Genetic Mapping Studies in the AS/AGU rat

This thesis is submitted in partial requirement for the degree of Master of Science at the University of Glasgow

by
Maurice Canham

Institute of Biomedical and Life Sciences
Division of Molecular Genetics
University of Glasgow

September 1997

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

The research reported in this thesis is my own and original work except where otherwise stated and has not been submitted for any other degree.

Summary

The AS/AGU rat arose by spontaneous mutation from a closed colony of AS rats in the Laboratory of Human Anatomy, The University of Glasgow, in 1990. These rats had locomotor abnormalities and appeared to have a striatal dopamine deficiency syndrome, symptoms and pathology reminiscent of human Parkinson's disease. Test crosses between AS/AGU and AS rats indicated that the mutant locus, *agu*, was inherited in an autosomal recessive fashion with no evidence of maternal or X-linked involvement.

Two reference strains, BN and F344, were used to set up backcrosses with AS/AGU, namely, [AS/AGUxF1(AS/AGUxBN)] and [AS/AGUxF1 (AS/AGUxF344)]. Previously, 75 microsatellites markers spanning the rat genome were tested for informativeness between AS/AGU and the strains BN and F344. In this study, 19 other microsatellites were typed on these strains, 9 of which displayed size differences and could therefore be mapped relative to *agu*. Two candidate chromosomes had been thought to harbour agu, one of which was eliminated in this report. Initially, four markers from the other chromosome had all shown linkage with *agu* but none were close enough to commence chromosome walking to the disease locus. This report describes the analysis of pre-existing microsatellites and the search for new ones associated with a gene family believed to be adjacent to *agu*. A total of twelve markers, associated with this gene family, were analysed for size or sequence variation between AS/AGU and reference strains. None proved informative. It was concluded that members of this gene family had not diverged sufficiently between these strains.

Another reference strain, DA, has been shown to display size differences from the AS strain from members of this gene family and other markers believed to be near *agu*. A backcross program using this reference strain is now underway. To date, another gene-based marker has been shown to be less than 0.5 cM from *agu* and physical mapping is now underway to track down a gene harbouring this mutation. The eventual discovery of such a gene and its role in normal physiology will prove valuable in the understanding of basal ganglia disorders.

Acknowledgements

Firstly, I have to thank my supervisors, Dr. Sutcliffe and Prof. Davies, for any assistance they have given me during my laboratory experience. Thanks to Dr. Sutcliffe for an extra-speedy reading of my thesis. I acknowledge receipt of a BBSRC studentship for funding my studies.

I want to take this opportunity to pay special homage to the main suppliers in this and other projects, the rats. If it were not for the sacrificing of their short-lived lives, science would be barren, diseases would remain uncured and nobody would get rich.

I'm indebted to my assesors, Dr. B.L. Cohen and Dr. J.A. Dow for having faith in my abilities. Thanks to Dr. R. Wilson for explaining to me some key genetics concepts and therefore enabling me to interpret important results. Also Margaret Adam for teaching me to face the things that should not be.

The prep-room ladies, the wash-room ladies and secretaries deserve thanks for keeping the wheels of the department turning. Irene and Lynn, the ancillary staff in my lab, I thank for putting-up with my orders at short notice and of course, for being nice. Special thanks to Maria, Tam and Ed (fellow nightshift workers and drinking buddies), Hugh and David Livingstone for being bitter and twisted, Rufus the rat for being cool and any other humans or rodents in the department who have helped me during my sentence. I also thank people outside the department who have helped to keep me sane over the last three years; all the barstaff at G.C.U.S.A., mad Athar for being more mental than me, and all my pals back in Derry.

Deepest appreciations are for my family for their encouragement, support and love throughout all my years at university.

Welcome to our fortress tall
Take some time to show you around
Impossible to break these walls
For you see the steel is much too strong
Computer banks to rule the world
Instruments to sight the stars

Possibly I've seen too much
Hanger 18 I know too much

Foreign lifeforms inventory
Suspended state of cryogenics
Selective amnesia's the story
Believed foretold but who'd suspect
The military intelligence
Two words combined that can't make sense

Possibly I've seen too much
Hanger 18 I know too much

Dave Mustaine (1990)

Contents **Page**

Chapter 1

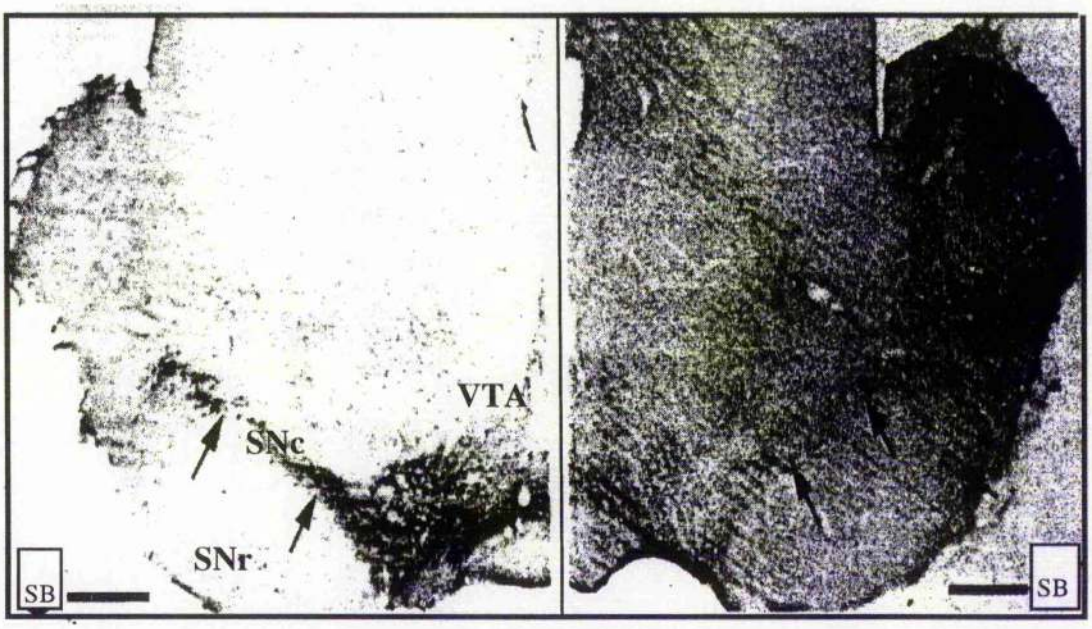
Introduction

The AS/AGU rat

In 1990, a spontaneous mutation in a closed breeding colony of Albino-Swiss (AS) rats, in the laboratory of Human Anatomy, Glasgow University, produced a new strain, AS/AGU, with movement disorders. These rats, although in good general health, fertile and with a life expectancy of 18 months, were seen to have an unsteady gait and hind limb rigidity. They walked with raised tails and a very wide gait and had a tendency to fall over every few steps. Although they retained inquisitive behaviour, they had difficulty in initiating movement. The movement disorder became progressively worse with age on assessment with open field and specific locomotor tests. When dropped from a height of 50cm, the mutant rats failed to execute a mid-air righting reflex and on walking down an inclined ramp, they consistently fell off before reaching the bottom. The mutant phenotype could be detected by visual inspection at 4-5 weeks of age with 100% precision (Clarke and Payne, 1994).

A cross between AS and AS/AGU rats was used to produce F1 offspring and subsequently an F1 x AS/AGU backcross. This backcross yielded 49% normal and 51% mutant progeny. Reciprocal crosses between AS and AS/AGU gave the same 1:1 ratio of normal to mutant rats. These observations implied an autosomal recessive pattern of inheritance consistent with a single mutant locus, *agu*, present in the AS/AGU strain. There was no evidence of X-linkage, maternal inheritance or karyotype abnormalities (Campbell et al, 1996).

At the gross anatomical level, there was no obvious differences in size, shape or orientation between the AS and AS/AGU rat brains. No cerebellar abnormalities have been detected. Tyrosine hydroxylase (TH) immunostaining of sections through the mesencephalon of 12 month old AS and AS/AGU rats revealed 60% deficits of dopaminergic cells in the substantia nigra pars compacta (SNc) with lesser depletion in the ventral tegmental area (VTA) and the ventral striatum in the mutant strain (Clarke and Payne, 1994). Figure 1.1 shows an example of the difference in TH immunostaining of the mesencephalon between AS and AS/AGU rat brains. High performance liquid chromatography analysis of micropunch biopsies showed 30-40% reduction of dopamine levels in the dorsal and lateral caudate-putamen of the striatum in 12 month old AS/AGU rats and in mutant (ASxAS/AGU)F1 x AS/AGU backcross progeny (Campbell et al, 1996). Analysis of the uptake of radiolabelled 2-deoxyglucose (2-DG) in 45 regions of the brain of 12 month old AS and AS/AGU rats showed a 20-25% reduction of 2-DG uptake in



AS AS/AGU

Fig 1.1 Low power photomicrographs of 70μm sections through the mesencephalon of AS and AS/AGU rats. Arrows indicate areas of the substantia nigra where the most obvious differences in tyrosine hydroxylase immunoreactive cell numbers between AS and AS/AGUrats. SNc = substantia nigra pars compacta, SNr = substantia nigra pars reticulata and VTA = ventral tegmental area. Scale bars (SB) = 300μm. Taken from Clarke and Payne (1994).

the SNc, the subthalamic nucleus and the ventrolateral nucleus of the AS/AGU strain. This finding was indicative of decreased local metabolic activity in these regions of mutant individuals (Lam et al, submitted). Significant deficits of noradrenergic and serotenergic neurons were observed in the locus cœruleus and raphe nuclei of AS/AGU rats, respectively (Scott et al, 1994. Stewart et al, 1994).

Recent work has been carried out in an attempt to reverse some key aspects of the *agu* phenotype. Treatment of AS/AGU rats with L-dopa did not correct the disordered hindlimb gait but did increase their frequency and periods of movement, enabled them to successfully walk down an inclined ramp and to execute a mid-air righting reflex (in preparation). The ability to improve these defects with L-dopa highlighted the fact that there is a dopamine deficiency in AS/AGU rats. Bilateral foetal midbrain transplants into the striatum of AS/AGU rats improved the locomotor defect to the same extent as with L-dopa treatment (in preparation). This result showed that AS/AGU rats have a striatal dopamine deficiency syndrome. Indeed, the areas of the brain severely affected by this mutation, to date, constitute specifically the nigrostriatal pathway in rats (Beckstead et al, 1979).

The characteristics of locomotor impairment, regional specificity of neuronal loss and dopamine depletion and the ability to improve these defects with L-dopa administration and foetal midbrain transplants in this rat showed clear signs reminiscent of human Parkinson's disease.

Parkinson's disease

(i) *Symptoms and Pathology*

Paralysis agitans or Parkinson's disease (PD), is one of the most common neurodegenerative disorders of the elderly. The disease was named after James Parkinson who first described it as the shaking palsy in 1817 (Eldridge and Rocca, 1992). This progressive syndrome usually occurs late in life with at least two of the four cardinal signs; tremor, rigidity, bradykinesia/akinesia and postural instability. It generally begins between the ages of 50 and 70 and may include other symptoms, such as, decreased eye blinking, excessive drooling related to infrequent swallowing and dementia occurring in approximately one third of patients. The characteristic 'pill rolling' tremor generally increases with stress and disappears with movement and during sleep (Eldridge and Rocca, 1992).

The major pathological finding in PD is the degeneration of the neuromelanin containing cells in the SNc. Cells survive in this region contain less neuromelanin granules. Also present in the remaining cells of the substantia nigra are cytoplasmic inclusions, Lewy bodies, which contain protein and sphingomyelin (Eldridge and Rocca, 1992). The dopamine content of the substantia nigra and the terminal regions of the nigrostriatal pathway (putamen and dorsal caudate nucleus) in post-mortem brains of PD patients are extremely low (≤10% of normal values) and this is correlated with the loss of dopaminergic neurones in the substantia nigra and degeneration of nerve terminals in the striatum (Kish et al, 1988). Cells present in the ventral striatum, locus coerulus and raphe nuclei appear less affected. The degeneration of the SNc and its efferent connections to the dorsal striatum results in a failure of the skeletomotor, ocularmotor and cognitive functions (Kish et al, 1988. Martin, 1996). Indeed it has been shown that the severity of motor functions in PD patients correlated strongly with depression and intellectual impairment (Mayeux et al, 1981). Another more recent study concluded that cognitive impairment was largely independent of motor dysfunction, probably due to deficits in another affected dopamine system, the mesolimbic pathway (Cooper et al, 1991). The dopamine system invovled in this pathway is the ventral tegmental area which lies just dorsomedial to the substantia nigra (Martin, 1996).

There are approximately twenty other disorders that exhibit symptoms of parkinsonism which are listed in table 1.1. The common denominator of these syndromes is the neuronal dysfunction of the basal ganglia. Striatonigral degeneration is most similar to Parkinson's disease with cell depletions in the SNc, the putamen and the globus pallidus, though this disorder begins several decades earlier and there is an absence of resting tremor and Lewy bodies (Eldridge and Rocca, 1992). Parkinson's disease is a specific disorder with unknown ætiology, though there is evidence of environmental and genetic factors.

(ii) *Environmental Factors*

Several investigations of the relationship between smoking and PD suggested that smoking is negatively correlated with the disease. This negative association has yet to be explained (Baumann et al, 1980). Other environmental factors have been suggested to cause PD, namely, people living in rural areas near an industrialised city who would therefore be prone to pesticide poisoning from rotenone which is MPTP like in nature (see below) (Eldridge and Rocca, 1992). It has also been

Primary

Parkinson's disease (Idiopathic)

Secondary

- Infectious and Postinfectious
  - Encephalitis lethargica
  - Leutic enchephalitis
- Toxins
  - 1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)
  - Manganese
  - Carbon monoxide
  - Carbon disulphide
  - Cyanide
  - Methanol
- Drugs
  - Resperine
  - Tetrabenazine
  - Phenothiazine
  - Haloperidol
  - α-Methyldopa
- Brain Tumors
- Brain trauma
- Syringomesencephalia

With Associated Disorders

- Dementia
  - Alzheimer's disease
  - Pick's disease
  - Wilson's disease
  - Huntington's disease
  - Parkinson's dementia of Guam
  - Normal pressure hydrocephalus
  - Creutzfeldt-Jacob disease
  - Hallervorden-Spatz disease
- Basal cell ganglia calcification
  - Hyproparathyroidism
  - Pseudohypothyroidism
- Orthostatic hypotension
- Shy-Drager syndrome
- Ophthalmoplegia
- Progressive supranuclear palsy
- Amyotrophy
  - Motor neuron disease
  - Spinocrebellar-nigral disease
- Atrophy of multiple neurological systems
  - Striatonigral degeneration
  - Olivopontocerebellar atrophy
  - Corticodentonigral degeneration with neural achromosia

Table 1.1 Classification of parkinsonism. Taken from Eldridge and Rocca (1992).

observed that people living in more dense areas or drinking well water were prone to developing PD, whereas people who lived in villages, raising pigs and chickens, as in China, were less prone. Due to these conflicting findings and an overall lack of a consistent geographical distribution of PD, it has been proposed that a genetic factor or factors, may be influencing or directly causing the triggering of this disease (Eldridge and Rocca, 1992).

(iii) *Genetic Factors*

Work by Duvoisin and colleagues in 1981 on twin studies of PD, in a US wide survey, found little if not any concordance in twelve monozygotic twin pairs (Duvoisin et al, 1981). These twelve pairs were the first twelve consecutively encountered in the study and due to the lack of concordance, it may seem feasible that genetic factors do not play a role. It must be noted that these studies suffered from a number of problems: some subjects were not personally examined; diagnostic criteria were not completely standardised and up to twenty six years may elapse before twins become concordant. Somatic mutation in a PD locus could also explain these results. ^{18}F L-6 fluorodopa positron emission tomography (PET) was used to examine concordance in twins for nigral dysfunction during PD (Burn et al, 1992).. Striatal uptake of ^{18}F-dopa assessed by PET, reflected the capacity of the striatum to form and store exogenous dopamine. The mean ^{18}F uptake was reduced in the putamen to 38% and in the caudate nucleus to 66% of normal levels in affected individuals. Analysis of monozygotic twins using this technique, showed concordance of 45% and analysis of dizygotic twins showed concordance of 29% with a 2SD threshold. This method of analysis indicated a reduced amount of dopamine activity in the non affected twin compared to normal individuals, due to a reduced number of dopaminergic cells.

Several personality traits have been found to be associated with PD. Patients are described as moralistic, law abiding, contentious , averse to risk taking and being introverted. Twin studies have shown that the affected twin was more nervous, quiet, self-controlled and depressed (Paulson and Dadmehr, 1991). However, this personality may be due to a lack of cells in a particular part of the brain which inspire the will to be more extroverted and active.

Evidence for a genetically predetermined number of neurons, of particular neurochemical class, has been gathered by the comparison of two strains of mice. Examination of the basal ganglia showed 20% more dopaminergic neurones present

in BALB/cJ mice compared to CBA/J mice (Baker et al, 1980). These morphological differences were paralleled by variations in behaviour. Treatment with *d*-amphetamine, which enhances stereotypical patterns of motor behaviour in the rat (sniffing, grooming and gnawing), induced more inquisitive behaviour in the BALB/cJ strain than the CBA/J strain. Contrary to these findings, McGreer and co-workers have shown that this may not be the case in human PD. HLA-DR staining of positive microglia phagocytosing dopamine neurones in post-mortem PD brains detected that there are losses six fold greater than in normal brains (McGreer et al, 1988). This indicated a highly active neuropathological process at the time of death, as opposed to an initial low number of dopaminergic cells.

The cytochrome P450 mono-oxygenase gene, *CYP2D6*, which is present on chromosome 22q13, has inconsistently been associated with idiopathic PD (Diederich et al, 1996). In a recent report the *CYP2D6* locus did not show linkage with PD (Wilhelmen et al, 1997). However, the same report did show that two microsatellite markers linked to *CYP2D6* were associated with PD. It was concluded that an unidentified locus for susceptibility to Parkinson's disease may be in linkage disequilibrium with microsatellite markers close to the *CYP2D6* locus.

There is evidence for transmission of PD in some large families. Lewy body positive PD has been documented to be inherited in an autosomal dominant fashion with reduced penetrance in one large Italian family (Golbe et al, 1990). Genetic markers on chromosome 4q21-23 were found to be linked to the PD phenotype in this family (Polymeropoulos et al, 1996). Recently a single base mutation in the α-synuclein gene, present in this chromosomal region, was detected in the Italian family and three other unrelated Greek families with PD (Polymeropoulos et al, 1997). The α-synuclein gene codes for a presynaptic protein thought to be involved in neuronal plasticity. The resultant ALA53THR change is predicted to lead to the disruption of an α-helix and continuation of ß-sheet. ß-sheets are thought to lead to the self-aggregation of proteins which might cause the formation of amyloid like structures (Polymeropoulos et al, 1997). A Rat homologue of this protein, *SYN1*, has been shown to be expressed at high levels in the limbic system and at intermediary levels in the basal ganglia (Maroteaux et al, 1991).

A juvenile form of PD, absent of Lewy bodies, exists showing clear signs of an autosomal recessive mode of transmission (Takahashi et al, 1994). No candidate loci have been implicated to date.

A rare autosomal recessive disorder, Aceruloplasmiemia, exhibiting symptoms similar to PD has been shown to be caused by mutations in the ceruloplasmin gene which has been mapped to chromosome 3 (Harris et al, 1995. Bost et al, 1995). Ceruloplasmin is a copper containing protein found in blood plasma. Despite the decreased serum concentration in Wilson's disease, ceruloplasmin has no essential role in copper transport (Miyajima et al, 1987). The protein has ferroxidase activity and catalyses the reaction:

$$4Fe(II) + 4H^+ + 4O_2 \rightarrow 4Fe(III) + 2H_2O$$

It is believed that mutations in this protein, which lead to aceruloplasminemia, cause an accumulation of iron in the retina and basal ganglia (Klomp et al, 1996). Expression studies have detected ceruloplasmin transcripts in glial cells associated with brain microvasculature surrounding the substantia nigra (Klomp et al, 1996). The location of ceruloplasmin in brain is in conjunction with the high storage of iron in the macrophages and glia surrounding the SNc and striatum (Gelman, 1995. Poirier et al, 1991). Affected patients have been shown to have a significantly higher level of lipid peroxidation compared to controls which implies free-radical mediated tissue damage (Miyajima et al, 1996).

Gathering the evidence to date, there appears to be no single gene responsible for PD. It is therefore important to gain an understanding of the biochemical mechanisms which lead to substantia nigra-specific necrosis. Analysis of the effects of exogenous and endogenous toxins has shone some light on PD pathology.

(iv) *Exogenous and endogenous toxins*

N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), a by-product of the synthesis of the opiate meperidine, produces in humans, symptoms and pathology similar to idiopathic PD. Dopamine neurons of the SNc are damaged and there are Lewy body-like cellular inclusions. Dopamine markers are depleted in the nigrostriatal tract and to a lesser extent in the mesolimbic pathway (Furtado and Mazurek, 1991). Initially, MPTP is oxidised by monoamine oxidase-B (MAO-B), in glia, producing, N-methyl-4-phenylpyridine (MPP^+) (Ransom et al, 1987). This metabolite then elicits its cell-specific death due to its uptake by the dopamine neuronal uptake system (Javitch et al, 1985). Indeed, the difficulty in producing MPTP-induced nigrostriatal-specific neurotoxicity in rats is due to the low density of MPTP binding sites in the substantia nigra and caudate-putamen (Javitch et al, 1984. Javitch et al, 1985. Zuddas et al, 1994). Once taken into the cytoplasm,

MPP^+ crosses into mitochondria via an energy-dependent uptake system where it binds to and inhibits complex I (NADH dehydrogenase) of the electron transport chain (Ramsey et al, 1991). This interference with oxidative phosphorylation leads to a decrease in ATP production which can cause cell death (Chen et al, 1991). Dopamine has also been reported to inhibit complex I through direct interaction of the electron transport chain (Shachar et al, 1995). That MPTP-induced parkinsonism is a good model of idiopathic Parkinson's disease is strengthened by the observation that Parkinson's patients also exhibit complex I deficiencies (Parker et al, 1989).

MPTP has been shown to cause neuronal damage by increasing the release of dopamine into the straitum by a sodium dependent mechanism (Strinathsinghji et al, 1988. O'Dell et al, 1993). Autoxidation of dopamine generates cross-linked quinones and neuromelanin which yield highly reactive hemiquinones and cytotoxic oxygen radicals, hydrogen peroxide (H_2O_2) and superoxide anions ($^{\cdot}O_2^-$) (Coyle and Puttfarcken, 1993). Oxidation of dopamine by MAO-B also generates H_2O_2 which slowly decomposes to highly reactive hydroxyl radicals ($^{\cdot}OH$), a process that is accelerated in the presence of ferrous iron (Fe^{++}) by the Fenton reaction (Coyle and Puttfarcken, 1993. Gelman, 1995). Lipid peroxidation results from excesses of reactive oxygen species which ultimately leads to a decrease in membrane integrity and cell death (Coyle and Puttfarcken, 1993). The process of lipid peroxidation itself can generate $^{\cdot}OH$ molecules by the formation of conjugated dienes which therefore accelerates the process of free radical induced cytotoxicity (Coyle and Puttfarcken, 1993). Diseases leading to excess iron deposition or improper iron storage in the basal ganglia (e.g. Aceruloplasminemia), could also cause neuronal cell death by these mechanisms (Gelman, 1995. Harris et al, 1995). Indeed the examination of postmortem brains of idiopathic Parkinson's patients found a 176% increase in the levels of total iron content in the SNc compared to age-matched controls. There were no significant differences in iron levels in the putamen, globus pallidus and hippocampus between Parkinsonian and normal brains (Sofic et al, 1988).

6-hydroxydopamine (6-ODA) is a well studied neurotoxic agent which selectively lesions dopamine systems in rats when administered intraventricularly (Poirier et al, 1991). There is evidence that this catecholamine also induces its cytotoxicity by the generation of free radicals due to the high levels of intracellular calcium observed (Kumar et al, 1995). Oxidative stress promotes calcium release from the endoplasmic reticulum (Kim et al, 1988) and mitochondria (Sandri et al, 1991) which can in turn catalyse free radical producing reactions, summarised in fig. 1.2.

1: Ca^{++} + Phospholipase A2 + O_2 ⟶ Arachadonic acid + $^{\cdot}O_2^-$

2: Ca^{++} + Calpain + Xanthine dehydrogenase ⟶ Xanthine oxidase

Xanthine + O_2 $\xrightarrow{\text{Xanthine oxidase}}$ Uric acid + $^{\cdot}O_2^-$ + H_2O_2

3: Ca^{++} + Calmodulin + Nitric oxide synthetase ⟶ NO

NO + $^{\cdot}O_2^-$ + H^+ ⟶ ONOOH ⟶ $^{\cdot}OH$ + NO_2

Fig. 1.2 Mechanisms of calcium induced free radical production. NO = nitric oxide and ONOOH = peroxymitrite. Adapted from Coyle and Puttfarcken (1993).

Intracellular calcium can also be increased by the additional effects of MPTP and dopamine. Excess electrons from the electron transport chain, due to dopamine and MPTP-induced ATP depletion, may cause neuronal depolarisation and secondary activation of voltage dependent NMDA, AMPA and kainate receptors and other voltage dependent calcium channels (Schulz et al, 1995. Coyle and Puttfarcken, 1993. Savidge and Bristow, 1997. Alagarsamy et al, 1995, Regan and Choi, 1994). Activation of these channels can also lead to more immediate cytotoxicity by allowing the massive influx of Na^+ and Cl^- ions, resulting in osmotic lysis of neurons (Coyle and Puttfarcken, 1993).

That excessive activation of NMDA receptors can lead to neuronal injury (Regan and Choi, 1994), by these mechanisms (Schulz et al, 1995), is strengthened by the observation that NMDA antagonists have been shown to protect substantia nigra neurons in MPTP-induced degeneration in mice and primates (Brouillet and Beal, 1993. Lange et al, 1993). NMDA antagonists have also been shown to counteract side effects of the commonly used dopamine replacement drug, L-dopa, in MPTP treated mice and marmosets (Fredriksson et al, 1994. Kupsch et al, 1992). Although L-dopa greatly counteracts Parkinsonian symptoms, its long term use becomes less beneficial as the side effects of dyskinesia become more apparent (Blanchet et al, 1997). It is thought that autoxidation of L-dopa, leading to free radical production, may exacerbate the symptoms of Parkinson's disease in the long term (Basma et al, 1995).

Superoxide dismutase (SOD), catalase and glatathione peroxidase (GPx) are enzymes responsible for the detoxification of $^{\cdot}O_2^-$ and H_2O_2. Copper/zinc SOD (Cu/ZnSOD), found in the cytoplasm, extracellular Cu/ZnSOD (E-Cu/ZnSOD), found in plasma and extracellular fluid, and manganese SOD (MnSOD), found in the mitochondria, all catalyse the dismutation of $^{\cdot}O_2^-$ to H_2O_2, thereby reducing the risk of $^{\cdot}OH$ formation through the Haber-Weiss reaction. Catalase and GPx then reduce H_2O_2 to water, thereby reducing the risk of $^{\cdot}OH$ formation through the Fenton reaction (Coyle and Puttfarcken, 1993. Basma et al, 1995). In concert, these enzymes reduce the amount of $^{\cdot}OH$, the most reactive and noxious free radical species. Reduced activity of GPx has been reported in MPTP-treated mice, though this was prevented by pre-treatment with the antioxidants, α-tocopherol and ß-carotene (Yong et al, 1986). Low levels of reduced glutathione, a substrate necessary for GPx, have also been reported in idiopathic Parkinsonian brains, though the reduction maybe an effect as opposed to a cause (Riederer et al, 1989). Deficiencies in glutathione reductase, the enzyme responsible for keeping glatathione in its reduced state, may explain the observed reductions of GPx in PD and MPTP-treated brains.

The most convincing evidence so far for a link between neurological disorders and free radical formation is the association found between familial amyotropic lateral sclerosis (fALS) and mutations in the Cu/ZnSOD gene. Eleven different misense mutations in the Cu/ZnSOD gene, present on chromosome 21q22.1, have been linked to this autosomal dominant disorder, a degenerative disease of motor neurons in the central nervous system (Rosen et al, 1993). The three most frequently occurring mutations were shown to lead to a 50% decrease in Cu/ZnSOD activity in red blood cells of fALS patients (Deng et al, 1993). X-ray crystallographic analysis was used to determine the structural consequences of the resultant amino acid changes, Ala^4 to Val, His^{43} to Arg and Glu^{100} to Gly on the dimeric Cu/ZnSOD protein. These amino acids were shown to be crucial for ß-barrel folding and dimer contact (Deng et al, 1993). It was concluded that the dominant effect of these amino acid changes may reflect their localisation to positions where both subunits are effected. In contrast to a reduction of SOD activity leading to ALS, expression of high levels of wildtype human Cu/ZnSOD, was shown to cause a motor neuron disorder in mice (Bar-Peled et al, 1996. Peled-Kamar et al, 1997). Because Cu/ZnSOD has evolved to be one of the fastest enzymes known ($V_{max} \sim 2x10^9\ M^{-1}\ s^{-1}$), elevation of its activity or its overexpression, without a concomitant increase in GPx activity, would lead to an accumulation of H_2O_2 (Peled-Kamar et al, 1997. Getzoff et al, 1992). Excess intracelluler H_2O_2 could then lead to Fenton chemistry

as described above. Besides catalysing the dismutation of $^{\cdot}O_2^-$ to H_2O_2, Cu/ZnSOD can also generate $^{\cdot}OH$ molecules from H_2O_2, which would also explain biological damage associated with gain-of-function SOD activity (Yim et al, 1990).

(v) Treatment

The main aim of drug therapy in the treatment of PD is to restore the balance between dopamine and acetylcholine. This can be achieved by increasing the level of dopamine, blocking the action of acetylcholine or a combination of both. In most newly diagnosed patients improvements can be achieved by careful introduction of one or more of these drugs, the main features of which are listed in table 1.2. The most commonly prescribed drug, L-dopa, has been in use since the 1960s and can eleviate most symptoms of PD (Godwin-Austin, 1994). Long term use can however cause dyskinesis possibly due to its free radical producing properties as described above. Seligiline, a MAO-B inhibitor, has been shown to slow down the process of neurodegeneration and rescue dying neurons from MPTP-induced Parkinsonism (Tatton and Greenwood, 1991). The dopamine receptor agonist, bromocriptine, has also been shown to protect against neuronal death, possibly due to its free radical scavenging activity (Liu et al, 1995).

A revival of a old surgical procedure used in the 1940s which was abandoned on the introduction of L-dopa, has bought hope for some intractable PD patients. The procedure, termed pallidotomy, involves the lesioning of the internal globus pallidus (GP). This neurosurgical procedure is used to prevent abnormal output signals from the GP due to the depletion of striatal dopamine which results in abnormal movements characteristic of parkinsonism. There are however varying degrees of success with this technique (Godwin-Austin, 1994. Martin, 1996).

Transplantation of foetal dopamine neurons into the brains of patients with PD has been shown to ameliorate symptoms of the disease. However, the symptomatic relief is not complete, possibly due to the poor survival of the transplantation tissue during the procedure (Nakao et al, 1994. Nakao et al, 1995. Barkats et al, 1997). Similar procedures using rat and human dopamine neurons into the brains of 6-ODA-treated rats showed only 5-20% survival of the transplanted tissue (Frodl et al, 1994). Significant improvements have been made by treating the grafted tissue with lazeroids before implantation into rat brains (Nakao et al, 1994). These compounds were believed to improve the survival of grafted tissue due to their ability to potently

Generic Name	Proprietary Names	Description
Lævodopa	Madopar, Sinemet	Used since the '60s, it replaces dopamine but can cause nausea, involuntary movements and changes in mental state. Madopar and Sinemet also contain the compound carbidopa, a decarboxylase inhibitor which minimises the metabolism of L-dopa before it reaches the brain.
Selegiline	Deprenyl, Eldepryl	Prevents the breakdown of dopamine by inhibiting MAO-B and may slow down neurodegeneration. Acts as a mild stimulant and may therefore interfere with sleep.
Bromocriptine Lisuride Pergolide	Parlodel Revanil Celance	These dopa agonists are usually taken in conjunction with L-dopa. Have a longer lasting effect than L-dopa alone. May give rise to nausea and sickness.
Benzhexol Benztropine Orphenadrine Procyclidine Methixene	Artane, Pipanol Cogentin Dispal Kemadrin Tremonil	These anticholinergics have a mild effect on the symptoms. Useful for the younger patients in the early stages of the disease. May cause blurring of vision and a dry mouth. Not used as often nowadays.
Amantadine	Symmetrel	This drug allows dopamine to stay longer at it's site of action. It has few side effects and it may be useful in the control of tremor. However effects are short lived.
Apomorphine		For patients who experience fluctuations in their symptoms despite usual medication. Injections of this drug of this drug has a direct effect on the site of action of dopamine. Doperidone, 'Motilium', has to be taken to combat the severe nausea since apomorphine is a very strong dopamine agonist. Only patients responding to treatment with Sinemet or Madopor can take this drug.

Table 1.2 List of drugs commonly prescribed for Parkinson's patients. Taken from Godwin-Austin (1994).

inhibit iron-dependent lipid peroxidation and thereby suggesting that free radicals may be generated during mechanical dissociation of tissue. Based on this premise, a study was conducted to graft murine mesencephalic tissue, which overexpresses human Cu/Zn SOD, into the brains of 6-ODA lesioned rats. The transgenic tissue survived 4 times longer than the nontransgenic tissue (Nakao et al, 1995). The survival of the transgenic and non transgenic tissues was only monitored over a short period (5 weeks) by which time the correlation between the presence of the transgene and graft tissue survival was beginning to weaken. Perhaps over the long term this technique may not be totally beneficial as overexpression or increased activity of SOD can lead to free radical generation in itself (Peled-Kamar et al, 1997. Bar-Peled et al, 1996. Yim et al, 1990). A more strategic approach, addressing the balance of all detoxifying enzymes, may generate improved long term benefits for grafted tissue.

A possible new approach for the drug-treatment of PD patients is the finding that repeated injections of glial-cell-line-derived neurotrophic factor (GDNF) has been shown to protect and repair mesencephalic dopamine neurons in MPTP-treated mice and axotomy-induced degeneration in rats (Tomac et al, 1995. Beck, et al, 1995). Treatment of 6-ODA lesioned rats with injections of GDNF, or a recombinant adeno-associated virus vector expressing GDNF, found no change in striatal dopamine levels or uptake (During and Leone, 1997). Another study used an adenoviral vector expressing human GDNF cDNA in 6-ODA lesioned rats. A significant survival effect at 6 weeks was obtained using this recombinant vector system (During and Leone, 1997). The two conflicting results could possibly be explained due to a better efficiency of the adenoviral vector system to express GDNF. The ability to arrest neurodegeneration using GDNF could have an enormous impact in the treatment of patients with PD.

The *weaver* mutant mouse

(i) *Symptoms and Pathology*

The only previously described inheritable model which bears similarities to the locomotor abnormalities and nigrostriatal degeneration seen in human PD is the *weaver* mutation in mouse (Schmidt et al, 1982). Behaviourally, *weaver* mice differ from PD in that there is an absence of akinetic rigidity and ataxia is involved, i.e. inaccuracies in speed, force and distance of movement. These differences are

due to the profound loss of granule cells in the cerebellum, which is totally spared in PD. Other regions of the brain affected in PD, the locus cœrulus and raphe nuclei, are normal in *weaver* mice (Bandmann et al, 1996). The *weaver* mutation, which maps to mouse chromosome 16 (Reever et al, 1989), results from a Gly to Ser amino acid change in the G-protein-gated inward rectifier K^+ channel (*GIRK2*) (Tong et al, 1996. Silverman et al, 1996). This channel also conducts the flow of Na^+ ions and in its mutant state is also permeable to Ca^{++}. Therefore, the degenerative *weaver* phenotype could occur due to increased influxes of Na^+ and Ca^{++} leading to a collapse of ion gradients and eventual osmotic-induced or free radical-induced cell death (Tong et al, 1996. Silverman et al, 1996).

(ii) *Comparisons with the AS/AGU rat*

The AS/AGU rat bears similarities with the *weaver* mouse with respect to the loss of dopaminergic neurons in the substantia nigra (Clarke and Payne, 1994). Where these two mutant rodents differ is in the additional cerebellar abnormalities and the observation that the locus cœrulus and raphe nuclei are not affected in the *weaver* mouse (Bandmann et al, 1996. Scott et al, 1994. Stewart et al, 1994). The distribution of the specific morphological and biochemical changes seen in the AS/AGU rat appear more similar to human PD than those seen in the *weaver* mouse. Mapping and characterisation of the gene responsible for the phenotype present in the AS/AGU strain would therefore prove valuable in the understanding of striatonigral-specific neurodegeneration seen in disorders like PD.

Strategy for Mapping the locus responsible for causing the locomotor defects and striatalnigral pathology in the AS/AGU rat

From the initial test crosses carried out between AS and AS/AGU rats, it was apparent that the mutant locus, *agu*, was inherited in an autosomal recessive fashion (Campbell et al, 1996). Crossing the AS/AGU strain with a suitably divergent reference strain and then backcrossing the resultant F1 progeny to AS/AGU would provide a means for analysing the cosegregation of markers with the *agu/agu* genotype, which would be present in 50% of the backcross progeny.

The type of markers now commonly used for genomic mapping projects due to their abundance, their tendency to be highly polymorphic and wide distribution

throughout mammalian genomes are microsatellites. They consist of relatively short runs of repeated sequences with repeat unit lengths of 6bp or less (Beckmann and Weber, 1992). The most common type of microsatellite populating mammalian genomes is the $(AC)_n/(GT)_n$ repeat, with an average of one in every 30kbp of DNA (Beckmann and Weber, 1992). Analysis of length polymorphisms are carried out by use of the polymerase chain reaction (PCR) using primers which anneal to unique sequences flanking the repeat unit. Relative sizes of PCR products can be detected by polyacrylamide or agarose gel electrophoresis (Litt and Luty, 1988). The number of microsatellite markers (gene-based and anonymous) available for the rat genome has grown considerably over recent years due to contributions from several key research groups because of the importance of this animal as an experimental model for the study of biomedical research and genetic disease (Jacob et al, 1995. Serikawa et al, 1992. Levan et al, 1990. Goldmuntz et al, 1993). Perhaps the main landmark of rat genetics in recent years has been the production of the first complete rat genetic map in 1995 consisting of 432 molecular markers (Jacob et al, 1995).

Three reference strains have been co-analysed with the AS/AGU strain in order to detect rates of polymorphisms between them using available microsatellite markers and therefore determine which reference strain(s) would be suitable for construction of a backcross. Seventy-five microsatellites, spanning the rat genome, were examined in the inbred strains, AS/AGU, PVG, BN and F344. AS/AGU differed from BN, F344 and PVG at 62%, 47% and 43% of loci, receptively (Shiels et al, 1995). The strains BN and F344 were therefore used to construct backcrosses with AS/AGU as follows, [AS/AGUxF1(AS/AGUxBN)] and [AS/AGUxF1(AS/AGUx F344)].

Aims of My Project

As part of the *agu* mapping project, my aims were as follows:

1, To contribute to the analysis of microsatellite markers using the available rat strains:.
2, To help confirm or reject any possible linkage of *agu* to microsatellite markers using the two backcross programs, [AS/AGUxF1(AS/AGUxBN)] and [AS/AGUx F1(AS/AGUx F344)].
3, To find any new microsatellites near the *agu* locus and therefore strengthen the rat genetic map.

Chapter 2

Materials and Methods

Chemical and Reagent Supplies

Acrylamide	Anachem
Adenosine triphosphate	Boehringer Mannheim
Agarose	GibcoBRL
Antibiotics	Sigma
Bacterial media	Difco, Oxoid
DNA markers	GibcoBRL
Enzymes: AmpliTaq, FS	Perkin Elmer
Restriction endonucleases	GibcoBRL
Ribonuclease A	Sigma
SAP	USB
T7 DNA polymerase	Pharmacia Biotech
All other enzymes	Promega
General chemicals	B.D.H., Fisons, Sigma
IPTG	Sigma
MetaPhor agarose	F.M.C.
Nylon membranes	M.S.I.
Organic chemicals	Aldrich, B.D.H., Fisons, Sigma
Sephadex G-25 (NAP columns)	Pharmacia Biotech
Radiochemicals	New England Nuclear
X-Gal	Boehringer Mannheim
X-ray film	Fuji XR

Bacterial Strains (*E. coli*)

DH10B: F' *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80d*lacZ*ΔM15 Δ*lacX74* *endA1 recA1 deoR* Δ(*ara-leu*)7697 *araD139 galUK nupG rpsL* (Grant et al, 1990).

F3: DH10B [F::Tn10d-Cam] (Kimmerly et al, 1994).

NS3145: F'[*lacI*^{q}] *recA1 hsdR mcrAB* λimm434*nin5XI-cre* (Southard-Smith and MacDonald, 1993).

