Israel 1977-1981. Clin.Genet. 1987; 32: 106-108.

Mandell AJ, Smith CK. Hereditary sensory radicular neuropathy. Neurology 1960; 10, 627-630.

Mardy S, Miura Y, Endo F *et al.* Congenital insensitivity to pain with anhidrosis: novel mutations in the TRKA (NTRK1) gene encoding a high-affinity receptor for nerve growth factor. Am.J.Hum.Genet. 1999; 64: 1570-1579.

Marques W, Sweeney JG, Wood NW *et al.* Central nervous system involvement in a novel connexin 32 mutation affecting identical twins. J.Neurol Neurosurg.Psychiatry 1999; 66: 803-804.

Marrosu MG, Vaccargiu S, Marrosu G *et al.* Charcot-Marie-Tooth disease type 2 associated with mutation of the myelin protein zero gene. Neurology 1998; 50: 1397-1401.

Martin-Subero JI, Gesk S, Harder L *et al.* Recurrent involvement of the REL and BCL11A loci in classical Hodgkin lymphoma. Blood 2002; 99: 1474-1477.

Mashima T, Naito M, Fujita N *et al.* Identification of actin as a substrate of ICE and an ICE-like protease and involvement of an ICE-like protease but not ICE in VP-16-induced U937 apoptosis. Biochem.Biophys.Res.Commun. 1995; 217: 1185-1192.

Mashima T, Naito M, Noguchi K *et al.* Actin cleavage by CPP-32/apopain during the development of apoptosis. Oncogene 1997; 14: 1007-1012.

McCormack EA, Rohman MJ, Willison KR. Mutational screen identifies critical amino acid residues of beta-actin mediating interaction between its folding intermediates and eukaryotic cytosolic chaperonin CCT. J.Struct.Biol. 2001; 135: 185-197.

Meier C, Moll C. Hereditary neuropathy with liability to pressure palsies. Report of two families and review of the literature. J.Neurol 1982; 228: 73-95.

Melki R, Vainberg IE, Chow RL *et al.* Chaperonin-mediated folding of vertebrate actin-related protein and gamma-tubulin. J.Cell Biol. 1993; 122: 1301-1310.

Mersiyanova IV, Perepelov AV, Polyakov AV *et al.* A new variant of Charcot-Marie-Tooth disease type 2 is probably the result of a mutation in the neurofilament-light gene. Am.J.Hum.Genet. 2000; 67: 37-46.

Miller R. Linkage mapping of plant and animal genomes. Dear P.H. Genome mapping---a practical approach. pp27-pp47. 1997. New York, United States, Oxford University Press Inc.

Mukasa T, Urase K, Momoi MY *et al.* Specific expression of CPP32 in sensory neurons of mouse embryos and activation of CPP32 in the apoptosis induced by a withdrawal of NGF. Biochem.Biophys.Res.Commun. 1997; 231: 770-774.

Murray TJ. Congenital sensory neuropathy. Brain 1973; 96: 387-394.

Nabetani A, Hatada I, Morisaki H, Mukai T. Chromosomal assignment and imprinting tests for the mouse delta subunit of the cytosolic chaperonin containing TCP-1 (Cct4) gene to proximal chromosome 11. Genomics 1996; 34: 246-249.

Nagarajan R, Svaren J, Le N *et al.* EGR2 mutations in inherited neuropathies dominant-negatively inhibit myelin gene expression. Neuron 2001; 30: 355-368.

Nelis E, Van Broeckhoven C, De Jonghe P *et al.* Estimation of the mutation frequencies in Charcot-Marie-Tooth disease type 1 and hereditary neuropathy with liability to pressure palsies: a European collaborative study. Eur.J.Hum.Genet. 1996; 4: 25-33.

Nicholson G, Corbett A. Slowing of central conduction in X-linked Charcot-Marie-Tooth neuropathy shown by brain stem auditory evoked responses. J.Neurol Neurosurg.Psychiatry 1996a; 61: 43-46.

Nicholson GA, Dawkins JL, Blair IP *et al.* The gene for hereditary sensory neuropathy type I (HSN-I) maps to chromosome 9q22.1-q22.3. Nat.Genet. 1996b; 13: 101-104.

Nicholson GA, Yeung L, Corbett A. Efficient neurophysiologic selection of X-linked Charcot-Marie-Tooth families: ten novel mutations. Neurology 1998; 51: 1412-1416.

Nilsson S, Helou K, Walentinsson A *et al.* Rat-mouse and rat-human comparative maps based on gene homology and high-resolution zoo-FISH. Genomics 2001; 74: 287-298.

Nordborg C, Conradi N, Sourander P *et al.* A new type of non-progressive sensory neuropathy in children with atypical dysautonomia. Acta Neuropathol.(Berl) 1981; 55: 135-141.

Nukada H, Pollock M, Haas LF. The clinical spectrum and morphology of type II hereditary sensory neuropathy. Brain 1982; 105 (Pt 4): 647-665.

Ohta M, Ellefson RD, Lambert EH *et al.* Hereditary sensory neuropathy, type II: clinial, electrophysiologic, histologic, and biochemical studies of a Quebec kinship. Arch Neurol. 1973; 29, 23-37.

Oppenheim RW. Cell death during development of the nervous system. Annu.Rev Neurosci. 1991; 14: 453-501.

Othmane KB, Johnson E, Menold M *et al.* Identification of a new locus for autosomal recessive Charcot-Marie- Tooth disease with focally folded myelin on chromosome 11p15. Genomics 1999; 62: 344-349.

Ott J. Analysis of Human Genetic Linkage. pp. 302. 1991. Baltimore, MD, John Hopkins University Press.

Ouvrier RA, McLeod JG, Conchin TE. The hypertrophic forms of hereditary motor and sensory neuropathy. A study of hypertrophic Charcot-Marie-Tooth disease (HMSN type I) and Dejerine-Sottas disease (HMSN type III) in childhood. Brain 1987; 110 (Pt 1): 121-148.

Pareyson D. Charcot-marie-tooth disease and related neuropathies: molecular basis for distinction and diagnosis. Muscle Nerve 1999; 22: 1498-1509.

Parman Y, Plante-Bordeneuve V, Guiochon-Mantel A *et al.* Recessive inheritance of a new point mutation of the PMP22 gene in Dejerine-Sottas disease. Ann.Neurol. 1999; 45: 518-522.

Parrish JE, Nelson DL. Methods for finding genes. A major rate-limiting step in positional cloning. Genet.Anal.Tech.Appl. 1993; 10: 29-41.

Pearson J, Dancis J, Axelrod FB *et al.* The sural nerve in familial dysautonomia. J.Neuropath.Exp.Neurol. 1975; 34, 413-424.

Peracchia C, Wang XG, Peracchia LL. Is the chemical gate of connexins voltage sensitive? Behavior of Cx32 wild-type and mutant channels. Am.J.Physiol 1999; 276: C1361-C1373.

Pinksy L, DiGeorge AM. Congenital familial sensory neuropathy with anhidrosis. J.Pediat. 1966; 68, 1-13.

Polzar B, Peitsch MC, Loos R *et al.* Overexpression of deoxyribonuclease I (DNase I) transfected into COS- cells: its distribution during apoptotic cell death. Eur.J.Cell Biol. 1993; 62: 397-405.

Previtali SC, Quattrini A, Fasolini M *et al.* Epitope-tagged P(0) glycoprotein causes Charcot-Marie-Tooth-like neuropathy in transgenic mice. J.Cell Biol. 2000; 151: 1035-1046.

Priest JM, Fischbeck KH, Nouri N *et al.* A locus for axonal motor-sensory neuropathy with deafness and mental retardation maps to Xq24-q26. Genomics 1995; 29: 409-412.

Raeymaekers P, Timmerman V, Nelis E *et al.* Duplication in chromosome 17p11.2 in Charcot-Marie-Tooth neuropathy type 1a (CMT 1a). The HMSN Collaborative Research Group. Neuromuscul.Disord. 1991; 1: 93-97.

Rafel E, Alberca R, Bautista J, Navarrete M, Lazo J. Congenital insensitivity to pain with anhidrosis. Muscle Nerve 1980; 3: 216-220.

Reiter LT, Murakami T, Koeuth T *et al.* A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element. Nat.Genet. 1996; 12: 288-297.

Reiter LT, Hastings PJ, Nelis E *et al.* Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. Am.J.Hum.Genet. 1998; 62: 1023-1033.

Ri Y, Ballesteros JA, Abrams CK *et al.* The role of a conserved proline residue in mediating conformational changes associated with voltage gating of Cx32 gap junctions. Biophys.J. 1999; 76: 2887-2898.

Riley CM, Day RL, Greeley DM *et al.* Central autonomic dysfunction with defective lacrimation: report of five cases. Pediatrics 1949; 3, 468-478.

Roa BB, Warner LE, Garcia CA *et al.* Myelin protein zero (MPZ) gene mutations in nonduplication type 1 Charcot-Marie-Tooth disease. Hum.Mutat. 1996; 7: 36-45.

Rogers T, Chandler D, Angelicheva D *et al.* A novel locus for autosomal recessive peripheral neuropathy in the EGR2 region on 10q23. Am.J.Hum.Genet. 2000; 67: 664-671.

Rommelaere H, Van Troys M, Gao Y *et al.* Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven related subunits. Proc.Natl.Acad.Sci.U.S.A 1993; 90: 11975-11979.

Roobol A, Sahyoun ZP, Carden MJ. Selected subunits of the cytosolic chaperonin associate with microtubules assembled in vitro. J.Biol.Chem. 1999; 274: 2408-2415.

Rosemberg S, Marie SK, Kliemann S. Congenital insensitivity to pain with anhidrosis (hereditary sensory and autonomic neuropathy type IV). Pediatr.Neurol. 1994; 11: 50-56.

Rossi A, Paradiso C, Cioni R *et al.* Charcot-Marie-Tooth disease: study of a large kinship with an intermediate form. J.Neurol. 1985; 232: 91-98.

Rubtsova SN, Kondratov RV, Kopnin PB *et al.* Disruption of actin microfilaments by cytochalasin D leads to activation of p53. FEBS Lett. 1998; 430: 353-357.

Sandre-Giovannoli A, Chaouch M, Kozlov S *et al.* Homozygous Defects in LMNA, Encoding Lamin A/C Nuclear-Envelope Proteins, Cause Autosomal Recessive Axonal Neuropathy in Human (Charcot- Marie-Tooth Disorder Type 2) and Mouse. Am.J.Hum.Genet. 2002; 70.

Satterwhite E, Sonoki T, Willis TG *et al.* The BCL11 gene family: involvement of BCL11A in lymphoid malignancies. Blood 2001; 98: 3413-3420.

Scherer SS, Chance PF. Myelin genes: getting the dosage right. Nat.Genet. 1995; 11: 226-228.

Scherer SS, Xu YT, Nelles E *et al.* Connexin32-null mice develop demyelinating peripheral neuropathy. Glia 1998; 24: 8-20.

Senderek J, Bergmann C, Quasthoff S *et al.* X-linked dominant Charcot-Marie-Tooth disease: nerve biopsies allow morphological evaluation and detection of connexin32 mutations (Arg15Trp, Arg22Gln). Acta Neuropathol.(Berl) 1998; 95: 443-449.

Senderek J, Hermanns B, Bergmann C *et al.* X-linked dominant Charcot-Marie-Tooth neuropathy: clinical, electrophysiological, and morphological phenotype in four families with different connexin32 mutations(1). J.Neurol Sci. 1999; 167: 90-101.

Sereda M, Griffiths I, Puhlhofer A *et al.* A transgenic rat model of Charcot-Marie-Tooth disease. Neuron 1996; 16: 1049-1060.

Shapiro L, Doyle JP, Hensley P *et al.* Crystal structure of the extracellular domain from P0, the major structural protein of peripheral nerve myelin. Neuron 1996; 17: 435-449.

Shizuya H, Birren B, Kim UJ *et al.* Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc.Natl.Acad.Sci.U.S.A 1992; 89: 8794-8797.

Silander K, Meretoja P, Pihko H *et al.* Screening for connexin 32 mutations in Charcot-Marie-Tooth disease families with possible X-linked inheritance. Hum.Genet. 1997; 100: 391-397.

Silverman FN, Gilden JJ. Congenital insensitivity to pain, a neurologic syndrome with bizarre skeletal lesions. Radiology 1959; 72, 176-190.

Slaugenhaupt SA, Blumenfeld A, Gill SP *et al* . Tissue-specific expression of a splicing mutation in the IKBKAP gene causes familial dysautonomia. Am.J.Hum.Genet. 2001; 68: 598-605.

Smeyne RJ, Klein R, Schnapp A *et al.* Severe sensory and sympathetic neuropathies in mice carrying a disrupted Trk/NGF receptor gene. Nature 1994; 368: 246-249.

Spiryda LB. Myelin protein zero and membrane adhesion. J.Neurosci.Res. 1998; 54: 137-146.

Stephenson DA, Mercola M, Anderson E *et al.* Platelet-derived growth factor receptor alpha-subunit gene (Pdgfra) is deleted in the mouse patch (Ph) mutation. Proc.Natl.Acad.Sci.U.S.A 1991; 88: 6-10.

Street VA, Goldy JD, Golden AS *et al.* Mapping of Charcot-Marie-Tooth disease type 1C to chromosome 16p identifies a novel locus for demyelinating neuropathies. Am.J.Hum.Genet. 2002; 70: 244-250.

Swanson AG. Congenital insensitivity to pain with anhydrosis: a unique syndrome in two male siblings. Arch Neurol. 1963; 8, 299-306.

Swanson AG, Buchan GC Alvord EC. Anatomic changes in congenital insensitivity to pain: absence of small primary sensory neurons in ganglia, roots and Lissauer's tract. Arch Neurol. 1965; 12, 12-18.

Szpirer C, Szpirer J, Klinga-Levan K *et al.* The rat: an experimental animal in search of a genetic map. Folia Biol.(Praha) 1996; 42: 175-226.

Tabaraud F, Lagrange E, Sindou P *et al.* Demyelinating X-linked Charcot-Marie-Tooth disease: unusual electrophysiological findings. Muscle Nerve 1999; 22: 1442-1447.

Tamari I, Goodman RM, Sarova I, Hertz M, Adar R, Zvibach T. Autosomal recessive peripheral sensory neuropathy in 3 non-Ashkenazi Jewish families. J.Med.Genet. 1980; 17: 424-429.

Tezel G, Wax MB. The mechanisms of hsp27 antibody-mediated apoptosis in retinal neuronal cells. J.Neurosci. 2000; 20: 3552-3562.

Thiagalingam S, Foy RL, Cheng KH, Lee HJ, Thiagalingam A, Ponte JF. Loss of heterozygosity as a predictor to map tumor suppressor genes in cancer: molecular basis of its occurrence. Curr.Opin.Oncol. 2002; 14: 65-72.

Timmerman V, De Jonghe P, Spoelders P *et al.* Linkage and mutation analysis of Charcot-Marie-Tooth neuropathy type 2 families with chromosomes 1p35-p36 and Xq13. Neurology 1996; 46: 1311-1318.

Timmerman V, De Jonghe P, Ceuterick C *et al.* Novel missense mutation in the early growth response 2 gene associated with Dejerine-Sottas syndrome phenotype. Neurology 1999; 52: 1827-1832.

Topilko P, Schneider-Maunoury S, Levi G *et al.* Krox-20 controls myelination in the peripheral nervous system. Nature 1994; 371: 796-799.

Tournev I, King RH, Workman J *et al.* Peripheral nerve abnormalities in the congenital cataracts facial dysmorphism neuropathy (CCFDN) syndrome. Acta Neuropathol.(Berl) 1999; 98: 165-170.

Tsui LC, Buchwald M, Barker D *et al.* Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. Science 1985; 230: 1054-1057.

Valentijn LJ, Baas F, Wolterman RA *et al.* Identical point mutations of PMP-22 in Trembler-J mouse and Charcot- Marie-Tooth disease type 1A. Nat.Genet. 1992; 2: 288-291.

van Engeland M, Kuijpers HJ, Ramaekers FC *et al.* Plasma membrane alterations and cytoskeletal changes in apoptosis. Exp.Cell Res. 1997; 235: 421-430.

Vance JM, Barker D, Yamaoka LH *et al.* Localization of Charcot-Marie-Tooth disease type 1a (CMT1A) to chromosome 17p11.2. Genomics 1991; 9: 623-628.