TG1: *supE hsd* Δ5 *thi* Δ(*lac-proAB*) F'[*traD36 proAB' lacI*^{q} *lacZ*ΔM15] (Gibson, 1984).

Bacterial Growth Media

L-agar: 1% w/v tryptone, 0.5% w/v yeast extract, 0.02% w/v thymine, 0.88M NaCl, 1.5% w/v agar

L-broth: 1% w/v tryptone, 0.5% w/v yeast extract, 0.1% w/v glucose, 0.02% w/v thymine, 0.88M NaCl

2x YT-broth: 1% w/v bacto-tryptone, 1% w/v yeast extract, 0.88M NaCl

Antibiotics

	Stock Solution	Working Concentration
Ampicillin	100mg/ml in water	100µg/ml
Chloramphenicol	100mg/ml in ethanol	10µg/ml
Kanamycin	100mg/ml in water	25µg/ml
Streptomycin	100mg/ml in water	100µg/ml

Library

A rat genomic P1 library, supplied by Du Pont NEN, was constructed as previously described (Southard-Smith and MacDonald, 1993). The vector used was pA*d*10-*Sac* BII (Pierce et al, 1992).

Vector

Plasmid subclones were prepared using pUC-18 (Yannish-Perron et al, 1985).

Solutions

Media and solutions were sterilised by autoclaving at 120°C, 15 p.s.i. for 15 minutes or by filtration through a 0.2µm pore filter (Acrodisc™, Gelman Sciences).

Birnboim-Doly (BD) 1: 50mM glucose, 25mM Tris, 10mM EDTA, (pH8.0).

Birnboim-Doly (BD) 2: 0.2M NaOH (fresh), 1% SDS.

Birnboim-Doly (BD) 3: 60ml 5M Potassium acetate, 11.5ml glacial acetic acid and 28.5ml ddH_2O, (pH4.8).

Chloroform: mixed with iso-amyl alcohol in a 24:1 ratio, stored at r.t. in the dark.

Denaturing Solution: 1.5M NaCl, 0.5M NaOH.

100x Denhardt's solution (Denhardt, 1966): 2g Ficoll, 2g polyvinylpyrrolidone, 2g BSA (Fraction V, fat free), ddH_2O to 100ml.

5x DNA loading buffer: 0.25% xylene cyanol, 0.25% bromophenol blue, 15% Ficoll-400.

Formamide: mixed with 10% w/v amberlite for 30 minutes and filtered. Stored at -20°C in the dark.

Neutralizing Solution: 1.5M NaCl, 0.5M Tris.HCl (pH7.0).

Oligo Hybridisation Buffer: 5ml 100X Denhardts, 25ml 20x SSC, 1.6ml salmon sperm DNA, 1ml sodium pyrophosphate, 5ml 1M sodium phosphate buffer pH6.8, 20ml deionized formamide. Made upto 100ml with ddH_2O. Stored at -20°C.

Phenol: All phenol used in DNA preparation was buffered against 1M Tris-HCl, pH8.0. For colouration, 0.1% w/v hydroxyquinoline was added. Phenol was then stored at -20°C in the dark for upto 3 months.

RNaseA Solution: 100mg/ml ddH_2O, boiled for 10 minutes. Stored at -20°C.

1x TBE (Tris borate buffer): 10.8g Tris base, 5.5g Boric acid, 4mls 0.5M EDTA (pH8.0), ddH_2O to 1000ml.

1x TAE (Tris acetate buffer): 0.04M Tris-acetate, 0.001M EDTA

TE buffer: 10mM Tris (pH7.5), 1mM EDTA (pH8.0).

20x SSC: 175.3g NaCl, 88.2g NaCitrate, ddH_2O to 1000ml.

STOP loading buffer: 95% deionized formamide containing 0.1% w/v xylene cyanol, 0.1% bromophenol blue and 10mM $Na_2EDTA.2H_2O$. This was then stored at -20°C in the dark.

TMG-glycerol: 10mMTris-HCl pH 8.0, 10mM $MgCl_2$, 0.1% gelatin, 30% glycerol.

Transformation Buffer: 10mM MES (2 -[N- Morpholino] ethanesulphonic acid), 100mM RhCl, 45mM MnCl, 10mM CaCl, 3mM hexammine cobalt (III) Cl.
X-gal: 2% w/v in dimethylformamide.

Use of IPTG and X-gal

To freshly prepared dry agar plates, 40µl each of IPTG (0.1M) and X-gal (2%) were plated with the aid of a glass spreader aseptically. When bacteria tranformed with pUC18 were grown on thes plates they turned blue due to the presence of the ß-galactosidase gene on the vector. When tranformed with recombinant plasmids, the cells turned white due to the presence of the multiple cloning site within the ß-galactosidase gene.

Growth and Storage Conditions

Liquid cultures were grown overnight at 37°C with vigorous shaking in l-broth or 2xYT with the required antibiotic. Plate cultures were incubated upside down overnight at 37°C. Stocks of bacterial strains were made by adding glycerol to 1ml of culture to make a final concentration of 15% w/v. These were stored at -70°C. In the case of bacteria containing the P1 plasmid, 1ml of an overnight culture was centrifuged briefly to collect the cells which were then resuspended in TMG-glycerol. This was also placed at -70°C for long term storage. For short term storage, bacteria could be left on sealed plates upside down at 4°C.

F' Episome Mediated Transfer of the P1 Plasmid

To 1.8ml of l-broth, 100µl of a single colony overnight culture containing a P1 clone and 100µl of an F3 overnight culture were added and shaken with a speed of 30 r.p.m. for 3 hours at 30°C. After mating, this was diluted to 1 in 1111 of which 30µl were plated out onto an agar plate containing kanamycin and chloramphenicol. Cells which had grown on an agar plate containing kanamycin and chloramphenicol were due to the successful mating between NS3145 and F3 strains. For the second round of mating, two overnight cultures were set up, one containing a single colony from the previous mating and the other containing DH10B. To 1.8ml of l-broth, 100µl of each overnight was added and shaken with a speed of 30 r.p.m. for 3 hours at 30°C as before. After mating, 100µl of this was plated onto an agar plate conataining kanamycin, chloramphenicol and streptomycin. Succesfully mated cells were grown up in culture in order to make long term stocks.

Plasmid Purification

Both methods are modifications from the method of Birnboim and Doly, (1979).

(i) *Small scale preparation*

From an overnight culture of cells, 1.5ml were transferred to an eppendorf tube and centrifuged for 1 minute at top speed. The supernatant was removed and a further 1.5ml of cells were added and centrifuged as before. The supernatant was removed and the cell pellet was resuspended in 200µl of BD-1. To this, 300µl BD-2 were added and mixed gently by repeated inversions. After a 5 minute incubation on ice, 300µl of ice cold BD-3 were added to neutralize. The contents were centrifuged for 10 minutes to remove chromosomal DNA and cell debris. The supernatant was transferred to a fresh tube and 400µl of chloroform/iso-amyl alchohol (24:1) were added. The two layers were mixed by inversion and the contents centrifuged for 1 minute. DNA was precipitated and collected by the addition of 0.7x volume of isopropanol, mixing by inversion and centrifugation at r.t. for 10 minutes. The supernatant was aspirated and the pellet of DNA was washed with 70% ethanol, air dried and resuspended in 32µl of TE. The plasmid DNA was then treated with 2µl of RNAse A at 37°C for 30 minutes.

For sequence-grade purity, 8µl of 4M NaCl and 40µl of 13% PEG were added, mixed well and incubated on ice for 20 minutes. DNA was collected by a 15 minute centrifugation at 4°C. The pellet was washed in 500µl of 70% ethanol, air dried for 10 minutes and resuspended in 20µl of ddH_2O.

(ii) *Large scale preparation (P1 DNA preparation)*

From an overnight culture, 50ml were transferred to a Nalgene tube and centrifuged at 5,000 r.p.m. for 10 minutes. The supernatant was discarded and the pellet of cells was resuspended in 3ml of ddH_2O and centrifuged as before. The cell pellet was resuspended in 3.5ml of BD-1. To this 4ml of BD-2 were added and the contents were mixed by inversion and incubated on ice for 5 minutes. After the solution became clear and viscous, 4ml of BD-3 were added to neutralize and the contents were mixed gently by inversion and incubated on ice for 20 minutes. The contents were centrifuged at 12,000 r.p.m. for 20 minutes to get rid of chromosomal DNA and cell debris. DNA was collected by the addition of an equal volume of isopropanol to the supernatant, gentle mixing and centrifugation at 12,000 r.p.m. for 20 minutes. The supernatant was discarded, and the pellet resupended in 3ml of sterile ddH_2O and 300µl of 3M Na acetate. To remove any protein contaminants, an equal volume of phenol:chloroform (1:1) was added and the layers were mixed by vigorous vortexing. The two phases were separated by centrifugation at 12,000 r.p.m. for 5 minutes. To the aqueous phase, 2 volumes of 100% ethanol were added. DNA was collected by centrifugation at 12,000 r.p.m. for 10 minutes and the pellet was washed with 70% ethanol, air dried and resuspended in 50µl of TE. The plasmid DNA was then treated with 2µl of RNAse A at 37°C for 30 minutes.

Transformation of *E.coli*

(i) *Preparation of competent TG1 cells*

To 10ml of 2xYT, 200µl of a fresh overnight culture was added and incubated with vigorous shaking at 37°C. Once growth had reached mid log phase (O.D.600nm= 0.5), the cells were transferred to a nalgene tube and pelleted at 5,000 r.p.m. for 10 minutes in a Beckman JA-21 centrifuge. The supernatant was discarded and the cells resuspended in 2ml of transformation buffer. The cells were then pelleted again as before. The supernatant was discarded and the cells were resuspended in

1ml of transformation buffer. After a 15 minute incubation on ice, 34µl of dimethylformamide were added. After a 5 minute incubation on ice with gentle mixing, 34µl of 2 ß-mercaptoethanol were added. After a further 10 minute incubation on ice with gentle mixing, another 34µl of dimethylformamide were added and a final 5 minute incubation on ice rendered the cells competent for transformation.

(ii) *Transformation by heat shock*

In a 1.5ml eppendorf tube, 5-50ng of DNA were gently mixed and incubated on ice for 45 minutes with 200µl of freshly prepared competent cells. The cells were then heat shocked at 42°C for 3 minutes and replaced onto ice. The transformants were incubated in 1ml of 2xYT at 37°C for 45 minutes before being grown on selective plates overnight where individuals could then be analysed.

Animal Work

All rats were bred and maintained by Mr. N.K.Bennett (Laboratory of human anatomy, University of Glasgow). The albino Swiss (AS) strain originated from the department of anatomy, University of Glasgow, where it has remained in a closed colony since 1974. Brown Norway (BN) and F344 strains originated from Harlan Olac. Rats were sacrificed by aneasthetising with carbon dioxide followed by cervical dislocation. Spleens were removed using fresh tweezers and stored at -70°C.

Genomic DNA Preparation

Using a Puregene™ isolation kit, DNA was prepared from frozen spleen tissue. To 600µl of prechilled cell lysis solution, 10-20mg of tissue were added and quickly homogenised. After a 30 minute incubation at 65°C, 3µl of RNaseA solution was added and the lysed tissue was incubated for 30 minutes at 37°C. To the RNaseA treated cell lysate, 200µl of chilled protein precipitation solution were added and vortexed vigorously for 20 seconds. Precipitated proteins were pelleted by centrifugation at 14,000 r.p.m. for 3 minutes and the supernatant was transferred to a fresh tube containing 600µl of isopropanol. DNA was precipitated and collected by gentle inversions and centrifugation at 14,000 r.p.m. for 1 minute.

The DNA pellet was washed in 70% ethanol and resuspended in 100µl of DNA hydration solution.

Agarose Gel Electrophoresis

Horizontal gels were prepared by boiling together predetermined proportions of agarose and 1x running buffer containing 50mg/ml ethidium bromide. After cooling to 50°C, the gel was poured into perspex formers to a depth of about 6 mm. DNA samples were loaded after addition of 0.2 volumes DNA loading buffer. A size marker was always simultaneously electrophoresed with any samples. This, along with the fact that electrophoretic mobility is inversely proportional to $\log^{10}$ molecular weight, allowed accurate sizing of fragments once they had been separated. The gel was electrophoresed in the appropriate running buffer for between 2 and 20 hours at a voltage of between 1-3V/cm depending on the degree of resolution required. The DNA was visualized using a UV transilluminator and photographed through a UVP camera using a Mitshubishi video copy processor loaded with a Mitshubishi K65 HM thermal film pack.

MetaPhor Agarose Gel Electrophoresis

During the course of this study, most size differences between PCR products were resolved by using MetaPhor agarose as opposed to polyacrylamide. This quicker, cleaner and safer method could detect a difference as low as a 2bp and resolve a difference of 4bp amongst DNA fragments ranging in size between 75bp and 350bp.

This high resolution agarose was electrophoresed in the same manner as normal agarose. In order to perform with maximum efficiency however, the following extra measures were taken. The typical concentration of this agarose used was 4%, though a value as high as 6% was at times also necessary. MetaPhor agarose was dissolved in pre-chilled 1x running buffer by very gentle heating to avoid frothing. After the mix was cast and set, the gel was chilled for 30 minutes before use. It was also ideal for the gel to be electrophoresed in chilled running buffer and that the temperature of this to remain cool throughout. Due to the fact that ethidium bromide travels in the opposite direction to DNA, causing slight retardation, MetaPhor gels were stained after and not during electrophoresis.

Agarose Gel Electrophoresis of Large DNA Molecules

DNA molecules greater than 25kb no longer run at rates determined by their sizes but all run at about the same rate called the limiting mobility. If the electric field is reversed, DNA must assume a new confirmation before it starts travelling in the opposite direction. If the field is reversed for a short time only, DNA will not travel much in the reverse direction. The time taken for DNA to travel in the forward direction again is determined by its size, larger DNA molecules take a longer time to change confirmation. This technique, called field inversion gel electrophoresis (Carle et al, 1986) slows down large molecules that would otherwise run at limiting mobility.

Normal agarose gels (0.8%) were precast as before. A PPI-200 programmable electrophoresis controller (MJ research) was connected between the gel tank and the power supply and depending on the size range of DNA to be resolved, different built-in programs were selected according to the manufacturers instructions. The running buffer was recirculated by connecting a peristaltic pump (Genetic Research Instruments) to the gel tank. Gels were run in the cold room at 5-8 volts/cm for 18 hours and visualised and photographed as before.

Denaturing Polyacrylamide Gel Electrophoresis

Analysis of single stranded fragments of DNA was carried out on polyacrylamide gels containing high concentrations of urea. The presence of urea denatures DNA molecules and therfore ensures that they migrate at rates almost completely independent of their base composition and sequence.

(i) *Preparation of glass plates*

Two suitable glass plates were scrubbed thoroughly with warm water and detergent then rinsed well. After washing with isopropanol, the plates were left to dry and one was treated with Sigmacote. Gel spacers were placed into position and the glass plates were carefully clamped together.

(ii) *Preparation of gel*

An 8% polyacrlamide/7M Urea gel was prepared by mixing 25.2g of urea, 6ml of 10x TBE and 23ml of ddH_2O under hot water. After the urea had dissolved, 12ml of acrylamide (19:1), 30µl of TEMED and 700µl of 10% ammonium persulphate were added. Quickly but carefully, the gel mixture was poured between the glass plates with the aid of a syringe and needle and a shark-tooth comb was fixed into place. Once the gel had polymerized, the comb area was cleaned to remove excess acrylamide.

(iii) *Running of gel*

The gel was pre-run at 60W for 30 minutes in order to bring the gel to optimal temperature and establish a voltage gradient throught it. Between 3-5µl of DNA samples were loaded into individual wells. The gel was run at 60W until the marker dyes had migrated the desired distance. In an 8% acrylamide gel, xylene cyanol and bromophenol blue comigrated with DNA at sizes of 160 and 45 nucleotides respectively.

(iv) *Analysis of gel*

After electrophoresis the gel plates were separated and the gel transferred to 3MM paper (Whatman), dried under vacuum at 80°C and subject to autoradiography by placing it in direct contact with X-ray film in a cassette with an intensifying screen at -70°C.

Non-denaturing Polyacrylamide Gel Electrophoresis

Sepearation of single stranded fragments of DNA due to their base composition and sequence, as well as size, was carried out in the absence of any denaturant. Electrophoresis with non-denaturing gels was similar to denaturing gels with the following exceptions: A 6% acrylamide gel was prepared by mixing 58ml ddH_2O, 4ml 10x TBE, 8ml glycerol, 12ml acrylamide (19:1 or 37.5:1), 800µl of 10% ammonium persulphate and 22µl of TEMED; The gel was run at 30W for 5-6 hours at 4°C.

Southern blotting

DNA was transferred onto a nylon membrane using a method based on that described by Southern (1975). Following electrophoresis, the gel was soaked in denaturing solution for 30 minutes. After a brief rinse in ddH_2O, the gel was immersed in neutralising solution for a further 30 minutes. A membrane was placed on the gel and 20 X SSC was allowed to soak through by capillary action overnight, taking the DNA with it. The membrane was then rinsed in 2 X SSC and baked at 80°C for 2 hours to fix the DNA to the membrane.

Transfer of Colonies onto Nylon Membranes

Bacterial colonies to be anaylsed were duplicated onto 2 fresh plates in a grided fashion. After growth at 37°C for 12 hours, nylon membranes were placed on top of the colonies for 1 minute. The filters were carefully lifted off and placed colony side up on 3MM paper soaked in denaturing solution for 7 minutes. The filters were then placed onto 3MM paper soaked in neutralising solution for 3 minutes. This last step was repeated using fresh neutralising solution. Cell debris was removed by gentle shaking in 2x SSC. After the membranes were air dried they were baked for 2 hours at 80°C.

Hybridisation of DNA Immobilised on Nylon Membranes

(i) *Preparation of Radiolabelled probes*

During the coarse of this study, all probes used were oligonucleotides of 30 bases or less. All oligonucleotides were end-labelled with [γ^{32}P]ATP using T4 polynucleotide kinase in the forward labelling reaction. The following components were incubated at 37°C for 30 minutes; 10pmol of DNA, 4µl of 5x forward reaction buffer, 20µCi of [γ^{32}P]ATP, 1µl of T4 polynucleotide kinase, ddH_2O to 20µl. After incubation, the labelled oligo was purified from unincorporated radiation by passing it through NAP Sephadex G-25 columns in the presence of 0.1M sodium phosphate buffer pH 6.8. The desired eluted fraction was identified by comigration with bromophenol blue, the undesired fraction comigrating with phenol red.

(ii) *Measurement of Specific activity*

Before using any labelled oligo, it was necessary to know if the labelling reaction had worked well. Before the labelled probe was purified, 0.5µl was spotted onto a small piece of 3MM paper (A). After purification, another 0.5µl was spotted onto each of 2 more pieces of 3MM paper (B and C), one of which (C) was washed for two 15 minute periods in 0.5M Na_2HPO_4 pH 7. The 3 samples were analysed on a scintillator using Cherenkov counting. The reading obtained for sample C was used to calculate the specific activity since this represented incorporated radiation only. A value of 10^6 cpm/µg DNA or greater was assumed to be suitable for use.

(iii) *Hybridisation of radiolabelled probe to DNA immobilised on nylon membranes*

Baked membranes were prehybridised in order to block any non-specific sites. This was carried out by placing the membrane in a sealed tube with 10ml of oligo hybridisation buffer and incubating at 42°C for 1-2 hours. A freshly labelled and purified probe (10^6 cpm/ml oligo hybridisation buffer) was added to the tube and incubation was continued for a further 12-18 hours.

(iv) *Washing of membranes*

After hybridisation, membranes were removed from the sealed tube and excess and weakly hybridised probe was removed by a series of washes. This was carried out in steps of increasing stringency. Two 15 minute washes at room temperature followed by two washes at 55°C in 6x SSC were sufficient to remove all excess radiation. The membranes were then covered with plastic sheeting and exposed onto X-ray film for 24-72 hours at -70°C with the aid of an intensifying screen.

Cleavage of DNA Molecules with Restriction Endonucleases

Digests were carried out with 0.5-5µg DNA, 10 units of enzyme and the appropriate reaction buffer in a final reaction volume of 20µl. The samples were incubated at 37°C for 90 minutes. The reaction was stopped by heat denaturing at 65°C for 10 minutes. Digested DNA molecules were usually purified by solvent extractions before further manipulations.

Dephosphorylation of Vector Ends

In order to prevent recircularisation of vector DNA during ligation, the terminal phosphate groups of its cohesive ends were removed. Once vector DNA had been digested to completion, the volume of the reaction was increased to 50µl with ddH_2O and 1µl of SAP (shrimp alkaline phosphatase) was added. Incubation at 37°C was allowed to continue for a further 20 minutes. The reaction was terminated by decontaminating the DNA using solvent extractions. The DNA was concentrated and resupended in a suitable volume of TE or ddH_2O.

Purification of DNA by Solvent Extraction

Sequential extractions with phenol/chloroform (1:1) and chloroform were used to remove contaminating proteins such as enzymes from DNA samples. This was carried out in large volumes to reduce any loss of DNA. Equal volumes of the DNA solution and organic solvent were mixed by vortexing and the phases were separated by brief centrifugation. The aqueous phase, containing DNA, was carefully removed and kept leaving behind proteins which had partitioned into the organic phase.

Concentration of DNA

To a DNA solution, $^1/_{10}{}^{th}$ volume of 3M sodium acetate (pH 4.8) and 2 volumes of ethanol were added. This was left at -70°C for 30 minutes. DNA was pelleted by centrifugation at 14,000 r.p.m. for 30 minutes. The pellet was washed in 70% ethanol and resuspended in either TE or ddH_2O.

Alternatively, 0.7 volume of isopropanol was added to the DNA solution. After brief mixing, DNA was collected by centrifugation at 14,000 r.p.m. for 30 minutes. The DNA pellet was washed in 70% ethanol, air dried and resuspended as described above. Isopropanol precipitation was favoured when large volumes became inconvienient.

DNA Ligation

DNA ligations were performed using T4 DNA ligase in the presence of the accompanying buffer and 1mM ATP, as recommended by the manufacturer. Generally, 50ng vector was ligated with the insert at a molar ratio of insert:vector of 3:1 in a total volume of 10µl. Ligations were incubated overnight at 15°C.

Polymerase Chain Reaction

Throughout the course of this study the main diagnostic procedure used for the detection of specific sequences and allelic size and sequence analyses was the polymerase chain reaction (PCR). All PCR experiments performed were made more effeicient by adding the polymerase enzyme at a high temperature. This 'hot start' technique reduced misprining, a common cause for the generation of artifacts.

A standard 10µl reaction consisted of: 20ng template DNA; 1mM $MgCl_2$; 1x Thermo Buffer; 125µM each of dCTP, dGTP, dATP and dTTP; 0.25mM of each primer. This was overlaid with a drop of mineral oil. Reaction tubes were placed in a Perkin Elmer Cetus 9600 thermocycler and heated to 100°C for 10 minutes. At a temperature of 80°C, 1 unit of *Taq* polymerase was added. A typical cycling programme entailed 20-35 cylces of, a denaturing step at 94°C for 15 seconds, an annealing step at between 53°C and 70°C for 30 seconds and an extension step at 72°C for 30 seconds. A final extension step at 72°C for 2 minutes occured before samples were cooled to 4°C. The samples were resuspended in loading buffer before being loaded onto the appropriate agarose gel for electrophoretic analysis.

Radioactive PCR

(i) *Addition of an end-labelled primer*

One of the primers (2.5 pmol) to be used in the PCR experiment was incubated for 30 minutes at 37°C with 1µl of T4 polynucleotide kinase, 10mCi of [γ^{32}P]ATP, 2µl of 5x forward reaction buffer and ddH_2O in a 10µl reaction. The reaction was terminated by incubation at 65°C. Purification of the labeled primer in preparation for PCR was not necessary.

A PCR experiment, set up as described above, was spiked with the labeled primer in a ratio of 30:1 unlabeled to labeled (Brook et al, 1992). The cycling programme used was the same as that described above. One quarter of each sample was resuspended in STOP buffer, heat denatured at 80°C for 2 minutes and loaded onto a denaturing polyacrylamide gel.

(ii) *Single-strand conformational polymorphism assay*

A 10µl PCR for this assay consisted of: 0.5µM of each primer; 5µM each of dATP, dCTP, dGTP and dTTP; 20ng of template DNA; 1x thermo buffer; 1mM $MgCl_2$; 1µCi of [α^{32}P]ATP; 5% deionised formamide and 1 unit of *Taq* polymerase. The cycling programme used was the same as described as above.

A 2µl aliquot of each completed PCR was diluted with 30µl of 0.1% SDS/10mM EDTA solution. After a brief vortex, 2µl of each diluted PCR was added to 2µl of STOP buffer, heated to 80°C for 2 minutes and loaded onto a non-denaturing polyacrylamide gel.

Double Stranded DNA Sequencing

This method, originally described by Chen and Seeburg (1985), is based on the Sanger method of chain terminating sequencing (Sanger et al, 1977). The T7 sequencing kit supplied by Pharmacia was used to sequence RNA-free plasmid DNA.

(i) *Annealing of primer*

This first step was modified from the standard dideoxy sequencing protocol to compensate for the high GC content of the primers used in these studies (Margolis et al, 1993).

To a volume of 32µl containing 3µg of plasmid DNA, 8µl of 2M NaOH were added, vortexed briefly and incubated at r.t. for 10 minutes. To this 7µl of 3M sodium acetate, 4µl of ddH_2O and 120µl of 100% ethanol were added. DNA was precipitated and collected by a 15 minute incubation on dry ice followed by centrifugation at 14,000 r.p.m. for 15 minutes. The DNA pellet washed in 70% ethanol and resuspended in 11µl. This was heated to 100°C for 2 minutes with 1µl containing 300ng of primer and 2µl of annealing buffer. The temperature was

gradually lowered to 80°C over a 15 minute period and the reaction was placed on ice for 5 minutes.

(ii) *Sequencing reactions*

The remainder of the procedure was caried out exactly according to the instructions supplied with the Pharmacia T7 sequencing kit. Working solutions of dNTPs, ddNTPs, labelling mix and T7 polymerase enzyme were supplied in the kit. Samples were analysed on a denaturing polyacrylamide gel.

Automated Sequencing

Most of the sequencing was performed on the ABI 373 DNA sequencer (Perkin Elmer) in conjuntion with the ABI PRISM dye terminator cycle sequencing ready reaction kit. Sequencing of double stranded DNA was obtained with plasmids isolated by the alkaline lysis/PEG precipitation method as described previously. PCR products were cleaned up through spin columns (Pharmacia) according to the manufacturers instructions before being sequenced.

(i) *Cycle sequencing reaction*

A typical 20µl reaction contained: 8µl of Terminator ready reaction mix; Between 300ng and 500 ng of double stranded DNA or between 30ng and 180ng of PCR product; 3.2pmole of primer and ddH_2O. Reactions were placed in a Perkin Elmer 9600 thermocycler. A cycle sequencing programme entailed 25 cylces of, a 10 second period at 94°C, a 5 second period at 50°C and a 4 minute period at 60°C. At the end of this, the samples were stored at 4°C awaiting purification.

(ii) *Purification of extension products*

Each 20µl sequencing reaction was transfered to a new tube containing 50µl of 100% ethanol and 2µl of 3M sodium acetate. The contents were gently vortexed and centrifuged at 14,000 r.p.m. for 15 minutes. The DNA pellet was washed in 250µl of 70% ethanol and resuspended in 4µl of STOP buffer. Samples were loaded on a 6% denaturing polyacrylamide gel which was electrophoresed in an ABI 373 DNA sequencer according to the maufacturers instructions.

Computing Software

The design of PCR primers was aided with MacVector software (New Haven, CT) version 4.1. Alignments and multiple alignments of sequences were performed with the facilities supplied with the GCG sequence analysis package (Wisconson, USA) version 7 which was accessed via UNIX. Nucleotide and amino acid databases (National Centre for Biotechnology Information) were accessed via the internet with Netscape version 2.

Chapter 3

Typing of microsatellites in AS/AGU, PVG, BN and F344 rats

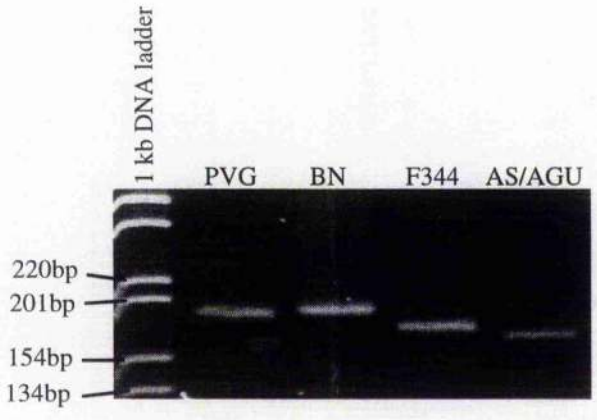
Introduction

Previously, 75 microsatellites had been typed in the rat strains AS/AGU, PVG, BN and F344 (Shiels et al, 1995). The reference strains chosen to set-up backcrosses for the mapping of the *agu* locus were BN and F344 as they differed from AS/AGU at 62% and 47%, respectively. It was therefore necessary to analyse more microsatellites in order to increase the number of anchor points throughout the rat genome to which *agu* could be linked.

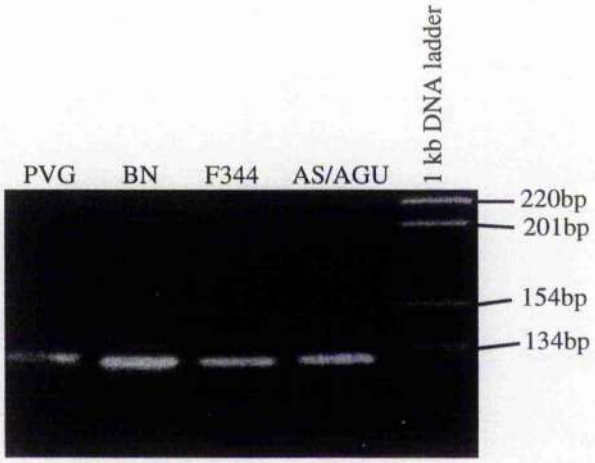
Typing of microsatellites

Microsatellite loci were analysed for strain variation by the polymerase chain reaction (PCR) using commercially available primers. Optimisation of conditions was necessary to yield single products to avoid any confusion should any microsatellites be used for genotyping. This was primarily done by altering the annealing temperature of the reaction. Optimisation of PCR experiments did not require alteration of Mg^{++} concentration as had been required for other microsatellite loci (Shiels et al, 1995). Microsatellites from 19 loci were typed in the rat strains AS/AGU, PVG, BN and F344. PCR products were electrophoresed using high resolution MetaPhor agarose which is capable of detecting size variations as low as 2bp. Figure 3.1 shows examples of highly polymorphic and non-polymorphic microsatellites.

Despite detecting differences as low as 2bp, MetaPhor agarose appeared to have resolution limitations in terms of clearly separating two products on the same gel lane. Figure 3.2 shows the results of typing the microsatellite locus *D14Mgh2* on the four rat strains. Although it was evident that there was a size difference between AS/AGU and BN, this system failed to resolve the two bands amplified from an F1 DNA template of these two strains. An alternative buffer with higher resolution properties was used in conjunction with MetaPhor agarose but this also failed to resolve the size difference in the F1. Polyacrylamide gel analysis of the PCR products did resolve this difference clearly in the F1 sample. Although the detection system used employed the end labelling of one primer with [γ^{32}P]ATP, shadow bands could still be seen in each lane. This was probably due to the amplification of non-target regions in the genome (Litt et al, 1993). Despite the presence of these shadow bands, resolution was unaffected and this microsatellite locus could therefore be used for genotyping progeny from the backcross, [AS/AGUxF1(AS/AGUxBN)].



a



b

Fig. 3.1 Electropherograms of gels (4% MetaPhor, 1xTBE) showing the PCR products from four rat DNA templates PVG, BN, F344 and AS/AGU using primer sets for the loci, *D14Mit1* (a) and *LSN* (b). *D14Mit1* is highly polymorphic for these four rat strains while *LSN* shows no size differences.

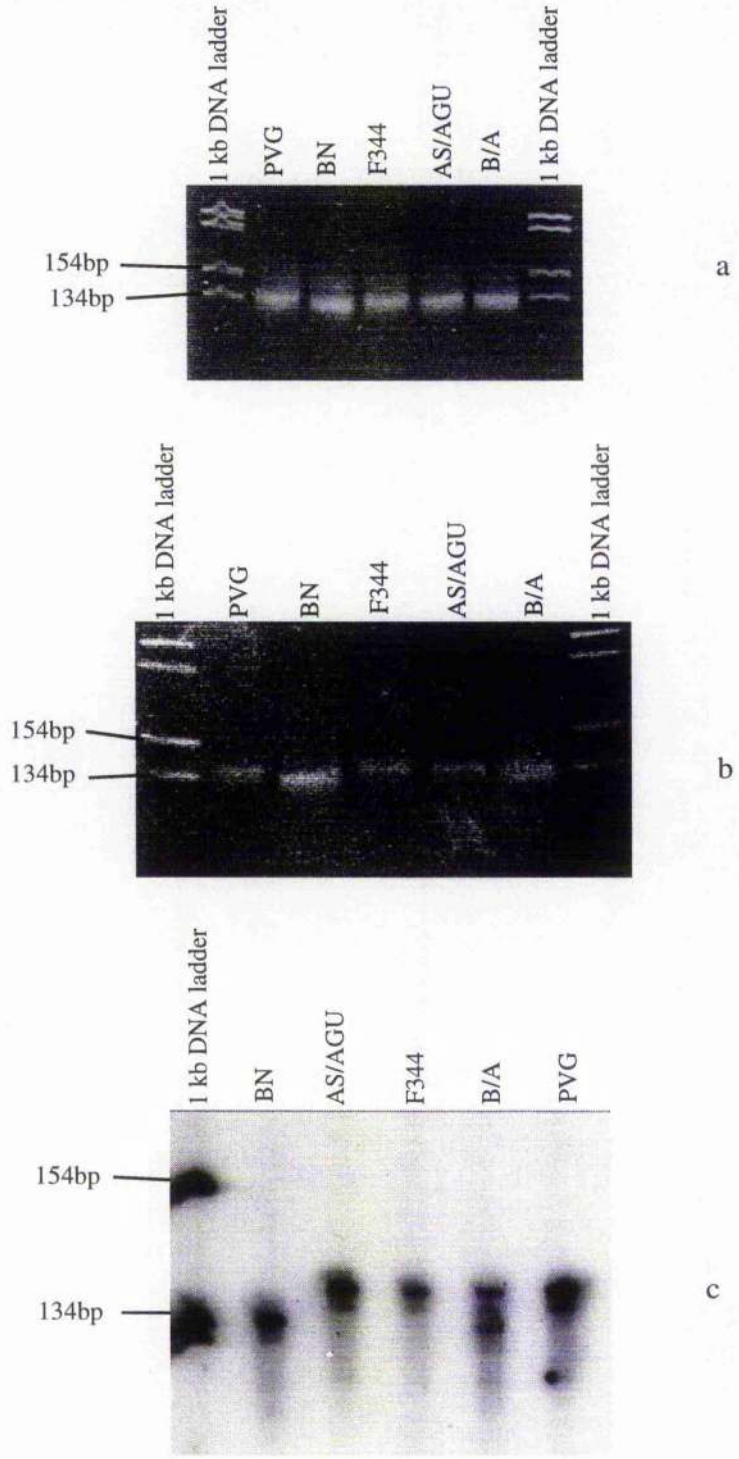


Fig. 3.2 Electropherograms of gels showing the PCR products from four rat DNA templates PVG, BN, F344, AS/AGU and an F1 DNA template, BN/ASAGU (B/A) using the primer set for the locus, *D14Mgh2*. a, Electrophoresis on 5% MetaPhor using 1xTBE indicates a size difference between BN and AS/AGU. F1 DNA between BN and AS/AGU is not resolved to two bands. b, Electrophoresis on 5% MetaPhor using 1xTAE again detects this size difference but does not resolve the F1 DNA. c, Electrophoresis on 8% polyacrylamide using 1xTBE resolves the size differece clearly in the F1 DNA template.

Table 3.1 shows the results of the typing experiments using the 19 microsatellite loci on the four rat strains. Of these microsatellites, nine were found to show size differences between AS/AGU, BN and/or F344. These could therefore be genotyped on the [AS/AGUxF1(AS/AGUxBN)] and [AS/AGUxF1(AS/AGUx F344)] backcrosses to help map the *agu* locus on the rat genome.

Locus	Primers	C'some	Anneal. Temp. (°C)	Approx expected allele Size (bp)	Approx. observed allele size (bp)	Rel. Size
D5Mit9	MR735	5	55	130	130	P>F=A>B
D3Mit3	MR991	3	57	155	155	P=F=A>B
A2UG	A2UG	5	57	230	230	P=B=F=A
D4Mit19	MR1209	4	56	140	140	P=F>A>B
D1Mgh22	MIGF-2	1	52	200	200	P>F=A>B
D7Mit12	PTBZRA	7	51	130	130	P>F=B>A
D10Mgh8	MHCG	10	54	180	180	A>P=F>B
D13Mit5	MR120	13	51	250	250	P=B=F=A
D12Mit5	MR1053	12	65	130	130	P=B=F=A
LSN	Leukos1	1	70	130	130	P=B=F=A
D2Mit20	PKLG	2	53	110	110	P>B=F=A
D5Mit2	MR980	5	53	200	200	P=B=F=A
D5Mit6	MR589	5	53	210	210	B=F=A>P
D5Mit11	MR1138	5	53	145	145	P=B=F=A

Locus	Primers	C'some	Anneal. Temp. (°C)	Approx expected allele Size (bp)	Approx. observed allele size (bp)	Rel. Size
D14Mit6	MR5	14	51	190	190	P>B=F=A
D14Mit1	MR1041	14	64	180	180	B>P>F>A
D14Mit8	MR1692	14	52	140	140	B>P=F=A
D14Mgh1	PGMUT	14	52	105	105	P=B=F=A
D14Mgh2	MR5333	14	54	130	130	B<P=F=A

Table 3.1 List of microsatellite markers used to determine relative sizes of alleles between the rat strains, BN (B), PVG (P), F344 (F) and AS/AGU (A). All primers were obtained from Research Genetics (Huntsville, Alabama) and designed by Jacob et al (1995) except *LSN* (designed by Serikawa et al. 1992).

Chapter 4

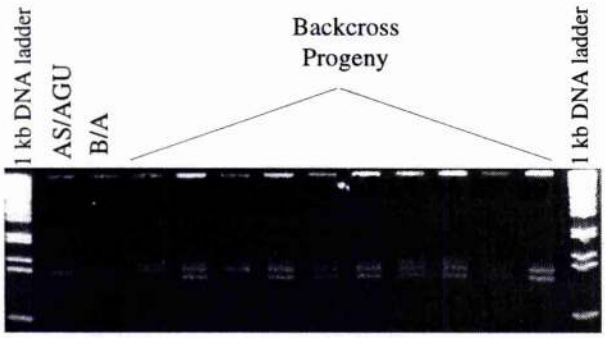
Is *agu* on Chromosome 14 ?

Introduction

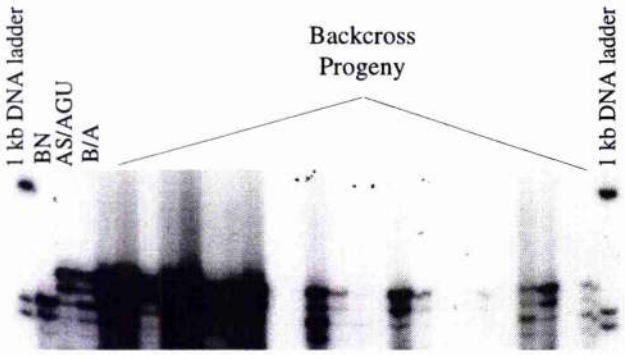
Progeny from the backcross [AS/AGUxF1(AS/AGUxBN)] were of three different colour types, albino, patched (hooded) and black (with white ventral patches). There appeared to be some association of the hooded colour type with the *agu* mutant phenotype (R. W. Davies, R. G. Sutcliffe and P. Shiels, personal communication). Since the locus controlling the hooded phenotype was known to be located on chromosome 14, it was desired to check for any possible linkage of molecular markers and the hooded locus itself to the *agu* phenotype. To help accomplish this, two genetic maps were constructed, the first consisting of five molecular markers based on 75 samples and the second was developed with the addition of the hooded locus based on 39 samples. Chi-square analyses were performed between the genotypes of these six markers and the phenotypes (mutant or normal) of the animals to determine any possible of linkage of *agu* to chromosome 14. Construction of the two linkage maps would help to confirm or reject linkage of *agu* to this chromosome should any of these markers give significant chi-square values.