Venkatesh B, Si-Hoe SL, Murphy D *et al.* Transgenic rats reveal functional conservation of regulatory controls between the Fugu isotocin and rat oxytocin genes. Proc.Natl.Acad.Sci.U.S.A 1997; 94: 12462-12466.

Villanova M, Timmerman V, De Jonghe P *et al.* Charcot-Marie-Tooth disease: an intermediate form. Neuromuscul.Disord. 1998; 8: 392-393.

Walsh D, Li Z, Wu Y, Nagata K. Heat shock and the role of the HSPs during neural plate induction in early mammalian CNS and brain development. Cell Mol.Life Sci. 1997; 53: 198-211.

Warner LE, Hilz MJ, Appel SH *et al.* Clinical phenotypes of different MPZ (P0) mutations may include Charcot- Marie-Tooth type 1B, Dejerine-Sottas, and congenital hypomyelination. Neuron 1996; 17: 451-460.

Warner LE, Shohat M, Shorer Z *et al.* Multiple de novo MPZ (P0) point mutations in a sporadic Dejerine-Sottas case. Hum.Mutat. 1997; 10: 21-24.

Warner LE, Mancias P, Butler IJ *et al.* Mutations in the early growth response 2 (EGR2) gene are associated with hereditary myelinopathies. Nat.Genet. 1998; 18: 382-384.

Warner LE, Svaren J, Milbrandt J *et al.* Functional consequences of mutations in the early growth response 2 gene (EGR2) correlate with severity of human myelinopathies. Hum.Mol.Genet. 1999; 8: 1245-1251.

Watanabe TK, Bihoreau MT, McCarthy LC *et al.* A radiation hybrid map of the rat genome containing 5,255 markers. Nat.Genet. 1999; 22: 27-36.

Willison, K. R. Composition and function of the eukayotic cytosolic chaperonin-containing TCP1. Bukau, B. molecular chaperones and folding catalysts. pp. 555-pp. 571. 1999. Amsterdam, The Netherlands, Harwood Academic Publishers.

Wilson JM, Tarr GE, Kelley WN. Human hypoxanthine (guanine) phosphoribosyltransferase: an amino acid substitution in a mutant form of the enzyme isolated from a patient with gout. Proc.Natl.Acad.Sci.U.S.A 1983; 80: 870-873.

Woon PY, Osoegawa K, Kaisaki PJ *et al.* Construction and characterization of a 10-fold genome equivalent rat P1- derived artificial chromosome library. Genomics 1998; 50: 306-316.

Wu-Baer F, Lane WS, Gaynor RB. Identification of a group of cellular cofactors that stimulate the binding of RNA polymerase II and TRP-185 to human immunodeficiency virus 1 TAR RNA. J.Biol.Chem. 1996; 271: 4201-4208.

Xue D, Shaham S, Horvitz HR. The Caenorhabditis elegans cell-death protein CED-3 is a cysteine protease with substrate specificities similar to those of the human CPP32 protease. Genes Dev. 1996; 10: 1073-1083.

Yaffe MB, Farr GW, Miklos D *et al.* TCP1 complex is a molecular chaperone in tubulin biogenesis. Nature 1992; 358: 245-248.

Yan WX, Archelos JJ, Hartung HP *et al.* P0 protein is a target antigen in chronic inflammatory demyelinating polyradiculoneuropathy. Ann.Neurol 2001; 50: 286-292.

Zhao C, Takita J, Tanaka Y *et al.* Charcot-Marie-Tooth disease type 2A caused by mutation in a microtubule motor KIF1Bbeta. Cell 2001; 105: 587-597.

Zilkha-Falb R, Barzilai A, Djaldeti R *et al.* Involvement of T-complex protein-1delta in dopamine triggered apoptosis in chick embryo sympathetic neurons. J.Biol.Chem. 2000; 275: 36380-36387.

Appendix I

Reagent A (1l):	109.4g sucrose
	1.02g $MgCl_2.6H_2O$
	10 ml 1M Tris
	10 ml Triton X100
10% SDS	100g sodium dodecyle sulfate
	900 ml np H_2O
CVS buffer (100mL)	4ml 4M NaCl
	10ml 0.25M EDTA
	86ml Nanopure H_2O
Proteinase K solution	40mg Proteinase K
	12ml 20mM Tris-HCl,
	2mM EDTA
	adjust to pH 8.5 with HCl
1x TE	10mM Tris-HCl, pH 8.0
	1mM EDTA
Orange G loading dye (for 10 mL)	2.5ml 1% orange G
	3ml Glycerol
	4.5ml Sterile water

SSCP loading blue

255µl deionized formamide

30µl bromophenol blue/xylene blue

15µl glycerol

377 sequencing/genescan gel mixture

(total volume 50mL):

18g Urea

5.3ml 40% 19:1 acrylamide,

bisacrylamde solution

27.5ml np water

0.5g Amberlite resin

Stirring to dissolve the urea, filter 5ml of 10x TBE through 0.45µm Whatman filter paper and then, filter the gel mixture

373 sequencing gel mixture

(total volume 80mL):

40g Urea

12ml 40% 19:1 acrylamide:

bisacrylamide solution

30ml np water

1g Amberlite resin

Stirring to dissolve the urea, filter 5ml of 10x TBE through 0.45µm Whatman filter paper and then, filter the gel mixture

20X SSC

3M NaCl

0.3M Sodium Citrate Dihydrate

Add distilled water to make 1 liter.

Hybridisation solution, total 50 mL

20ml	12.5% Dextran Sulphate
5ml	10% SDS
12.5ml	np water
12.5ml	4 M NaCl

Buffer 1

0.1M	Maleic acid
0.15M	NaCl
pH	7.5

Buffer 2

Blocking stock solution diluted 1:10

in buffer 1

Buffer 3

100mM	Tris-HCl
100mM	NaCl
50mM	$MgCl_2$
pH	9.5

Buffer 4

10mM	Tris-HCl

1mM	EDTA
pH	8.0

NBT-solution

75mg/ml of nitroblue tetrazolium salt in dimethylformamide to make up 70% (v/v)

X-phosphate-solution

50mg/ml	5-bromo-4-chloro-3-indolylphosphate, toluidin salt in dimethlyformamide

Colour-substrate solution (freshly prepared)

45µl	NBT-solution
35µl	X-phosphate-solution

add buffer 3 to 10ml

5X MOPS buffer

0.2M	MOPS (pH 7.0)
50mM	sodium acetate
5mM EDTA	(from a stock of 0.5M EDTA [pH 8])

RNA sample buffer

10.0ml	deionized formamide
3.5ml	37% formaldehyde
2.0ml	5X MOPS buffer

RNA loading buffer

50%	glycerol
1mM	EDTA
0.4%	bromophenol blue
1mg/ml	ethidium bromide

Hybridization solution

5X	SSC,
2%	blocking agent
0.1%	LRS (N-Lauroylsarcosine sodium)
0.02%	SDS (sodium dodecyl sulfate)
50%	deionized formamide

Appendix II

Table 1. Primer sequences for sequencing *Otx1*

Fragment	Forward primer	Reverse primer
Intron 1 (≅ 1.2 kb)	(106)-ATGGACCTCCTGCACCCTTC - (125)	(223) –GAGTCTTTGCGAACAGCG-(206)
Intron 1 (fragment 1 include exon 1)	(26)-TCCACCCGGCTGTTAGCATG-(45)	(I275) –TTAGCAAGCGGACACAAG-(I258)
Intron 1 (fragment 2)	(I208)-GAAACTGCAGCCTACAAC-(I225)	(I429)-ACTCAGGTAAATAAGTAGGATC-(I450)
Intron 1 (fragment 3)	(I447)-CCTACTTATTTACCTGAGTTTG-(I426)	(223)-GAGTCTTTGCGAACAGCG-(206)
Exon 2	(I331)-TCATGGTCCACCGCTGACTC-(I350)	(283)-TGGACTCGGGCAGGTTGATCTTG-(261)
Exon 3	(299)-TCAAGAACCGCCGAGCCAAG-(318)	(712)-ACATAGGGTAGGAGGTCG-(695)
	(618)-AGTGCCAGAGCCATTAGG-(635)	(1079)-TCCTTATAGTCCAGACAGTCG-(1059)
	(917)-ACCACCACCACCACCATCAC-(936)	(1184)-CAGCAGCAACATGATAGGGAC-(1164)
3' untranslated region	(MOUSE)-CGCCTTTGCCTTGGCTTTTCC	(MOUSE)-GCTGTCCAGAAAGACTTGAAG

I = intron

Table 2. Degenerate primer sequences for sequencing *KIAA0127*

Fragment	Forward primer	Reverse primer
1	426-AGATATATGTTGGGTAAAGG -445	883-AGTYTRGTNGGNGCYGYNGG-864
2	766-AYCTNGGNAGYACNACNC –783	1262-NGTYTTNAGNAGRTCRTC-1245
3	1160-YGCNGAYATHGAYACNTCN-1178	1512-TACTCGAATGAAGGCACTC-CAA-1533
AI044638 (Rat EST clone)	1334-CATCATGGAGGTACTTGTTG- GC–1355	1512-TACTCGAATGAAGGCACTC-CAA-1533

Table 3. Primer sequences for sequencing *Bcl11A*

Bcl11A	Forward primer	Reverse primer
Fragment 1	181-GAGTCTCCTTCTTTCTAACC –200	529- AAGGCGGCTTATCCACAC –512
Fragment 2	473- AACGGAAACAATGCAATG –490	920- CCTGAAGGGATACCAACC –903
Fragment 3	841- CAGAACACTCATGGATTAAG –860	1241- GCCAGCTCTCTAAGTCTC –1224
Frament 4	1168- AGGGTGCTGCGGTTGAAT-C –1186	1569- GGACGATTTGTGCATGTG –1552
Fragment 5	1489- AAGTGCAACCTGTGCGAC –1506	1945- GTGCTGCATGGAGCTGAG –1927
Fragment 6	1793- AGAGCGAGAGGGTGGAC-TAC –1812	2253- GAGCCACTGCGAATACAC –2236
Fragment 7	2168- TCTCCAAGCGCATCAAGC –2185	2582- GCTCTTGAGCTTCCATCC –2565
Rat EST -AA818595	TATGTCAATGTTAAGCACAC	AGGCGGATGGAAGCTCAAG

Table 4. Primer sequences for sequencing the cDNA of rat *Cct4*

Cct4	Forward sequence	Reverse sequence
Fragment 1	118- TCCATCCGTCCTCCTGAACC –137	563- GGAAATGATGGTTGGATG –546
Fragment 2	461- AGAAGCAGGAGATGGCAC –478	913- AACTGAATAAGCCCAATC –896
Fragment 3	804- GGTGGGACAATWGATGAC –821	1232- ACCCAGCATGTCAGCAGTR –1214
Fragment 4	1169- CTGTAAGACAATTGGAACC –1187	1506- AGGACTCCATGCCACTCAG –1488
Fragment 5	1423- TTATTGCAGGAGGTGGTG –1440	1785- CAGATTATCGAGTRTTTACCAC –1764
Rat EST clone (UI-R-CW0s-ccg-b-09-0-UI.s1)	1562- AATTGCTGGCCTGAATCC-1579	1673- TTCTCTCCATACTCGTGC –1656

Appendix III

Sequence of the rat microsatellite markers and BLAST searching results

D14Rat34

```
GA x5J8B7W4T6T:1..500000 /organism=Mus musculus /order=1
              /ga_uid=196000025308379 /len=1142318
          Length = 500000

Score = 44.1 bits (22), Expect = 0.072
 Identities = 56/66 (84%), Gaps = 1/66 (1%)
 Strand = Plus / Plus


Query: 47..106
ttatgttaaacaaagttataacaggagacccagagagtcatcatggaatggtatgtaatg
|||||| || |||| || |||||||||||||| || | | ||||| ||||||||||||||
ttatgtcaagcaaaattctaacaggagacccggaaatt-atcatataatggtatgtaatg
Sbjct:497976..498034

Query:   107    gtaaag 112
                |||||| 
Subject: 498035 gtaaag 498040


 Score = 85.7 bits (43), Expect = 2e-14
 Identities = 71/81 (87%)
 Strand = Plus / Plus


Query: 239..298
aatgagggagaatagagaataatataaaaagattgtcaaaaccatcaggtgacataaatg
||||||||| ||||||||||||||| |||| |||||||||||||||||| || ||||||
aatgagggaagatagagaataatattaaaatattgtcaaaaccatcaggtaacttaaatg
Sbjct:498405..498464

Query: 299    ngnaaatcatagtttaatacc 319
                  |||| ||||||||||||
Sbjct: 498465 aagaaattatagtttaatacc 498485
```

D14Rat130

```
GA x5J8B7W4498:500001..1000000 /organism=Mus musculus /order=2
/ga_uid=196000025284904 /len=1726612      Length = 500000
 Score =  147 bits (74), Expect = 1e-32
 Identities = 135/154 (87%), Gaps = 1/154 (0%)
 Strand = Plus / Minus
```

```
Query: 8..67
tctacccaccagtggatacattcttaagggaaaccgactctcccactcccagcatctacc
|||||||||||||  || |||||||||| |||||| ||||||||| ||||||||| | |||
tctacccaccagtaaatgcattcttaacggaaacagactctccctctcccagcacccacc
Sbjct: 49830..49771

Query: 68..127
aactgagaagaactcctcaactaggggtgagacttcatgctgaattctcctctccatgct
|||||  ||||| |||| ||||||| ||| | ||||||||||||||||||||||||||
aactgccaagaa-tcctaaactaggagtggggcttcatgctgaattctcctctccatgct
Sbjct: 49770..49712

Query: 128   gggcttttgtctggcatgagcacaggctgtgttc 161
              ||| |||||||||||||||| ||| ||| |||||
Sbjct: 49711 gggattttgtctggcatgagtacaagctatgttc 49678

 Score = 54.0 bits (27), Expect = 1e-04
 Identities = 52/59 (88%), Gaps = 1/59 (1%)
 Strand = Plus / Minus

Query: 314..372
tgaatatatcattgaggtggaaattaggatggggacataggtgagattttcaaagagaa
|||||||| ||| |||||||||||||||| || |||||  |||||||||| |||||||
tgaatata-catagaggtggaaattaggaaggagacatcagtgagattttccaagagaa
Sbjct: 49576..49519
```

D14Rat21

```
GA_x5J8B7W4498:1000001..1500000 /organism=Mus musculus /order=3
/ga_uid=196000025284904 /len=1726612      Length = 500000
 Score =  347 bits (175), Expect = 4e-93
 Identities = 268/300 (89%), Gaps = 12/300 (4%)
 Strand = Plus / Minus

Query: 11..70
ccaacattccacctccacctgcaccggcttgttccacctgaaagcttaccaagtgaaagc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ccaacattccacctccacctgcaccggcttgttccacctgaaagcttaccaagtgaaagc
Sbjct:430067..430008


Query: 71..130
aggagccatgtgaggacaccttaattatacccacccagaaacttgcttacggagagccag
|||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
aggagccatgtgaggacaccttaattatacccacccagaaacttgcttacagagagccag
Sbjct:430007..429948


Query: 131..184
tagggctccaacgttagagattagctcttactctgctgtattaaaaacaagaa------c
|||||||||||| |||||||||||||||| | |||||| | | |||| || ||      |
tagggctccaatgttagagattagctcttgcactgctgaaatgaaaagaaaaaaaaaccc
Sbjct:429947..429888


Query: 185..244
acaaattaagcatgaaacatgtgctttngttcttccttttatttgacccccacccctcct
||||||||||  | |||||||||||||  |||||||| |||||||       ||||||||
acaaattaagtgtaaaacatgtgcttttcttcttcctgttatttg------ccccctcct
```

```
Sbjct:429887..429834


Query: 245..304
tcctggcaattcaagggcactactgggtaacaaaatcttccagagtaccagtctctcaga
| ||| | ||||||||||||||||||||||||||||||||||||||||||| |||||||||
ttctgtctattcaagggcactactgggtaacaaaatcttccagagtaccaatctctcaga
Sbjct:429833..429774
```