Genetic linkage of five molecular markers on rat Chromosome 14

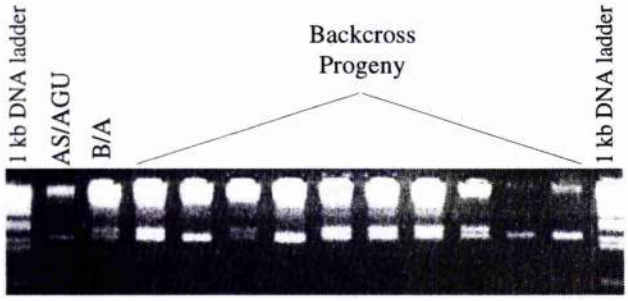
Five molecular markers, *CSNA*, *IGFBP*, *D14Mgh2*, *D14Mit1* and *D14Mit8*, were chosen for this study as they were previously shown to be located on and span the entirety of chromosome 14 (Jacob et al, 1995). The genotypes of these were analysed in 75 backcross progeny of the rat cross [AS/AGUxF1(AS/AGUxBN)]. The markers *CSNA* and *IGFBP* were originally mapped along with three others to chromosome 14 (Serikawa et al, 1992). On a separate map, the markers, *CSNA*, *D14Mgh2*, *D14Mit1* and *D14Mit8* were placed on the same chromosome with thirteen other new ones (Jacob et al, 1995). Genotype analysis was performed by PCR and the products obtained were resolved by gel electrophoresis. For accurate genotyping, adequate resolution was obtained by electrophoresis on 4% MetaPhor agarose gels except in one case, *D14Mgh2*. Despite a size difference detected between the two parental PCR products on a 4% MetaPhor gel using these primers, resolution using even 5% MetaPhor gels was not adequate to categorise a backcross PCR product as being either homozygous (one band) or heterozygous (two bands). Therefore it was necessary to use polyacrylamide gel electrophoresis in this case. Figure 4.1 shows electropherograms of PCR experiments of each of the five primer sets on parental and some backcross progeny samples.



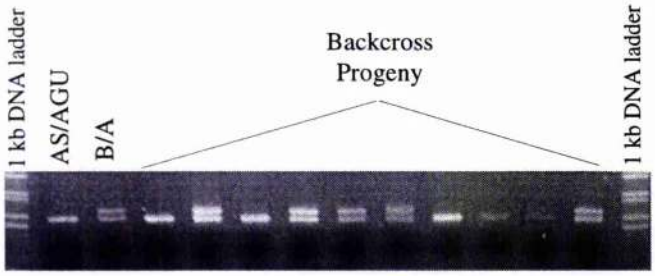
CSNA



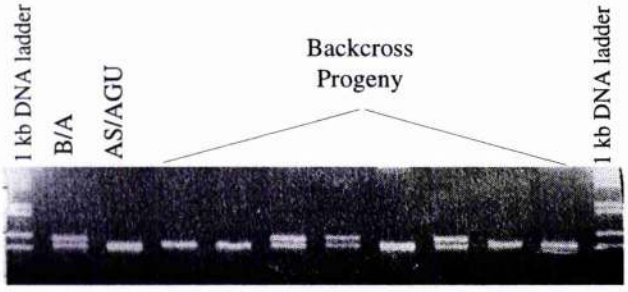
D14Mgh2



D14Mit1

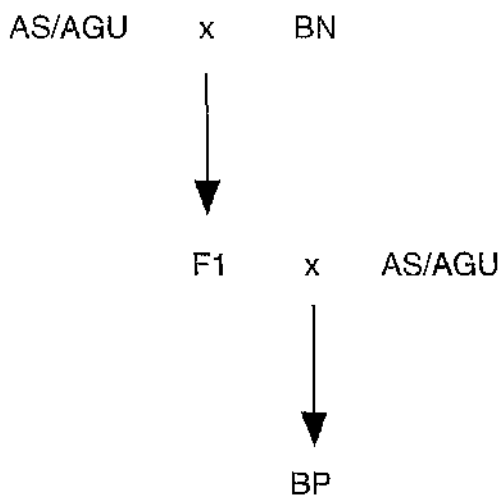


IGFBP



D14Mit8

a



Animal	Locus				
	D14Mit1	*CSNA*	*D14Mit8*	*D14Mgh2*	*IGFBP*
AS/AGU	160	150	140	136	155
BN	180	130	155	134	170
F1	160/180	150/130	140/155	136/134	155/170
BP	160/160 or 160/180	150/150 or 150/130	140/140 or 140/155	136/136 or 136/134	155/155 or 155/170

b

Fig. 4.1 a (opposite), Electropherograms of gels (1xTBE) showing the results of PCR based genotyping on BN and AS/AGU parental strains, an F1 between AS/AGU and BN (B/A) and some backcross progeny from the cross [AS/AGU xF1(AS/AGUxBN)] using the following sets of primers, *CSNA*, *D14Mit1*, *IGFBP*, *D14Mit8* and *D14Mgh2*. Products were electrophoresed on 4% MetaPhor gels except those generated using the primers *D14Mgh2* which were electrophoresed on 8% polyacrylamide (7M Urea) for increased resolution purposes. b (above), Diagram depicting breeding program of backcross and table showing allele sizes (in base pairs) for each animal at each locus. BP = Backcross progeny.

Table 4.1 shows, for the 75 backcross DNA samples, the genotypes of the five molecular markers which are described as being either homozygous for the AS/AGU allele (a/a) or heterozygous for the AS/AGU and BN alleles (a/b). For each of the primers used on the 75 backcross DNA samples, the numbers of homozygous and heterozygous PCR products obtained did not differ significantly from the expected 50:50 ratio. Genetic distances between pairs of markers is shown below in table 4.2.

	D14Mit1	*CSNA*	*D14Mit8*	*D14Mgh2*	*IGFBP*
D14Mit1	x	13.33 ± 3.9	39 ± 5.6	37.3 ± 5.6	53.3 ± 5.8
CSNA	x	x	30.66 ± 5.3	37 ± 5.6	50.7 ± 5.8
D14Mit8	x	x	x	28 ± 5.2	29.3 ± 5.3
D14Mgh2	x	x	x	x	20 ± 4.6
IGFBP	x	x	x	x	x

Table 4.2 Genetic distances between pairs of markers on chromosome 14. The values shown are expressed as the percentage of recombination. The standard error for each value is shown and was calculated using the formula,

$$SE^2 = \frac{f \times (1-f)}{n}$$

Where f = the recombination frequency and n = the number of samples.

From this data, two points of marker order rivalry emerged. In the first case, the two markers *D14Mit8* and *D14Mgh2*, which were 28cM from each other, were seen to be 39cM and 37.3cM from *D14Mit1* respectively. In the second case, the two markers *D14Mgh2* and *IGFBP*, which were 29cM from each other, were seen to be 28cM and 20cM from *D14Mit8* respectively. These discrepancies in marker order became apparent because double recombinants between markers could not be taken into account.

Analysis of the data by grouping the markers in sets of three was undertaken to reveal the consequent double recombinants and hence establish a marker order. The sets of three chosen in order to remedy marker order rivalry were: *D14Mit1*, *D14Mit8* and *D14Mgh2*; *D14Mit1*, *CSNA* and *D14Mit8*; *CSNA*, *D14Mit8* and *D14Mgh2*; *D14Mit8*, *D14Mgh2* and *IGFBP*..The least number of double recombinants within a set of three indicated the correct order for the number of samples used. From this, the genetic distances were calculated between sets of three markers believed to be adjacent to one another. The calculations of these were as follows,

Animal	*D14Mit1*	*CSNA*	*D14Mit8*	*D14 Mgh2*	*IGFBP*	Phenotype
4	a/b	a/b	a/b	a/b	a/b	normal
5	a/b	a/b	a/b	a/b	a/b	mutant
6	a/a	a/a	a/b	a/a	a/b	mutant
7	a/b	a/b	a/a	a/b	a/b	mutant
9	a/a	a/a	a/b	a/b	a/b	normal
10	a/a	a/b	a/b	a/b	a/b	normal
11	a/b	a/b	a/b	a/b	a/b	normal
12	a/a	a/a	a/a	a/a	a/a	normal
13	a/b	a/b	a/b	a/b	a/b	normal
14	a/b	a/a	a/a	a/b	a/b	mutant
15	a/b	a/b	a/b	a/a	a/a	mutant
18	a/b	a/b	a/b	a/b	a/b	normal
19	a/b	a/b	a/b	a/b	a/a	normal
20	a/a	a/a	a/a	a/a	a/a	normal
21	a/b	a/b	a/b	a/b	a/b	normal
22	a/a	a/b	a/b	a/b	a/a	normal
23	a/b	a/b	a/b	a/b	a/b	mutant
24	a/b	a/b	a/b	a/b	a/b	mutant
25	a/a	a/a	a/b	a/b	a/b	mutant
26	a/b	a/b	a/a	a/a	a/a	mutant
29	a/b	a/b	a/b	a/b	a/b	mutant
34	a/b	a/b	a/b	a/b	a/a	normal
36	a/b	a/b	a/a	a/a	a/a	normal
37	a/b	a/b	a/b	a/b	a/b	mutant
38	a/a	a/a	a/a	a/a	a/b	mutant
39	a/b	a/b	a/b	a/b	a/a	mutant
40	a/b	a/b	a/b	a/a	a/a	normal
41	a/b	a/b	a/a	a/b	a/a	mutant
42	a/a	a/a	a/a	a/a	a/a	mutant
43	a/a	a/a	a/a	a/a	a/a	mutant
44	a/a	a/a	a/a	a/a	a/a	normal
45	a/a	a/a	a/a	a/a	a/b	mutant
46	a/a	a/a	a/a	a/a	a/a	mutant
47	a/a	a/a	a/b	a/a	a/a	normal
48	a/b	a/b	a/a	a/b	a/a	mutant
49	a/a	a/b	a/b	a/a	a/b	normal
50	a/b	a/b	a/a	a/b	a/a	normal
51	a/b	a/b	a/a	a/a	a/a	mutant
52	a/b	a/b	a/a	a/a	a/a	normal
53	a/b	a/b	a/b	a/b	a/b	mutant
54	a/a	a/a	a/a	a/a	a/a	normal
56	a/a	a/a	a/b	a/b	a/b	normal
58	a/b	a/b	a/b	a/b	a/b	normal
60	a/a	a/a	a/a	a/a	a/a	mutant
63	a/a	a/a	a/a	a/a	a/a	normal
64	a/a	a/a	a/a	a/a	a/a	normal
65	a/b	a/b	a/b	a/b	a/b	normal
66	a/b	a/b	a/b	a/b	a/a	normal

Animal	*D14Mit1*	*CSNA*	*D14Mit8*	*D14 Mgh2*	*IGFBP*	Phenotype
67	a/a	a/a	a/a	a/a	a/a	mutant
68	a/b	a/b	a/b	a/a	a/a	normal
69	a/b	a/b	a/a	a/a	a/a	normal
70	a/a	a/a	a/a	a/a	a/a	normal
71	a/b	a/b	a/b	a/a	a/a	normal
72	a/b	a/b	a/a	a/a	a/a	normal
73	a/b	a/b	a/a	a/b	a/a	normal
74	a/a	a/b	a/b	a/b	a/b	normal
75	a/b	a/b	a/b	a/b	a/b	mutant
76	a/a	a/a	a/a	a/a	a/a	normal
77	a/a	a/a	a/a	a/b	a/b	mutant
78	a/b	a/a	a/b	a/b	a/b	normal
79	a/a	a/a	a/b	a/b	a/b	normal
80	a/b	a/b	a/b	a/a	a/a	normal
81	a/a	a/a	a/b	a/b	a/b	normal
82	a/a	a/a	a/a	a/b	a/b	normal
83	a/b	a/a	a/a	a/a	a/a	mutant
84	a/b	a/a	a/b	a/a	a/a	normal
85	a/a	a/b	a/b	a/a	a/a	mutant
86	a/a	a/a	a/b	a/b	a/b	normal
87	a/b	a/b	a/b	a/a	a/a	mutant
88	a/b	a/b	a/b	a/b	a/a	mutant
90	a/b	a/b	a/b	a/a	a/a	mutant
91	a/a	a/b	a/b	a/a	a/b	mutant
92	a/a	a/a	a/b	a/b	a/b	normal
93	a/a	a/a	a/b	a/b	a/b	mutant
95	a/b	a/b	a/b	a/b	a/b	mutant

Table 4.1 Genotypes of five molecular markers on 75 backcross progeny of the cross, [AS/AGUxF1(AS/AGUxBN)]. Genotypes are expressed as being either homozygous for the AS/AGU allele (a/a) or heterozygous for the AS/AGU and BN alleles (a/b).

1:

	D14Mit1 a/a		*D14Mit1* a/b		
	D14Mit8 a/a	*D14Mit8* a/b	*D14Mit8* a/a	*D14Mit8* a/b	
D14Mgh2 a/a	15	5	7	9	36
D14Mgh2 a/b	2	11	5	21	39
Totals	17	16	12	30	75

Possible Marker Order	Consequent double recombinants
(I) *D14Mit1-D14Mit8-D14Mgh2*	5+5 =10
(ii) *D14Mit8-D14Mgh2-D14Mit1*	2+9 =11
(iii) *D14Mgh2-D14Mit1-D14Mit8*	7+11 =18

2:

	D14Mit1 a/a		*D14Mit1* a/b		
	CSNA a/a	*CSNA* a/b	*CSNA* a/a	*CSNA* a/b	
D14Mit8 a/a	17	0	2	11	30
D14Mit8 a/b	10	6	2	27	45
Totals	27	6	4	38	75

Possible marker orders	Consequent double recombinants
(i) *D14Mit1-CSNA-D14Mit8*	0+2=2
(ii) *CSNA-D14Mit8-D14Mit1*	10+11=21
(iii) *D14Mit8-D14Mit1-CSNA*	2+6=8

Distances

D14Mit1 to *CSNA*,	10/75 =13.333 cM ± 3.9
CSNA to *D14Mit8*,	23/75 =30.666 cM ± 5.3

3:

	CSNA a/a		*CSNA* a/b		
	D14Mit8 a/a	*D14Mit8* a/b	*D14Mit8* a/a	*D14Mit8* a/b	
D14Mgh2 a/a	16	3	6	5	30
D14Mgh2 a/b	3	9	10	23	45
Totals	19	12	16	28	75

Possible marker orders	Consequent double recombinants
(i) *CSNA-D14Mit8*-D14Mgh2	3+5=8
(ii) *D14Mit8-D14Mgh2-CSNA*	3+10=13
(iii) *D14Mgh2-CSNA-D14Mit8*	6+9=15

Distances	
CSNA to *D14Mit8*,	23/75 =30.666˙ cM ± 5.3
D14Mit8 to *D14Mgh2*,	21/75 =28 cM ± 5.2

4:

	D14Mit8 a/a		*D14Mit8* a/b		
	D14Mgh2 a/a	*D14Mgh2* a/b	*D14Mgh2* a/a	*D14Mgh2* a/b	
IGFBP a/a	20	4	10	6	40
IGFBP a/b	2	4	3	26	35
Totals	22	8	13	32	75

Possible marker orders	Consequent double recombinants
D14Mit8-D14Mgh2-IGFBP	4+3=7
D14Mgh2-IGFBP-D14Mit8	2+6=8
IGFBP-D14Mit8-D14Mgh2	10+4=14

Distances	
D14Mit8 to *D14Mgh2*,	21/75=28cM ± 5.2
D14Mgh2 to *IGFBP*,	15/75 =20cM ± 4.6

The resultant marker orders and distances obtained from these calculations were used to generatea linkage map of all these molecular markers and is shown in fig. 4.2.

Determination of unknown pigmentation genes present in the strain AS/AGU

A percentage of progeny from the backcross [F1(AS/AGU x BN) x AS/AGU] were of different coat colour and pattern from their parents and grandparents. The strain AS/AGU, like most rat strains, is albino (c/c), i.e. skin, hairs and eyes lack the pigments eumelanin (black) and pheomelanin (yellow). Despite the pigment

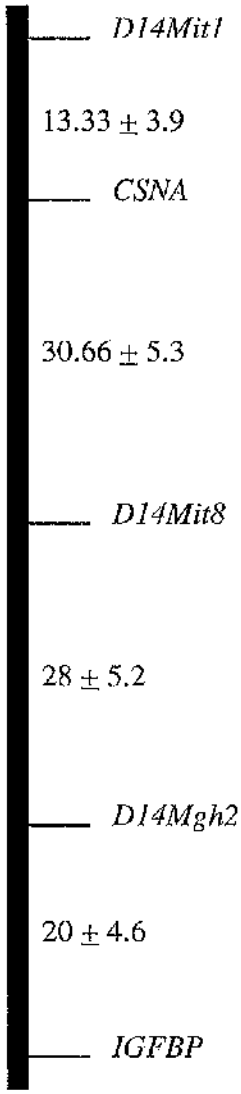


Fig. 4.2 Genetic linkage map of rat chromosome 14 using five molecular markers on 75 backcross progeny from the cross, [AS/AGUxF1(AS/AGUxBN)]. Distances are expressed in centiMorgans and standard errors are shown.

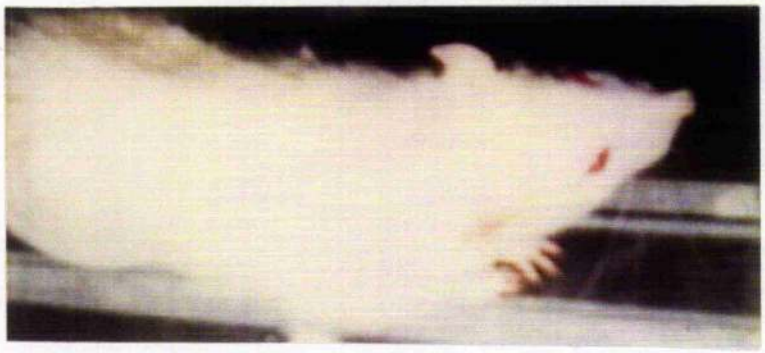
genotype, albinos lack any pigmentation due to the fact that albinism is epistatic to all other coat colour determinants. The reference strain that was used BN, was nonagouti brown (b/b) which is recessive to black (B/B) (King and Castle. 1935). About half of these rats however had an additional small white mark on their chest. Figure 4.3 shows photographs of an AS/AGU and two BN rats. All F1 progeny were primarily brown all over except for a large white patch extending from the chest to the pubic region on their ventral side. Some of the F1 rats however were a darker shade of brown, almost approaching black, than others. Figure 4.4 shows photographs of two F1 rats. Of the 313 backcross offspring produced, 134 were albino, 100 were black and had a white patch on their underside and 79 were white with a black hood present on the head, the shoulders and the mid dorsal strip. Figure 4.5 shows photographs of the three forms of backcross rats.

The ratio of black:hooded:albino rats obtained in the backcross progeny, differed significantly from the expected ratio of 1:1:2 respectively. A chi-square value of 9.61 was obtained which was greater than the critical value allowed in order to assume that the numbers obtained differed from the expected numbers due to chance alone. It may be concluded that these 313 backcross animals were not a good representative sample for analysis of coat colour genetics of the rat cross [AS/AGUxF1(AS/AGUxBN)].

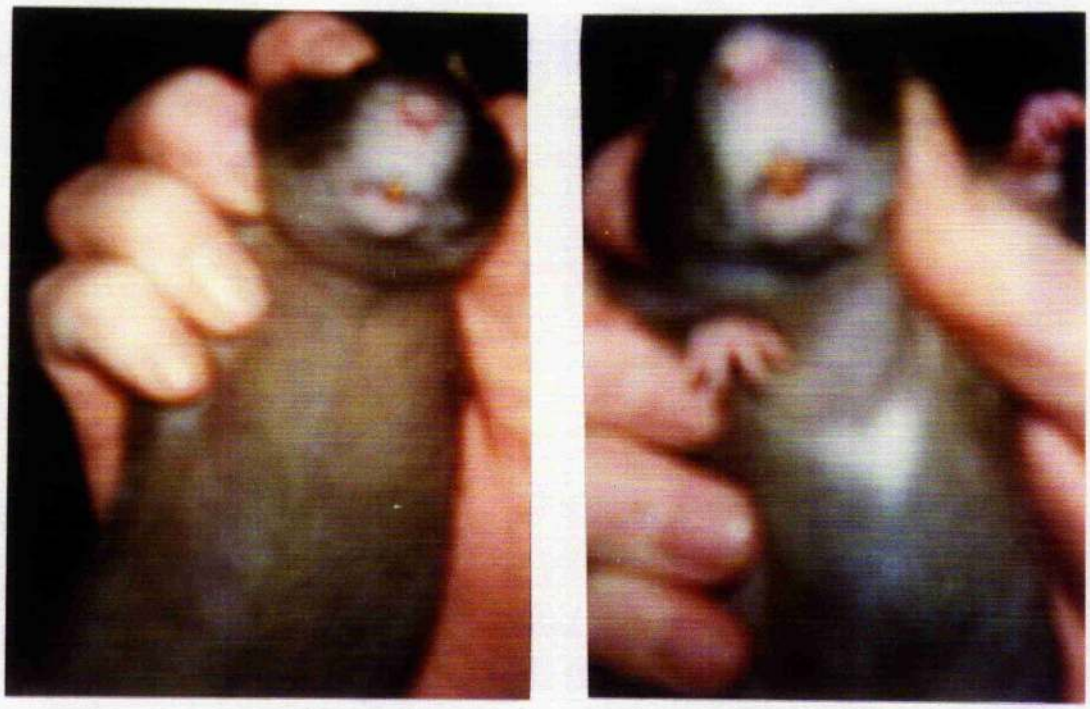
The hooded phenotype which arose in the backcross was due to the action of the hooded locus which is present on chromosome 14. This locus is responsible for pattern varieties in the rat by delaying migration of melanoblasts during skin differentiation. Five alleles of the hooded locus have been recorded to date, hooded (h), self (H), Irish (h^i), restricted (H^{re}) and notch (h^n) (Castle, 1951, Gumbreck et al , 1972).

The restricted allele (H^{re}) is dominant over all other alleles of the hooded series. Heterozygotes with h, h^i and H cause extensive restriction in the normal patterns of these alleles. H^{re}/h males become sterile after three months of age due to their testes being devoid of spermatagonial cells, spermatocytes, spermatids and spermatazoa. Female heterozygotes are normal. H^{re} homozygotes however die around birth due to an arrest of metabolism and a cessation of haematapoesis.

The hooded allele (h) constitutes pigment limited to the head, the shoulders and the mid dorsal strip. The width of the back stripe varies in a quantitative way from being extremely broad to being narrow and interrupted. Castle and Philips (1914) introduced a graded scale for hooded rats where variation in the plus direction



a



(Self) b (Irish)

Fig 4.3 a, An AS/AGU rat. b, Two BN rats, one displaying the Self phenotype (*H/H*) and the other displaying the Irish (h^i /h^i) phenotype. Of the BN rats used for setting up crosses, half were Self and half were Irish.

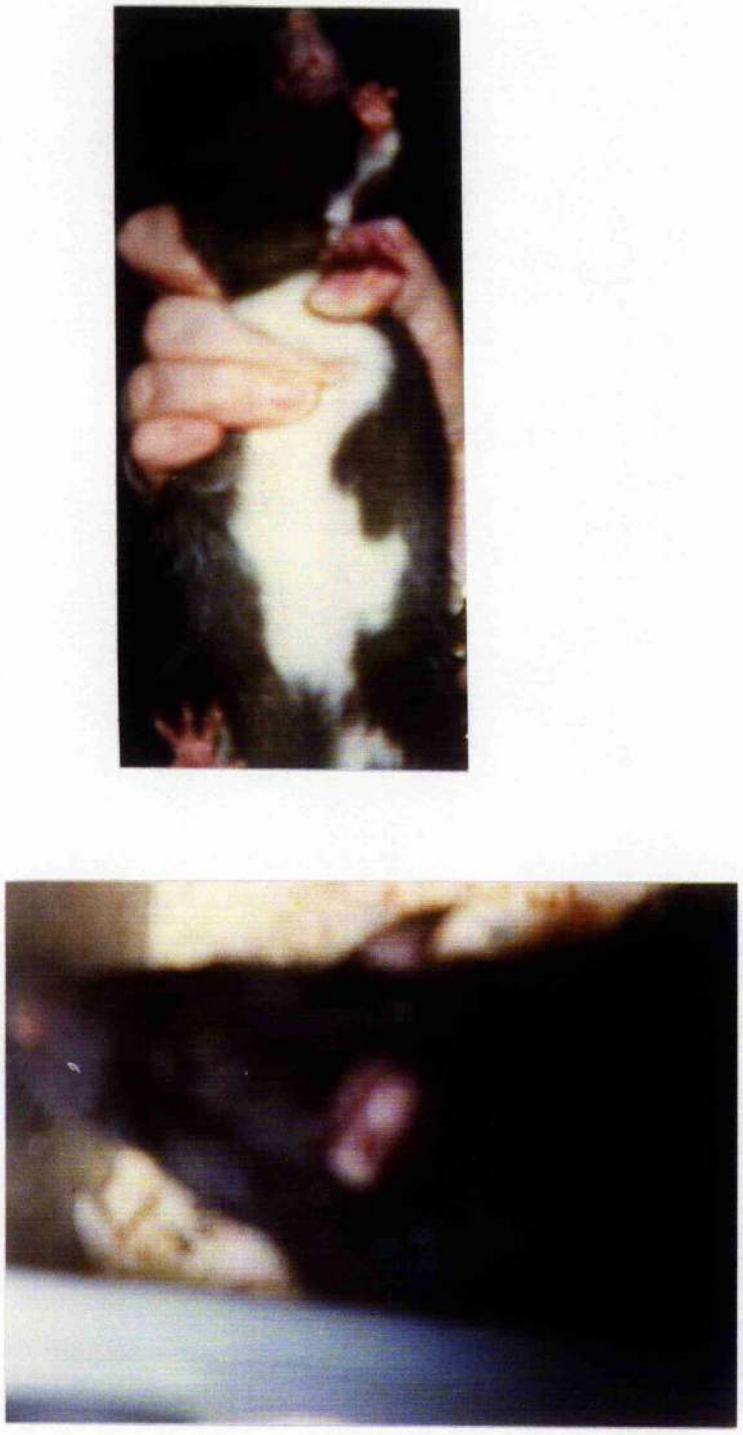


Fig. 4.4 Two F1 (AS/AGUxBN) rats. All F1 rats displayed the same pattern phenotype, brown all over with a large ventral white patch extending from the chest to the pubic region. The pattern genotype of these animals was either *H/h* or h^i/h, depending on whether the BN parent was Self or Irish (see fig. 4.3). The second F1 rat shown exemplifies that some F1 rats were a darker shade of brown, almost approaching black.

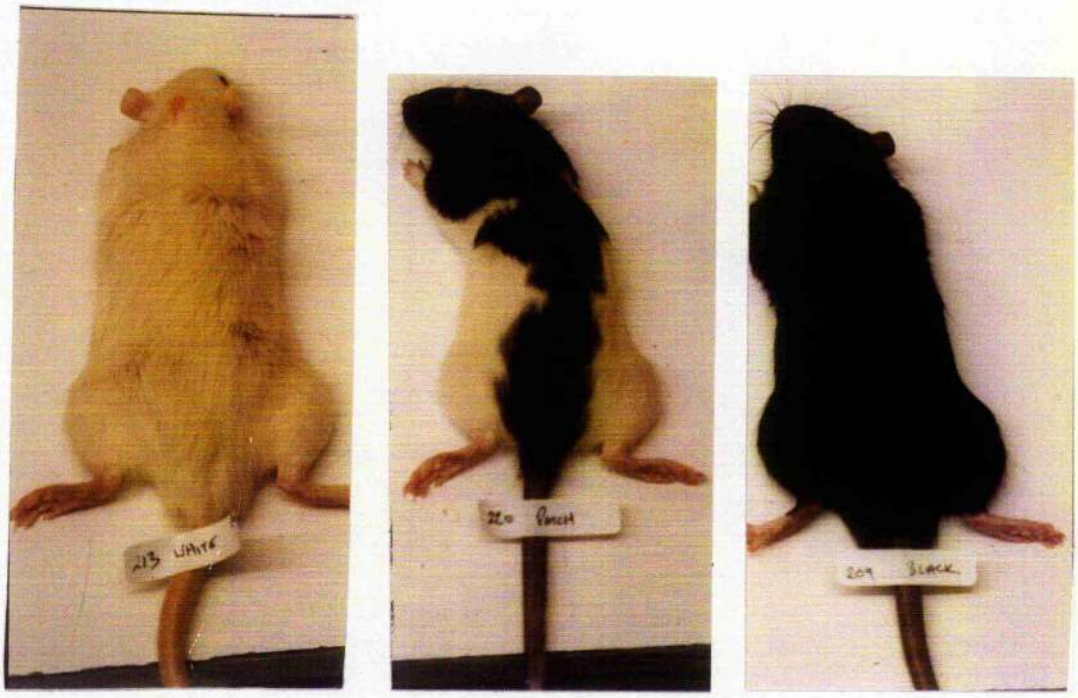


Fig. 4.5 Backcross progeny from the cross [AS/AGUxF1(AS/AGUxBN)]. The albino rat has inherited two copies of the *c* allele and therefore displays neither colour nor pattern. The hooded rat has inherited two copies of the *h* allele from the AS/AGU strain. The black rat has inherited a *H* or h^i allele from the BN strain, depending on the BN grandparent used (see fig. 4.3) and a *h* allele from the AS/AGU strain. All black rats produced from the backcross had a white ventral patch similar to that seen in the F1 rats (see fig. 4.4). The table below shows the numbers of albino, hooded and black rats produced and whether they were normal or if they were homozygous for the *agu* mutation.

General Total	Normal	*agu/agu*
134 albino	61 albino	73 albino
79 hooded	50 hooded	29 hooded
100 black	58 black	42 black
313	169	144

broadens and lengthens the hood whereas variation in the minus direction narrows and shortens the hood (fig. 4.6). Variation of the hooded pattern is due to modifying factors, plus acting or minus acting, which become associated with the gene for hooded, leaving that gene itself unchanged. The hooded modifier locus (*Hm*), which is linked to *h* on chromosome 14, is responsible for the variation of the hooded pattern and has two alleles of it's own, long (Hm^l) and short (Hm^s), the former being dominant over the latter.

Notch (h^n) is recessive to all other alleles of the hooded series. The small amount of pigment present is limited to the head. Castle (1951) showed with extensive breeding experiments that this was a new allele of hooded as opposed to an accumulation of minus acting modifying factors acting on *h*.

Irish (h^i) rats are pigmented all over except for a usually quite small white patch on the chest.. As with the h^n allele, h^i is a distinct allele of the series as opposed to an accumulation of plus acting modifying factors acting on *h*.

Self (*H*) rats are pigmented all over and therefore lack any form of pattern. Heterozygotes with *h*, h^i and h^n have a large white patch on their ventral side extending from the chest to the pubic region which may be mistaken for h^i.

The albino rats that arose in the backcross of these studies inherited two copies of the *c* allele from the strain AS/AGU through the breeding programme. The hooded rats that arose in the backcross progeny inherited two copies of the *h* allele from the AS/AGU strain and inherited pigment (*C*) from the BN strain. Most of the hooded rats were of the grade +1.5 to +2 on the Castle and Philips scale. Only a few rats differed being of the grade -0.5 (not photographed). The colour of the hood of all of these rats was pure black which indicates that the AS/AGU strain has the genotype, *B/B*. If hooded backcross individuals inherited a *b* allele from AS/AGU, some individuals would have coinherited a *b* allele from the BN strain and therefore would have had a brown coloured hood which was not the case. The black rats that arose in the backcross progeny had again inherited at least one *B* allele from AS/AGU. As far as pattern is concerned, all the black rats had the same phenotype, i.e. a white patch on the underside. The genotypes of these rats were either h/h^i or *h/H* depending on whether the BN grandparent was either coloured all over (*H/H*) or whether it was a hooded Irish rat (h^i/h^i). As with these rats, the F1 progeny had the same pattern phenotype, coloured all over with a white patch on the underside, but differing genotypes. Again they were either h/h^i or *h/H*. The phenotypes of these heterozygotes agree with other previous breeding experiments

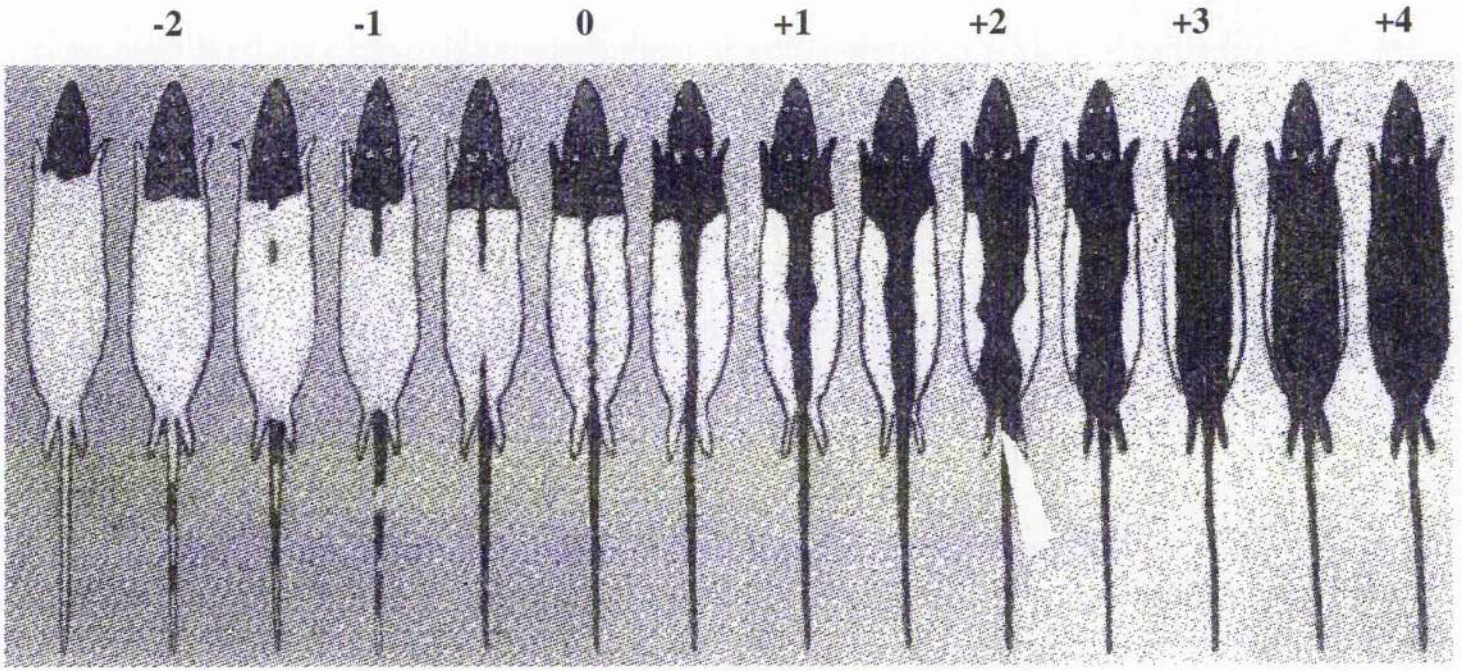


Fig. 4.6 Graded scale for hooded rats. Taken from Castle and Phillips (1914).

by Doncaster (1906) who showed that albino rats can carry colour determinants for brown, some for black and others for brown and black and also that the same is true for pattern, i.e. albinos can carry the full pattern (*H*), the piebald pattern (*h*) and the hooded Irish pattern (h^i) .

One observation concerning colour did conflict with the fact that brown is recessive to black. About half of the F1 rats were brown and the other half were darker, almost approaching black. The colour of the BN parent, brown, had the genotype *b/b*. The colour of all the backcross progeny with pigment was black indicating that the genotype of the AS/AGU strain was *B/B*. Therefore all the F1 progeny, *B/b*, should have been black. This apparent lack of complete dominance that arose in the F1 progeny was restored in the backcross progeny. The brown locus, present on chromosome 5, is not linked to any other known coat colour genes which are present on different chromosomes (Szpirer et al, 1996). One other point of note was that the pattern phenotypes of the BN parent individuals were not all the same (fig. 3.3). About half of them were self patterned and the other half had an additional small white patch on their chests. These phenotypes are indicative of the self (*H/H*) and (h^i/h^i) genotypes respectively. It has been noted before that some BN strains have been described as expressing h^i (Hedrich, 1990). It may be possible that the two types of BN rat used in these studies were substrains of BN.

What was to be gained from all these observations is that hooded individuals from the backcross could be seen as representing an AS/AGU homozygous (a/a) genotype and the black rats from the backcross as an AS/AGU and BN heterozygous genotype (a/b). Therefore it was possible to incorporate the hooded locus into the chromosome 14 map with the other five molecular markers by using the phenotypes of the coloured individuals.

Genetic linkage of five molecular markers and one coat colour marker on rat chromosome 14

A linkage map was constructed with the *H* locus by three point cross calculations using the derived genotypes from the coloured backcross individuals and the genotypes of the molecular markers from the same progeny. The genotypes of 39 animals were analysed with these six markers. Table 4.3 displays the genotypes of the markers on the 39 samples as being either homozygous for the AS/AGU allele (a/a) or heterozygous for the AS/AGU and BN alleles (a/b).. Calculations for the marker orders and distances were as follows,

Animal	*D14Mit1*	*CSNA*	Coat Colour	*D14Mit8*	*IGFBP*	*D14 Mgh2*
4	a/b	a/b	black	a/b	a/b	a/b
9	a/b	a/b	black	a/a	a/b	a/b
10	a/a	a/a	hooded	a/b	a/b	a/b
11	a/a	a/b	black	a/b	a/b	a/b
20	a/b	a/b	black	a/b	a/b	a/b
22	a/a	a/a	hooded	a/a	a/a	a/a
24	a/b	a/b	black	a/b	a/b	a/b
25	a/a	a/a	black	a/b	a/b	a/b
36	a/b	a/b	hooded	a/a	a/a	a/a
39	a/b	a/b	black	a/b	a/a	a/b
40	a/b	a/b	black	a/b	a/a	a/a
42	a/a	a/a	hooded	a/a	a/a	a/a
43	a/a	a/a	hooded	a/a	a/a	a/a
44	a/a	a/a	hooded	a/a	a/a	a/a
45	a/a	a/a	hooded	a/a	a/b	a/a
46	a/a	a/a	hooded	a/a	a/a	a/a
47	a/a	a/a	hooded	a/b	a/a	a/a
48	a/b	a/b	black	a/a	a/a	a/b
49	a/a	a/b	black	a/b	a/b	a/a
50	a/b	a/b	hooded	a/a	a/a	a/b
51	a/b	a/b	black	a/a	a/a	a/a
52	a/b	a/b	hooded	a/a	a/a	a/a
54	a/a	a/a	hooded	a/a	a/a	a/a
63	a/a	a/a	hooded	a/a	a/a	a/a
64	a/a	a/a	hooded	a/a	a/a	a/a
66	a/b	a/b	black	a/b	a/a	a/b
67	a/a	a/a	hooded	a/a	a/a	a/a
69	a/b	a/b	hooded	a/a	a/a	a/a
70	a/a	a/a	hooded	a/a	a/a	a/a
71	a/b	a/b	black	a/b	a/a	a/a
72	a/b	a/b	black	a/a	a/a	a/a
73	a/b	a/b	black	a/a	a/a	a/b
76	a/a	a/a	hooded	a/a	a/a	a/a
80	a/b	a/b	black	a/b	a/a	a/a
85	a/a	a/b	black	a/b	a/a	a/a
87	a/b	a/b	black	a/b	a/a	a/a
88	a/b	a/b	black	a/b	a/a	a/b
90	a/b	a/b	black	a/b	a/a	a/a
91	a/a	a/b	hooded	a/b	a/b	a/a

Table 4.3 Genotypes of six markers on 39 backcross progeny from the cross, [AS/AGUxF1(AS/AGUxBN)]. Genotypes are expressed as being either homozygous for the AS/AGU allele (a/a) or heterozygous for the AS/AGU and BN alleles.