D14Rat95

```
GA_x5J8B7W3EF1:1..500000 /organism=Mus musculus /order=1
/ga_uid=196000025262561 /len=1989279   Length = 500000
Score = 99.6 bits (50), Expect = 2e-18
 Identities = 118/139 (84%), Gaps = 5/139 (3%)
 Strand = Plus / Plus

Query: 85..143
cttcatacatgtgttcttgctta-tcgggtggggctggaggagaggatttctgagggatc
|||| |||| ||||||||||||  | |||||||||||| |||||||||  |||| | |||
cttcgtacacgtgttcttgctttgtggggtggggctgggggagaggatggctgaagaatc
Sbjct: 85391..85450


Query: 144..199
agaagtctggtacagatagggcagtgtgggagactgctcagaacggcaag----aatagt
|||||||| |||||||||||||| |||||||| ||||  |||| |||||    ||||||
agaagtctagtacagatagggcaatgtgggagattgcttggaacagcaagaacaaatagt
Sbjct: 85451..85510


Query: 200   gcatggtacaccatacaga 218
              |||||||| ||||||||||
Sbjct: 85511 gcatggtataccatacaga 85529

 Score =  101 bits (51), Expect = 4e-19
 Identities = 100/115 (86%), Gaps = 1/115 (0%)
 Strand = Plus / Plus

Query: 302..361
tttgttgctttttctccttactcctgcgcaaagccatgggtcctcttttgctctgtattc
|||||||||||||||   |||||||| | ||||||| |||||| || |||||||| |||
tttgttgctttttctcacgactcctgcacgaagccataggtcctgttgtgctctgtgttc
Sbjct: 85752..85811


Query: 362..416
attgtttatgatttcctcagaacaccaaatccagacacaacaatgagcccctaag
|||||||| |||||  ||||||| |||||||||| ||||| ||||||||||||
attgtttacgatttaatcagaacgccaaatccagatacaac-atgagcccctaag
Sbjct: 85812..85865
```

D14Rat139

```
GA_x5J8B7W3EF1:1500001..1989279 /organism=Mus musculus /order=4
/ga_uid=196000025262561 /len=1989279       Length = 489279
```

```
 Score = 58.0 bits (29), Expect = 6e-06
 Identities = 41/45 (91%)
 Strand = Plus / Plus


Query: 2      catgtgactatggtctctaaagactcagtagcctagtggtgatgc 46
               ||||||||| ||||||  ||||||| ||||||||||||||||||
Sbjct: 154253 catgtgactctggtcttcaaagactgagtagcctagtggtgatgc 154297

 Score = 54.0 bits (27), Expect = 9e-05
 Identities = 83/101 (82%), Gaps = 3/101 (2%)
 Strand = Plus / Plus


Query: 74..133
aggttgtctgcacaatggctgaacctagcgtctagcaagcagtttctctgattgtcccaa
|||||  ||||||  |||||||||||| ||  |||||  ||    ||||||| ||||||||||
aggttatctgcaaaatggctgaacccagtgtctaacagctcgtttctcagattgtcccaa
Sbjct:154309..154368


Query: 134    tgtggtccctcagcagatggagcatggctgtgcttgatgct 174
               ||  |  |||||   ||  |||  ||||||||| ||||||||||
Sbjct: 154369 tgggatccct---caaatgaagcatggctatgcttgatgct 154406

 Score = 42.1 bits (21), Expect = 0.33
 Identities = 21/21 (100%)
 Strand = Plus / Plus


Query: 232    gttcataggtgagactgggta 252
               |||||||||||||||||||||
Sbjct: 154509 gttcataggtgagactgggta 154529

Score = 60.0 bits (30), Expect = 1e-06
 Identities = 54/62 (87%)
 Strand = Plus / Plus

Query: 250..311
gtagtgtgtggatgtcggaaaacagcctcaggtgtcagtccttccctagttggtttgacaca
||||||||||  ||  ||  ||  |  ||  ||||||||||||||||  |||||||  ||||||||||||
gtagtgtgtgaatatctgagagcaccctcaggtgtcagtctttccctacttggtttgacaca
Sbjct:154545..154606
```

D14Rat138

```
1.  GA_x5J8B7W3EF1:1..500000 /organism=Mus musculus /order=1
/ga_uid=196000025262561 /len=1989279     Length = 500000
Score = 42.1 bits (21), Expect = 0.48
 Identities = 21/21 (100%)
 Strand = Plus / Minus

Query: 6      tggaatgtgagtgtgggtggg 26
               |||||||||||||||||||||
Sbjct: 458517 tggaatgtgagtgtgggtggg 458497

 Score =  123 bits (62), Expect = 2e-25
 Identities = 145/170 (85%), Gaps = 2/170 (1%)
 Strand = Plus / Minus
```

```
Query: 33..92
aggaagtggccactggcttccactgagtcacagaggcttgagtgggggaaggctcccca
|||||||||||||| ||||||||||||||||||||  ||  |||| || ||||||||||
aaggaagtggccaccggcttccactgagtcacagaaactcaagtgagg-aaggctccca
Sbjct: 458481..458424


Query: 93..152
agccccacctaggtgtagcaggaagagcacggacaggaaggaggctggctcttacccaga
||  ||||| |||||||||||| |||||| || | |||||||||| |||  ||||||||
aggtccacccaggtgtagcagggagagcaagggctggaaggaggcaggcatttacccagc
Sbjct:458423..458364


Query: 153    ggatcctgggaggccaagcaggcaatagatacaaatgctgtcctgtgaaa
202
               |||||||||||||||||||||| | | ||||||| ||| |||||||||
Sbjct: 458363 agatcctgggaggccaagcaggcagtcgctacaaatactgccctgtgaaa
458314

 Score = 48.1 bits (24), Expect = 0.008
 Identities = 35/39 (89%)
 Strand = Plus / Minus


Query: 537    ttcagaaagatccagancccttgcctttctttagtttga 575
               |||||||||||||||| |||||||| || | ||||||||
Sbjct: 458073 ttcagaaagatccagagcccttgccgttgtctagtttga 458035
```

D14Rat22

```
1.  GA_x5J8B7W498Y:500001..1000000 /organism=Mus musculus /order=2
/ga_uid=196000025290015 /len=3916428     Length = 500000
Score = 81.8 bits (41), Expect = 4e-13
 Identities = 41/41 (100%)
 Strand = Plus / Plus


Query: 26     cttaaactttcaaattttagttctagtcagaagtcttgact 66
               |||||||||||||||||||||||||||||||||||||||||
Sbjct: 322558 cttaaactttcaaattttagttctagtcagaagtcttgact 322598

 Score =  133 bits (67), Expect = 1e-28
 Identities = 85/91 (93%)
 Strand = Plus / Plus


Query: 64
actaatacatgaagtgcactttctgtaaggtagtcaattgagtaagcagagtaaatctgt
|||||||||||||| |||||||||||||| |||||||||||||| || ||||||||| |
actaatacatgaagagcactttctgtaagatagtcaattgagtaaacaaagtaaatctct
Sbjct:322609..322668


Query: 124    cctcctgtgccaaatctctggaaattagaga 154
               |||||||||||||||||||||||| ||||
Sbjct: 322669 cctcctgtgccaaatctctggaaattggaga 322699
```

```
 Score = 60.0 bits (30), Expect = 1e-06
 Identities = 80/93 (86%), Gaps = 4/93 (4%)
 Strand = Plus / Plus


Query: 250..308
accagaagttgtgagaatccactgagcaccacacgatggacaa-gtgagattaagaaagg
|||| ||||||||||||||||||||  |||  | |||| |||| ||||||||| ||||||
accaaaagttgtgagaatccactga--accgtatgatgaacaaagtgagatta-gaaagg
Sbjct: 322799..322855


Query: 309    tggtagcgagtgttaagaggagaggactgctgc 341
              ||||||| |||||||||| |||| ||| |||||
Sbjct: 322856 tggtagcaagtgttaagaagagaagaccgctgc 322888
```

OTX1

```
1.  GA_x5J8B7W498Y:1000001..1500000 /organism=Mus musculus
/order=3 /ga_uid=196000025290015 /len=3916428  Length = 500000

 Score =  258 bits (130), Expect = 6e-66
 Identities = 136/138 (98%)
 Strand = Plus / Minus


Query: 2..61
gccccctggatccgtcggggcgcctccacccggctgttagcatgatgtcttacctcaaac
|||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
gccccctggatccgtcggggcacctccacccggctgttagcatgatgtcttacctcaaac
Sbjct: 55674..55615


Query: 62
aacccccatacggcatgaacgggctgggcctagctggccctgccatggacctcctgcacc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
aacccccatacggcatgaacgggctgggcctagctggccctgccatggacctcctgcacc
Sbjct: 55614..55555


Query: 122   cttccgtgggctaccctg 139
              | ||||||||||||||||
Sbjct: 55554 cctccgtgggctaccctg 55537

Score =  291 bits (147), Expect = 5e-76
 Identities = 153/155 (98%)
 Strand = Plus / Minus


Query: 139..198
gccaccccgcggaagcagcgacgggagcgcaccaccttcacgcgctcacagctggacgtg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
gccaccccgcggaagcagcgacgggagcgcaccaccttcacgcgctcacagctggacgtg
Sbjct: 54678..54619
```

```
Query: 199..258
ctcgaggcgctgttcgcaaagactcgctacccagacatcttcatgcgcgaggaggtggct
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ctcgaggcgctgttcgcaaagactcgctacccagacatcttcatgcgcgaggaggtggca
Sbjct: 54618..54559


Query: 259   ctcaagatcaacctgcccgagtccagagtccaggt 293
              ||||||||||||||||| |||||||||||||||
Sbjct: 54558 ctcaagatcaacctgccagagtccagagtccaggt 54524

 Score =  940 bits (474), Expect = 0.0
 Identities = 549/574 (95%)
 Strand = Plus / Minus


Query: 290..349
aggtttggttcaagaaccgccgagccaagtgccgccagcagcagcagagcgggaatggaa
|||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
aggtttggttcaagaaccgccgcgccaagtgccgccagcagcagcagagcgggaatggaa
Sbjct: 53216..53157

Query: 350..409
cgaaaagccggccggtcaagaagaagtcgtctccggtgcgcgagagctcgggttccgaga
|||||| |||||||||||||||||||||||| || |||||||||||||||| ||||
cgaaaacccggccggtcaagaagaagtcgtctccagtacgcgagagctcgggttcagaga
Sbjct: 53156..53097


Query: 410..469
gcagcggccagttcacgccgcccgccgtatctagctctgcgtcttcgtctagctcagcgt
|||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
gcagcggccagttcacgccgcccgccgtatctagctctgcttcttcgtctagctcagcgt
Sbjct: 53096..53037


Query: 470..529
ccagtgcctccgctaacccggcggctgctgcggccgcgggcctgggtgggaacccggtgg
|||||||||||||| ||||| ||||||||||||||||||||||||||||||||| ||||
ccagtgcctccgccaaccccgcggctgctgcggccgcgggcctgggtgggaacccagtgg
Sbjct: 53036..52977


Query: 530..589
cagccgcgtcctctctgagcacgcctactgcttcgtccatctggagcccggcctccatct
|||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
cagcagcgtcctctctgagcacgcctactgcttcgtccatctggagcccggcctccatct
Sbjct: 52976..52917


Query: 590..649
ctcccggctcagcgccggcatccgtatcagtgccagagccattaggcgctccgagcaacg
| || |||||||||||| |||||||||| |||||||||||| | | ||||||||||||||
cgccgggctcagcgccgacatccgtatcggtgccagagccactggccgctccgagcaacg
Sbjct: 52916..52857


Query: 650..709
cctcatgcatgcagcgctcggtagccgcaggtgcggccactgccgcgacctcctacccta
|||| ||||||||||||||||||||||||||||||||||||||||  ||||||||||||
cctcgtgcatgcagcgctcggtagccgcaggtgcggccactgccgcagcctcctacccta
```

```
Sbjct: 52856..52797


Query: 710..769
tgtcctatggccagggcagaagctatggccagggctaccccgcgccttcctcttcttact
|||||||||||||||| |||||||||| ||||| |||||||||| ||||||||||||
tgtcctatggccagggcggaagctatggtcagggataccccgcgccctcctcttcttact
Sbjct: 52796..52737


Query: 770..829
ttggcggtgtagactgcagctcctaccttgcgcccatgcactctcatcaccacccgcacc
|||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
ttggcggtgtagactgcagctcctacctagcgcccatgcactctcatcaccacccgcacc
Sbjct: 52736..52677


Query: 830   agcttagccccatggcaccctcctccatggctgg 863
            |||||||||||||||||||||||||||||||||
Sbjct: 52676 agcttagccccatggcaccctcctccatggctgg 52643


 Score =  369 bits (186), Expect = 2e-99
 Identities = 210/218 (96%)
 Strand = Plus / Minus


Query: 951..1010
aggttatggaggctctgggctcgccttcaactctgccgactgcttggattacaaggagcc
|||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
aggttatggaggctctgggctcgccttcaattctgccgattgcttggattacaaggagcc
Sbjct: 52555..52496


Query: 1011..1070
cgccgccgccgctgcttcctctgcctggaaactcaacttcaactctcccgactgtctgga
|||||||||||||||||||||||||||||||||| ||||||||||||||||||||
cgccgccgccgctgcttcctctgcctggaaactcaatttcaactctcccgactgtctgga
Sbjct: 52495..52436


Query: 1071..1130
ctataaggaccaagcctcatggcggttccaggtcttgtgagcccaggagggaaaggaagt
|||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
ctataaggaccaagcctcgtggcggttccaggtcttgtgagcccaggagggaaaggaagt
Sbjct: 52435..52376


Query: 1131  ggagaagaaacgcaactaccttcgccctccgtggtccc 1168
            |||||||||||||||||||   |||||| |||||||
Sbjct: 52375 ggagaagaaacgcaactacctataccctccatggtccc 52338
```

D14Got73

```
1.  GA x5J8B7W498Y:1000001..1500000 /organism=Mus musculus /order=3
/ga_uid=196000025290015 /len=3916428      Length = 500000

 Score = 97.6 bits (49), Expect = 1e-18
 Identities = 78/88 (88%)
```

```
 Strand = Plus / Plus

Query: 9..68
tgcaagcactggagagctggaaagccaacagtgcaattcagaaccagccaacgtcctgan
|||||||||||||||||||||||  | || || |||||||||||||| || | |||||
tgcaagcactggagagctggaaagttagcaatgaaattcagaaccagcaaatggcctgag
Sbjct: 466473.. 466532

Query: 69     aagtactgaggctaatggtgatactctc 96
               |||||||||||||||| ||||||||||
Sbjct: 466533 aagtactgaggctaatgatgatactctc 466560


1.  GA_x5J8B7W498Y:1000001..1500000 /organism=Mus musculus /order=3
/ga_uid=196000025290015 /len=3916428     Length = 500000
 Score = 56.0 bits (28), Expect = 4e-06
 Identities = 31/32 (96%)
 Strand = Plus / Plus

Query: 12     taatggaaataggagagctggctttaaatctg 43
               |||||||||||||| ||||||||||||||||
Sbjct: 466262 taatggaaataggaaagctggctttaaatctg 466293

 Score = 40.1 bits (20), Expect = 0.26
 Identities = 36/40 (90%), Gaps = 1/40 (2%)
 Strand = Plus / Plus

Query: 92     gccaaagtatatgtagccctaaagt-aatacctagcatgt 130
               |||||||||| |||||||||| | |||||||| |||||
Sbjct: 466362 gccaaagtatacgtagccctaaattcaatacctaccatgt 466401
```

D14Rat131

```
GA_x5J8B7W4TBK:1000001..1500000 /organism=Mus musculus /order=3
/ga_uid=196000025308531 /len=2644602      Length = 500000
 Score = 56.0 bits (28), Expect = 3e-05
 Identities = 37/40 (92%)
 Strand = Plus / Minus

Query: 36     ttctgtaaactctgtggcctcttactggtctcctagtggt 75
               |||| ||||||||||||||||||||||||| ||| ||||
Sbjct: 259678 ttctataaactctgtggcctcttactggtcttctaatggt 259639
```