1:

	D14Mit1 a/a		*D14Mit1* a/b		
	CSNA a/a	*CSNA* a/b	*CSNA* a/a	*CSNA* a/b	
Black (a/b)	1	3	0	16	20
Hooded (a/a)	14	1	0	4	19
Totals	15	4	0	20	39

Possible marker orders	Consequent double recombinants
(i) *D14Mit1-CSNA-H*	1+0=1
(ii) *CSNA-H-D14Mit1*	1+4=5
(iiii) *H-D14Mit1-CSNA*	0+3=4

Distances

D14Mit1 to *CSNA*,	4/39 =10.26cM ± 4.9
CSNA to *H*,	6/39 =15.4cM ± 5.7

2:

	CSNA a/a		*CSNA* a/b		
	Black	Hooded	Black	Hooded	
D14Mit8 a/a	0	12	5	4	21
D14Mit8 a/b	1	2	14	1	18
Totals	1	14	19	5	39

Possible marker orders	Consequent double recombinants
(i) *CSNA-H-D14Mit8*	0+1=1
(ii) *H-D14Mit8-CSNA*	2+5=7
(iii) *D14Mit8-CSNA-H*	1+4=5

Distances

CSNA to *H*,	6/39 =15.4cM ± 5.77
H to *D14Mit8*,	8/39 =20.5cM ± 6.5

3:

	Black		Hooded		
	D14Mit8 a/a	*D14Mit8* a/b	*D14Mit8* a/a	*D14Mit8* a/b	
IGFBP a/a	4	9	15	1	29
IGFBP a/b	1	6	1	2	10
Totals	5	15	16	3	39

Possible Marker Orders	Consequent double recombinants
(i) *H-D14Mit8-IGFBP*	1+1 =2
(ii) *D14Mit8-IGFBP-H*	1+9 =10
(iii) *IGFBP-H-D14Mit8*	4+2 =6

Distances

H to *D14Mit8*,	8/39 =20.5cM ± 6.5
D14Mit8 to *IGFBP*,	12/39 =30.8cM ± 7.4

4:

	D14Mit8 a/a		*D14Mit8* a/b		
	IGFBP a/a	*IGFBP* a/b	*IGFBP* a/a	*IGFBP* a/b	
D14Mgh2 a/a	16	1	7	2	26
D14Mgh2 a/b	3	1	3	6	13
Totals	19	2	10	8	39

Possible Marker Orders	Consequent double recombinants
(i) *D14Mit8-IGFBP-D14Mgh2*	1+3 =4
(ii) *IGFBP-D14Mgh2-D14Mit8*	3+2 =5
(iii) *D14Mgh2-D14Mit8-IGFBP*	7+1 =8

Distances

D14Mit8 to *IGFBP*,	12/39 =30.8cM ± 7.4
IGFBP to *D14Mgh2*,	9/39 =23.1cM ± 6.7

Figure 4.7 depicts a linkage map of chromosome 14 using these markers. By comparing this second map with the first (fig. 4.2), it could be seen that the marker of *D14Mit8-D14Mgh2-IGFBP* in the first map became *D14Mit8-IGFBP-D14Mgh2* in the second. This change of order was due to the number of samples used: in the first map more samples were used, revealing more double recombinants, giving a different order in this case. It is probable that the order of these three markers is more true in the first map due to the larger number of samples analysed, although analysis with more samples would be desirable to confirm this. The total genetic distance covered in these two maps also differed due to numbers of samples analysed, i.e. 92cM with 75 samples (fig. 4.2) and 100cM with 39 samples (fig. 4.7).

Comparison of the two chromosome 14 maps with previous studies

Genetic Length Coverage	Type of Cross	No. of markers	No. of samples	Reference
92cM	Backcross, Female F1	5	75	Fig. 4.2
100cM	Backcross, Female F1	6	39	Fig. 4.7
95cM	Combined linkage data up to 1993	19	-	Yamada et al, (1994)
43cM	F2 (BNxSHR)	14	46	Jacob et al, (1995)

Table 4.4 Comparison of the genetic length coverage of chromosome 14 obtained in figures 4.2 and 4.7 with previous studies.

As for the genetic length of chromosome 14 covered, the value obtained from the map in fig. 4.2, 92cM, was more reasonably close to that obtained by Yamada's map, 95cM (combined from previous studies) (Yamada et al, 1994), than that of Jacob's, 43cM (based on 14 markers) (Jacob et al, 1995). The difference in genetic length of this chromosome in Jacob's map and in these studies could be due to the different types of crosses that were set up. F2 (SHRxBN) progeny were used to construct a genetic map by Jacob and thus that map would represent be a sex average value. The two maps constructed here was based on the female genome

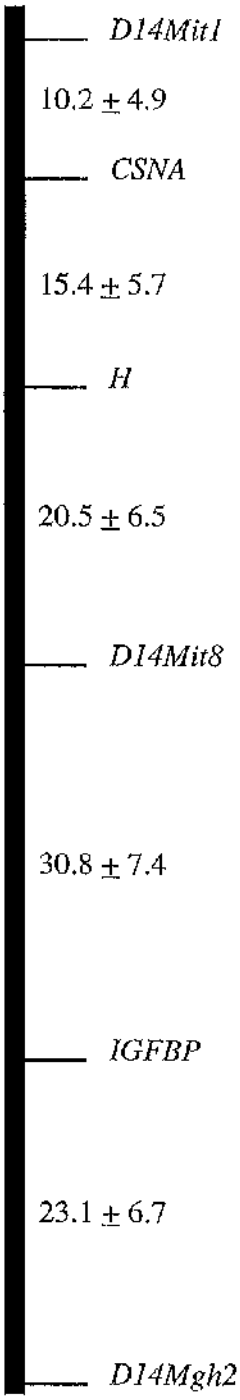


Fig. 4.7 Genetic linkage map of rat chromosome 14 using six markers on 39 backcross progeny from the cross, [AS/AGUxF1(AS/AGUxBN)]. Disatnces are expressed in centiMorgans and standard errors are shown.

since the backcross progeny were derived from female F1 parents. Since female genetic length is greater than male genetic length in mammals by a factor of about 1.8 (Jacob et al, 1995) the length of this chromosome in these studies would be greater than the sex average value obtained by Jacob. This value would equate to 55.3cM if it were based on female genetic length which would still be quite different from the 92cM value obtained here. Numbers of samples, or numbers of markers used here and by Jacob could not explain difference in length because the two values obtained here were 92 and 100 cM based on 75 and 39 samples, respectively and the value obtained by Jacob, 43cM was based on 46 samples. The absence of mapping functions used in calculating the genetic distances between markers on chromosome 14 in this study would also contribute to the observed differences in total distance covered from other maps. The maps generated here are indeed non-additive due to this fact. It was decided not to use mapping functions to attempt to generate a more true representation of true real map distance because such functions are less reliable when large recombination frequencies are involved.

The markers *IGFBP*, *CSNA* and *H* were mapped relative to each other on two separate linkage groups of the chromosome on the map of Yamada (Yamada et al, 1994). The markers *CSNA*, *D14Mit1*, *D14Mit8* and *D14Mgh2* were mapped relative to each other on the map of Jacob (Jacob et al, 1995). The construction of the maps in these studies has combined these markers onto the same map relative to each other. No comparisons of marker order could be made between the maps in these studies and in the Yamada map since it was divided into two linkage groups. The order of the four mentioned markers on the Jacob's map did agree with the maps here.

The *agu* mutation is not present on chromosome 14

Chi-square analyses were performed to determine possible linkage of any of the six markers mapped to chromosome 14 to the *agu* locus. Results of these calculations are shown the table 4.4. The chi-square values for *hooded* and *D14Mit8* were statistically significant to suggest that independent segregation of these markers from *agu* was not occurring. Calculating the genetic distances of *hooded* and *D14Mit8* from *agu* gave cM values of 51 (± 3.7) and 41 ((± 5.7) respectively. A value greater than 50cM for the hooded locus indicated no linkage. Despite the apparent linkage of *agu* to *D14Mit8*, with a recombination distance of ~41cM, the other markers of chromosome 14 had recombination values greater than 50%.

Marker	Number of Samples	Chi-square values	Level of significance
D14Mit1	75	3.02	0.5
CSNA	75	7.3	0.1
Hooded	179 *	10.25	0.05
D14Mit8	75	10.8	0.05
D14Mgh2	75	1.5	0.5
IGFBP	75	3.6	0.3

Table 4.4 Results of Chi-square analysis of six markers to the *agu* phenotype. *A total of 313 animals were produced from the backcross. 179 of these were coloured and were therefore used to determine possible linkage of *hooded* to *agu*.

From fig. 4.2 it could be seen that the markers immediately flanking *D14Mit8* (*CSNA* and *D14Mgh2*) were closer than 41cM. Therefore these markers should also have shown to be linked to *agu* . From the second map produced (fig 4.7), the markers immediately flanking *D14Mit8* (*hooded* and *IGFBP*) were also closer than 41cM. Therefore these markers should also have shown linkage to *agu*.

Final Discussion

Two linkage maps of rat chromosome 14 were produced. The first was based on 75 samples and encompassed five molecular markers. The second was constructed with the addition of a coat colour marker, hooded, based on 39 samples. Inclusion of the hooded locus into this map was made possible once the genotypes of the coloured backcross individuals were determined. Chi-square analyses indicated possible linkage of *agu* to *hooded* and *D14Mit8*. The recombination value between *agu* and *hooded* was greater than 50% which therefore ruled out any linkage of *agu* to *hooded*. Theoretically, a value greater than 50% is not possible butt the use of small sample numbers has probably generated sueh a spuriously high number. A recombination value of 41% between *D14Mit8* and *agu* further suggested linkage between these two. Lack of linkage of the markers immediately flanking *D14Mit8* suggested that agu could not be linked to this marker. Taking all these observations into consideration it could concluded that it was very unlikely that the *agu* locus was present on chromosome 14.

The next chapter describes linkage of *agu* to chromosome 1 and the search for informative markers associated with the carcinoembryonic gene family which is believed to be close to the mutant locus.

Chapter 5

Analysis of CEA associated microsatellites in the Rat

Introduction

The *agu* locus was initially linked to four markers on rat chromosome 1, *R33* (*kal*), *D1Mit1, D1Mgh7* and *R191* using the backcross programs, [AS/AGUxF1 (AS/AGUxBN)] and [AS/AGUxF1(AS/AGUxF344)].

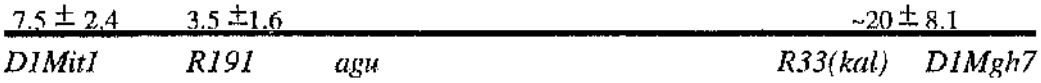


Fig. 5.1 A diagram showing the relationship of four markers on rat chromosome 1 to the *agu* locus based on the genotyping of 120 animals. The markers *D1Mit1*, *R33* and *D1Mgh7* were mapped using the cross [AS/AGUxF1(AS/AGUxBN)]. The marker *R191* was mapped using the cross [AS/AGUxF1(AS/AGUxF344)]. The numbers shown above the markers indicate the distances expressed in cM from *agu*.

Figure 5.1 shows a linkage map of the four markers linked to *agu*. The eventual physical mapping of the *agu* gene would require a marker starting point of between 0.1 and 0.5 cM (approximately 100 to 500 kb) from this locus. Due to the lack of useful starting points for this process it was necessary to examine other microsatellites to establish a more refined linkage map flanking *agu*. Within the linkage map shown in figure 4.1 all other gene-based and anonymous markers known to lie there (Jacob et al, 1995 Serikawa, et al, 1992) had been exhausted. Analysing syntenic relationships between rat and other species offered a method of revealing new genes and possibly new markers which could lie in the interval shown in the figure above.

As part of the human and mouse genome mapping projects, genetic and physical mapping studies have shown that mouse and human chromosomes resemble each other within syntenic blocks which can contain up to several hundred genes (Copeland et al, 1993). With the recent advancement of the rat gene map comparisons of gene locations to human and mouse have progressed. The central part of rat chromosome 1, including the markers *R191* and *R33*, has been shown by G-banding pattern similarity, chromosome morphological similarity and gene homologies to be syntenic to regions on mouse chromosome 7 and human chromosome 19 (Levan et al, 1991, Yamada et al, 1994). Analysis of the mouse genome database (MGD, Jackson Laboratory) revealed specifically which genes

from these syntenic regions are homologous between human, mouse and rat and are listed in table 5.1. The carcinoembryonic antigen gene family present in this syntenic block seemed a good candidate for microsatellite analysis. In humans this family contains many members (Olsen et al, 1994) and is described below. It was also shown that two members of a CEA family, CGM3 and CGM4, are located in this syntenic region in the rat (Serwikawa et al, 1992, Jacob et al, 1995). With the potential for of may of these gene members being present near the *agu* locus it was desired to map CEA genes to *agu* to aid in the genetic map and possibly the physical map of this region. Literature and sequence databases were searched to find any useful informative markers in rat CEA genes or in putative mouse or human homologues in the syntenic region

The Carcinoembryonic Anitgen family in human, mouse and rat

The Carcinoembryonic Antigen (CEA) family consists of a group of glycoproteins ranging in size from 20kd to 200kd. CEA was the first member to be described in colonic tumours as well as in some foetal tissues. Due to the relatively high concentration in sera of cancer patients, CEA is widely used as a marker in postoperative surveillance of such patients. It is gaining importance for immunolocalisation of tumours and metastases. Strong evidence shows that CEA appears to function as a cell adhesion molecule, being anchored to the cell membrane (Tynan *et al* ,1992). This gene family, due to the presence of IgV-like domains is a member of the immunoglobulin superfamily. Is divided into two subgroups on the basis of sequence similarity, as to whether the protein is membrane bound or secreted and location of expression. Proteins of the CEA subgroup contain hydrophobic stretches and are membrane bound. Nine members are present in this group to date, including CEA itself, NCA (non-specific cross-reacting antigen), BGP (biliary glycoprotein) and six genes referred to as CEA-Gene family Members (CGM 1,2,6,7,8 and 9). Members of the PSG (Pregnancy Specific Glycoprotein) subgroup are expressed almost uniquely in the placenta and all appear to be secreted. The PSG subgroup contains eleven members (Thompson and Zimmerman, 1992).

All CEA family members have a similar gene organisation. The first exon, denoted L, codes for the 5' untranslated region (UTR) and 21 amino acids of the signal peptide. Exon 2 (L/N) codes for the remaining 13 amino acids of the signal peptide and the first domain of the mature protein, the N-domain which is homologous to the immunoglobulin variable domain (IgV). Depending on the gene, this domain is

Species	Symbol	Chrom	cM Pos	Cyto-band
Human	*ATPA3*	19		q13.2
Mouse	*Atpa3*	7	4.0	
Rat	*Atpa3*	1		
Human	*ATP4A*	19		q13.1
Mouse	*Atp4a*	7	9.0	
Rat	*Atp4a*	1		
Human	*CALM3*	19		q13.3
Mouse	*Calm3*	7	4.5	
Rat	*Calm3*	1		
Human	*CEA*	19		q13.2
Mouse	*Cea*	7	2.5	
Rat	*Cea*	1		
Human	*CEBPA*	19		q13.1
Mouse	*Cebpa*	7	12.0	
Rat	*Cebpa*	1		
Human±	*CYP2B*	19		q13.2
Mouse	*Cyp2b9*	7	6.5	
Rat	*Cyp2b1*	1		
Human	*DBP*	19		q13
Mouse	*Dbp*	7	23.0	
Rat	*Dbp*	1		
Human	*GPI*	19		q13.1
Mouse	*Gpi1*	7	11.0	
Rat	*Gpi*	1		
Human	*GRIK5*	19		q13.2
Mouse	*Grik5*	7	6.5	
Rat	*Grik5*	1		
Human	*KLK1*	19		q13.3-13.4
Mouse	*Klk1*	7	23.0	
Rat	*Klk1*	1		
Human	*PEPD*	19		q13.1
Mouse	*Pep4*	7	15.0	
Rat	*Pepd*	1		
Human	*PRKCG*	19		q13.4
Mouse	*Pkcc*	7	2.0	
Rat	*Prkcg*	1		

Table 5.1 Locus homologies between human, mouse and rat emphasising the syntenic relationship between chromosomes 19, 7 and 1, respectively. Taken from the mouse genome database (MGD) at the Jackson Laboratory.

followed by different numbers of immunoglobulin constant -like domains, ranging from zero (e.g. in CGM-7) to six in CEA. The PSG subgroup have a short hydrophilic C-domain whilst the CEA subgroup have a more hydrophobic one, indicating membrane anchorage (Thompson and Zimmerman, 1992).

Members of the PSG subfamily are synthesised in large amounts in the placenta during pregnancy. They appear to function as immunosurveillance molecules during this time (Teglund *et al* , 1994). They can be detected after 2-3 weeks of gestation with serum concentrations 200-400 mg/ml. Low PSG serum concentrations have been associated with pregnancy related complications and antisera against them induced abortion in pregnant monkeys. Most of them contain an Arg-Gly-Asp (RGD) motif, which has been shown to function as adhesion recognition signals for several integrins (Teglund *et al* , 1994).

Gene organisation within the CEA subgroup was achieved by pulsed-field gel mapping and cosmid contig assembly and analysis by hybridisation of various CEA family cDNA clones to a human chromosome 19 cosmid library. Differential hybridisation using 5' and 3' gene probes together with fine restriction mapping showed the orientation of individual genes (Tynan *et al* , 1992). Analysis of gene organisation within the PSG subgroup was more difficult to ascertain due to the higher degree of homology between members of the PSG subgroup than members of the CEA subgroup. The use of several complementary approaches, including restriction fragment fingerprinting, specific probe hybridisation, complete and partial digest restriction mapping and FISH has generated a complete cosmid map spanning the 700kb PSG region (Olsen *et al* ,1994). Also in this work seven novel genes were found interspersed within the PSG region, (CGM12-18). CGM12 is part of the CEA subgroup and appears to be a psuedogene. CGM13-18 appear to form a new subgroup showing 94-99% identity to each other but only 70-80% to other members of either the CEA or PSG subgroups. They lack an exon for an IgV-like N-terminal domain.

Twenty nine genes have now been identified in the CEA family, 12 in the CEA subgroup, 11 in the PSG subgroup and 6 in the new subgroup (CGM13-18). The entire gene family was shown to be contained within a region located at position 19q13.1-13.2 between the CYP2A and the D19S15/D19S8 markers by high resolution FISH (fluorescence *in situ* hybridisation) (Thompson and Zimmermann, 1992). The gene organisation of the family spanning 1.3Mb is shown in fig. 5.2.

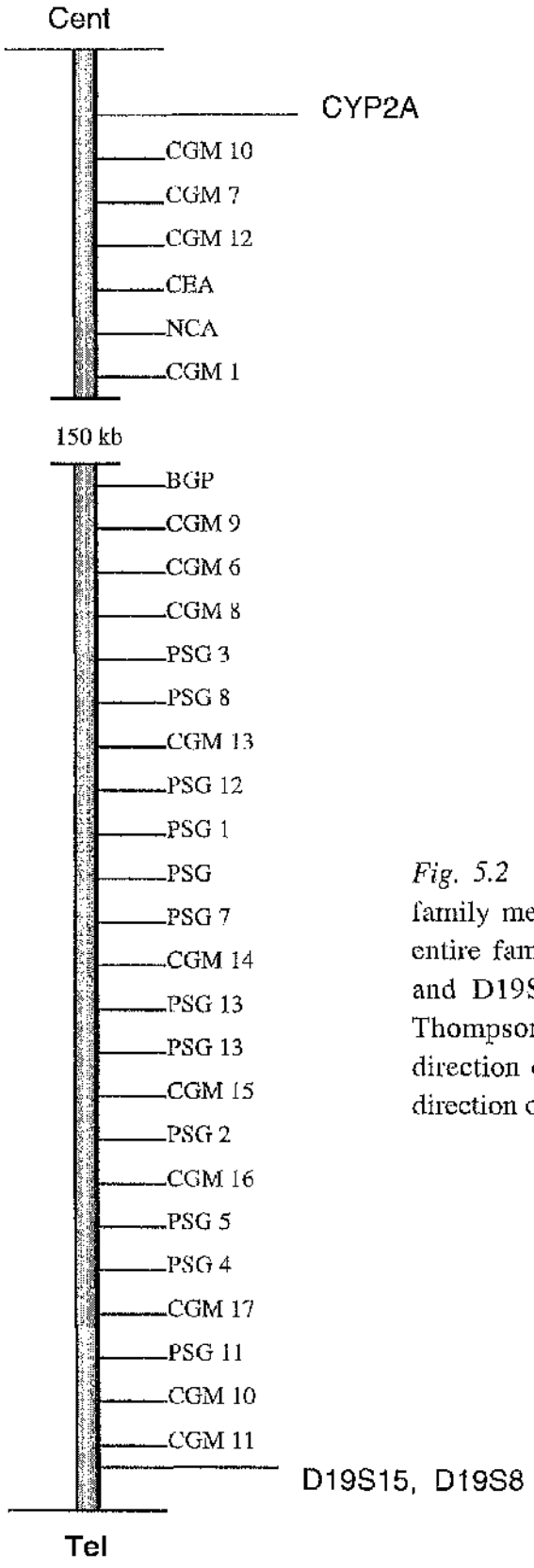


Fig. 5.2 Gene organisation of human CEA/PSG family members on chromosome 19q13.2. The entire family lies between the markers CYP2A and D19S15/D19S8 (Olsen et al., 1994 and Thompson et al, 1992). Tel indicates the direction of the telomere and cent indicates the direction of the centromere.

Several different rodent genes homologous to human CEA members have been found to exist. One of these genes, mouse BGP1 has two allelic forms, mmCEA1 (MHVR) and mmCEA2, the former acting as a receptor for the mouse hepatitis virus A59 with the latter being resistant (Dveksler et al, 1993),. These allelic forms generate many splice variants, the products of which are largely expressed in the liver (Turbide et al, 1991). A second mouse BGP gene, BGP2 (Nédellec et al, 1994), also acts as a MHV receptor and shows sequence homology with mouse BGP1. The rat gene encoding for an ecto-ATPase (Lin and Guidotti, 1988), which is identical to rat cell-CAM 105 (Aurivillius et al, 1990), shows a high degree of sequence homology with these mouse genes and human BGP. Like human BGP, rat ecto-ATPase/C-CAM 105, is found mainly in the epithelia of bile ducts (Odin et al, 1988). It appears therefore that these rodent genes are the equivalent of the human CEA subgroup of this family.

A second set of rodent genes appear to be analogues of the human PSG subgroup due to their exclusive expression in the placenta, the presence of RGD or RGD-like domains (except for *Cea*-4) and the lack of hydrophobic regions which imply secretion. Some of these rodent genes show very different gene structure to human CEA members due to the presence of numerous IgV-like domains as opposed to one. It was initially thought that there were five separate rat CEA genes when five IgV-N-terminal domain exons were isolated from a rat genomic library (Kodelja et al, 1989). The full cDNA of rnCGM1 was isolated and found to have five IgV-like domains and one IgC domain and dispelled the theory that there is only one IgV like domain per CEA protein, which is the case in all human CEA proteins (Rebstock et al, 1990). The same case was true for other rat genes expressed in the placenta, rnCGM3 and rnCGM4 which have three and at least two IgV-like domains respectively (Chen et al, 1994, Rebstock et al, 1990). The two mouse genes, *Cea-2* and *Cea-6* each have three IgV-like domains and also have hydrophilic C domains. The mouse genes *Cea-3* and *Cea-4* are also believed to belong to this group, showing exclusive expression in the placenta and having high homology at the nucleotide and amino acid level to *Cea-2* respectively (Rudert et al, 1992). *Cea-2* and *Cea-4* are separated by only 16kb which suggests that the mouse CEA family is tightly clustered as in human. The genes *Cea-5* and rnCGM2 have greater similarity than between other mouse and rat genes. They appear to form a third group as they are more distantly related to the other rodent genes and it is thought that they may represent the counterparts to human CGM 7 which is the most distantly related human CEA subgroup member (Rudert et al, 1992). Figure 5.3 displays the relationship between human and rodent CEA families.

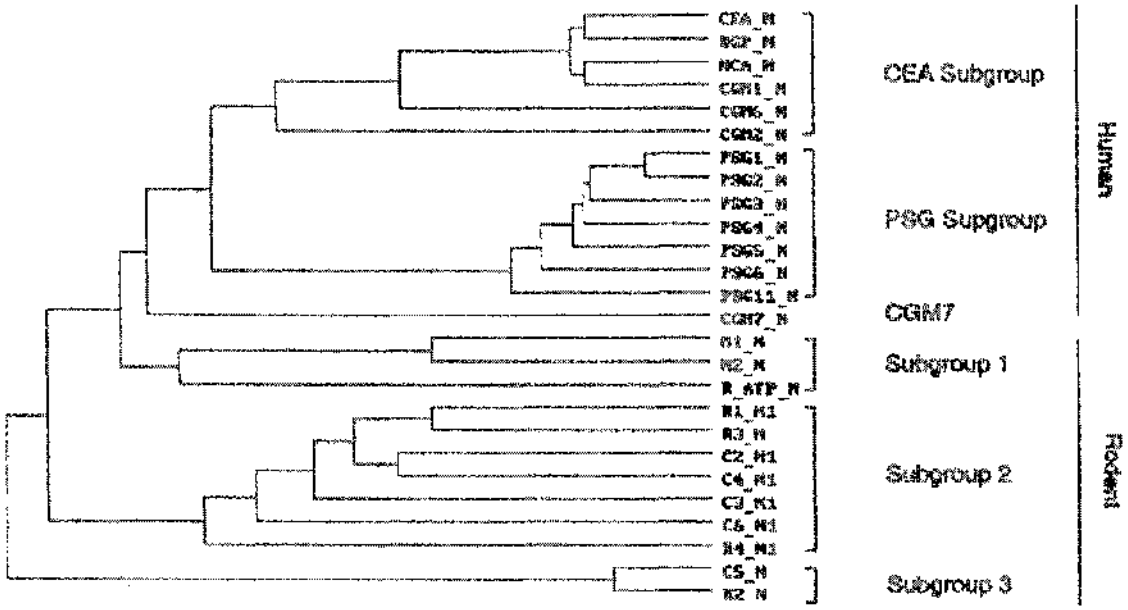


Fig.5.3 Relationship of human and rodent CEA family members. The amino acid sequences of the N-domain exons were compared and displayed as a dendrogram. The length of the branches reflect inversely the degree of similarity between the compared sequences. C = *Cea*, M = mmCGM, R = rnCGM and R_ATP = rat ecto-ATPase. Taken from Rudert et al (1992).

Rat CEA gene sequences

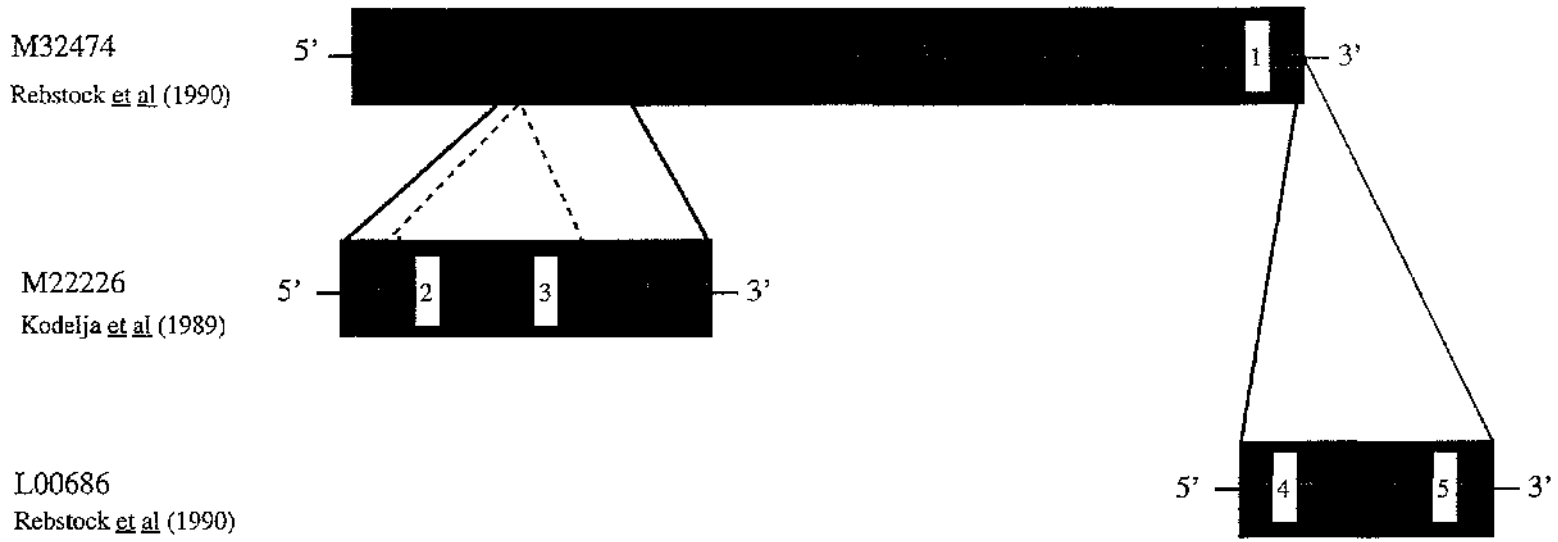
Genbank and EMBL databases were screened, using the STRINGSEARCH protocol of the Genetics Computer Group Sequence Analysis Software Package of the University of Wisconsin version 7, for any rat nucleotide sequences under the name CEA or CGM (CEA gene family member). Primers flanking any repeat units and other sequences of use were designed using MacVector Software version 4.1 (New Haven, CT) and constructed commercially (BRL, Life Technologies).. Primers were designed to be gene specific and therefore amplified 5'UTR, intronic or 3'UTR sequences. Checking that PCR products generated from each primer set were the correct was initially done by size analysis. PCR products generated from screening the P1 library using these primer sets were confirmed by sequencing (see chapter 6). Sequences and thermocycling parameters of each set of primers are described in the figure legends.

(I) rat CEA 1

Three sequences referring to rnCGM1 were obtained with the accession numbers, M22226, M32474 and L00686. Figure 5.4 summarises these rat CEA sequences and primers designed flanking microsatellites and for SSCP analysis. The sequence M2226, obtained by screening a rat genomic λ-phage library with human NCA and CEA5 cDNA clones (Kodelja et al, 1989), was in the form of an anonymous exon containing an N-terminal, IgV-like domain. This sequence contained two simple repeat sequences which are shown in fig. 5.5. Primers flanking both repeat units in this sequence were, designed, constructed and used to amplify the repeat units to check for any possible allelic differences amongst the rat strains, BN, AS/AGU and F344. Figure 5.6 shows electropherograms of the PCR products using the two primer sets, rnCGM1-rep1 and rnCGM1-rep2 on the DNA templates of the three rat strains and the two F1 DNA templates, BN/ASAGU and F344/AS/AGU. Electrophoresis of the products on 4% MetaPhor agarose gave no indication of size difference of the two microsatellites between the rat template DNAs. None of these markers were therefore informative for genotype analysis of the backcross programs. Another set of primers have recently been designed to amplify the first repeat unit of rnCGM1 called D1Arb32 (Ding et al, 1996). These however were not used as they were almost identical to the primers designed here.

The sequences M32474 and L00686 referring to the complete cDNA sequence of rnCGM1 including part of the 3'UTR and the putative remaining 3'UTR part of rnCGM1 respectively, were obtained by screening an 18 day rat placenta cDNA

(a)



(b)

Microsatellite	SSCP region	Primers designed	Reference	Size/sequence variations between BN, AS/AGU and F344 ?
1		D1Arb33	Ding et al (1996)	BN=AS/AGU=F344
2		rnCGM1-rep1 *	This Study	BN=AS/AGU=F344
		D1Arb32	Ding et al (1996)	Not Done
3		rnCGM1-rep2	This Study	BN=AS/AGU=F344
	4	rnCGM1-3'-1	This Study	BN=AS/AGU=F344
	5	rnCGM1-3'-2	This Study	BN=AS/AGU=F344

Fig. 5.4 (opposite) a, Diagram of sequences representing rat CGM1. M32474 = complete cDNA sequence of rnCGM1, M22226 = exon X of rnCGM1 and L00686 = putative remaining part of 3' UTR of rnCGM1. E = exon and I = intron. S = signal peptide, N1 -N5 = IgV domains, IgC = IgC domain, 3' = part of 3' UTR. Dashed lines show exon X containing the intron between N1 and N2. White boxes show positions of microsatellites or regions to be analysed by SSCP. b, Summary of primers designed and tested flanking microsatellites and regions for SSCP analysis. in sequences representing rat CGM1. * denotes primers used to isolate a P1 clone containing this gene (see chapter 6).

```
   1  GAATTCACTC CTCAGCTCTC ACAGCATAGA TGGACATACA GACTCCTGAA
  51  GGCTCTTCTC TTCCCTCCAC ACTGGTGTGT GTCACGTACC TGTAGTGTGC
 101  ACACTGGGAC ATGTACCTTC CCAAACCCTC ACGAACAATA CAGAAATATT
 151  AAATTACACT TGAATATAAT TATTTTTATG TGCTATAAAC ATGGAAATTA
 201  TGTAGACAAA CCCAGAGATA TCTTTTCTTC CTTCCTTCCT TCCTTCTTCC
 251  TTCCTTCCTT CCTTCCTCTT TTTCCATACT AGTTTCTGAG ATTTTTTGAG
 301  GAACTGAACC TTCCAAAAAG ACCATACCAA TCCCTGTCCT CAAAAAGCCT
 351  TTTTTATTCT AATGGACTGG AAATCATTGT ATCCAGAGGA GAAAGTCAAT
```

a

```
 301  GAACTGAACC TTCCAAAAAG ACCATACCAA TCCCTGTCCT CAAAAAGCCT
 351  TTTTTATTCT AATGGACTGG AAATCATTGT ATCCAGAGGA GAAAGTCAAT
 401  GATTTAGTGG AACCATAAAT AGAACAGAAA ACATTCAGGA AGTGAGGATT
 451  GTATGGAGGA GGAAAAAGAG GAGGAGGAGG AGGAAGAGGA GGAGGAGGAG
 501  GAGGAGGACC GAGAGCCGGT TCTCCACTCA CCAGACACTT TATGGAAAGA
 551  GTGATATGGG GACACCTGAG TAGAGGATTC CACAGAGAGG AAATGACACC
 601  CTTTGAGGTT CTGAGGGCAT GGAGGTCATG CTGCTCACCT CCATTAAGGG
 651  TGCATCCTAC CTACAGGCTG AGGGATGCTC ACACCTGCTC AGGATTGTCA
 701  ACTTTTCTCT CTTCCCTTCT AGCCTCCCTC TTAACCTGCT GGCTCCTGCC
 751  CACCACTGCC CAAGTCTCCA TTGAATCCTT ACCACCCCAG GTGGTTGAAG
 801  GAGAAAATGT TCTTCTACGT GTTGACAATT TGCCAGAGAA TCTCATAGCC
 851  TTTGTCTGGT ACAAAGGGCT GACAAACATG AGCCTCGGAG TTGCACTGTA
 901  TTCACTAACC TATAACGTAA CTGTGACGGG ACCTGTGCAC AGTGGTAGAG
 951  AGACATTGTA CAGCAATGGG TCCCTGTGGA TCCAAAATGT CACCCAGAAG
1001  GACACAGGAT TCTACACCCT ACGAACCATA AGTAATCATG GAGAAATTGT
1051  ATCAAATACA TCCCTGCACC TTCATGTGTA CTGTAAGTAA TTCTTTGTGA
1101  ATTC
```

b

Fig. 5.5 a, Partial nucleotide sequence of rnCGM1 (accession M2226, Kodelja et al, 1989) showing the first repeat unit which is underlined. Sequences in bold type are primers designed for the amplification of the repeat unit (rnCGM1-rep1). b, Partial nucleotide sequence of rnCGM1 (accession M2226, Kodelja et al, 1989) showing the second repeat unit which is underlined. Sequences in bold type are primers designed for the amplification of the repeat unit (rnCGM1-rep2). Boxed sequence shows intron-exon border.

library with the rat genomic clone for rnCGM1 (Rebstock et al, 1990).. Within the 3'UTR of the sequence, M32474, was a [(CCTT)8] repeat unit. Primer sequences for the amplification of this microsatellite were recently designed by Ding et al, (1996) called D1Arb33. This primer set was constructed and used to test for an allelic size difference between the rat strains. Figure. 5.7a shows the electropherogram of the PCR products using primers D1Arb33 on the DNA templates of the three rat strains and the two F1s. Electrophoresis of these products on 4% MetatPhor hinted on a size difference between the parental DNA templates but would not resolve any possible difference in the F1 DNA templates. Figure 5.7b shows an electropherogram of the same PCR products with end labelled primer, electrophoresed on 8% polyacrylamide. This method however could not resolve any size differences either.

For the putative remaining 3' UTR sequence of rnCGM1, accession L00686, two sets of primers were designed and constructed to cover its entirety (441nucleotides) for single-strand conformational polymorphism (SSCP) shown in fig. 5.8. It is known that in eukaryotic genomes, the 3' untranslated region of genes is largely not conserved compared to coding regions and therefore offers the greatest potential for sequence variation which can be identified with an expressed sequence (Levitt, 1991). SSCP involves electrophoresis of radiolabelled PCR products through a non-denaturing polyacrylamide gel which detects sequence differences, which may arise from point mutations, deletions compensated by insertions or rearrangements, that do not alter the size of the PCR product. The first set of primers, rnCGM1-3'-1, covered bases 68 to 305 of the sequence. Figure 5.9a shows an electropherogram of the PCR products using primers rnCGM1-3'-1 on the three rat strains and the two F1s. No sequence difference could be seen. The second set of primers, rnCGM1-3'-2, covered bases 173 to 409 of the sequence. Figure 5.9b shows an electropherogram of the PCR products using primers rnCGM1-3'-2 on the DNA templates of the three rat strains and the F1s. An apparent sequence difference could be seen between BN and AS/AGU. This initial result could not be repeated, even after numerous attempts, on any of the backcross progeny DNAs from the cross [AS/AGUxF1(AS/AGUxBN)]. It is possible that the bands with the apparent sequence difference were artifactual as they seemed distant from the main banding cluster putatively representative of the desired PCR products. It was therefore accepted that SSCP using the primers CAE1-3'-2 was also uninformative.

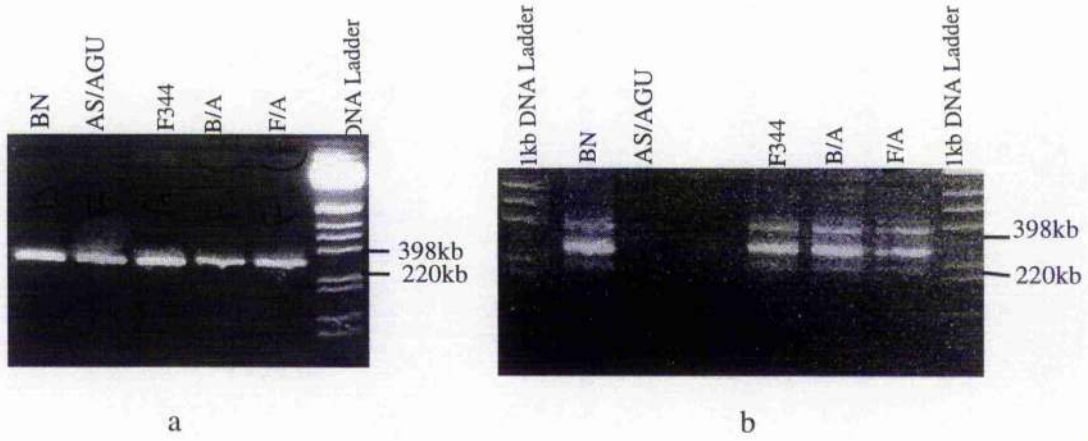


Fig. 5.6 Electropherograms of gels (4% MetaPhor, 1xTBE) showing the PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) using the following primer sets. a, rnCGM1-rep1, forward (5' TGGTGTGTGTCACGTACCTGTAGTG 3') reverse (5' CAGGGA TTGGTATGGTCTTTTTGG 3') PCR conditions were 25 cycles of 94°C for 15'', 53°C for 30'' and 72°C for 30''. b, rnCGM1-rep2, forward (5' AAGACCATA CCAATCCCTGTCCTC 3') reverse (5' AGGTGTCCCCATATCACTC TTTCC 3') PCR conditions were 25 cycles of 94°C for 15'', 55°C for 30'' and 72°C for 30''.