D14Rat132

```
GA_x5J8B7W4TBK:1000001..1500000 /organism=Mus musculus /order=3
/ga_uid=196000025308531 /len=2644602      Length = 500000
 Score =  139 bits (70), Expect = 2e-30
 Identities = 112/126 (88%)
 Strand = Plus / Minus

Query: 193..252
acatcccatttcctttggagacatttatttcaactcttgcttatgatgctattgaattct
|||||||||| || |||||||||||||| |||||||||| |  ||||||||| |||||
acatcccattttctctggagacatttattccaactcttgctgacaatgctattggattct
Sbjct:296240..296181
```

```
Query: 253..318
cccaggaacctctttgatatcaaggctctgtgagactcttttgttgactgttcctccatcatctgt
|||||||||||||||   |  ||||||||||||||| |||||||||||||||||||||||||| ||||
cccaggaacctcttcagtgacaaggctctgtgagtctcttttgttgactgttcctccatcacctgt
Sbjct:296180..296115
```

Appendix IV

1. The end-sequence of BACs with DNA covering the sequence of marker *D14Got72* (Sp6 and T7 are primers of both end of vector)

RPCIB657O1162-Sp6

67..499

```
Gggtaggccggcggacgcctggccgtcgacatttaggtgacactatagaaggatccgcggaattc
caggaaaacaaaaataaacaagtagaagcccatagagaggagacacaaaaatccctgaaagaatt
ccaggaaaacacaatcaaacagttgatggaattaaaaatggaaatagaagcaatcaagaaagaac
acatggaaacaaccctggatagagaaaaccaaaagaagagacaaggagctgtagatacaagcttc
accaacagaatacaagagatggaagagagaatatcaggagcagaagattccatagaaatcattga
ctcaactgtcaaagataatgtaaagcggaaaaagctactggtccaaaacatacaggaaatccagg
actcaatgagaagatcaaacctaaggataataggtatagaag
```

RPCIB657O13359-Sp6

60..501

```
cgtcgacatttaggtgacactatagaaggatccgcggaattcagtttgtaagtattttgttgaga
gttttgtcatctaagaaaattggtctataatgctccgttgttgttgttttaaaacaaaatctgtt
cttccatatgtgtttcagaaactaatagcaggaccagcctctgaggatagttggcttttatgagc
tgccaccaggaagcaggctaggagaccctgtcattactgccatctgaaacagtccatacagcact
agctagcatgacctggagctcactgtgtagaccagaacagccttaaacccgcagagggttccctc
ttcctacctcctaagctggggttaaaggactgcaccgtcatgcccagctgacgagagactccata
tgtgagtcatgcagttcatttcgggtcttctatagatgttcttcatttc
```

RPCIB657B21324-Sp6

85..420

```
gatccgcggaattccaagacagccaggaccacacagtaagacctcctcccaaaaaacaaacaaac
aaacaaacaacaacaaaacccaaccctaaaaccaaaacaaagtcacggcctgctctacccttagc
tgcctgtccgtttggacccagattaaaagatagccctgttgtaatggccttgtaatgtaagaatc
attccagaactgccttgtgaatgtgcagatggtgctatgtattattttggccccttcccccctca
aaaacagtattatgtacatgtattttctgaactttgttgtctattcagtnggtgatttaaaataa
aaagat
```

RPCIB657O1162-T7

60..510

```
tctatgggcttctacttgtttatttatgatttcctggaattctttcagggatttttgtgattcct
ctctgtaggcttctacatgttctctaagggagttcttcacatctttcttgaagtcctccggcatc
ataatcaaatatgattttgatactagatcttgcttttctggtgtgtttggatattccgtgtttgc
tttggtgggagaattgggctctgatgatgccatgtagtcttggtttctgttgcttgggttcctgc
gcttgcctctcgccatcagattatctctagtgttaccttgttctgctatttctgacagtggctag
actgtcctataagcctgtgtgtcaggagtgctgtaacctgttttcctgttttctttcagccagtt
atggggacagagtgttctgctttcggtcgtgtagtttttcctatctacaggtcctca
```

RPCIB657O13359-T7

40..480

```
ccttttaccttagcctacaaaagctggtgttgcagaccaccagcatgcctagcattgaaaagggc
ttcttagtgaggcagcagagtgtaactatgataaaaacacattgtacgtatggagttctcaataa
aaaccattatagttagataaaaacagcagaaacttttaaaatgacgcttatgaatttcacccatg
taaacagttccaacctttccccgctgtaccaggatacagatgacaaagctccttatcatgtacac
aggtctcacttccatggctgtgaagaaagtgattctaaaggccaggaaatccatactatgcagat
tagtccaaggctccgaaagcttcaggtcctgcacccgtttttttaaggtttatttaacatgtctg
tctgtgtgtaccgttgttcacatgtgtgtgcagatgctctcataggtca
```

RPCIB657B21324-T7

60..440

```
gtggaggttcaggtcacctcgcaccctggcaggaccgcctctgcgcacgcgcacagtccataact
gcggcgccgccccgcggacagagccgctctaggcacgcagtgttgtccggtctcgttggcaccgg
cactcacgccgcccccccacccccgtacgctggggccggacgctgttctgcgtgctggctccacg
ctcttctcacactacgtcacgaaaaaggcgtgggccgcgctcccccagaaagctaagcgtcccac
cctcggccgtcaccgcctccgtccgtggcccacaatcccctagcgcgggggctggtttccggtcg
gggggctcggacccttcaccttcccgaagcctgccatggtgctggctggcaacacc
```

2. Search result of the end-sequence of BACs

RPCIB657O13359-Sp6

Sequence Report: GA_x5J8B7W498Y:2000001..2500000

sequence location - chr.11 17836454 - 21752882

```
1.  GA_x5J8B7W498Y:2000001..2500000 /organism=Mus musculus
/order=5/ga_uid=196000025290015 /len=3916428    Length = 500000
 Score =  168 bits (85), Expect = 2e-39
 Identities = 106/113 (93%)
 Strand = Plus / Plus


Query: 106..165
ttgttgttttaaaacaaaatctgttcttccatatgtgtttcagaaactaatagcaggacc
|||||||||||||||||||||||||||| ||  ||||||||||||||||||||||||||
ttgttgttttaaaacaaaatctgttctttcaggtgtgtttcagaaactaatagcaggacc
Sbjct: 138997..139056


Query: 166..218
agcctctgaggatagttggcttttatgagctgccaccaggaagcaggctagga
|| |||||||||||| ||| ||||||| |||||||||||||||||||||||||
agtctctgaggatagtgggcctttatgaactgccaccaggaagcaggctagga 139109
Sbjct: 139057..139109

 Score = 77.8 bits (39), Expect = 4e-12
```

```
 Identities = 135/166 (81%), Gaps = 2/166 (1%)
 Strand = Plus / Plus


Query: 272..331
cctggagctcactgtgtagaccagaacagccttaaacccgcagagggttccctcttccta
|||||| |||||||||||||||||   |||||||| | ||||||| | | | ||  |
cctggaactcactgtgtagaccagaatgaccttaaacacacagagggctacttgtttttg
Sbjct: 139172..139231


Query: 332..389
cctcctaag--ctggggttaaaggactgcaccgtcatgcccagctgacgagagactccat
|||||  |||  |||||||||||||| | |||| |||||  ||||||| |||| ||| |||
cctcccaagtgctggggttaaaggagtacaccatcatgtacagctgatgagaaacttcat
Sbjct: 139232..139291


Query: 390    atgtgagtcatgcagttcatttcgggtcttctatagatgttcttca 435
              ||||||||||||||||  |||| ||| || ||||| ||||| ||||
Sbjct: 139292 atgtgagtcatgcagtatattttggggctcctatatatgtttttca
139337
```