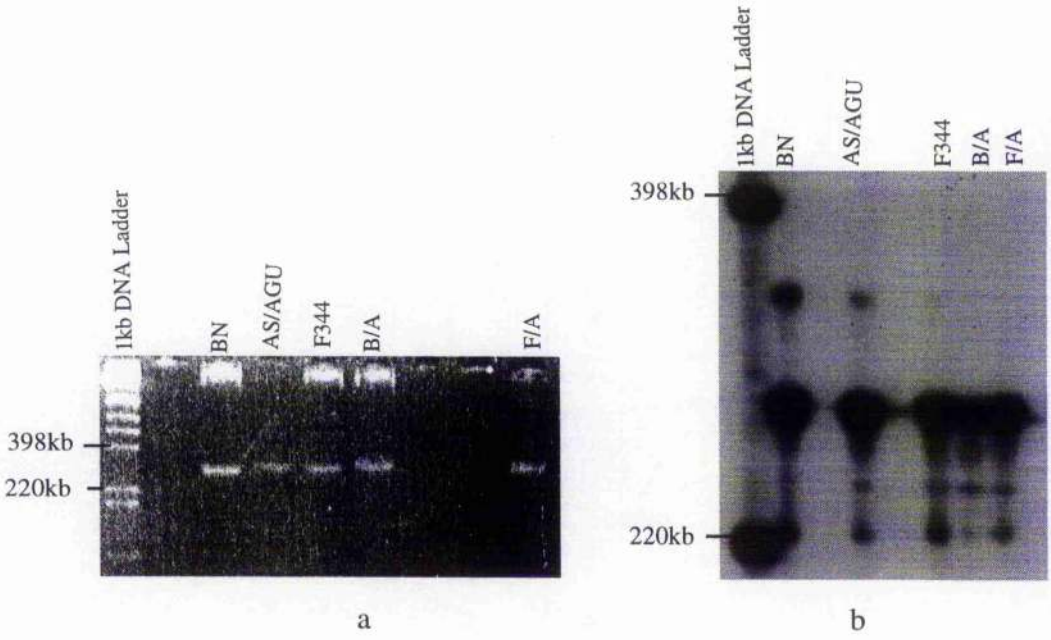


Fig. 5.7 Electropherograms of gels (1xTBE) showing the PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (BA) and F344/AS/AGU (FA) using the primer set, D1Arb33 desigend by Ding et al, (1996). a, electrophoresis on 4% MetaPhor. b, electrophoresis on 8% polyacrylamide, 7M urea.

```
  1 gaattcatga cacctgcaga catgtcttac agccaattga ttagccacca
 51 tttgtctgtt tgtgctgcct aattgctctg tctatggctc ctgagggatt
101 actgagcaga tgtggtcaca gcagcatcct tgatttttgt cttagagtga
151 aatctgttag tgtccctttt gagtactggg ttgcctatgt gttggtctcc
201 acgtgaggcc cttaacttaa ctgggatagt ctatccctaa ctgcctattt
251 catcactaaa aactcttcac tttacctggg tggcaaatgc ctttaatccc
301 agcactcaag agacagagac aggaagctct tggttagttc aaggccagcc
351 tggtctacag ttcctatcct gagttccagg atagccaagg ctacacagag
401 aaaccctgtc ttgaaaaaac aaaataaaac aaatgcaatt c
```

Fig. 5.8 Nucleotide sequence of the putative remaining 3'UTR sequence of rnCGM1 (acce ssion L0686, Rebstock et al, 1990) showing two sets of primers desinged and used for SSCP. The first set, rnCGM1-3'-1, are shown as underlined. The second set, rnCGM1-3'-2, are shown in italics.

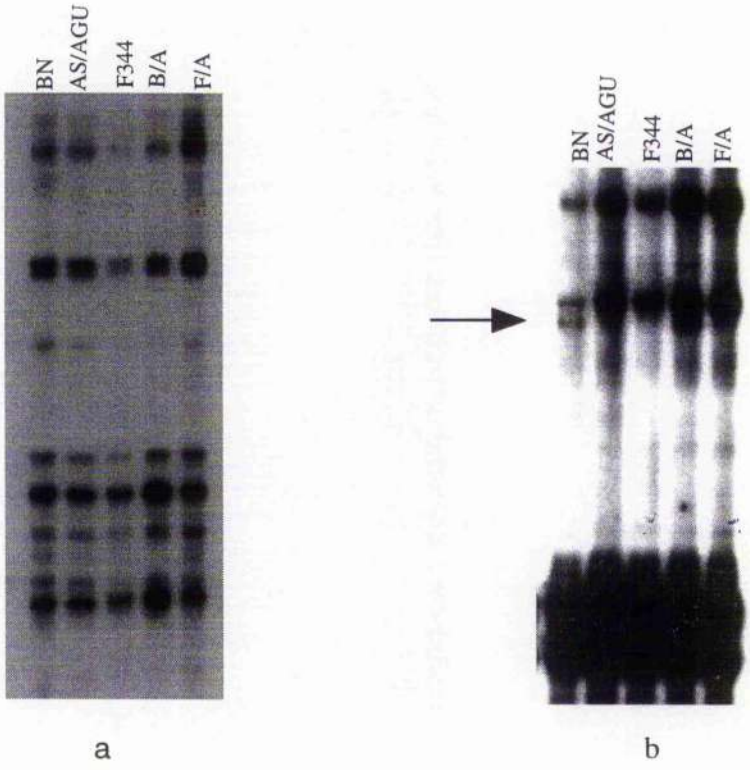


Fig. 5.9 Electropherograms of gels (6% acrylamide, 0.5xTBE) showing PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) using the following primer sets. a, rnCGM1-3'-1, forward (5' CCTAATTGCTCTGTCTATGGCTCC 3') reverse (5' GTGCTGGATTAAAGGCATTTG 3') PCR conditions were 25 cycles of 94°C for 15'', 55°C for 30'' and 72°C for 30''. b, rnCGM1-3'-2, forward (5' GTACTGGGTT GCCTATG TGTTGG 3') reverse (5' ACAGGGTTTCTCTGTGTAGCCTTG 3') PCR conditions were 25 cycles of 94°C for 15'', 54°C for 30'' and 72°C for 30''. Arrow indicates position of possible strain difference.

(ii) rat CEA 2

One sequence referring to the gene rnCGM2 was obtained with the accession number M2227. This sequence, obtained in the same way as M22226, was in the form of an anonymous exon containing an N-terminal domain and is shown in fig. 5.10. No repeat units or intronic sequences long enough for SSCP analysis could be found.

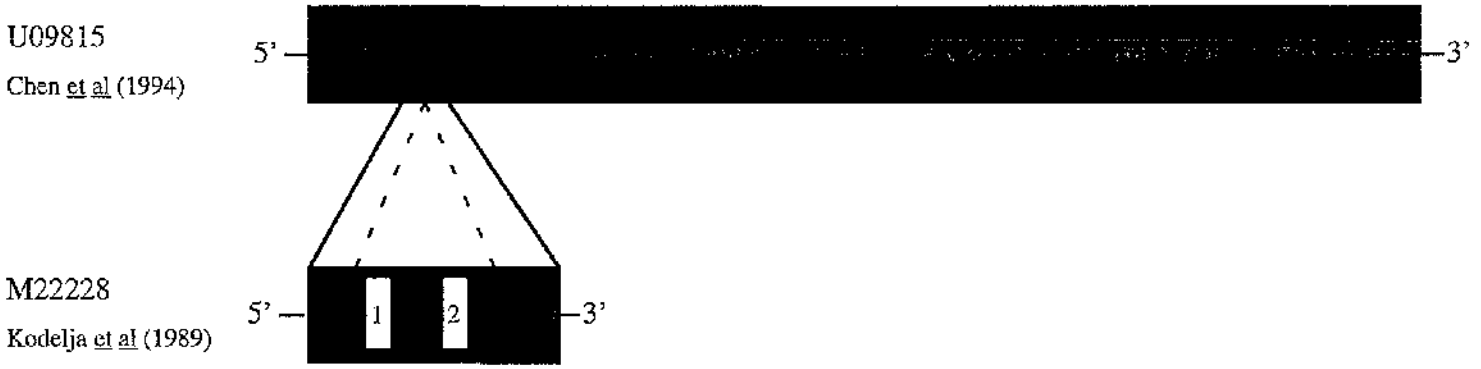
(iii) rat CEA 3

Two sequences referring to rnCGM3 were obtained with the accession numbers, M22228 and U09815. Figure 5.11 summarises these rat CEA sequences and primers designed flanking microsatelites. The sequence M2228, obtained in the same way as M22226, was in the form of an anonymous exon containing an N-terminal domain. This sequence also contained two simple repeat sequences which are shown in fig. 5.12. Primers could only be designed flanking the second repeat unit which were used to amplify it to check for any possible allelic differences amongst the rat strains. Figure 5.13 shows an electropherogram of the PCR products using the primer set rnCGM3-rep2 on the DNA templates of the three rat strains and the two F1 DNA templates. Electrophoresis of the products on 4% MetaPhor agarose gave no indication of size difference of the two microsatellites between the rat template DNAs. Although no primers could be designed for the first repeat of this sequence, primers were available for this microsatellite called CEAC or D1Mgh5 (Jacob et al, 1995). This marker also proved uninformative for the two backcross programmes (M. B. Duran Alonso, personal communication). Another set of primers have been used to amplify this microsatellite called D1Arb36 (Ding et al, 1996) but were not used due to their similarity with D1Mgh5.

The sequence, U09815, referring to the complete cDNA sequence of rnCGM3. was obtained by screening a 17 day rat placental cDNA library (Chen et al, 1994) and finding that it matched up with the genomic clone containing the sequence M22228. This cDNA contained 1137 bases of 3'UTR. It was desired to develop a method of checking the whole 3'UTR for SSCP. By digesting the DNA into three or four parts with appropriate restriction endonucleases, the fragments could be labelled and run through a non-denaturing polyacrylamide gel. Any sequence differences arising from a particular fragment would result in that fragment being sequenced and the location of the polmorphism in the 3'UTR would be known. However no

Fig. 5.10 Diagram of sequence representing rnCGM2. No microsatellites or intronic regions long enough for SSCP analysis were present. E = exon and I = intron The short lines above and below represent the primers rnCGM2-1 which were designed for the isolation of a P1 clone containing this gene (see chapter 6).

(a)



(b)

Microsatellite	Primers designed	Reference	Size variations between BN, AS/AGU and F344 ?
1	D1Mgh5	Jacob et al (1995)	BN = AS/AGU = F344 †
	D1Arb36	Ding et al (1996)	Not Done
2	rnCGM3-rep2	This Study	BN = AS/AGU = F344

Fig. 5.11 a, Diagram of sequences representing rat CGM3. U09815 = complete cDNA sequence of rnCGM3, M22228 = exon X of rnCGM3. E = exon and I = intron. S = signal peptide, N1 -N3 = IgV domains, IgC = IgC domain, 3' = 3' UTR. Dashed lines show exon X containing the intron between S and N1. White boxes show positions of microsatellites. b, Summary of primers designed and tested flanking microsatellites in sequences representing rat CGM3. † experiment carried out by M. Alonso. * denote primers used for the isolation of a P1 clone containing this gene (see chapter 6).

```
  51  ACTGAATTAT TAAATCACAC TTGAATATAT TGATTTCCCT TTGCTCTGAG
 101  CCTGGGCACT ATGTAGATAA GTCCATGGAA ATATTAATCT TTCCTTCCTT
 151  CCTTCCTTCC TTCCTTCCTT CCTTCCTTCC TTCCTTCCTT TCTTCCTTCT
 201  TCTAGTTCTT TTCACGTTTT CCCTTTTCTT TTTCTCTCCA ATTTGTTTCT
 251  AATCTATTTT CAGGAACTGA ACCTTCCAAA AAGATGATTC CAGTCCCTGT
```

a

```
 301  CCTCACAAAG CCCTTTTCTT GTGGACTGGA AGTCAGAGTA TCCAGAGAAA
 351  GGCAATGGTT TAATGGAACC TCAAACAGAA CAGAAAACAA TTCTGAGAGT
 401  GAGCATTGCA TGAGGAAGAG GAGGAACGGG AAGAGGAGGA GGAAGAGGAG
 451  GAGGAAGAGG AGGAGGAAGA GGAATGGGAA GAGGAGGAGG AAGAGGAGGA
 501  AGAGGAGGAA GAGGTCAGAC AGCTGCTTCA CCTCTCACCA GACACTCTAT
 551  GGGAAGAATG ATATGGGGAC ACCTGAGTAG AGGATTCCTG GAGAGGAAAT
 601  GACAGCTTTT GAGTCTTTGA GGGCATGGAG GTCATGCTGC TCACCTCCAT
 651  TAAGGGTGCA TCCTACCTAC AGGCTGAGGG ATGCTCACAC CTGCTCAGGA
 701  TCGGTGACTT TTTTCTCTTC CCTTCTAGCC TCACTTTTAA CCTGCTGGCT
 751  CCTGCCCACC ACTGCCCACG TCACCCTCAA GTCCTCACCG CCCCAGGTGG
 801  TTGAAGGAGA AAACGTTCTT CTAAGTGCTG ACAATCTGCC AGAGAACATT
 851  ATAGCTTTCG CCTGGTACAA AGGGGAGACC GACATGAACC GTGGAATTGC
 901  ACTGTATTCA CTGAGGTATA CTGTAAGTTT GACGGGGCCT GTGCACAGTG
 951  GTCGAGAGAC ATTGTACAGC GACGGGTCCC TGTGGATCAA AAATGTCACC
1001  CAGGAGGACA CAGGATTTTA TACCTTTCGA ATCATAAATA ATCATGGAAA
1051  AATTCAATCA AATACAACCC TGTTCCTTCA CGTGAAATGT AAGTAACTCT
1101  TTGTGAACTG TGGGTTTTGG GTGGTGTCCT TCCACTAGAC ACATAGAAGT
1151  ATCAGGCCAG GGCTGTGTCT CCCTTCCCCC TGCAG
```

b

Fig. 5.12 a, Partial nucleotide sequence of rnCGM3 (accession M22228, Kodelja et al, 1989) showing the first repeat unit which is underlined. No suitable primers could be designed for the amplification of this repeat unit. b, Partial nucleotide sequence of rnCGM3 (accession M22228, Kodelja et al, 1989) showing the second repeat unit which is underlined. Sequences in bold type are primers designed for the amplification of the repeat unit (rnCGM3-rep2). Boxed sequence shows the intron-exon boarder.

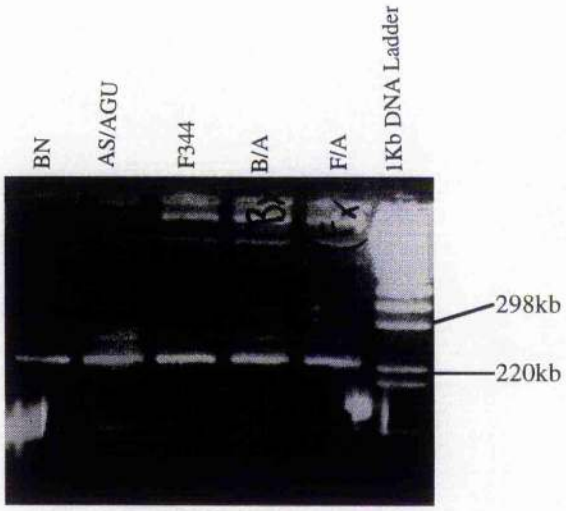


Fig. 5.13 Electropherogram of gel (4% MetaPhor, 1xTBE) showing the PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) using the primer set rnCGM3-rep2, forward (5' TTGTGGACTGGAAGTCAGAGTATCC 3') reverse (5' AGT GTCTGGTGAGAGGTGA AGCAG 3') PCR conditions wre 25 cycles of 94°C for 15'', 56°C for 30'' and 72°C for 30''.

suitable primers or restriction endonucleases could be designed to initiate this method.

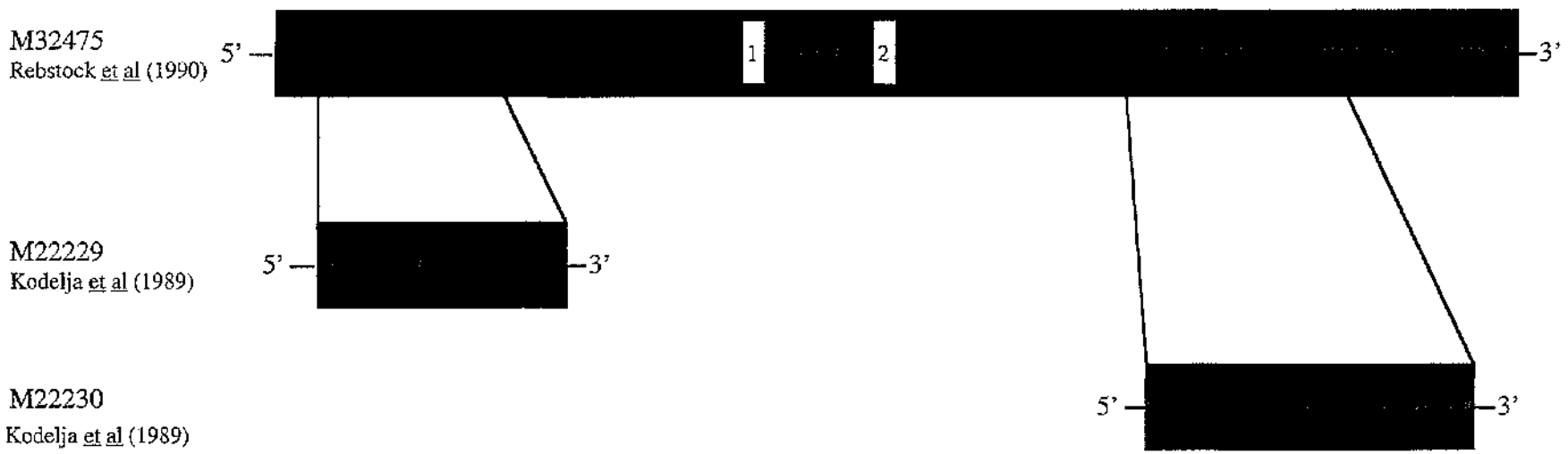
(iv) rat CEA 4

Two sequences referring to rnCGM4 were obtained with the accession numbers, M22229 and M32475. Figure 5.14 summarises these rat CEA sequences and primers designed flanking microsatelites and for SSCP analysis. The sequence M22229, was again obtained in the same way as M22226 and was referred to as an anonymous exon containing an N-terminal domain. Present on the genomic clone containing this exon, was another N-terminal domain exon (sequence M22230), which was assumed to belong to another rat CEA gene, CEA5. No repeat units were detected in either of these sequences. The sequence M32475, referring to exons 2 and 3 of rnCGM4, was obtained by further analysing the original genomic clone containing the sequences M22229 and M22230 (Rebstock et al, 1990). Since the theory of one IgV-like domain per CEA gene had been dispelled, it was now thought that the anonymous exon sequences, M22229 and M22230, present on the same genomic clone, were two exons of the same gene, CEA4. These sequences encompass exons 2 and 3 of the sequence M32475. Within the intron, between exons 2 and 3, was a CCA trinucleotide repeat. Primers were available for this microsatellite which were called CEAR or R100 (Serikawa et al, 1992). This marker proved uninformative for the two backcross programs (M. B. Duran Alonso, personal communication). A set of primers were designed, constructed and used to amplify part of the intron between these two exons for SSCP. Figure 5.15 shows the intronic sequence on which these primers, rnCGM4-intr-1, were constructed. Figure 5.16 shows an electropherogram of the PCR products using primers rnCGM4-intr-1 on the three rat strains and the two F1s. No sequence difference could be seen.

Use of mouse primers flanking a microsatellite from the gene mmCEA2 to amplify putative homologues in the rat

Mouse CEA/PSG database sequences were searched for any repeat units to facilitate in finding other rat microsatellites in putative gene homologues. The sequence, x53084, referring to the mouse gene mmCEA2, was found to have a $(CA)_{19}$ repeat unit. Primers were designed flanking this repeat, to possibly amplify the rat DNA templates. Figure 5.17 an electropherogram of the PCR products using the primers for the repeat unit in mmCEA2 on the rat DNA templates and mouse C57/BL6 DNA

(a)



(b)

Microsatellites	SSCP regions	Primers designed	Reference	Siz/sequence variations between BN, AS/AGU and F344
1		R100	Serikawa et al (1989)	BN = AS/AGU = F344 †
	2	rnCGM4-int-1 *	This Study	BN = AS/AGU = F344

Fig. 5.10 (opposite) a, Diagram of sequences representing rat CGM4. M32475 = exons 2 and 3 of rnCGM4, M22229 = exon X of rnCGM4 and M22230 = second exon X of rnCGM4. E = exon and I = intron.. White boxes show positions of microsatellites or regions to be analysed by SSCP. b, Summary of primers designed and tested flanking microsatellites and regions for SSCP analysis. in sequences representing rat CGM4. † experiment carried out by M. Alonso. * denote primers used for the isolation of a P1 clone containing this gene (see chapter 6).

2351 ACTGGAGAAG AGTTTCAGGT TTGTAGGGCT GAAAACACTA AAGTATAGGG
2401 GCTCATCATC ATCATCATCA CCACCGCCAT CACCACCACC ACCACCACCA
2451 CCACCACCAC CACCACCACC ACCACCACCA CCATCATCAT CATCATCATG
2501 AGGCTCTTGG TAAATAAGAA GAAGCAGGGG GAGGAGGAGA TTATTGTCAA
2551 CCCACAGTTC ACCATCAATG AGCCCAGTGT TCTGAAGACT GAGGTTCTCA
2601 GCTGTGATGC CCCAAATAAG AAACCAAGCT GGTGTTGATC AGTGACATGG
2651 CTCAGTGGAT CTGGGTGTTT GCTTCATGTC TGACAACCTG AGAACCAGTG
2701 AACACAAGTT GTCCCTGACC TCCACCTAGG GACGGCGTTT TGCACCCAAC
2751 ACAGACACAC TGAGGCATGC CCTTGCACAT GAACTCATAC ACCAATATAA
2801 TAAGCAAATG CATAAAAATT ATAGCAAATG GAAGCAGTCA ACACTGTATT
2851 CCCAAACATA CTAATTTGTT AAATAAATCC ATGGCCATGT ATTCATTCAT
2901 TCATTCATTC ACTCATTCAT TTACTCTCCA AGATATTTGA GTTTTCTTTT
2951 GCAGTCTTTT TTTTTTAAAA GATAATATAA GACAAATCCC AGTTCTCATT
3001 ATTCCCTAGC CCTAGACTGG AAGACGACCA GTGAAGAAAG CTAGAAGGCG
3051 AATCAGTCAC TAAAGGACAA GAAACAAAAG AGTCAGAGTG TGACGGTCGG
3101 GAGGCTTCAC CCCAACACCC ATCGACTGAC ACTGAGGGTG AGCAGGGATC
3151 TGAGGACGGT GAGGCAGGGC CA**TGTTGACA CCTGAGGAGA GAGCAG**CATA
3201 GAGAGGAAAT GACAAGTGAG GGGCGCGGAG TGCATGGAGG TAATGCACTG
3251 ACCTCCACTA GCTAGGGCAG GGAGACTCCC ACACCTCAGC TGACCACTGG
3301 ACACAGCTGC TCGGACTCAG GCACCATCTT AGCCAAATAC TAAAGTCCTG
3351 ATGTTGACGG ATCTCTCTTC CCTTCTAGCC TCTCTTTTCA TCTGTGGGC**G**
3401 **TCCTTTTAAC CCTGCCAAGC T**CACTATTGA ATCAGTGCCG CCCAGTGTTG

Fig. 5.15 Partial nucleotide sequence of the intron betweewn exons 2 and 3 of the sequence M32475, representative of rnCGM4. Sequences underlined indicate the primers used by Serikawa et al (1989), called R100 (CEAR). Sequences in bold type are the primers, rnCEA4-intr-1, used for SSCP analysis of this intron.

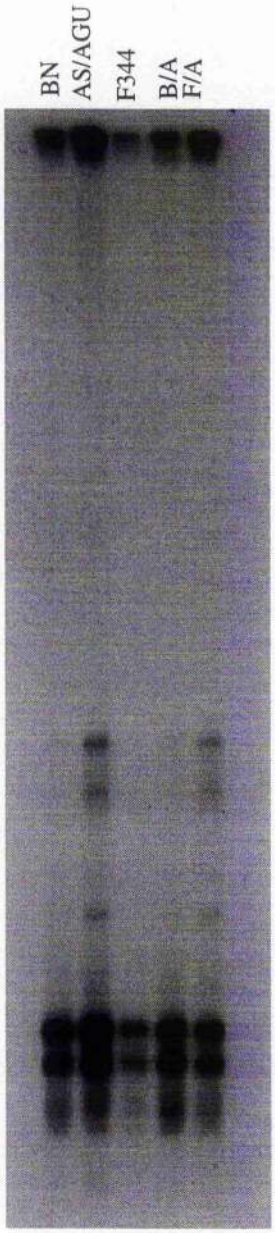


Fig 5.16 Electropherogram of gel (6% acrylamide, 0.5xTBE) showing SSCP analysis of PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) using the primer set rnCGM4-int-1, for the intonic region of the sequence M32475, forward (5' TGTTGACACCTGAGGAGAGA GCAG 3') reverse (5' AGCTTGGCAGGGTTAAAGGAC 3'). PCR conditions were 25 cycles of 94°C for 15'', 57°C for 30'' and 72°C for 30''.

as a control. Two sets of bands were produced from the rat DNAs putatively representing the desired product, although both product sizes were different from that seen in mouse. Electrophoresis on 4% MetaPhor hinted on a size difference among the rat PCR products, which was indeed the case when the products were run on 8% polyacrlyamide. From the cross [AS/AGUxF1(AS/AGUxBN)], 18 DNA samples were genotyped using these primers to test for possible linkage to the *agu* locus. The data from the genotyping is shown below,

mmCEA2	Normal	Mutant
a/a	2	2
a/b	9	5

where a/a = homozygous for the AS/AGU allele and a/b = heterozygous for the AS/AGU and BN alleles.

Chi-square analysis of this data gave a value of 7.33 which is less than the 5% critical value of 7.815. There was no significant variance from random segregation and therefore it could be concluded that the polymorphic PCR product in the rat using the mouse mmCEA2 primers was not clearly linked to *agu* and may not be located on chromosome 1 .

The PCR product in the rat giving rise to the strain difference was analysed by sequencing. The sequence obtained, when compared to the Genbank nucleotide database, was not related to any CEA members in any organism. The mouse alpha-1 antitrypsin gene, segment 1, was the closest sequence match with 50 % identity over 93 bases. This gene has been linked to mouse chromosome 12 (Krauter et al, 1986),, an area which is not syntenic to rat chromosome 1 (Copeland et al, 1993). Ideally, this polymorphic PCR product could be localised to a chromosome by using a somatic cell hybrid panel which consists of individual or groups of rat chromosomes on a mouse background. Comparison of the mouse mmCEA2 gene to rat CEA sequences, showed that there was 83% homology over 1300 bases to the rat ecto-ATPase gene. Present on this rat gene was a large AC repeat. Primers flanking this repeat were designed, constructed and used to test for a strain difference between the rat strains. No size difference was observed however (N. Craig, personal communication).

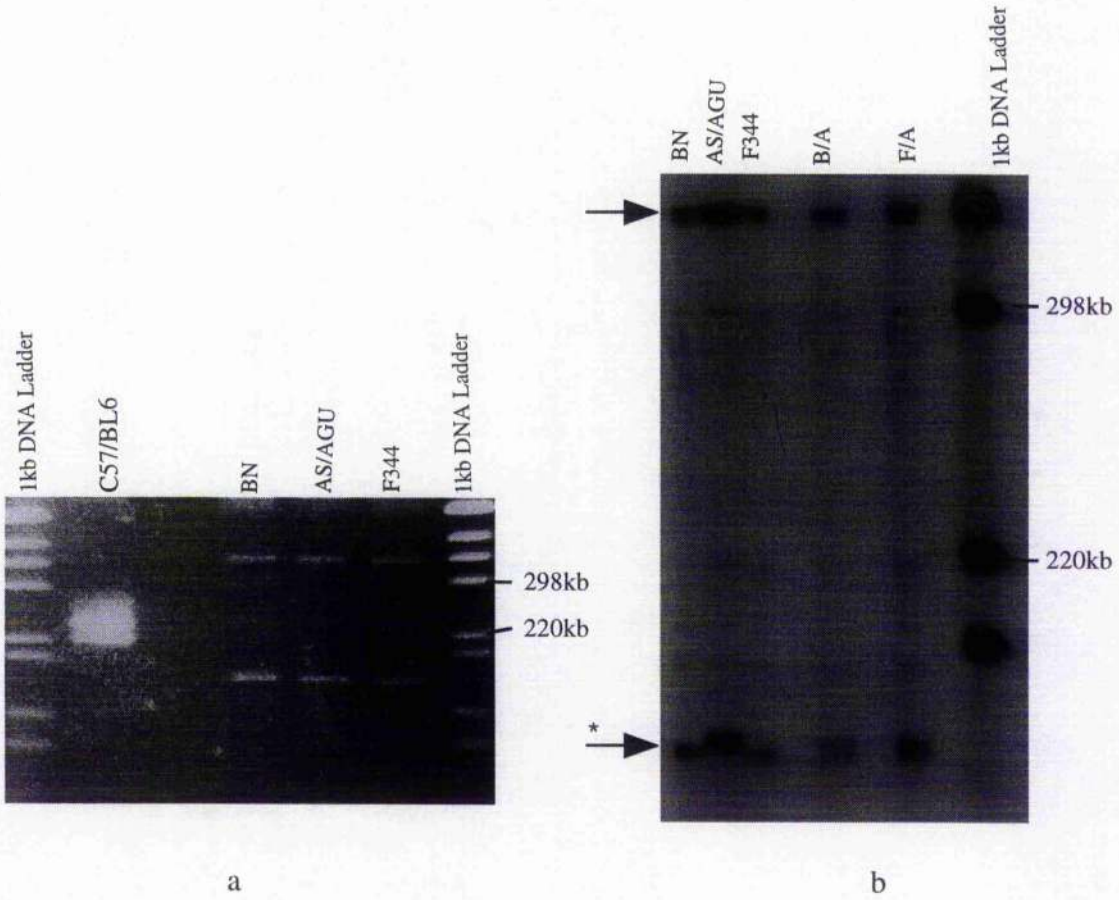


Fig. 5.17 Electropherograms of gels (1xTBE) showing the PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) and also on the mouse template C57/BL6, using the primer set designed flanking a CA repeat in the mouse gene mmCEA2, forward (5' GAA GTGAGCATAAT CTGTCCGTCTG 3') reverse (5' GGGTTGAGTTTTCTGTAGGCTCC 3'). PCR conditions were 25 cycles of 94°C for 15'', 56°C for 30'' and 72°C for 30''. a, electrphoresis on 4% MetaPhor. b, electrophoresis on 8% polyacrylamide, 7M urea. Arrows indicate the two putative rat PCR products using the mouse primers. The * shows a size difference.

Final Discussion

In total, ten CEA associated markers were analysed to established whether or not they could be used in genetic mapping using the two backcross programs established for the *agu* project . None of these proved informative even though they were based on 5'UTR, intronic or 3'UTR sequences.. One reason for this is that CEA family members are highly conserved between the strains BN, AS/AGU and F344. The next chapter describes the search for new CEA associated microsatellites.

Chapter 6

Searching for new CEA associated microsatellites in the rat

Introduction

All of the available microsatellites associated with rat CEA genes had proved uninformative using the two backcross programs. To obtain additional microsatellites coupled to these family members, a rat genomic P1 library was simultaneously screened for CEA gene sequences and for the presence of various repeat units. Sequences flanking any repeat units found, would be used for the design of primers to analyse the new microsatellites.

The Bacteriophage P1 system

The P1 vector system was developed to facilitate the cloning of large fragments of genomic DNA (~95kb), intermediary between the size accommodated by λ-cosmids and YACs (yeast artificial chromosomes) (Sternberg, 1990). Although a very efficient cloning vector, the λ-cosmid system can accommodate inserts up to only 40kb. The P1 vector's advantage over YAC clones is the greater efficiency of cloning and the larger amounts of insert DNA that can be recovered from transformed cells (Sternberg, 1990). Also, isolation of YACs is much more difficult than plasmid preparations. The preferred P1 vector for genomic library construction is pA*d*10-SacBII due to the ability to positively select for clones containing inserts (Pierce et al, 1992). This vector is a derivative of the originally developed P1 pA*d*10 (Sternberg, 1990). Figure 6.1 shows a diagram of the features and cloning rationale of the vector pA*d*10-SacBII. During the production of clones the vector is cut at two sites, *Sca*I and *Bam*HI, to generate two vector arms which flank the insert DNA during ligation. Packaging of the P1 genome is by a progressive headful mechanism. The 4.4kb arm contains the *pac* site which initiates packaging of the ligated product into phage heads and the 27kb arm contains the genetic information for selection and replication. The presence of *lox*P sites on each arm enables the entire ligated molecule to cyclize and be maintained as a single copy plasmid when injected into *Cre*+ bacteria. Cells containing P1 molecules can be selected for due to the presence of the kanamycin resistance gene on the vector. When foreign DNA is inserted into the *Bam*HI cloning site, the proximity of the *sac*B gene and its promoter are interrupted. This stops expression of the *sac*B gene and allows bacteria to grow on plates containing sucrose. P1 clones that do not contain inserts will express the gene which in the presence of sucrose produces levan, a compound deadly to E. coli. The discrimination between P1 molecules with

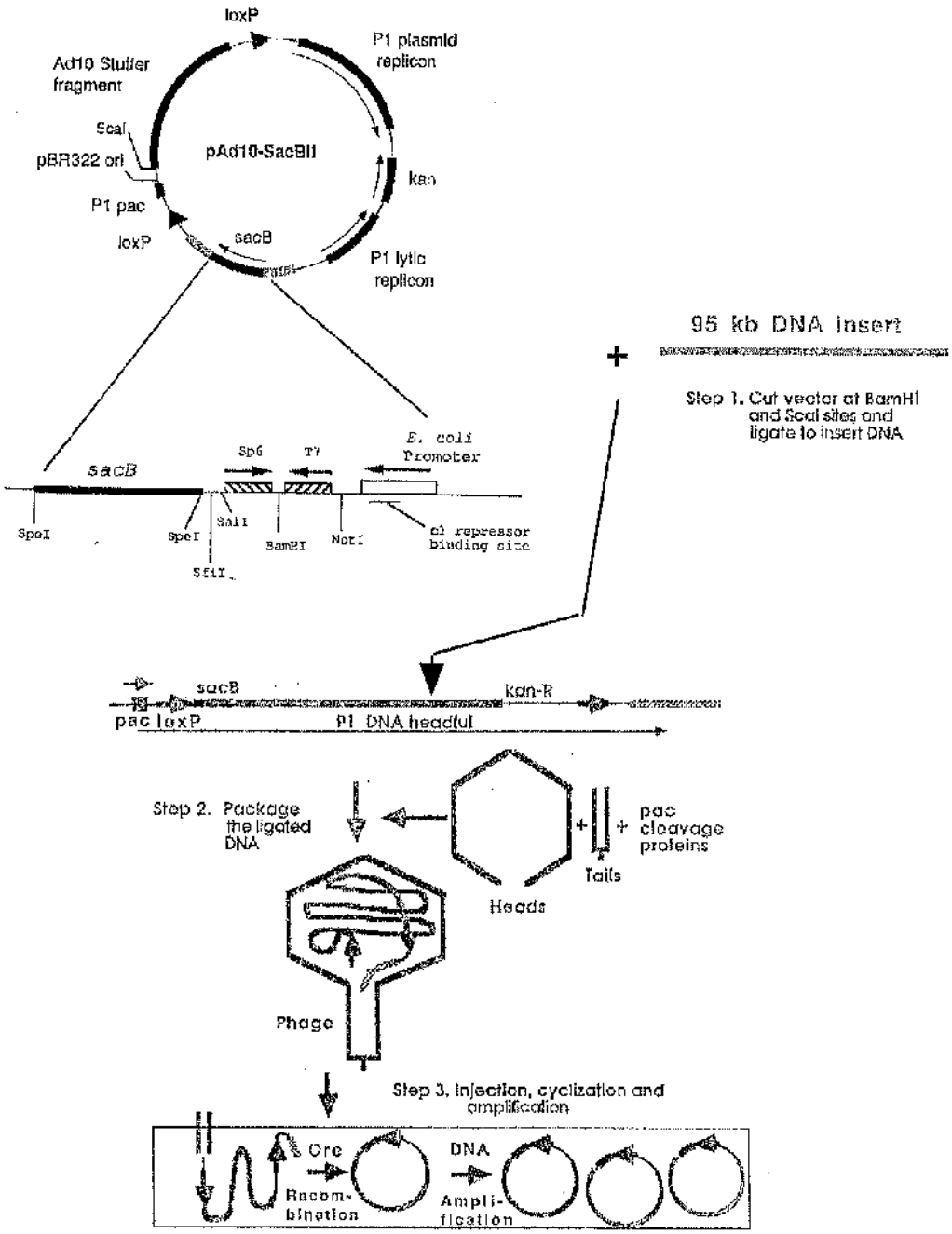


Fig 6.1 Diagram showing various features and the cloning scheme of the P1 vector pΛ*d*10-SacBII. The major portion of the vector is derived from pA*d*10 (Sternberg, 1990). The two domains are separated by *lox*P sites. One domain contains the *pac* sequence (packaging of DNA), an origin of replication and a stuffer fragment. The other contains a kanamycin resistance gene, a plasmid replicon, an IPTG inducible lytic replicon, and a positive selection *sac*BII cassette. The cassette codes for the enzyme levansucrase which , in the presence of sucrose produces levan, a substance lethal to *E. coli*. Also present in the cassette is the BamH1 cloning site which separates the *sac*B from its promoter and is flanked by Sp6 and T7 promoter elements and rare cutting restriction sites. The P1 vector DNA is cleaved by the enzymes ScaI and BamHI to generate two vector arms which flank the insert DNA. When ligated, packeged and injected into a *Cre* + host, site specific recombination between the two *lox*P sites occurs which cyclizes the molecule. From Pierce and Sternberg (1993).

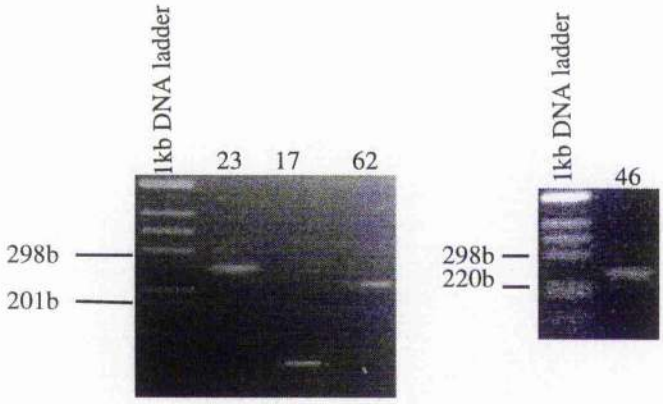
inserts and those without when grown on sucrose is approximately 70-fold (Pierce et al, 1992). The vector also contains the Sp6 and T7 promoter sites for RNA probe production to facilitate in the characterisation of the clone and to aid in chromosome walking.

The Rat P1 Genomic Library

A rat P1 genomic library was constructed by Southard-Smith and colleagues at the University of Dallas, Texas (Southard-Smith et al, 1994, Southard-Smith and MacDonald, 1993). It was originally created to clone rat kalekrein genes and has a coverage of approximately 2.4 genomes. The library was kindly made available to us by Dr. MacDonald as a set of 209 superpools stored in glycerol, containing an average of 550 clones per pool (range =230-1200). Lysates from each superpool were prepared for screening by PCR.

Screening of the rat P1 library for CEA sequences

The first 80 of the 209 superpools were screened by PCR for the presence of the rat genes rnCGM1, rnCGM2, rnCGM3 and rnCGM4. Primers for detecting genes 1, 3 and 4 were those used previously for analysing microsatellites and SSCP, rnCGM1-rep1, rnCGM3-rep2 and rnCGM4-int-1, respectively. For rnCGM2, a set of primers was designed (rnCGM2-1) solely for isolating a P1 clone as this gene did not contain any repeat units or the potential for SSCP analysis. Figure 6.2 shows the results of the PCR experiment using all of these primers on the superpools. Superpools 23, 17, 62 and 46 were identified as containing the rat CEA genes 1, 2, 3 and 4, respectively. The sizes of the PCR products from these superpools were as expected. To confirm that the PCR products generated from the superpools were representative of the rat CEA genes, they were purified, sequenced and aligned using the BESTFIT program (GCG) to their respective gene sequences shown in fig.6.3. Alignments for genes 1, 2 and 4 had high percentages of similarity (>90%). Some errors in the sequencing procedure may have been responsible for bases which did not align, particularly through the repeat region in rnCGM1. Taking this matter into consideration with the observation that these PCR products were of the expected size, it was considered that these primer sets were amplifying the desired genes in the P1 library and they were therefore used to isolate single clones. The product generated from superpool



PCR primers designed to rnCGM gene:	1	2	3	4
Expected PCR product size (base pairs):	264	123	228	249
Observed PCR product size (base pairs):	264	123	228	249
Percentage similarity of PCR product sequences to respective rnCGM genes:	90.1	97.5	64.2	98.3

Fig. 6.2 Electropherograms of gels (4% MetaPhor agarose, 1xTBE) showing PCR products from the superpools containing rat CEA genes. Superpool 23 was identified using the primers rnCGM1-rep1. Number 17 was identified using the primers rnCGM2-1, forward (5' AATTCT TGTTG GAGAGTGAGTGGG 3') reverse (5' CTCATTGGTGCAGGGAAAC TTC 3') PCR conditions were 25 cycles of 94°C for 15'', 55°C for 30'' and 72°C for 30''. Number 62 was identified using the primers rnCGM3-rep2. Number 46 was identified using the primers rnCGM4-int-1.