RPCIB657O13359-T7

Sequence Report: GA_x5J8B7W498Y:2000001..2500000
sequence location - chr.11 17836454 - 21752882

```
1.  GA_x5J8B7W498Y:2000001..2500000 /organism=Mus musculus
/order=5/ga_uid=196000025290015 /len=3916428
          Length = 500000

 Score =  204 bits (103), Expect = 3e-50
 Identities = 255/300 (85%), Gaps = 8/300 (2%)
 Strand = Plus / Minus


Query: 40..97
ccagcatgcctagcattgaaaagggcttcttagtgag--gcagcagagtgtaactatgat
|||||| ||||||| ||||||||||||||||||||| |  | |||||| ||||| ||||||
ccagcacgcctagctttgaaaagggcttcttagtgggaaggagcagattgtaaatatgat
Sbjct: 278513..278454


Query: 98..153
aaaaacacattgtacgtatggagttctcaataaa---aaccattatagttagat-aaaaa
|||||||||||||| |||||| |||||||| ||   ||| |||||||||| || |||||
caaaacacattgcacatatggaattctcaattaataaaactattatagttaaatgaaaaa
Sbjct: 278453..278394


Query: 154..212
cagcagaaacttttaaaatgacgctt-atgaatttcacccatgtaaacagttccaacctt
|| |||||||||||||||| |  || |||||||||||| |||||||||||||||| |
caacagaaacttttaaaatggcattttatgaatttcaccaatgtaaacagttccaactta
Sbjct: 278393..278334
```

```
Query: 213..272
tccccgctgtaccaggatacagatgacaaagctccttatcatgtacacaggtctcacttc
| |||  |||||||| ||||| |  ||||||||| || ||||||||| ||||||||||
ttcccattgtaccagaatacaaactacaaagctctttctcatgtacagaggtctcactt-
Sbjct: 278333..278275


Query: 273..332
catggctgtgaagaaagtgattctaaaggccaggaaatccatactatgcagattagtcca
 ||||||||| |||||||||||| ||||||| ||||||||| |||| ||| ||||||||
tatggctgtgcagaaagtgattctgaaggccaagaaatccatgctatacaggttagtcca
Sbjct: 278274..278215


 Score = 42.1 bits (21), Expect = 0.25
 Identities = 33/37 (89%)
 Strand = Plus / Minus

Query: 403    gttgttcacatgtgtgtgcagatgctctcataggtca 439
              ||||| ||||||||||||||||  ||||| ||||||
Sbjct: 278118 gttgtacacatgtgtgtgcagattatctcagaggtca 278082
```

RPCIB657B21324-Sp6

Sequence Report: GA_x5J8B7W498Y:2000001..2500000
sequence location - chr.11 17836454 - 21752882

```
GA_x5J8B7W498Y:2000001..2500000 /organism=Mus musculus /order=5
              /ga_uid=196000025290015 /len=3916428
          Length = 500000

 Score = 91.7 bits (46), Expect = 2e-16
 Identities = 102/120 (85%), Gaps = 3/120 (2%)
 Strand = Plus / Minus


Query: 140..196
gtttggacccagattaaaagatagccctgttgtaatggccttgtaatgtaaga---atca
||||||||  ||||||||||  ||| ||||||||||||||||||||||| |||||   ||||
gtttggacacagattaaaaagtaggcctgttgtaatggccttgtaatataagaattatca
Sbjct: 270266..270207


Query: 197..256
ttccagaactgccttgtgaatgtgcagatggtgctatgtattattttggccccttcccc
|||||||||||||| |||||||| ||  |||||  | |||| |||||||||  |||||||
ttccagaactgcctcgtgaatgtacaagtggtgtgaggtatcattttggccttttcccc
Sbjct: 270206..270147
```

RPCIB657B2124-T7

Sequence Report: GA_x5J8B7W498Y:2000001..2500000
sequence location - chr.11 17836454 - 21752882

```
1.  GA_x5J8B7W498Y:2000001..2500000 /organism=Mus musculus
/order=5/ga_uid=196000025290015 /len=3916428
          Length = 500000

 Score = 50.1 bits (25), Expect = 9e-04
 Identities = 43/49 (87%)
 Strand = Plus / Plus


Query: 155..203
gtacgctggggccggacgctgttctgcgtgctggctccacgctcttctc
||||||||| |||||| ||||| |||||||||   ||||||||||||
gtacgctggagccggaggctgtcctgcgtgctgttcccacgctcttctc
Sbjct: 117527.. 117575


Score = 48.1 bits (24), Expect = 0.003
 Identities = 39/44 (88%)
 Strand = Plus / Plus


Query: 88     gccgctctaggcacgcagtgttgtccggtctcgttggcaccggc 131
               |||||||||||||||||| |  |||| |||||||||| ||||||
Sbjct: 117465 gccgctctaggcacgcagcgcggtcctgtctcgttggtaccggc 117508


 Score = 42.1 bits (21), Expect = 0.21
 Identities = 24/25 (96%)
 Strand = Plus / Plus


Query: 32     ggaccgcctctgcgcacgcgcacag 56
               ||||||||||||| ||||||||||
Sbjct: 117433 ggaccgcctctgcacacgcgcacag 117457
```

Appendix V

Nucleotide sequence of rat *Cct4*

ATG CCG GAG AAC GTA GCT TCC CGA AGC GGG CCG CCC GCA GCG GGG CCG
GGC AAC CGC GGG AAA GGC GCC TAC CAG GAC CGC GAC AAA CCG GCC CAG
ATC CGC TTC AGC AAC ATT TCC GCG GCC AAA GCG GTT GCC GAT GCC ATT
AGA ACA AGC CTT GGA CCT AAA GGA ATG GAC AAA ATG ATT CAA GAC GGA
AAA GGC GAT GTG ACC ATT ACC AAT GAT GGT GCC ACC ATT CTG AAA CAA
ATG CAA GTA TTA CAC CCA GCA GCC AGA ATG CTG GTG GAA CTG TCC AAA
GCT CAA GAC ATA GAA GCA GGG GAT GGC ACC ACG TCG GTC GTC ATC ATT
GCT GGC TCT CTC TTA GAC TCC TGT ACC AAG CTT CTG CAG AAA GGT ATA
CAT CCA ACC ATC ATT TCC GAG TCA TTC CAG AAA GCT TTG GAA AAG GGT
CTT GAA ATC CTT ACC GAC ATG TCT CGG CCT GTG CAA CTG AGC GAC AGA
GAA ACC TTG TTA AAT AGT GCA ACC ACT TCA TTG AAC TCA AAG GTT GTC
TCT CAG TAT TCA AGT CTG CTC TCT CCA ATG AGT GTC AAT GCA GTG ATG
AAA GTG ATC GAT CCA GCC ACG GCT ACC AGT GTC GAT CTT AGA GAT ATT
AAA ATA GTT AAG AAG CTT GGT GGG ACG ATA GAT GAC TGT GAG CTG GTG
GAA GGC CTT GTT CTC ACA CAG AAA GTG GCA AAT TCT GGC ATA ACG AGA
GTT GAG AAG GCT AAG ATT GGG CTT ATT CAG TTT TGC TTA TCT GCT CCT
AAA ACA GAT ATG GAT AAT CAG ATA GTA GTA TCT GAC TAC GCC CAG ATG
GAT CGA GTG CTT CGA GAA GAA AGA GCC TAT ATT TTA AAT TTG GTG AAG
CAA ATT AAG AAA ACA GGA TGT AAT GTC CTT CTC ATA CAG AAG TCT ATT
CTA AGA GAT GCA CTT AGT GAT CTT GCG TTA CAT TTT CTG AAT AAA ATG
AAG ATT ATG GTG GTT AAG GAC ATT GAA AGA GAA GAC ATT GAA TTC ATC
TGT AAG ACA ATT GGA ACC AAA CCA GTG GCT CAC ATT GAC CAG TTC ACT
CCT GAC ATG CTG GGT TCT GCT GAG TTA GCA GAG GAA GTC AGT TTG AAT
GGT TCT GGA AAA CTG TTC AAG ATT ACA GGC TGT ACA AGC CCA GGC AAA
ACA GTG ACA ATT GTC GTC CGA GGC TCT AAC AAA CTG GTG ATT GAG GAA
GCT GAG CGC TCC ATT CAC GAT GCT CTC TGT GTT ATC CGA TGC TTG GTG
AAG AAG AGA GCT CTT ATT GCA GGA GGT GGT GCT CCA GAA ATA GAG CTG
GCC CTC AGA CTG ACG GAG TAC TCC CGG ACA CTG AGT GGC ATG GAG TCC
TAC TGT GTT CGA GCT TTT GCA GAT GCC ATG GAA GTC ATT CCA TCT ACA
CTG GCT GAA AAT GCT GGC CTG AAT CCC ATC TCT ACT GTA ACA GAG CTA
AGA AAC CGC CAT GCC CAA GGA GAA AAA ACT ACA GGC ATT AAT GTC CGA
AAG GGT GGG ATC TCC AAC ATT CTG GAG GAA ATG GTT GTC CAG CCT CTG
TTG GTG TCA GTC AGT GCG TTG ACC CTA GCA ACT GAA ACT GTG CGG AGC
ATT CTG AAA ATT GAT GAT GTG GTA AAT ACT CGA TAA