The number of clones present in these superpools were as follows:

Superpool number	Number of clones
17	456
23	474
46	950
62	950

rnCGM1-rep1 product vs. rnCGM1 exon X

Percent Similarity: 90.164

```
118 TTCCCAAACCCTCACGAACAATACAGAAATATTAAATTACACTTGAATAT 167
    |.:  : :: ::||||        ||||  |  ||| |  ||:|||||:
  3 ttngtngnngnncacg........agaataagaaaaattaacntgaatan 44

168 AATTATTTTTATGTGCTATAAACATGGAAATTATCTAGACAAACCCAGAG 217
    ||||| ||| ||||||||||||||:|||||||||||||||| |||||
 45 aatta.tttaangtgctataaacntggaaattatgtagacaaa.ccagag 92

218 ATATCTTTTCTTCCTTCCTTCCTTCCTTCTCCCTTCCTTCCTTCCTTCCT 267
    ||| ||||||||||||||| |||||           | ||  ||  ||
 93 ata.cttttcttccttcttcttc............ttctctctctct 128

268 CTTTTTCCATACTAGTTTCCGAGATTTTTTGAGGAACTGAACCTTCCAAA 317
    || ||||      :||||||||| ||| ||||||| || | |||||:||
129 ctctttc.......ntttctgaga.tttttgaggaa.tgaa.cttccaaa 168
```

rnCGM2-1 product vs. rnCGM2 exon X

Percent Similarity: 97.561

```
 82 AAAATAATTCTTGTTGGAGAGTGAGTGGGGAGCCATGCAGACACGGGAGG 33
    ||  |||||||:|||||:|||||||||||||| ||||||||||||||||
  1 AATGTAATTCTTGTTGGAGAGTGAGTGGGGAGCCATGCAGACACGGGAGG 50

 32 AGAGAGACCCGTAC.AAAGGTCACTCCAGCTCC 1
    |||||:|||| || ||||||||||||||||||
 51 AGAGAGACCCGTACAAAAGGTCACTCCAGCTCC 83
```

rnCGM3-rep2 product vs. rnCGM3 exon X

Percent Similarity: 64.220

```
151 CCTTCCTTCCTTCCTTCCTTCCTTCCTTC.CTTCCTTCCTTTCTTCCTTC 199
    ||   ||  |||| |||| | |||:|| | |||      || ||||
  1 ccctctctctttctctcctctcatccntctcctcctcttcctcctcctct 50

200 TTCTAGTTCCTTTCACGTTTTCCCTTTC.........TTTTTCTCTCCA 240
    | |:  | || || :|||  ||| | |          | ||||| |
 51 tccnectectcttcgngttcctcctctgcacatgcaatgctcactctcaa 100

241 ATTCGTTTCCAATCTATTT 259
    | | ||||||  || ||:
101 aatggtttctgttctgttn 111
```

rnCGM4-int-1 product vs. rnCGM4 exons 2+3

Percent Similarity: 98.361

```
   4 GGCACGGAGTGTATGGAGGTAATGCACTGACCTCCACTAGCTAGGGCAGG 53
     || |||||||  ||||||||||||||||||||||:|||||||||||||||
3222 GGCGCGGAGTGCATGGAGGTAATGCACTGACCTCCACTAGCTAGGGCAGG 3271

  54 GAGACTCCCACACCTCAGCTGAGCACTGGACACAGCTGCTCAGGACTCAG 103
     |||||||||||||||||||||| ||||:|||||| |||| || |||| ||
3272 GAGACTCCCACACCTCAGCTGACCACTGGACACAGCTGCTC.GGACTCA. 3319

 104 GGCACCATCTTAGCCAAATACTAAAGTCCTGATGTTGACGGATCTCTCTT 153
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3320 GGCACCATCTTAGCCAAATACTAAAGTCCTGATGTTGACGGATCTCTCTT 3369

 154 CCCTTCTAGCCTCTCTTTTCATCTGTGGGCGTCCT 188
     |||||||||||||||| |||||||||||| ||||
3370 CCCTTCTAGCCTCTCTTTCATCTGTGGGCGTCCT 3404
```

Fig. 6.3 Diagram showing lineups of the sequenced PCR products from the superpools against gene sequences from the Genbank database. Primers used were rnCGM1-rep1, rnCGM2-1, rnCGM3-rep2 and rnCGM4-int-1. The PCR products generated using these primers on the superpools were lined up against sequences representing the rat CEA genes 1, 2, 3 and 4 respectively.

62 using the primers rnCGM3-rep2 showed only ~64% homology to the rnCGM3 gene sequence. This value appeared too low when taking sequencing procedure errors into account. It was therefore decided not to pursue the search of superpool 62 using these primers. No other superpools were identified using the primers rnCGM3-rep2 primers

Isolation and analysis of a single P1 clone containing rnCGM1

(i) *Isolation of the single P1 clone*

All stages of isolation were performed in the presence of 5% sucrose to select against or limit any loss of insert. Identified as containing rnCGM1, superpool 23 which contains 474 clones, was plated out in serial dilution's. Starting from plates containing the least amount of clones upwards, batches of 10 clones were pooled and screened by PCR using the primers rnCGM1-rep1. To check if this method created a suitable environment for PCR operations, colonies were simultaneously screened by PCR using a set of primers to amplify part of the kanamycin resistance gene which is present on all P1 clones. It has been reported that wooden toothpicks used to transfer colonies inhibit PCR when using less than 3 units of *Taq* polymerase (Lee and Cooper, 1995). To check if this may affect the experiments, a PCR was set up to amplify serial dilution's of a P1 colony using the kanamycin primers. Figure 6.4 shows an electropherogram of the PCR products obtained using 0.2 units of *Taq* polymerase. Colonies were either transferred with wooden toothpicks or plastic pipette tips. Products were obtained using both means of transfer in all of the dilution series. These results conflicted with those of Lee and Cooper, (1995), who may have used different type of toothpick or different quality of polymerase enzyme. A batch of 10 colonies and thence a single colony (called P1 rnCGM1) was isolated from superpool 23 using the rnCGM1-rep1 primers. Figure 6.5 shows the electropherogram of the multiplexed PCR products obtained. . Due to the success of this colony PCR method, kanamycin primers were no longer used for the rest of this study.

(ii) *Searching for other rat CEA gene sequences on P1rnCGM1*

P1rnCGM1 DNA was screened by PCR for the presence of other rat CEA genes, rnCGM2 and rnCGM4. Figure 6.6 shows the results of this experiment. Primers for rnCGM2 and rnCGM4 failed to amplify DNA from

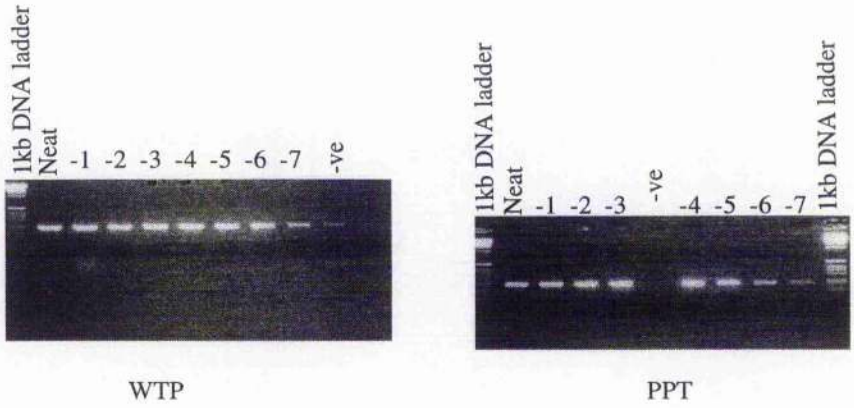


Fig. 6.4 Electropherograms gels (2.5% agarose, 1xTBE) showing PCR on serial dilutions (orders of magnitude shown) of a P1 colony using primer to amplify part of the kanamycin resistance gene. The primers were, forward (5' CGACTCGTCCAACAT CAATACAAC 3') reverse (5' CGCAATCACGAATGAATAACGG 3'). PCR conditions were 25 cycles of 94°C for 15'', 55°C for 30'' and 72°C for 30''. PCR was carried out using 0.2 units. WTP = wooden toothpick, PPT = plastic pipette tip.

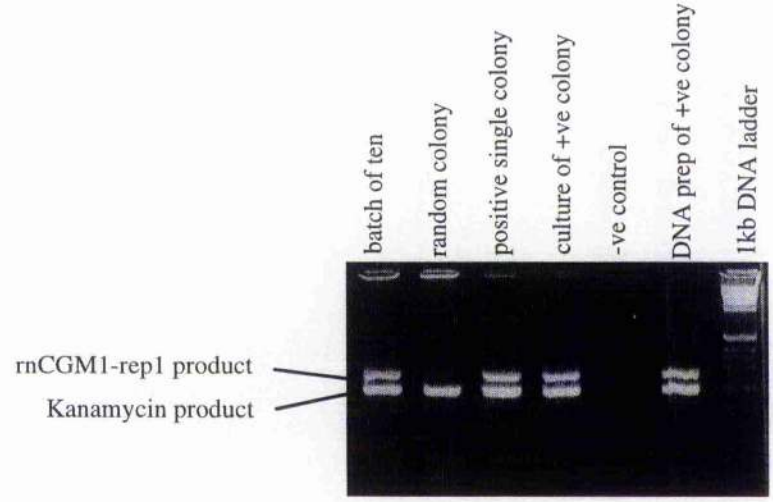


Fig 6.5 Electropherogram of gel (2.5% agarose, 1xTBE) showing identification of a sinlge P1 colony representative of the rat CGM1 gene. PCR products obtained by simutaneous screening using primers for the kanamycin gene and rnCGM1-rep1. This is an example of a multiplexed PCR experiment.

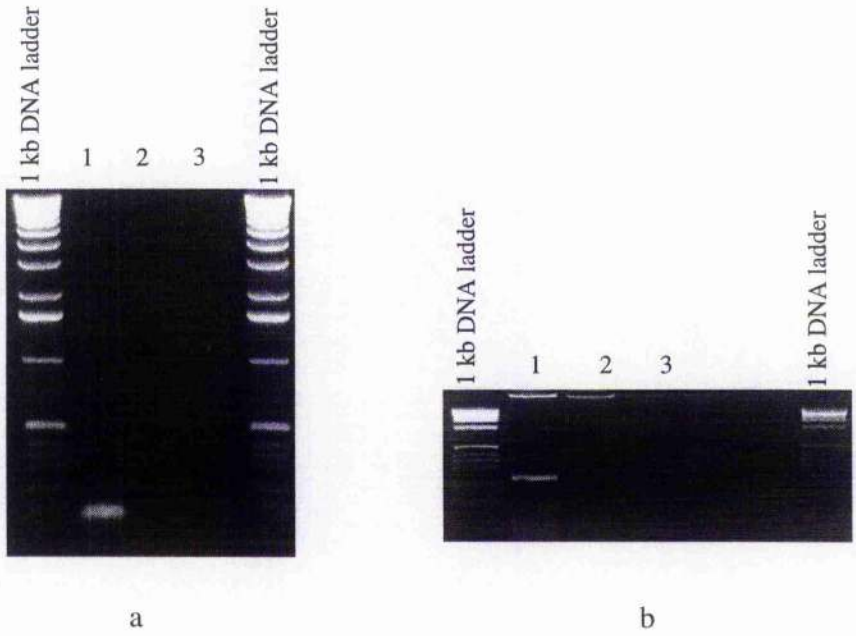


Fig. 6.6 Electropherograms of gels (2.5% agarose, 1xTBE) showing the search for rnCGM2 and rnCGM4 gene sequences on P1rnCGM1. a, rnCGM2-1 primers on; 1 = rat genomic DNA, 2 = nothing, 3 = P1rnCGM1. b, rnCGM4-int-1 primers on ; 1 = rat genomic DNA, 2 = nothing, 3 = P1rnCGM1.

P1rnCGM1. Therefore it was unlikely that these other rat CEA genes were present on this P1 clone.

(ii) *Preparation of P1rnCGM1 DNA and restriction analysis*

Plasmid DNA was isolated by a modified alkaline lysis method as described by Pierce and Sternberg (1993). To give an estimation of the size of the insert in P1rnCGM1, DNA was digested with the restriction enzymes *Bam*HI and *Bgl*II. Figure 6.7a shows the restriction patterns obtained. DNA was of poor quality due to the presence of bacterial DNA and there was an indication of a small insert due to the production of around only two bands in each digest. The 6kb *Bgl*II fragment present was derived from the P1 vector sequence. The enzymes used were six base pair cutters which cut every 4096 bases. Many more bands should have been produced if every P1 clone is assumed to have an average insert size of ~95kb. It was possible that this clone may have lost part of its insert.

(iii) *Influence of host strain on quality and quantity of P1 DNA*

To improve the quality and quantity of DNA, the clone was transferred out of the NS3529 Cre + host and into a Cre - host, DH10B. A two-step bacterial mating scheme was used which exploits γδ transposition events to transfer P1 genomic clones from NS3529 to DH10B (Kimmerly et al, 1994). In this new strain, the absence of the P1 recombinase, *Cre*, is thought to lessen the chances of unfavourable *lox* mediated recombinations. DH10B also carries a mutation in the endonuclease gene, *endA*. Strains that carry this mutant gene have been reported to yield better quality of DNA for sequencing purposes (Kimmerly et al, 1994). The vector has rare restriction sites (*Not*I, *Sfi*I and *Sal*I). If one of these sites is not present in the insert, then digestion with an enzyme that cleaves that site will linearize the plasmid. This would generate a restriction band indicating the size of the P1 clone. To enhance the resolution of the restriction patterns produced from the previous digestion and to give a better idea of the size of insert by digesting the clone with *Not*I, digest products were electrophoresed using field inversion gel electrophoresis (FIGE) (Carle et al, 1986). Figure 6.7b shows an electropherogram of the restriction patterns from the mated P1 clone using the enzymes *Bam*HI, *Bgl*II, *Eco*RI and *Not*I. The effect of the mating experiment had greatly improved the quality of the P1 DNA. In terms of quantity, there appeared not to be much difference in the amounts of DNA produced before and after the mating experiment. A possible explanation

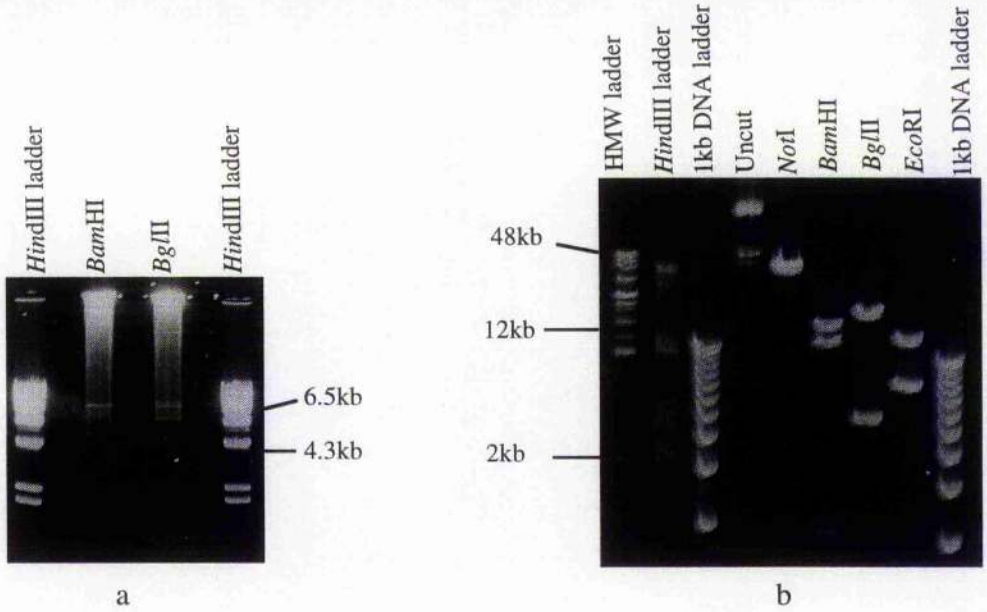


Fig. 6.7 a, Electropherogram of gel (0.8% agarose, 1xTBE) showing digests of P1rnCGM1 using *Bam*HI and *Bgl*II b, Electropherogram of gel (1% agarose, 0.5xTBE) showing FIGE of digests of P1rnCGM1, after transfer into DH10B, using *Not*I *Bam*HI, *Bgl*II and *Eco*RI. HMW = high molecular weight.

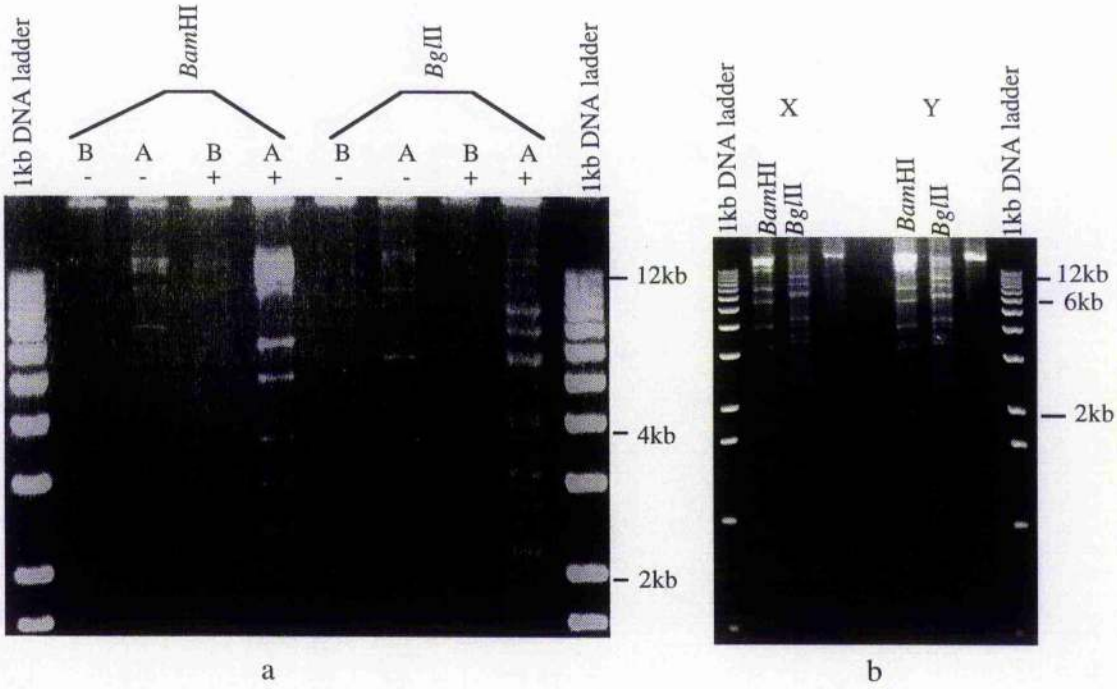


Fig. 6.8 a, Electropherogram of gel (0.8% agarose, 1xTBE) showing digests of P1rnCGM1 using *Bam*HI and *Bgl*II either before (B) or after (A) transfer into DH10B and either in the presence (+) or absence (-) of 5% sucrose during all stages of transfer and DNA preparation. b, Electropherogram of gel (0.8% agarose, 1xTBE) showing digests of P1rnCGM1 (X) and a second clone (Y) isolated from superpool 23 with the primers rnCGM1-rep1. Enzymes used were *Bam*HI and *Bgl*II.

for this is that once P1 plasmids have been transferred into DH10B they are no longer under *lac* control of the library host strain and thus the copy number cannot be elevated with the use of IPTG (Kimmerly et al, 1994). Electrophoresis using FIGE greatly enhanced the resolution of the restriction patterns. Only a few bands could be seen in the *Bam*HI, *Bgl*II and *Eco*RI digests as before. Totalling the bands in the *Bam*HI lane gave a value of around 44kb for the clone size. Restriction analysis with the *Not*I enzyme showed that the clone P1rnCGM1 had a total size of about 35kb. These two sizes probably differed due to an inaccuracy in electrophoresis. Taken together however, it could be deducted that this P1 clone had a size of around 39kb. According to the information supplied with the library, each recombinant clone contains approximately 17kb vector sequence once cyclized. Therefore the clone P1rnCGM1 isolated has an insert of approximately 22kb. With such a small insert size the chances of finding any repeat units would also be small. The most common repeat unit in mammalian genomes is AC which occurs approximately once every 30kb (Beckmann and Weber, 1991).

Since all P1 molecules must be of a minimum size (approx. 85kb) in order to be packaged by the headful mechanism, it could be concluded that there was a partial loss of insert of P1rnCGM1, probably during the mating and/or DNA isolation stages. In order to prevent this it was necessary to grow this clone in 5% sucrose during these and all other stages. In case these methods still gave rise to a small insert within the clone, another P1 clone containing the gene rnCGM1 was re-isolated from superpool 23 as a backup.

(iv) *Effect of sucrose on insert size of P1rnCGM1*

At all stages of the mating procedure and DNA preparation, 5% sucrose was added to select against no or minimal loss of insert. Figure 6.8a shows the results of restriction digests using the enzymes *Bam*HI and *Bgl*II on P1rnCGM1 before and after mating, with 5% sucrose present at all times and with 5% sucrose present only during the clone isolation procedure. In both cases where the clone was not mated the quality and quantity of DNA was so poor that it could not be visualised. As in fig.6.7a, P1rnCGM1 had been transferred to DH10B but had only been grown in 5% sucrose during the initial isolation step. Again only a few bands could be seen for each digest reaction on this occasion. When the mated clone had been grown in sucrose at all stages, many bands were produced from each digest, characteristic of a clone containing a large insert. Adding up the sizes of the bands in the *Bam*HI lane

gave a value of approximately 91.2kb. Adding up sizes of the bands in the *Bgl*II lane gave a value of approximately 89.9. These two similar values indicated that the size of the P1rnCGM1 insert was approximately 73kb.

This result showed that sucrose was selecting against the loss of insert. Due to this, all future work on P1 clones was done in the presence of 5% sucrose at all times. Another clone containing rnCGM1 had been isolated from superpool 23 and from the restriction patterns seen in fig. 6.8b this was identical to P1rnCGM1. Due to the size of the insert in this clone detected under these refined conditions it was decided to search this clone for repeat units.

(v) *Screening P1rnCGM1 for repeat units*

Digests and Southern blots of P1rnCGM1 were set up with positive controls for screening with (AC)10 and (CTG)10 repeat probes. Control for the AC hybridisation experiment was a subclone known to contain an AC repeat and was derived from a rat P1 clone containing the Grik5 gene (supplied by N.Craig). The control for the CTG hybridisation experiment was a cDNA clone containing 26 CTG repeats and was derived from a 12.5 day whole mouse embryo cDNA library (kindly supplied by T. Dunlop). Fig 6.9a shows restriction digests of P1rnCGM1 using the enzymes *Bam*HI, *Bgl*II and *Eco*RI. The resultant Southern blot of these digests hybridised with an (AC)10 probe is shown in fig.6.9b. The probe had hybridised to the control subclone as expected. From the digests of P1rnCGM1, three *Bam*HI fragments (1.8kb, 4kb and 6kb), two *Bgl*II fragments (3kb and 5kb) and two *Eco*RI fragments (6kb and 8kb) were all detected using this probe. Figure 6.10a shows restriction digests of P1rnCGM1 using the enzymes *Bam*III and *Bgl*II. These digests were Southern blotted with the control containing a CTG repeat and hybridised with a (CTG)10 probe shown in fig.6.10b. This probe had hybridised to the control as expected. A 1.8kb *Bam*HI fragment and a 5.8kb *Bgl*II fragment from P1rnCGM1 were detected using this probe. Of particular interest is that an AC repeat and a CTG repeat appeared to be on the same 1.8kb *Bam*HI fragment.

Two methods were employed to obtain sequence flanking these microsatellites. Firstly, a direct sequencing method was used to sequence outward from the CTG repeat on the P1 clone using a set of degenerate anchored primers which were available (see below). Secondly, P1rnCGM1 was shotgun subcloned,

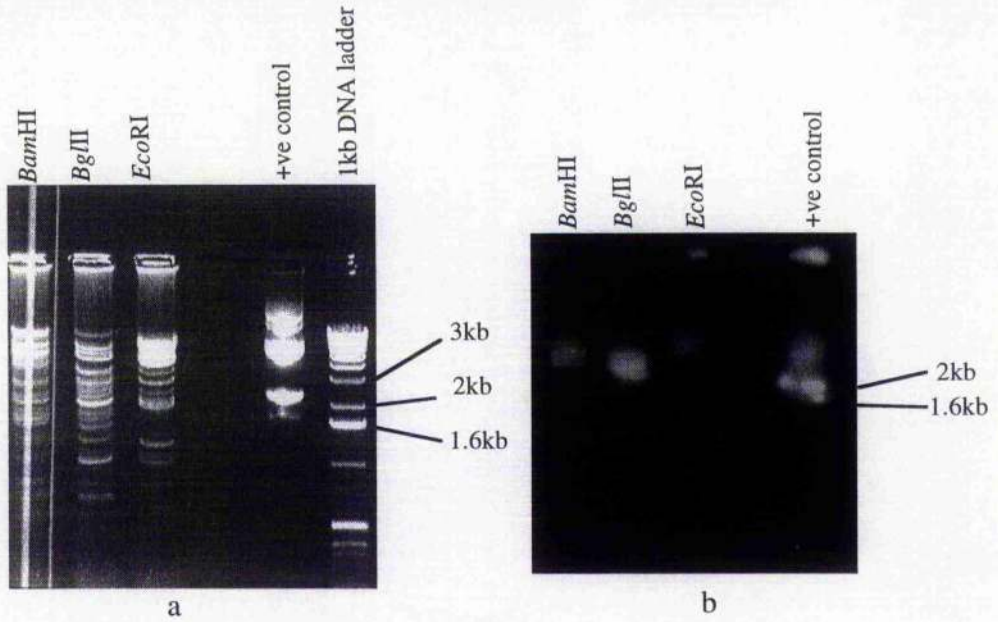


Fig. 6.9 a, Electropherogram of gel (0.8% agarose, 1xTBE) showing digests of P1rnCGM1 with *Bam*HI, *Bgl*II and *Eco*RI A positive control containing an AC repeat was also loaded (supplied by N. Craig). b, Southern blot of previous electropherogram hybridised with an $(AC)_{10}$ probe (specific activity = 2.1×10^7 cpm/µg DNA). Image here is inverted to help resolution.

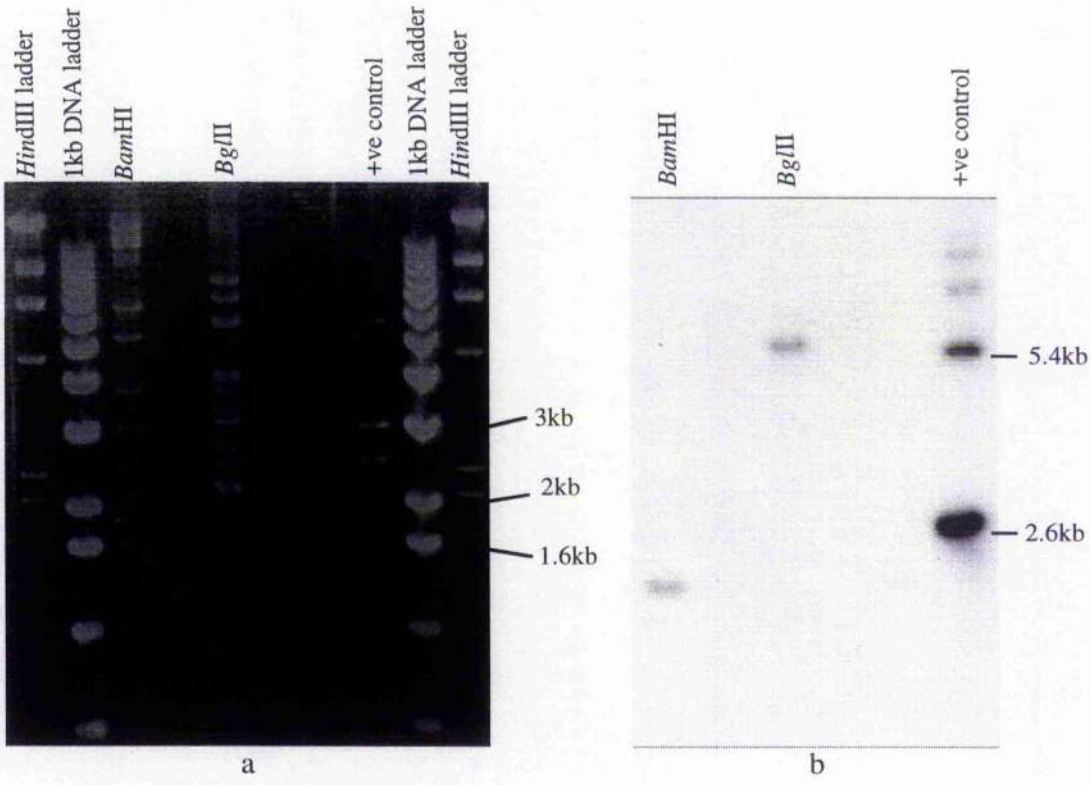


Fig. 6.10 a, Electropherogram of gel (0.8% agarose, 1xTBE) showing digests of P1rnCGM1 with *Bam*HI and *Bgl*II A positive control containing a CTG repeat was also loaded (supplied by T. Dunlop). b, Southern blot of previous electropherogram hybridised with a $(CTG)_{10}$ probe (specific activity.= 1.2×10^7 cpm/µg DNA).

using *Bam*HI, into pUC-18 in order to clone and therefore sequence the entire 1.8kb fragment that appeared to be containing an AC and a CTG repeats. If these methods failed in tracking down these two particular repeats, P1rnCGM1 could have been shotgun subcloned into pUC-18 using Sau3A in order to clone the other AC and CTG positive fragments.

(vi) *Direct sequencing from the CTG repeat*

A method was adopted to directly obtain sequence flanking the CTG repeat motif (Margolis et al, 1993).

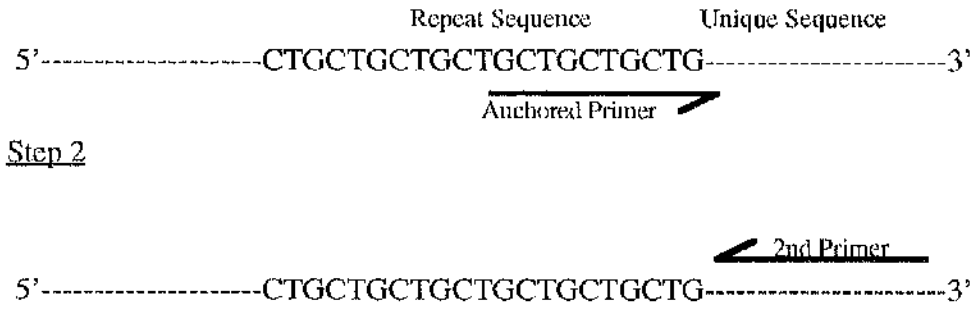


Fig. 6.11 Schematic diagram of method devised by Margolis et al, (1993) to rapidly obtain sequence flanking CTG repeats. The two stages of this procedure are described below.

Degenerate primers were used which were anchored on the interface between repeat sequence and unique sequence. This would enable sequencing outward from the repeat into the flank (step 1). A second primer could then be designed based on the complement of the unique sequence obtained and this would be used to sequence back through the repeat and into the other flank (step 2). Three sets of degenerate primers were used to encompass all possible combinations at the interface between a CTG repeat and unique sequence. These were: 1; 5'-(CTG)6CTX-3', 2; 5'-(TGC)6TGX-3' and 3; 5'-(GCT)6GCX-3', where X is any of three bases not extending the repeat (kindly supplied by T. Dunlop). Figure 6.12 shows that sequencing was unsuccessful using all the primers with P1rnCGM1 and the CTG positive control cDNA clone. This experiment was repeated a few times and still did not work. One possible explanation is that the degenerate primers may have lost their efficiency over

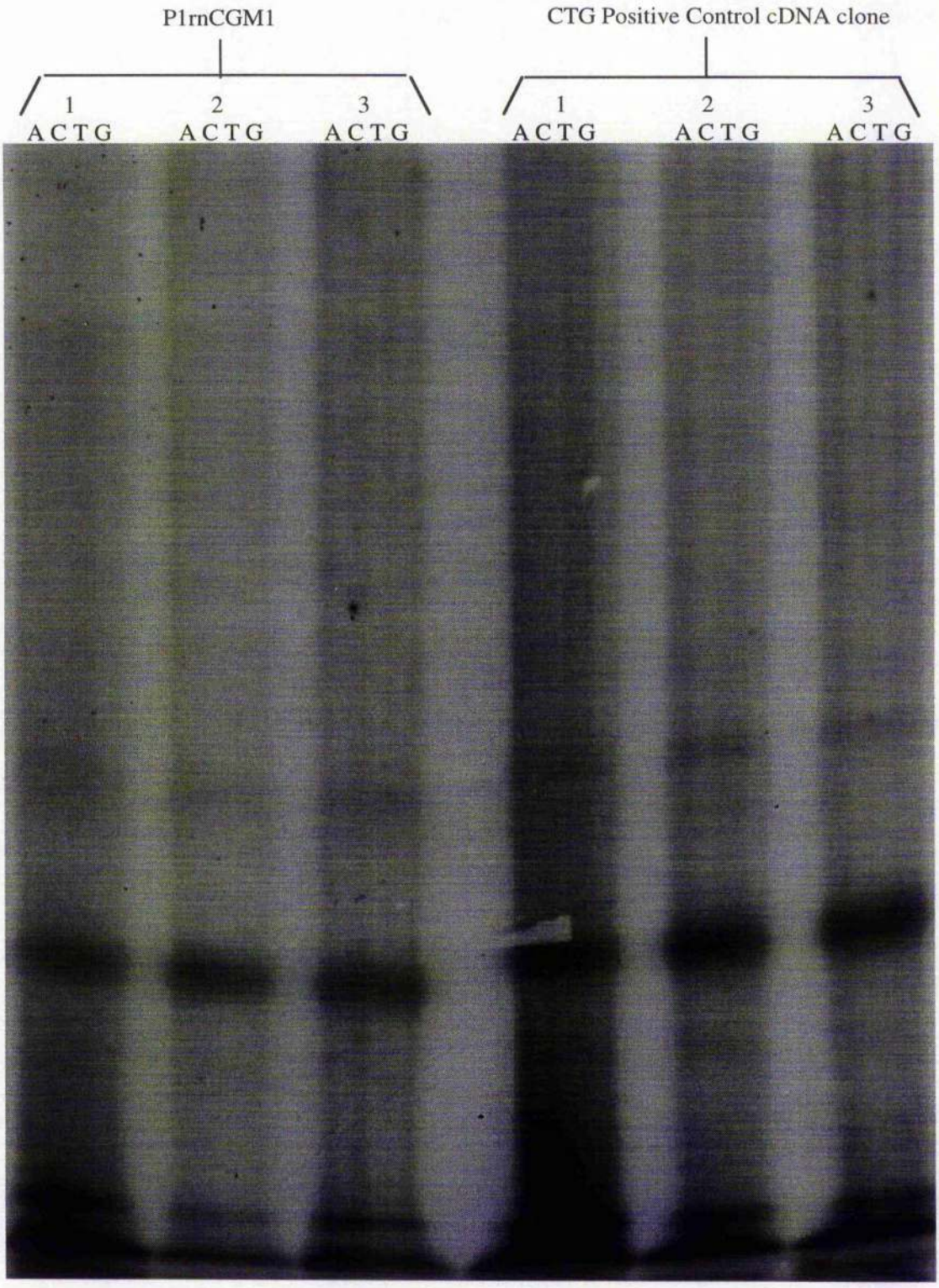


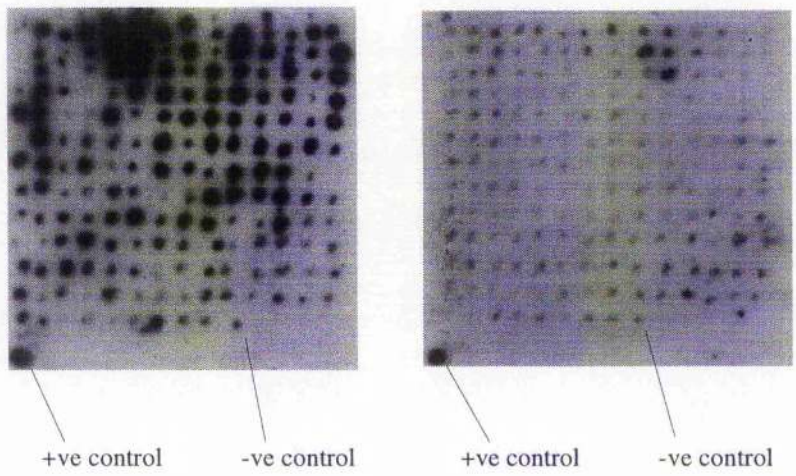
Fig. 6.12 Electropherogram of gel (6% acrylamide, 7M urea, 1xTBE) showing sequencing of P1rnCGM1 and a CTG positive control cDNA clone using the T7 sequencing kit (Pharmacia). Primers used were: 1, (CTG)6CTX; 2, (TGC)TGX; 3, (GCT)GCX, where X is any of three bases not extending the repeat. Primers were supplied by T. Dunlop.

time. It was therefore necessary to proceed with the second method of sequencing the repeat flanks as noted above.

(vii) *Shotgun subcloning of P1rnCGM1* Bam*HI fragments into pUC-18*

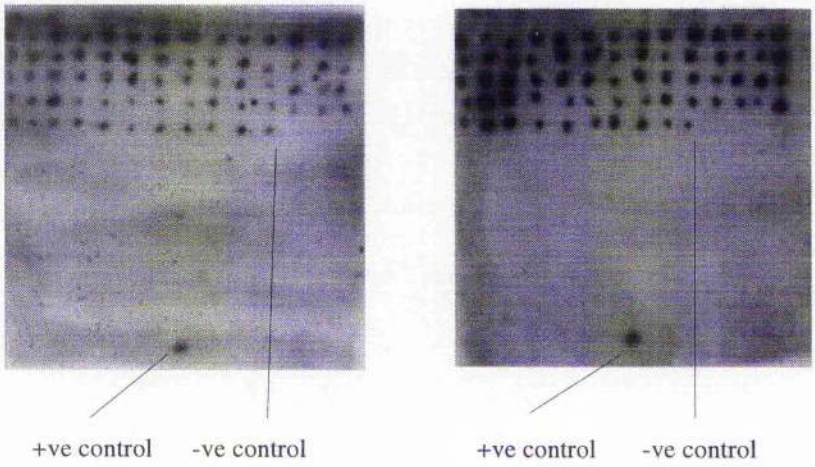
P1rnCGM1 was digested with *Bam*HI and subcloned into pUC-18. Two sets of these subclones and positive and negative controls were screened in duplicate by colony hybridisation, one set with an (AC)10 probe, the other with a (CTG)10 probe. The positive controls were those as described above. The negative controls were pUC-18 vector. Figure 6.13 shows the results of this screening procedure. Positive controls in each case were detected as expected with the probes. Negative controls were also detected faintly. Almost all of the subclones in both cases were detected at least faintly in duplicate. This non-specificity of probe binding was probably due to residual bacterial debris present on the nylon membranes. Not all signals had the same intensities between the duplicates in each case, again, probably due to non-specific binding of probe. Some of the subclones in each set were thought possibly to be true positives due to their similar intensities with positive controls and in some cases, similar intensities were duplicated. DNA from 22 possible AC positive subclones and 7 possible CTG positive subclones was prepared for Southern analysis. DNA from these subclones was digested with *Bam*HI and electrophoresed with positive and negative controls. Figure 6.14 shows results of these digests and the resultant Southern blot hybridised with an (AC)10 probe for subclones possibly containing an AC repeat. The positive control was detected as expected. The negative control was faintly detected probably due to the large amount of DNA present in this lane. None of the subclones were detected with the (AC)10 probe. Figure 6.15 shows the results of the 7 subclones possibly containing CTG repeats digested with *Bam*HI, Southern blotted and probed with a (CTG)10 probe. A signal was present for the positive control and no signal was present for the negative control as expected. A strong signal of equal intensity to the positive control signal was obtained for one of the subclones (No. 45). It appeared that this subclone, sc45P1rnCGM1, was not digested properly. Digestion with *Bam*HI was therefore repeated, the result of which is shown in fig. 6.15c. The size of insert was approximately 1.8kb which was in agreement with the size of fragment detected with the (CTG)10 probe in fig.6.10b. It was therefore appropriate to sequence the insert of this subclone.

P1rnCGM1 *Bam*HI subclones probed with (AC)10



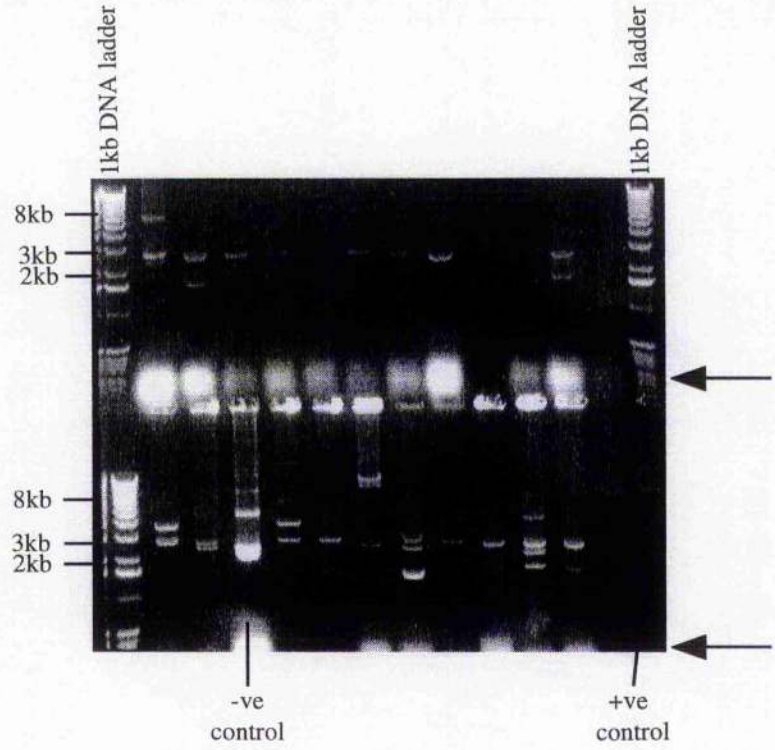
a

P1rnCGM1 *Bam*HI subclones probed with (CTG)10

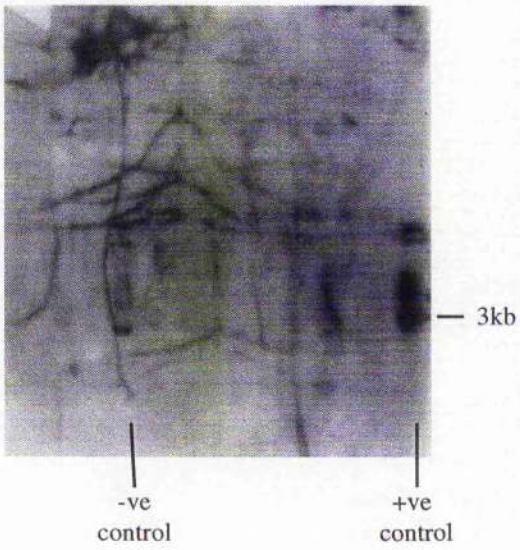


b

Fig. 6.13 a, Colony hybridisation of P1rnCGM1 *Bam*HI subclones with an (AC)10 probe (specific activity = 6.3×10^7 cpm/µg DNA). Positive (supplied by N. Craig) and negative (pUC-18) controls are shown. b, Colony hybridisation of P1rnCGM1 *Bam*HI subclones with an (CTG)10 probe (specific activity = 2.4×10^7 cpm/µg DNA). Positive (supplied by T. Dunlop) and negative (pUC-18) controls are shown.



a



b

Fig. 6.14 a, Electropherogram of gel (0.8% agarose, 1xTBE) showing *Bam*HI digests of 22 P1rnCGM1 subclones possibly containing AC repeats. Positive (supplied by N.Craig) and negative (pUC-18) controls were included as shown. RNA contamination shown by arrows. b, Southern blot of previous gel hybridised with an $(AC)_{10}$ probe (specific activity = 6.3×10^7 cpm/µg DNA). Positions of positive and negative controls are shown.

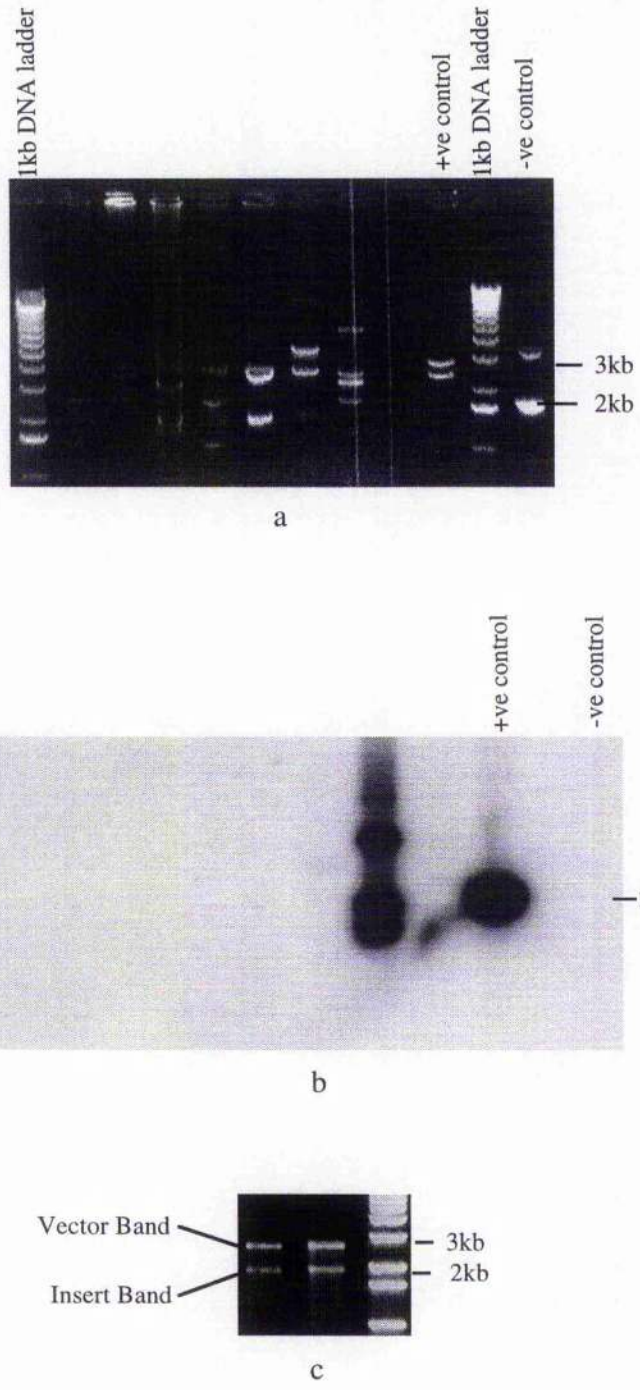


Fig. 6.15 a, Electropherogram of gel (0.8% agarose, 1xTBE) showing *Bam*HI digests of 7 P1rnCGM1 subclones possibly containing CTG repeats. Positive (supplied by T. Dunlop) and negative (pUC-18) controls were included as shown. b, Southern blot of previous gel hybridised with a $(CTG)_{10}$ probe (specific activity = 2.4×10^7 cpm/µg DNA). Positions of positive and negative controls are shown. c, Electropherogram of gel (0.8% agarose, 1xTBE) showing *Bam*HI digestion of the positive CTG subclone detected in b.

(viii) *Sequencing of a subclone containing a CTG repeat*

To initiate sequencing of the 1.8kb insert of sc45P1rnCGM1, universal forward and reverse sequencing primers which anneal either side of the *Bam*HI cloning site of pUC-18 were used in a cycle sequencing reaction. Approximately 350 bases of readable sequence was obtained from each end. Based on these sequences, two primers sc45F1 and sc45R1, were designed to sequence further into the insert. Again a further 350 bases of sequence was obtained from each end. In order to sequence the estimated remaining 400 bases of insert, a second set of primers (sc45F2 and sc45R2) were designed, based on the sequence obtained from the second round, to carry out this process. Figure 6.16 shows the entire composite sequence of the insert based on the forward direction. In total, 1965 bases were obtained which is close to the size estimated value of 1.8kb. Approximately 20 bases at each end of this sequence belonged to the cloning vector as the universal sequencing primer sites did not immediately flank the *Bam*HI cloning site.

A search of the Genbank database showed that 256 bases of this insert sequence (1662-1918) matched-up with 98% similarity to exon 3 of rnCGM1 of the rat. This was the best match obtained and is shown in fig.6.17. The few mismatches between these sequences were due to the presence of unreadable bases (N) which were produced when sequencing the insert of sc45P1rnCGM1. This region of the insert also matched -up to rnCGM1, exon 4 with 87% similarity and is also shown in fig.6.17. Conservation of IgV-like domain exons between and especially within CEA genes may account for the same stretch of sequence of sc45P1rnCGM1 to match-up with two different exons of rnCGM1. Other regions of the insert matched-up to a lesser degree with various CEA, BGP and PSG genes in rat, mouse and human. It is possible that this remaining stretch of insert sequence is complementary to an intron of rat CGM1 which has yet to be cloned or submitted to the Genbank database.

(ix) Analysis of two repeat units found on sc45P1rnCGM1

Two repeat regions were found in the insert of sc45P1rnCGM1. One contained an imperfect AC repeat interrupted by a region of dyad symmetry and the other was a CTGCTA hexanucleotide repeat followed by CTG and CTA triplet repeats. These were $(AC)_9$-ATGCAT-$(AC)_6$-AA-$(AC)_5$ and $(CTGCTA)_3$-$(CTG)_6$-$(CTA)_{12}$. None of these repeat units were the same as any of the

5′- NCNCTTGCCT GCCTGCAGGT CGACTCTAGA GGATCCTCAC ACACATCTCT
GTCTCCTTCC TCCTGAGAGC AAACTCCATC TTTGCAGGGC ACTGAGAACA
CAGGGCAGAC TGGGTGCCCA GCTGGTTCCT GGGTCACCCA NGGNAGTGCA
GAGTCCCCCT CCTTGTTGAT GCTGAGAAAG GCCAANGTAN AGCAGAGGGC
AACGTCTGGG TGAACAGCTG TTGTCTCAGG GAACAAGGCT GTTCAAAGGC
TTGTTGTTCA TAGTGTGTTC TGGAGGACTG AAGCATAAAG ANCCAAGGGG
GGAGGGTTTG TGGGAAGCAC TGANAATATA CAGGGCAGGA AACACAGCAT
TCTCTACCCC CCACACTTCA TGGGACTCCT GCCTGGGAGA TCGCTGAGCC
TANATCCGGT AAAACAAGAT AACTGGGGAT GTGCTGGACA TACTTTCTCA
NAGGGANGAT ACCTATGGTT TCATACACCA NTAAGGTACT CTGTACNCTC
TGANAAGGTT TTACCNCCTT CCANGGCCTA CTANTCATGG GTGANAANAC
AAACTNTACT CAGTGTGANA GGANAGCACT CACCTATTCT CTGTTTTCTC
CTCNGGGGCT CNATGACTTN AGAGGAGTCT CAGGGTTTTT AGGAAACAAC
ACTCNGTGGG ACNNAGGCTC TTGATGANNA AAGAGGGGAA GGAGGANATT
GTTCTCACCC ACAGTTCNTC ATCNTGAACC CNNTGTTCTG AANACTGAGG
TTCTCAGCTG TGATGCCCCA AATAAGTAAC CAAGCTGGTG TTGATCAGTT
GATGTCGNCA GNGGATATGG GTGTTAGCTT CATGTCTGAC CACTTGAGAA
CCAGTGAACA CAAGGTTGTT CCCTGACCTC CACCTAGGGA TGGTGTCCTG
CACCCACTAA CACACACACT GAGGCACAAC CTTGTAAAGA **AACACACACA**
CACACACATG CATACACACA CACACAAACA CACACACTGA GATAAATCCA
TGCTCTGAAA CTCACACACA AATATAATTC ATAAATAAGT AAAAATAAAG
AAAATGGAAG ATTCAATACT GTATCTCCAA ACTGATGAAT TTTGTAAATA
AATCTATGTA TTGGCAGTGT CCACGAACAT TTCTAAGCAT AACTACTGTT
ACTGCTACTG CTACTGCTAC TGCTGCTGCT GCTGCTGCTA CTACTACTAC
TACTACTACT ACTACTACTA CTACATGTAA GACATATCTG TGTTCTCATG
AGTCCAGCCA GTCCCTTGTG CCCTGAAAGA CAGATCTATA ATGAAAGCCA
GAAGATGAGG TCAGGTGTGG ACTAGAATCA CTAAAACAAC AGAAACTTAA
CAACAGAAAG TTGGAGAGTC AAGAAAGGGT GTGTAGGGGA AGAGTTGGGG
AGGCATCACC CCAACACCTC TNTAGTGTAT GCAGAGATCT GAGGACAGTG
AGGCGGGGCC ACATTGGCAC CCAAGAAAAG AGTACCATAG AGAGGAGATG
ACAAGTGAGG TAAAGCGCAT GGAGGTAAAG CACTGACCTC CACCAAGCTG
GGCAGGGAGA CTNTCATCCA CGCACCTCGC CTGGGTAATA GACACATCTT
TGGGTGCCAT CTTAGCCAAA TACAAAAGCC CTAATGTTGA TGGATCTCTG
TCTTCNTTCT ATCTCCACTT TGACCTGTGG ACGCGCTGCC ACCTNTGCTC
AGCTCAGTAT TGAATCAGTG CCGACCAGCA TCTNTAAAGG AGAAAGCGCT
NTTCTCCTTG CTCACAATCT CCCAGAGAAT CTCCGAGCCA TTTTCTGGTA
TAAAGGGGCG ATTGTGTTCA AGGACCTTGA GGTTGCTCGA TATGTAATAG
GCACAAATTC AAGTGTGCCG GGGCCTGCCC ACAACGGCAG AGAGACAATG
TACAGCAATG GATNCCCCGG GTACCGAGCT CGAATTCGTA ATCATGTCAT
AGCTGTTGCN CGCNG --3′

Fig. 6.16 Sequence of insert of sc45P1rnCGM1. Internal primers designed for continued sequencing are shown as underlined. Sequences in bold are two repeat units which are summarised as follows: $(AC)_9$-ATGCAT-$(AC)_6$-AA-$(AC)_5$ and $(CTGCTA)_3$-$(CTG)_6$-$(CTA)_{12}$.

(i)
Query: Insert of sc45P1rnCGM1.
Sbjct: Rat carcinoembryonic antigen-related protein (CGM1) gene, exon 3, clone lambda-rnCGM1-1.

Percentage Similarity = 98%

```
Query:  1662 TCTCCACTTTGACCTGTGGACGCGCTGCCACCTNTGCTCAGCTCAGTATTGAATCAGTGC 1721
             ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct:    61 TCTCCACTTTGACCTGTGGACGCGCTGCCACCTCTGCTCAGCTCAGTATTGAATCAGTGC 120

Query:  1722 CGACCAGCATCTNTAAAGGAGAAAGCGCTNTTCTCCTTGCTCACAATCTCCCAGAGAATC 1781
             |||||||||||| ||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct:   121 CGACCAGCATCTCTAAAGGAGAAAGCGCTCTTCTCCTTGCTCACAATCTCCCAGAGAATC 180

Query:  1782 TCCGAGCCATTTTCTGGTATAAAGGGGCGATTGTGTTCAAGGACCTTGAGGTTGCTCGAT 1841
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:   181 TCCGAGCCATTTTCTGGTATAAAGGGGCGATTGTGTTCAAGGACCTTGAGGTTGCTCGAT 240

Query:  1842 ATGTAATAGGCACAAATTCAAGTGTGCCGGGGCCTGCCCACAACGGCAGAGAGACAATGT 1901
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct:   241 ATGTAATAGGCACAAATTCAAGTGTGCCGGGGCCTGCCCACAACGGCAGAGAGACAATGT 300

Query:  1902 ACAGCAATGGATNCCCC 1918
             ||||||||||||  || |
Sbjct:   301 ACAGCAATGGATCCCTC 317
```

(ii)
Query: Insert of sc45P1rnCGM1.
Sbjct: Rat carcinoembryonic antigen-related protein (CGM1) gene, exon 4, clone lambda-rnCGM1-1.

Percentage Similarity = 87%

```
Query:  1669 TTTGACCTGTGGACGCGCTGCCACCTNTGCTCAGCTCAGTATTGAATCAGTGCCGACCAG 1728
             | ||| || ||   || || |||| | ||| ||||||| | | |||||||  || |||||
Sbjct:    97 TATGAGCTATGCTGGCCCTCCCACTTCTGCCCAGCTCACTGTCGAATCAGGCCCTACCAG 156

Query:  1729 CATCTNTAAACGAGAAAGCGCTNTTCTCCTTGCTCACAATCTCCCAGAGAATCTCCGAGC 1788
             | |   | |||||| ||||| | |||||||||||||| |||||| ||||||||||:|||
Sbjct:   157 CGTTGCTGAAGGAGCAAGCGTTCTTCTCCTTGCTCATAATCTCCCTGAGAATCTCCGAGC 216

Query:  1789 CATTTTCTGGTATAAAGGGGCGATTGTGTTCAAGGACCTTGAGGTTGCTCGATATGTAAT 1848
             ||||||||:|||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct:   217 CATTTTCTGGTATAAAGGGGCGATTTTGTTCAAGGACCTTGAGGTTGCTCGATATGTAAT 276

Query:  1849 AGGCACAAATTCAAGTGTGCCGGGGCCTGCCCACAACGGCAGAGAGACAATGTACAGCAA 1908
             ||||||||||||||||||.|||||||||||||| |||||||||||||| |||||||||||
Sbjct:   277 AGGCACAAATTCAAGTGTGCCGGGGCCTGCCCACAGCGGCAGAGAGACAATGCACAGCAA 336

Query:  1909 TGGATNCCCC 1918
             ||||| || |
Sbjct:   337 TGGATCCCTC 346
```

Fig. 6.17 Alingments of the insert of sc45P1rnCGM1 to its best two matches from the Genbank database.

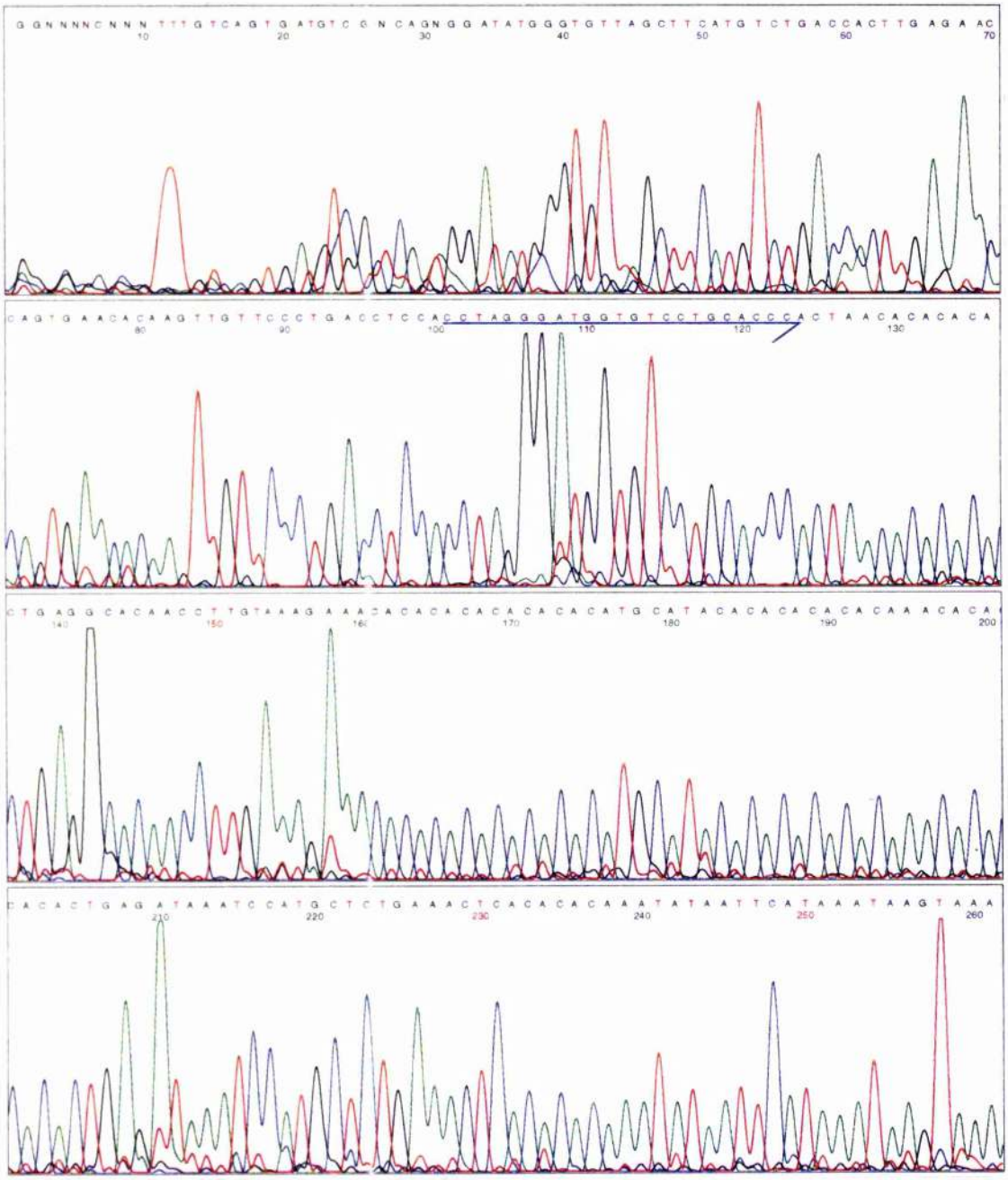


Fig. 6.18 Chromatogram showing partial sequence of the insert of sc45PrnCGM1 The AC and CTG/CTA repeat regions can clearly be seen. Sequences underlined in blue indicate the primers sc45-ac. Sequences underlined in black indicate the primers sc45-ctg.

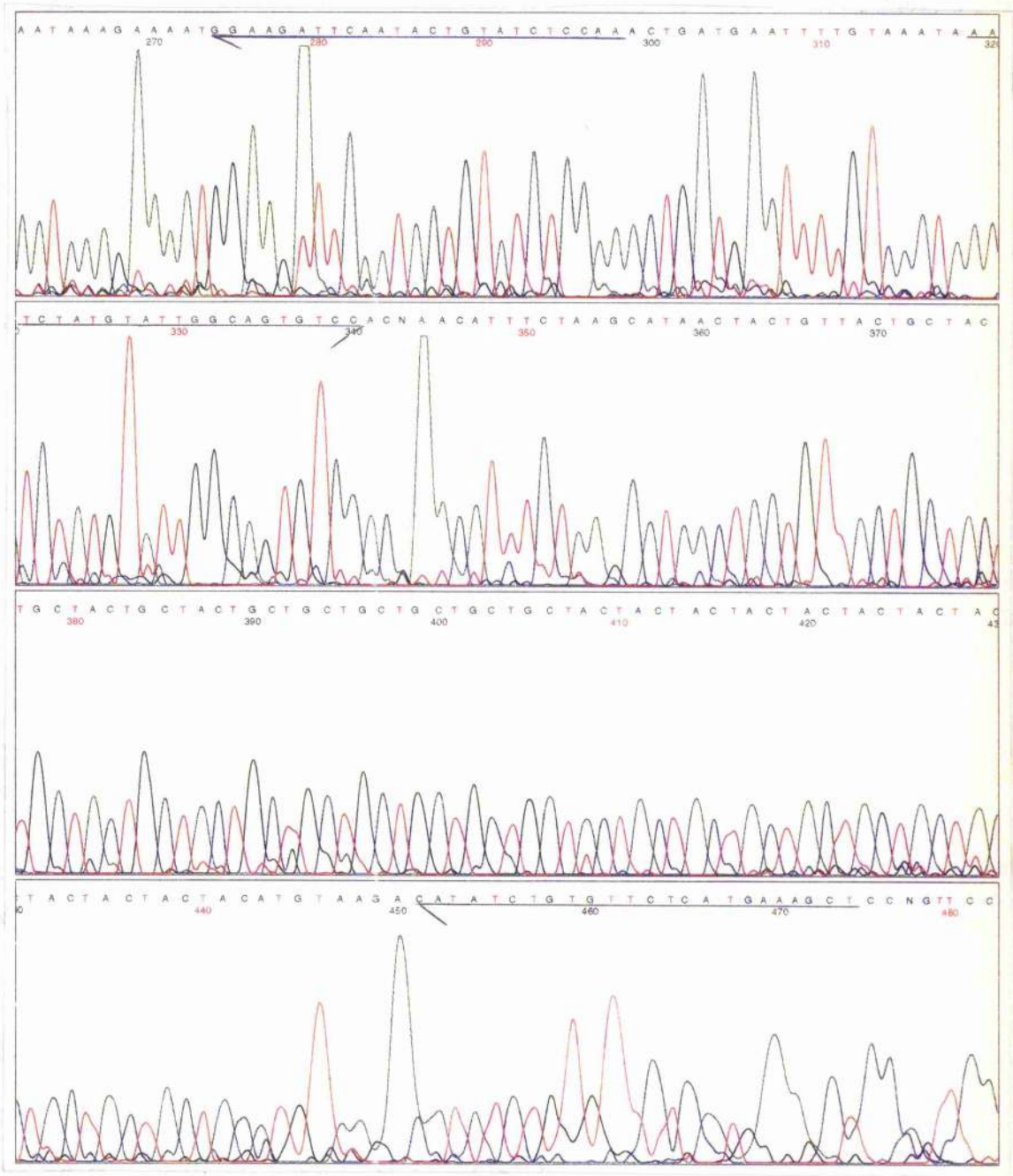


Fig. 6.18 continued

known microsatellites associated with rat CEA genes. Primers either side of these repeat units were designed, constructed and used to analyse possible allelic size and/or sequences differences amongst the rat strains BN, AS/AGU and F344. The primer sets were called sc45-ac and sc45-ctg. Figure 6.18 shows chromatograms of the repeat units and the flanking annealing sequences of these primers. Figures 6.19 and 6.20 show the results of the PCR experiment using the primers sc45-ac and sc45-ctg on the above mentioned rat strains. Two sets of bands were produced using sc45-ac primers, one set was of the expected size (196 bp) and the other set was possibly background. The products obtained using sc45-ctg primers were of the expected size (157 bp). MetaPhor agarose electrophoresis indicated no size difference between the rat strains using either set of primers (fig.6.19). Electrophoresis on 8% polyacrylamide also showed no evidence of size difference (fig. 6.20).

In a recent paper by Williams et al (1997), the use of SSCP analysis on uninformative microsatellites succeeded in detecting sequence variations within these repeat units between reference families of cattle. It was therefore decided to use the primers sc45-ac and sc45-ctg for SSCP analysis between the three rat strains. Figure 6.21 shows the result of this experiment. No identifiable variations could be seen

Isolation and analysis of a single P1 clone containing rnCGM2

Despite screening 3,200 colonies from superpool 17 by PCR with the primers, rnCGM2-1, no single colonies were identified. Rechecking that rnCGM2 was present on superpool 17 gave the same result as in fig.6.2. This superpool contained 456 clones and therefore screening 3,200 should have identified a single clone containing this gene seven times over. Screening the clones of this superpool by colony hybridisation with either the forward or reverse primers of rnCGM2-1 may isolate a single clone containing rnCGM2.

Isolation and analysis of a single P1 clone containing rnCGM4

(I) Isolation of the single P1 clone

Screening superpool 46 by PCR with the primers rnCGM4-int-1, a batch of ten clones and thence a single clone was isolated, shown in figure 6.22. This

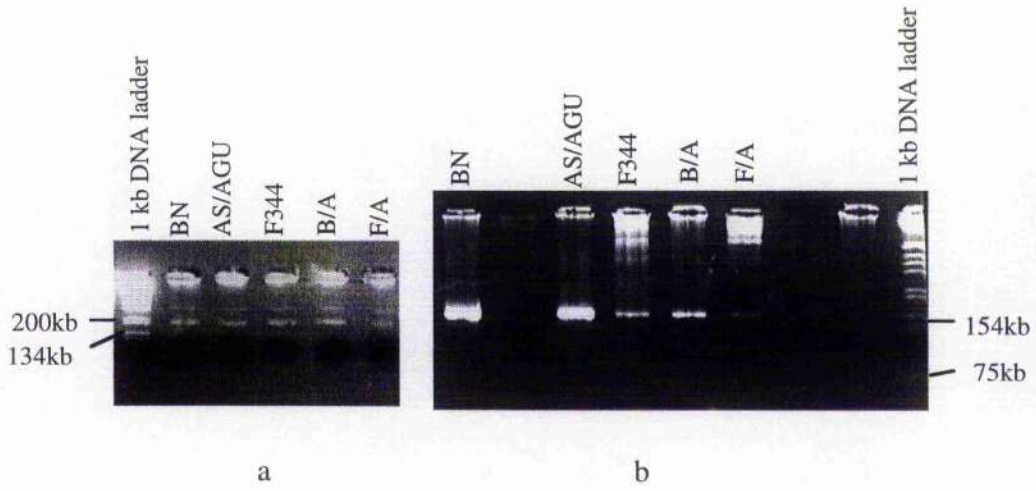


Fig. 6.19 Electropherograms of gels (4% MetaPhor, 1xTBE) showing the PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) using the following primer sets. a, sc45-ac, forward (5' CCTAG GGATGGTGTCCTGCACCC 3') reverse (5' TTGGAGATACAGTATTGAATCTTCC 3') PCR conditions were 25 cycles of 94°C for 15'', 54°C for 30'' and 72°C for 30''. b, sc45-ctg, forward (5' AATCTATGTATTGGCAGTGTCC 3') reverse (5' AGCTTTCATGAGAACAGATATG 3') PCR conditions were 25 cycles of 94°C for 15'', 55°C for 30'' and 72°C for 30''.

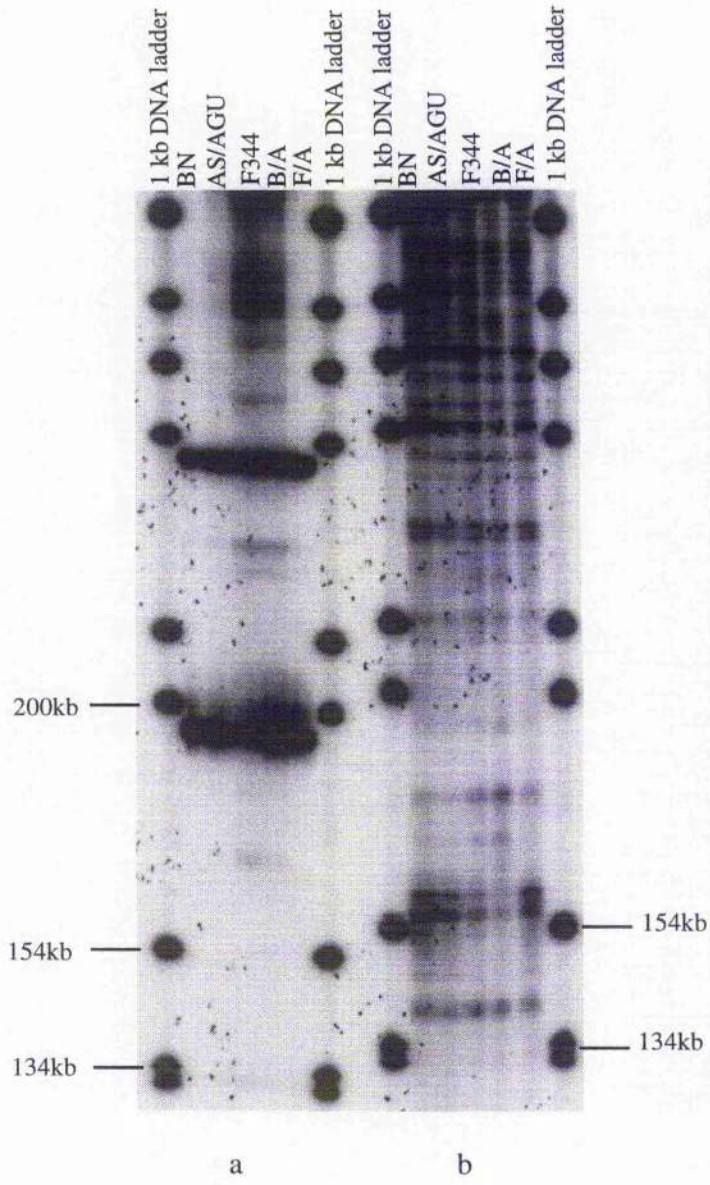


Fig. 6.20 Electropherograms of gels (8% acrylamide, 7M urea, 1xTBE) showing the PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) using the following primer sets. a, sc45-ac. b, sc45-ctg. Conditions for both sets of primers described in previous diagram.

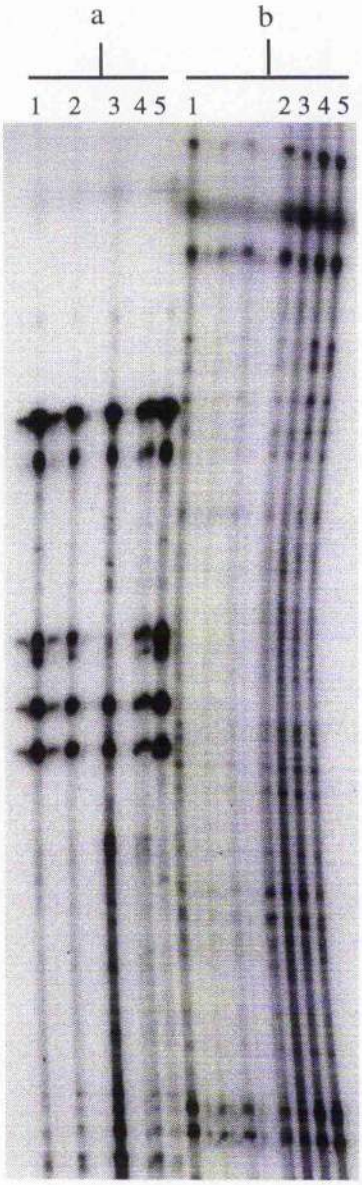


Fig. 6.21 Electropherograms of gels (6% acrylamide, 0.5xTBE) showing SSCP analysis of the PCR products from the three rat DNA templates BN, AS/AGU and F344 and F1 DNA templates BN/ASAGU (B/A) and F344/AS/AGU (F/A) using the following primer sets. a, sc45-ac. b, sc45-ctg.

clone, termed P1rnCGM4, was transferred out of the library host strain and into DH10B cells in the same manner as for P1rnCGM1, with 5% sucrose present at all stages. DNA was isolated from this clone and was digested with *Bam*HI and *Bgl*II to determine the size of insert. Figure 6.23 shows the result of these retriction digests on P1rnCGM4. Totalling the sizes of the fragments produced in the *Bam*HI and *Bgl*II lanes gave values of approximately 108 and 110 respectively. These two similar values indicated that this clone had an insert size of approximately 92kb since the vector sequence represented 17kb of this value.

(ii) *Searching for other rat CEA gene sequences on P1rnCGM4*

P1rnCGM4 DNA was screened by PCR for the presence of other rat CEA genes, rnCGM1 and rnCGM2. Figure 6.24 shows the results of this experiment. Primers for rnCGM1 failed to amplify DNA from P1rnCGM1. Therefore it was unlikely that this other rat CEA gene were present on this P1 clone. Primers for rnCGM2 did however indicate the presence of the rnCGM2 gene on P1rnCGM4. Southern analysis experiments to confirm the presence of rnCGM2 on P1rnCGM4 were unsuccesful probably due to inefficiency in probe construction. It would be desirable to repeat this experiment and possibly confirm the physical linkage of these two rat CEA genes as this would aid in the genetic and physical mapping of all of these family members.

(iii) *Screening P1rnCGM4 for repeat units*

Digests and Southern blots of P1rnCGM4 were set up with positive controls for screening with an (AC)10 repeat probe. Control for this experiment was a subclone known to contain an AC repeat and was derived from a rat P1 clone containing the Grik5 gene (supplied by N.Craig) as used for the analysis of P1rnCGM1. Fig 6.25a shows restriction digests of P1rnCGM4 using the enzymes *Bam*HI and *Bgl*II . The resultant Southern blot of these digests hybridised with an (AC)10 probe is shown in fig.6.25b. The probe had hybridised to the control subclone as expected. From the digests of P1rnCGM1, three *Bam*HI fragments (14kb, 9kb and 7kb), and six *Bgl*II fragments (2.8kb, 3kb, 4kb, 7kb, 9kb and 16kb) were all detected using this probe. This would indicate the presence of at least six AC repeat units on this clone. In order to obtain sequence of these repeats and their flanks, it was necessary to shotgun subclone P1rnCGM4 into pUC-18 using the four base-pair cutting enzyme *Sau*3A. This method would generate subclones containing

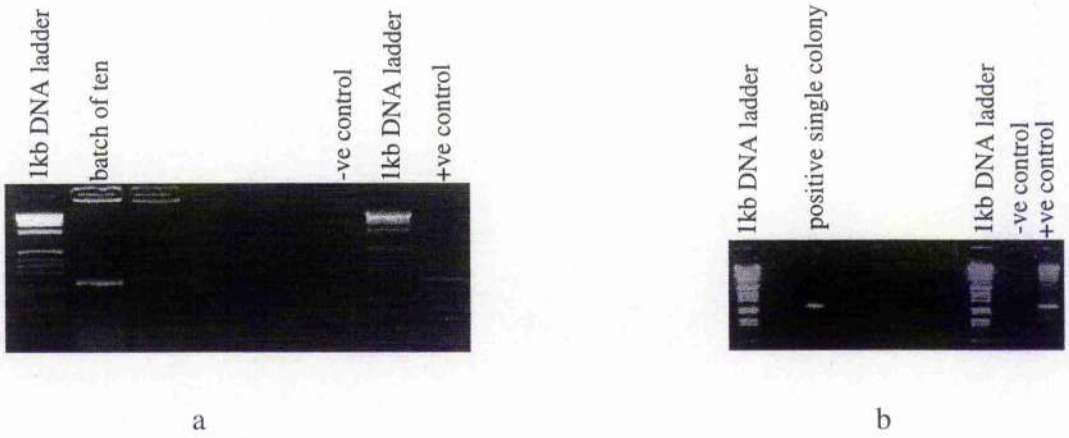


Fig. 6.22 Electropherograms of gels (2.5% agarose, 1xTBE) showing PCR products obtained using the primers rnCGM4-int-1 on, a, batches of 10 P1 clones from superpool 46 and b, single clones from superpool 46. The positive control template in each case was superpool 46.

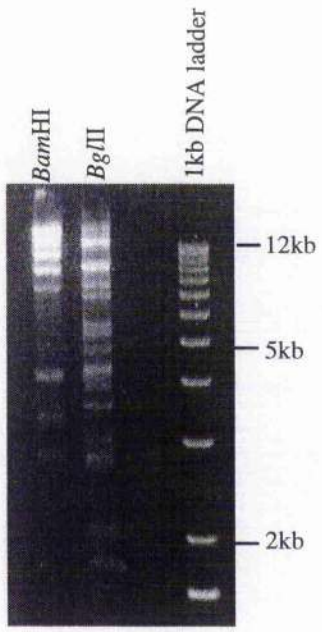


Fig. 6.23 Electropherogram of gel (0.8% agarose, 1xTBE) showing digests of P1rnCGM4 with *Bam*HI and *Bgl*II.

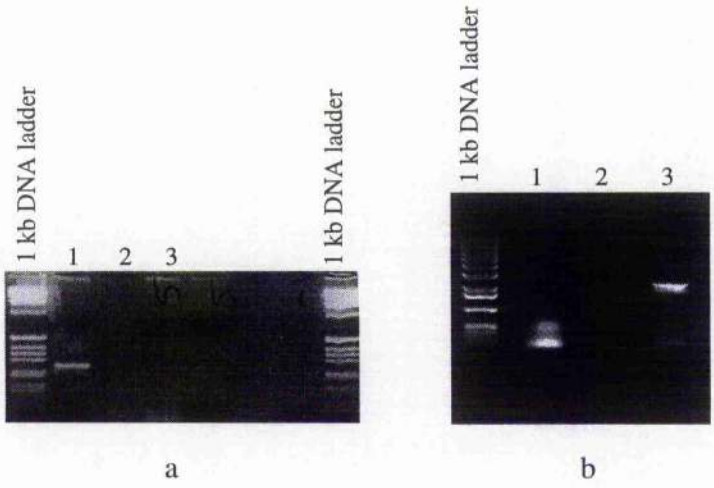


Fig. 6.24 Electropherograms of gels (2.5% agarose, 1xTBE) showing the search for rnCGM1 and rnCGM2 gene sequences on P1rnCGM4. a, rnCGM1-rep1 primers on; 1 = rat genomic DNA, 2 = nothing, 3 = P1rnCGM4. b, rnCGM2-1 primers on ; 1 = rat genomic DNA, 2 = nothing, 3 = P1rnCGM4.

inserts of around 256 bp. Inserts of any of these subclones found to contain one of the AC repeat units could rapidly and easily be sequenced.

(iv) *Shotgun subcloning of P1rnCGM4 Sau3A fragments into pUC-18*

P1rnCGM4 was digested with *Sau*3A and subcloned into pUC-18. Subclones produced were screened in duplicate by colony hybridisation with an (AC)10 probe. A positive control (described above) was present. The negative control was pUC-18 vector. Figure 6.26 shows the results of this screening procedure. The positive control was detected as expected with the probe. The negative control was not detected. None of the 150 subclones arrayed in duplicate gave a signal. P1rnCGM4 contained at least six AC repeat units which is equivalent to one repeat unit every 15.3kb. Screening 150 *Sau*3A subclones would have been the equivalent of screening 38.4kb of sequence and it was therefore surprising not to find AC containing subclones in this sample. This result was not statistically impossible though.

Attempts were made to screen more P1rnCGM4 subclones with the (AC)10 probe but were unsuccessful due to colony hybridisation problems. It would be desirable to repeat these experiments in order to isolate the AC repeat units present on P1rnCGM4.

Final Discussion

Two new microsatellites associated with rnCEA1 were found. As with the previously analysed microsatellites in this gene family, these were also noninfromative for the backcrosses in use. It would be desired to isolate and analyse the other AC repeat regions shown to be contained in the P1 clones, P1rnCGM1 and P1rnCGM4.

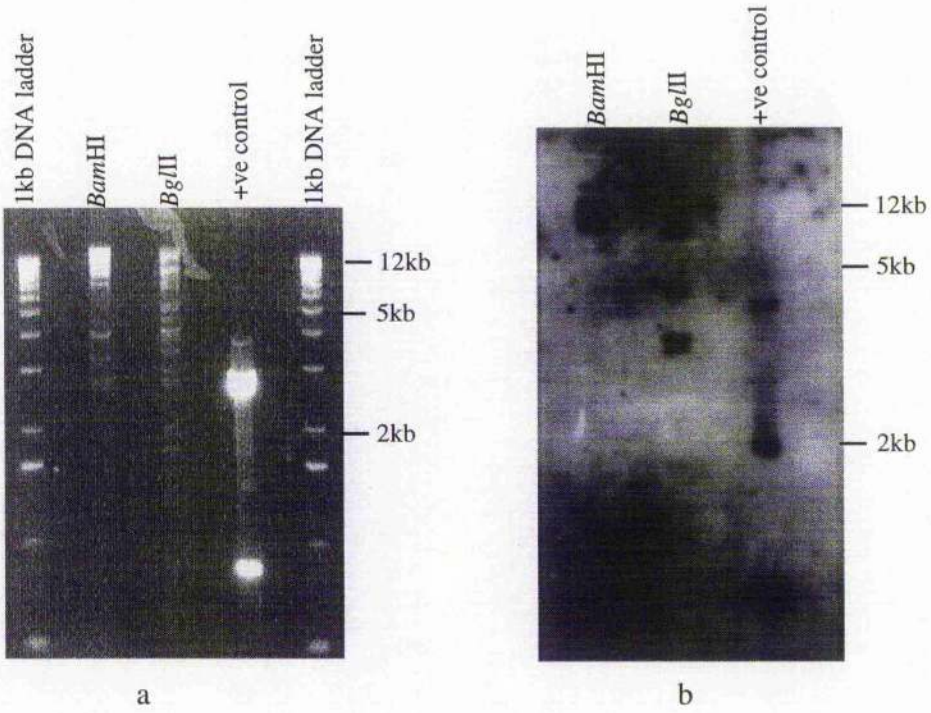


Fig. 6.25 a, Electropherogram of gel (0.8% agarose, 1xTBE) showing digests of P1rnCGM4 with *Bam*HI and *Bgl*II A positive control containing an AC repeat was also loaded (supplied by N. Craig). b, Southern blot of previous electropherogram hybridised with an $(AC)_{10}$ probe (specific activity.= 3.1×10^7 cpm/µg DNA).

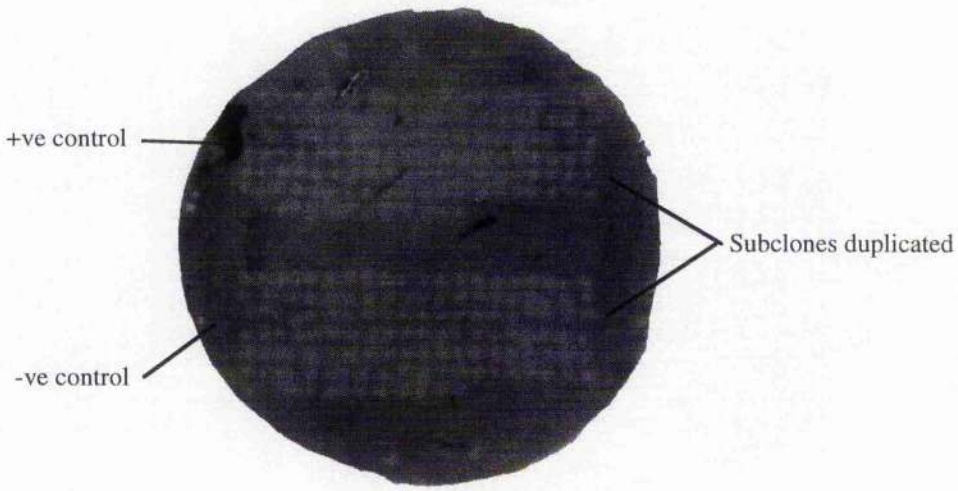


Fig. 6.26 a, Colony hybridisation of P1rnCGM4 *Sau*3A subclones with an $(AC)_{10}$ probe (specific activity = 1×10^7 cpm/µg DNA). Positive (supplied by N. Craig) and negative (pUC-18) controls are shown.

Chapter 7

Concluding remarks and Future Prospects

Discussion

During the course of this study it has been confirmed that CEA family members lie within the syntenic region of human chromosome 19 and mouse chromosome 7 and is shown in fig. 7.1 (Stubbs et al, 1996). It can be seen from this diagram that there appears to be a transposition of an 800kb region containing the PSG sub family of CEA in the human, although other CEA members are found in very similar positions between the two species. Another member of the rat CEA family, rnCGM1 has been mapped to the central part of chromosome 1 close to the marker R191 (D1Cep4) (Ding et al, 1996). This gene was shown to be clustered to the two previously mapped rat CEA genes, rnCGM3 and rnCGM4 (Ding et al, 1996). Results from these studies indicate that the fourth known rat CEA member, rnCGM2, may possibly be physically linked to rnCGM4 due to the identification by PCR of sequences representing these two genes present on the same P1 clone. The clustering of these genes in rat bears very similar resemblance to that seen in humans. In mouse however CEA/PSG family members appear to be more scattered throughout the syntenic region in chromosome 7 (Stubbs et al, 1996).

To achieve the mapping of CEA members to *agu*, two approaches could be taken. This first would be isolate and analyse other AC repeat regions shown to be contained in the P1 clones, P1rnCGM1 and P1rnCGM4. It would then be hoped that at least one of these repeats would prove informative. The second would be to set up a backcross using the reference strain DA and use microsatellites already isolated. It was shown by Ding et al, (1996) that the markers *D1Arb32*, *D1Arb33*, *R100* and *D1Arb36* generated size differences between BN and F344 and DA (BN=F344<DA). These markers, referring to the genes rnCGM1, rnCGM1, rnCGM4 and rnCGM3, respectively, would therefore prove informative should a cross between AS/AGU and DA be set up. It has since been shown that the primers developed in these studies, rnCGM1-rep1, which flank the marker *D1Arb32*, shows a size variation between DA and AS/AGU.

Although BN is the most divergent strain from AS (Canzian, 1997), there appears to high conservation of CEA genes between the two. Other microsatellite loci in the surrounding area of *agu* have also proved to be non-informative (M. Duran Alonso, personal communication). Ideally, a backcross using another species of rat as a reference strain would possibly generate a

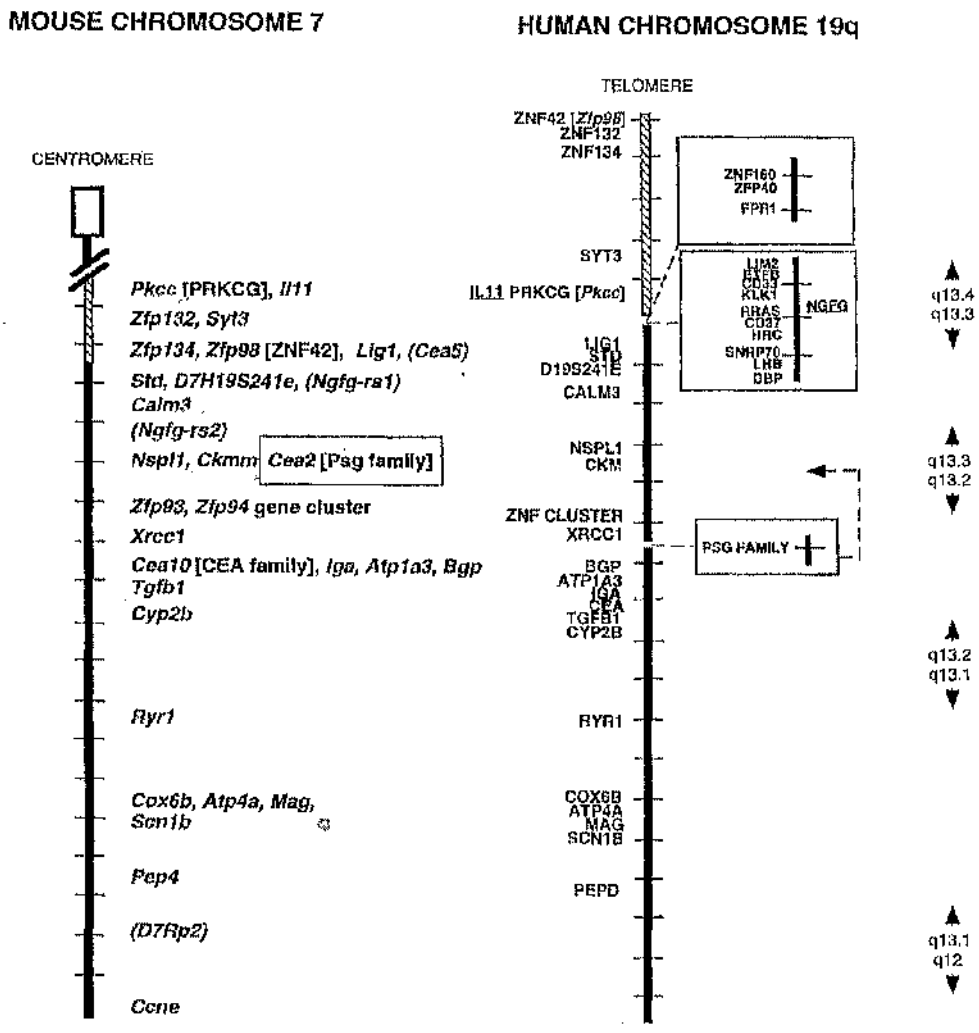


Fig 7.1 Diagram showing the syntenic relationship between between mouse chromosome 7 and human chromosome 19. The maps have been positioned to allow maximun alignement of the related human and mouse loci. Mouse loci in parenthesis have no known human equivalents. Scales are 1cM intervals in the mouse map and 1Mb intervals in the human map. The hatched scale on each map indicates a region of inversion between the two species. Taken from Stubbs et al, (1996).

higher rate of informativeness of microsatellite loci. It is known however that *rattus norvegicus* cannot breed with any of the subspecies of *rattus rattus*, even those harbouring the same number of chromosomes (n=42), namely; *rattus rattus flavipectus*, *rattus rattus diardii*, *rattus rattus tanezumi* and *rattus rattus rattus* (Baverstock et al, 1983. Hayashi et al, 1979). It would be interesting to know if *rattus norvegicus* could breed with another species of rat, *rattus exulans*, as this species appears genetically closer to *rattus norvegicus* than any other (Baverstock et al, 1983). A backcross program is underway using DA as a reference strain since quite a few previously uninformative loci near the *agu* region have been shown to exhibit allelic size differences between AS/AGU and DA.

Future work

To date, almost 2,000 backcross progeny have been generated from the crosses, [AS/AGUxF1(AS/AGUxBN)] and [AS/AGUxF1(AS/AGUx F344)]. The locus *Prkcg* has proven to be the closest marker to *agu* with only 1 recombinant in 2,000 animals. The next line of work will thus be the physical mapping of this region using yeast artificial chromosomes (YACs) and bacterial artificial chromosomes (BACs) in order to walk from *Prkcg* to *agu*. Once a single genomic clone of has been shown to contain agu, two strategies could then be employed to identify coding regions within it.

The first of these techniques uses a strategy to isolate exons from cloned genomic DNA and is called exon trapping (Buckler et al, 1991). Random segments of genomic DNA are inserted into an intron present in a mammalian expression vector . After transfection of this construct into COS-7 cells, exonic sequences are paired together with vector splice sites. Exons can then be amplified by using primers specific for sites flanking the splice donor and splice acceptor sites in the vector. If the PCR product is greater than a particular size (177bp), then an exon has been trapped. This technique has nee used to find the (CAG)n/(CTG)n triplet repeat expansion Huntingdon's disease chromosomes (Huntingdon's disease collaborative research group, 1993). False positives can frequently be encountered and not all exons contained within the genomic clone may be captured when using this method.

The second technique, termed hybrid selection, employs methods to select specific cDNA molecules by hybridisation to genomic regions (Lovett et al,

1991. Morgan et al, 1992). Large genomic clones (YACs or P1s) are immobilised on nylon membranes. cDNA inserts from a tissue specific library are amplified and hybridised to the nylon membrane containing the genomic DNA. It is important to pre-block the amplified cDNA molecules of immobilised genomic clone before hybridisation to reduce false positives. COT1 DNA is usually used as a blocking agent as it is highly enriched in repetitive sequences and therefore has more chance of blocking intronic sequences. Drawbacks are encountered when psuedogenes (transcribed or untranscribed) are co-selected which may be from a different genomic region to that desired.

In terms of the *agu* project, the hybrid selection method could be used to hybridise a SNc-specific cDNA library to a YAC or BAC containing the *agu* locus. Candidate cDNA molecules could then be sequenced and comparisons could be made between isolated cDNAs from affected (AS/AGU) and unaffected (AS) rats.

It may be possible that the *agu* mutation is not within expressed sequence. Mutations in promoters, enhancers or even intronic regions may have an effect on a gene or genes and therefore lead to the observed striatonigral degeneration seen in AS/AGU rats. A massive sequencing effort of a clones harbouring the agu locus, derived from AS and AS/AGU rats would then be required to identify a mutation should this be the case.

The eventual discovery of the agu mutation and its effect on brain metabolism will surely aid in the understanding of nigrostriatal degenerative diseases.

References

Alagarsamy, S. and Johnson, K.M. (1995). Voltage-dependent calcium channel involvement in NMDA-induced activation of NOS. *Neuroreport* **6**. 2250-2254.

Aurivillius, M., Hansen, O.C., Lazrek, M.B., Bock, E. and Obrink, B. (1990). The cell adhesion molecule Cell-CAM 105 is an ecto-ATPase and a member of the immunoglobulin superfamily. *FEBS Lett.* **264**. 267-269.

Baker, H., Joh, T.H. & Reis, D. (1980). Genetic control of the number of midbrain dopaminergic neurones in inbred strains of mice: Relationship to size and neuronal density of the striatum. *Proc. Natl. Acad. Sci.* **77**, 4369-4373.

Bandmann, O., Davies, M.B., Marsden, C.D. and Wood, N.W. (1996). The human homologue of the weaver mouse gene in familial and sporadic Parkinson's disease. *Neuroscience* **72**. 877-879.

Barkats, M., Nakao, N., Grasbon-Frodl, E.M., Bilang-Bleuel, A., Revah, F., Mallet, J. and Brundin, P. (1997). Intrastristal grafts of embryonic mesencephalic rat neurons genetically modified usinf an adenovirus encoding human Cu/Zn superoxide dismutase. Neuroscience 78. 703-713.

Bar-Peled, O., Korkotian, E., Segal, M. and Groner, Y. (1996). Constitutive overexpression of C/Zn superoxide dismutase exacerbates kainic acid-induced apoptosis of transgenic-Cu/Zn superoxide dismutase neurons. *Proc. Natl. Acad. Sci.* **93**. 8530-8535.

Basma, A.N., Morris, E.J., Nicklas, W.J. and Geller, H.M. (1995). L-dopa cytotoxicity to PC12 cells in culture is via its autoxidation. *J. Neurochem.* **64**. 825-832.

Baumann, R.J., Jameson, H.D., McKean, H.E., Haack, D.G. and Weisberg, L.M. (1980). Cigarette smoking and Parkinson's disease. 1. A comparison of cases with matched neighbours. *Neurology* **30**. 839-843.

Baverstock, P.R., Adams, M., Maxson, L.R. and Yosida, T.H. (1983). Genetic differentiation among karytypic forms of the black rat, rattus rattus. *Genetics* **105**. 969-983.

Beck, K.D., Valverde, J., Alexi, T., Poulson, K., Moffat, B., Vandlen, R.A., Rosenthal, A. and Hefti, F. (1995). Mesencephalic dopaminergic neurons protected by GDNF from axotomy-induced degeneration in the adult brain. *Nature* **373**. 339-341.

Beckstead, R.M., Domesick, V.B. and Nauta, W.J.H. (1979). Efferent connections of the substantia nigra and ventral tegmental area in the rat. *Brain Res.* **175**. 191-217.

Ben-Shachar, D., Zuk, R. and Glinka, Y. (1995). Dopamine neurotoxicity: Inhibition of mitochondrial respiration. *J. Neurochem.* **64**. 718-723.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**. 1513-1523.

Blanchet, P.J., Papa, S.M., Metman, L.V., Mouradian, M.M. and Chase, T.N. (1997). Modulation of levodopa-induced motor response complications by NMDA antagonists in Parkinson's disease. *Neurosci. Biobehav. Rev.* **21**. 447-453.

Bost, M., Berkaw, M., McBride, O.W., Chazot, G. and Arnaud, P. (1995). Restriction polymorphisms of the ceruloplasmin gene on chromosome 3. *Hum. Genet.* **96**. 239-240.

Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J-P., Hudson, T., Sohn, R., Zemelman, B., Snell, R.G., Rundle, S.A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., Juvonen, V., Johnson, K., Harper, P.S., Shaw, D.J. and Housman, D.E. (1992). Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3' end of a trancript encoding a protein kinase family member. *Cell* **68**. 799-808.

Brouillet, E. and Beal, M.F. (1993). NMDA antagonists partially protect against MPTP induced neurotoxicity in mice. *Neuroreport* **4**. 387-390.

Buckler, A.J., Chang, D.D., Graw, S.L., Brook, J.D., Haber, D.A., Sharp, P.A. and Housman, D.E. (1991). Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci.* **88**. 4005-4009.

Burn, D.J., Mark, M.H. & Playford, E.D. *et al.* (1992). Parkinson's disease in twins studied with ^{18}F-dopa and positron emission tomography. *Neurology* **42**. 1894-1900.

Campbell, J.M., Payne, A.P., Gilmore, D.P., Byrne, J.E., Russell, D., McGadey, J., Clarke, D.J., Davies, R.W. and Sutcliffe, R.G. (1996). Neostriatal dopamine depletion and locomotor abnormalities due to the Albino Swiss Rat *agu* mutation. *Neurosci. Letters* **213**. 173-176.

Canzian, F. (1997). Phylogenetics of the laboratory rat Rattus norvegicus. *Genome Res.* **7**. 262-267.

Carle, G.F., Frank, M. and Olsen, M.V. (1986). Electrophoretic separations of large DNA molecules by periodic inversion of the electric field. *Science.* **232**. 65.

Castle, W.E. (1951). Variation in the hooded pattern of rats, and a new allele of hooded. *Genetics* **36**. 254-266.

Castle, W.E. and Philips, J.C. (1914). Piebald rats and selection. *Carnegie Inst. Wash. Publ.* **195**.

Chan, P., DeLanney, L.E., Irwin, I., Langston, J.W. and Di Monte, D. (1991). Rapid ATP loss caused by 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine in mouse brain. *J. Neurochem.* **57**. 348-351.

Chen, E.Y. and Seeburg, P.H. (1985). Supercoil sequencing. A fast and simple method for sequencing plasmid DNA. *DNA* **4**. 165-170.

Chen, H., Chen, C.L. and Chou, J.Y. (1994). Characterization of two promoters of a rat pregnancy-specific glycoprotein gene. *Biochemistry* **33**. 9615-9626.

Clarke, D.J. and Payne, A.P. (1994). Neuroanatomical characterisation of a new mutant rat with dopamine depletion in substantia nigra. *Eur. J. Neurosci.* **6**. 885-888.

Cooper, J.A., Sagar, H.J., Jordan, N., Harvey, N.S. and Sullivan, E.V. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationshio to motor disability. *Brain* **114**. 2095-2122.

Copeland, N.G., Jenkins, N.A., Gilbert, D.J., Eppig, J.T., Maltais, L.J., Miller, J.C., Dietrich, W.F., Weaver, A., Lincoln, S.E., Steen, R.G., Stein, L.D., Nadeau, J.H. and Lander, E.S. (1993). A genetic linkage map of the mouse: current applications and future prospects. *Science* **262**. 57-66.

Coyle, J.T. and Puttfarcken, P. (1993). Oxidative stress, glutamate, and neurodegenerative disorders. *Science* **262**. 689-695.

Diederich, N., Hilger, C., Goetz, C.G., Keipes, M., Hentges, F., Vieregge, P. and Metz, H. (1996). Genetic variability of the CYP2D6 gene is not a risk factor for sporadic Parkinson's disease. *Ann. Neurol.* **40**. 463-465.

Deng, H-X., Hentati, A., Tainer, J.A., Iqbal, Z., Cayabyab, A., Hung, W-Y., Getzoff, E.D., Hu, Pi., Herzfeldt, B., Roos, R.P., Warner, C., Deng, G., Soriano, E., Smyth, C., Parge, H.E., Ahmed, A., Roses, A.D., Hallewell, R.A., Pericak-Vance, M.A. and Siddique, T. (1993). Amyotrophic lateral sclerosis and structural defects in Cu/Zn superoxide dismutase. *Science* **261**. 1047-1051.

Denhardt, D.T. (1966). A membrane filter technique for the detection of complementary DNA. *Biophys. Res. Commun.* **23.** 641.

Ding, Y.P., Remmers, E.F., Du, Y., Longman, R.E., Goldmuntz, E.A., Zha, H., Kotake, S., Cannon, G.W., Griffiths, M.M. andWilder, R.L. (1996). Genetic maps of polymorphic DNA loci on rat chromosome 1. *Genomics* **36**. 320-327.

Doncaster, L. (1906). On the inheritance of coat colour in in rats. *Proc. Camb. phil. Soc.* **13**. 215-227.

During, M.J. and Leone, P. (1997). Targets for gene therapy of Parkinson's disease: growth factors, signal transduction and promoters. *Exp. Neurol.* **144**. 74-81.

Duvoisin, R.C., Eldridge, R., Williams, A., Nutt, J. and Calne, D. (1981). Twin studies of Parkinson's disease. *Neurology (NY)* **31**. 77-80.

Dveksler, G.S., Dieffenbach, C.W., Cardellichio, C.B., McCuaig, K., Pensiero, M.N., Jiang, G.S., Beauchemin, N. and Holmes, K.V. (1993). Several members of the mouse carcinoembryonic antigen-related glycoprotein family are functional receptors for the coronavirus mouse hepatitis virus-A59. *J. Virol.* **67**. 1-8.

Eldridge, R. and Rocca, W.A. Parkinson's Disease. In. King, R.A., Rotter, J.I. and Motulsky, A.G. (eds). *The Genetic Basis of Common Diseases.* (1992). 775-791. Oxford University Pree, New York.

Fredriksson, A., Gentsch, C. and Archer, T. (1994). Synergistic interactions between NMDA-antagonists and L-dopa on activity in MPTP-treated mice. *J. Neural Transm. Gen. Sect.* **97**. 197-209.

Frodl, E.M., Duan, W.M., Sauer, H., Kupsch, A. and Brundin, P. (1994). Human embryonic dopamine neurons xenografted to the rat: effects of

cryopreservation and varying regional source of donor cells on transplant survival, morphology and function. *Brain Res.* **647**. 286-298.

Furtado, J.C.S. and Mazurek, M.F. (1991). MPTP-induced neurotoxicity and the quest for a preventative therapy for Parkinson's disease. *Can. J. Neurol. Sci.* **18**. 77-82.

Gelman, B.B. (1995). Iron in CNS disease. *J. Neuropath. Exp. Neurol.* **54**. 477-486.

Getzoff, E.D., Cabelli, D.E., Fisher, C.L., Parge, H.E., Viezzoli, M.S., Banci, L. and Hallewell, R.A. (1992). Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature* **358**. 347-351.

Gibson, T.J. (1984). Studies on the Epstein-Barr virus genome. PhD thesis. Cambridge University, England.

Grant, S., Jessee, J., Bloom, F. and Hanahan, D. (1990). Differential plasmid rescue from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc. Natl. Acad. Sci.* **87**. 4645-4649.

Godwin-Austin, R.B. (1994). *Parkinson's disease: a booklet for patients and their families.* Parkinson's Disease Society (UK), London.

Golbe, L.J., Di Iorio, G., Bonavita, V., Miller, D.C. and Duvoisin, R.C. (1990). A larger kindred with dominant Parkinson's disease. *Ann. Neurol.* **27**. 276-282.

Gumbreck, L.G., Stanely, A.J., Allison, J.E. and Easley, R.B. (1972). Restriction of coat colour in the rat with associated sterility in the male and heterochromia in both sexes. *J. Exp. Zool.* **180**. 333-350.

Harris, Z.L., Takahashi, Y., Miyajima, H., Serizawa, M., MacGillivray, R.T.A. and Gitlin, J.D. (1995). Aceruloplasminemia: Molecular characterisation of this disorder of iron metabolism. *Proc. Natl. Acad. Sci.* **92**. 2539-2543.

Hayashi, J-I., Yomekawa, H., Gotoh, O., Tagashira, Y., Moriwaki, K. and Yosida, T.H. (1979). Evolutionary aspects of variant types of rat mitochondrial DNAs. *Biochim. Biophys. Acta* **564**. 202-211.

Hedrich, H.J. (1990). *Genetic monitoring of inbred strains of rats.* Gustav Fisher Verlag, Stutgart.

Huntington's disease collaborative research group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**. 971-983.

Jacob, H., Brown, D.M., Bunker, R.K., Daly, M.J., Dzau, V.J., Goodman, A., Koike, G., Kren, V., Kurtz, T., Lernmark, A., Levan, G., Mao, Y-P., Pettersson, A., Pravenec, M., Simon, J.S., Szpirer, C., Szpirer, J., Trolliet, M.R., Winer, E.S. and Lander, E.S. 1995. A genetic linkage map of the laboratory rat, *Rattus Norvegicus. Nature Genetics* **9**. 63-69.

Javitch, J.A., Uhl, G.R. and Snyder, S.H. (1984). Parkinsonism-inducing neurotoxin, *N*-methyl-4-phenyl-1,2,3,6-tetrahydropyridine: Characterisation and localisation of receptor binding sites in rat and human brain. *Proc. Natl. Acad. Sci.* **81**. 4591-4595.

Javitch, J.A., D'Amato, R.J., Strittmatter, S.M. and Snyder, S.H. (1985). Parkinsonism-inducing neurotoxin, *N*-methyl-4-phenyl-1,2,3,6-tetrahydropyridine: Uptake of the metabolite *N*-mehyl-4-phenylpyridine by dopamine neurons explains selective toxicity. *Proc. Natl. Acad. Sci* **82**. 2172-2177.

Kim, R.S., Sukhu, B. and LaBella, F.S. (1988). Lipoxygenase-induced lipid peroxidation of isolated cardiac microsomes modulates their calcium-transporting function. *Biochim. Biophys. Acta.* **961**. 270-277.

Kimmerly, W.J., Kyle, A.L., Lustre, V.M., Martin, C.H. and Palazzolo, M.J. (1994). Direct sequencing of terminal regions of genomic P1 clones: A strategy for the design of sequence-tagged site markers. *GATA* **11**. 117-128.

Kish, S.J., Chang, L.J. and Marchandari, L. (1988). Uneven pattern of dopamine loss in the straiatum of patients with idiopathic Parkinson's disease. *N. Engl. J. Med.* **318**. 876-880.

Kumar, R., Agarwal, A.K. and Seth, P.K. (1995). Free radical generated neurotoxicity of 6-hydroxydopamine. *J. Neurochem.* **64**. 1703-1707.

Klomp, L.W., Farhangrazi, Z.S., Dugan, L.L. and Gitlin, J.D. Ceruloplasmin gene expression in the murine central nervous system. *J. Clin. Invest.* **98**. 207-215.

Kodelja, V., Lucas, K., Barnert, S., von Kleist, S., Thompson, J.A. and Zimmermann, W. (1989). Identification of a carcinoembryonic antigen gene family in the rat. Analysis of the N-terminal domains reveals immunoglobulin-like, hypervariable regions. *J. Biol. Chem.* **264**. 6906-6912.

Krauter, K.S., Citron, B.A., Hsu, M.T., Powell. D, and Darnell, J.E.Jr. (1986). Isolation and characterization of the alpha 1-antitrypsin gene of mice. *DNA* **5**. 29-36.

Lam, A.G., Campbell, J.M., Bennett, N.K., Payne, A.P., Davies, R.W., Sutcliffe, R.G. and McCulloch, J. (1997). Local cerebral glucose utilisation in the AS/AGU mutant rat. *In press*.

Lange, K.W., Loschmann, P.A., Sofic, E., Burg, M., Horowski, R., Kalveram, K.T., Wachtel, H. and Riederer, P. (1993). The competitive NMDA antagonist CPP protects substantia nigra neurons from MPTP-induced degeneration in primates. *Naunyn. Schmiedebergs Arch. Pharmacol.* **348**. 586-592.

Lee, A.B. and Cooper, T.A. (1995). Improved direct PCR screen for bacterial colonies: wooden toothpicks inhibit PCR amplification. *Biotechniques* **18**. 225-226.

Levan, G., Szpirer, J., Szpirer, C., Klinga, K., Hanson, C. and Islam, M.Q. (1991). The gene map of the norway rat (Rattus norvegicus) and comparitive mapping with mouse and man. *Genomics* **10**. 699-618.

Litt, M., Hauge, X. and Sharma, V. (1993). Shadow bands seen when typing polymorphic dinucleotide repeats: some causes and cures. *Biotechniques* **15**. 280-284.

Liu, X-H, Kato, H., Chen, T., Kato, K. and Itoyama, Y. (1994). Bromocriptine protects against delayed neuronal death of hippocampal neurons following cerebral ischemia in the gerbil. *J. Neurol. Sci.* **129**. 9-14.

Lovett, M., Kere, J. and Hinton, L.M. (1991). Direct selection: a method for the isolation of cDNAs encoded by large genomic regions. *Proc. Natl. Acad. Sci.* **88**. 9628-9632.

Margolis, R.L., Li, S-H. and Ross, C.A. 1993. A rapid method for sequencing trinucleotide repeats. *Nucleic Acids Res.* **21**. 4983-4984.

Maroteaux, L. and Scheller, R.H. (1991). The rat brain synucleins; family of proteins transiently associated with neuronal membrane. *Brain Res. Mol. Brain Res.* **11**. 335-343.

Martin, J.H. (1996). *Neuroanatomy: Text and Atlas.* Appleton and Lange, Stamford.

Mayeux, R., Stern, Y., Rosen, J. and Leventhal, J. (1991). Depression, intellectual impairment and Parkinson's disease. *Neurology (NY)* **31**. 645-650.

McGreer, P.L., Itagaki, S. & Akiyama, H. *et al.* (1988). Rate of cell death in parkinsonism indicates active neuropathological process. *Ann. Neurol.* **24**, 574-576.

Miyajima, H., Nishimura, Y., Mizoguchi, K., Sakamoto, M., Shimizu, T. and Honda, N. (1987). Familial apoceruloplasmin deficiency associated with blepharospasm and retinal degeneration. *Neurology* **37**. 761-767.

Miyajima, H., Takahashi, Y., Serizawa, M., Kaneko, E. and Gitlin, J.D. (1996). Increased plasma lipid peroxidation in patients with acerulopasminemia. *Free Radic. Biol. Med.* **20**. 757-760.

Morgan, J.G., Dolganov, G.M., Robbins, S.E., Hinton, L.M. and Lovett, M. (1992). *Nucleic Acids Res.* **20**. 5173-5179.

Nakao, N., Frodl, E.M., Widner, H., Carlson, E., Eggerding, F.A., Epstein, C.J. and Brundin, P. (1995). Overexpressing Cu/Zn superoxide dismutase enhances survival of transplanted neurons in a rat model of Parkinson's disease. *Nat. Med.* **1**. 226-231.

Nakao, N., Frodl, E.M., Duan, W.M., Widner, H. and Brundin, P. Lazaroids improve the survival of grafted rat embryonic dopamine neurons. *Proc. Natl. Acad. Sci.* **91**. 12408-12412.

Nédellec, P., Dveksler, G.S., Daniels, E., Turbide, C., Chow, B., Basile, A.A., Holmes, K.V.and Beauchemin, N. (1994). Bgp2, a new member of the carcinoembryonic antigen-related gene family, encodes an alternative receptor for mouse hepatitis viruses. *J. Virol.* **68**. 4525-4537.

O'Dell, S.J., Weihmuller, F.B. and Marshall, J.F. (1993). Methamphetamine-induced dopamine overflow and injury to striatal dopamine terminals: attenuation by dopamine D1 or D2 antagonists. *J. Neurochem.* **60**. 1792-1799.

Odin, P., Asplund, M., Busch, C. and Obrink, B. (1988). Immunohistochemical localization of cellCAM 105 in rat tissues: appearance in epithelia, platelets, and granulocytes. *J.Histochem. Cytochem.* **36**. 729-739.

Olsen, A., Teglund, S., Nelson, D., Gordon, L., Copeland, A., Georgescu, A., Carrano, A. and Hammarstrom, S. (1994). Gene organization of the pregnancy-specific glycoprotein region on human chromosome 19: assembly and analysis of a 700-kb cosmid contig spanning the region. *Genomics* **23**. 659-668.

Parker, W.D., Boyson, S.J. and Parks, J.K. (1989). Abnormalities of the electron transport chain in idiopathic Parkinson's disease. *Ann. Neurol.* **26**. 719-723.

Paulson, G.W. and Dadmehr, N. (1991). Is there a premorbid personality typical for Parkinson's disease? *Neurology* **41 (suppl. 2)**. 73-76.

Peled-Kamar, M., Lotem, J., Wirguin, J., Hermalin, A. and Groner, Y. (1997). Oxidative stress mediates impairments of muscle function in transgenic mice with elevated level of wild-type Cu/Zn superoxide dismutase. *Proc. Natl. Acad. Sci.* **94**. 3883-3887.

Pierce, J.C., Sauer, B. and Sternberg, N. (1992). A positive selection vector for high molecular weight DNA by the bacteriophage P1 system: improved cloning efficiency. *Proc. Natl. Acad. Sci. USA* **89**. 2056-2060.

Poirier, J., Kogan, S. and Gauthier, S. (1991). Environment, genetics and idiopathic Parkinson's disease. *Can. J. Neurol. Sci.* **18**. 70-76.

Polymeropoulos, M.H., Higgins, J.J., Golbe, L.I., Johnson, W.G., Ide, S.E., Di Iorio, G., Sanges, G., Stenroos, E.S., Pho, L.T., Schaffer, A.A., Lazzarini, A.M., Nussbaum, R.L. and Duvoisin, R.C. (1996). Mapping of a gene for Parkinson's disease to chromosome 4q21-23. *Science* **274**. 1197-1199.

Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejla, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., Stenroos, E.S., Chandrasekharappa, S., Athanassiadou, A., Papapertropoulos, T., Johnson, W.G., Lazzarini, A.M., Duvoisin, R.C., Di Iorio, G., Golbe, L.I. and Nussbaum, R.L. (1997). Mutation in the α-synuclein gene identified in famalies with Parkinson's disease. *Science* **276**. 2045-2048.

Ramsay, R.R., Krueger, M.J., Youngster, S.K., Gluck, M.R., Casida, J.E. and Singer, T.P. (1991). Interaction of 1-methyl-4-phenylpyridinium ion (MPP+) and its analogs with the rotenone/piericidin binding site of NADH dehydrogenase. *J. Neurochem.* **56**. 1184-1190.

Ransom, B.R., Kunis, D.M., Irwin, I. and Langston, J.W. (1987). Astrocytes convert the parkinsonism inducing neurotoxin, MPTP, to its active metabolite, MPP^+. *Neurosci. Lett.* **75**. 323-328.

Rebstock, S., Lucas, K., Thompson, J.A. and Zimmermann, W. (1990). cDNA and gene analyses imply a novel structure for a rat carcinoembryonic antigen-related protein. *J. Biol. Chem.* **265**. 7872-7879.

Reeves, R.H., Crowley, M.R., Lorenzon, N., Pavan, W.J., Smeyne, R.J. and Goldowitz, D. (1989). The mouse neurological mutant weaver maps within the region of chromosome 16 that is homologous to human chromosome 21. *Genomics* **5**. 522-526.

Regan, R.F. and Choi, D.W. (1994). The effect of NMDA, AMPA/kainate, and calcium channel antagonists on traumatic cortical neuronal injury in culture. *Brain Res.* **633**. 236-242.

Riederer, P., Sofic, E., Rausch, W.D., Schmidt, B., Reynolds, G.P., Jellinger, K. and Youdim, M.B. (1989). Transition metals, ferritin, glutathione, and ascorbic acid in parkinsonian brains. *J. Neurochem.* **52**. 515-520.

Rosen, D.R, Siddique, T., Patterson, D., Figelwicz, D.A., Goto, I., O'Regan, J.P., Deng, H., Rahman, Z., Krizus, A., McKenna-Yasek, D., Cayabyab, A., Gaston, S.M., Berger, R., Tanzi, R.E., Halperin , J.J., Herzfeldt, B., Van Den Burgh, R., Hung, M., Bird, T., Deng, G., Mulder, D.W., Smyth, C., Laing, N.G., Soriano, E., Pericak-Vance, M.A., Haines, J., Rouleau, G.A., Gusella, J., Horvitz, H.R. and Brown Jr., R.H. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**. 59-62.

Rudert, F., Saunders, A.M., Rebstock, S., Thompson, J.A. and Zimmermann, W. (1992). Characterization of murine carcinoembryonic antigen gene family members. *Mamm. Genome* **3**. 262-273.

Sandri, G., Panfili, E. and Ernster, L. (1990). Hydrogen peroxide production by monoamine oxidase inisolated rat-brain mitochondria: its effect on glutathione levels and Ca2+ efflux. *Biochim. Biophys. Acta.* **1035**. 300-305.

Sanger, F., Nicklen, S. and Coulsen, A.R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**. 5463-5467.

Savidge, J.R. and Bristow, D.R. (1997). Routes of NMDA- and K(+)-stimulated calcium entry in rat cerebellar granule cells. *Neurosci. Lett.* **229**. 109-112.

Schmidt, M.J., Sawyer, B.D., Perry, K.W., Fuller, R.W., Foreman, M.M. and Ghetti, B. (1982). Dopamine defficiency in the weaver mutant mouse. *J. Neurosci.* **2**. 376-380.

Schulz, J.B., Matthews, R.T., Muqit, M.M., Browne, S.E. and Beal, M.F. (1995). Inhibition of neuronal nitric oxide synthase by 7-nitroindazole protects against MPTP-induced neurotoxicity in mice. *J. Neurochem.* **64**. 936-939.

Scott, D.L., Campbell, J.M., Clarke, D.J., Gilmore, D.P. and Payne, A.P. (1994). Noradrenergic deficits in th locus cœrulus of AS/AGU rats. *Eur. J. Neurosci.* 7 **(suppl.)**. 96.

Serikawa, T., Kuramoto, T., Hilbert, P., Mori, M., Yamada, J., Dubay, C.J., Lindpainter, K., Ganten, D., Guénet, J.L., Lathrop, G.M. and Beckmann, J.S. 1992. Rat gene mapping using PCR analysed microsatellites. *Genetics* **131**. 701-721.

Shiels, P., Durán-Alonso, M.B., Davidson, A.O., Heeley, R.P., Dominiczak, A.F., Payne, A.P., Davies, R.W. and Stuteliffe, R.G. (1995). Optimized protocols for typing 75 microsatellite loci for AS, PVG, F344 and BN rats. *Mamm. Genome* **6**. 214-215.

Silverman, S.K., Kofuji, P., Dougherty, D.A., Davidson, N. and Lester, H.A. (1996). A regenerative link in the ionic fluxes through the weaver potassium channel underlies the pathophysiology of the mutation. *Proc. Natl. Acad. Sci.* **93**. 15429-15434.

Sirinathsinghji, D.J., Heavens, R.P. and McBride, C.S. (1988). Dopamine-releasing action of 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) and 1-methyl-4-phenylpyridine (MPP+) in the neostriatum of the rat as demonstrated in vivo by the push-pull perfusion technique: dependence on sodium but not calcium ions. *Brain Res.* **443**. 101-116.

Sofic, E., Riederer, P., Heinsen, H., Beckmann, H., Reynolds, G.P., Hebenstreit, G. and Youdim, M.B. (1988). Increased iron (III) and total iron content in post mortem substantia nigra of parkinsonian brain. *J. Neural. Transm.* **74**. 199-205.

Southard-Smith, M. and MacDonald, R.J. (1993). Isolating reiterated genes from a rat genomic library constructed with the bacteriophage P1 system. *DuPont Biotech Update* **8**. 36-39.

Southard-Smith, M., Pierce, J.C. and MacDonald, R.J. (1994). Physical mapping of the rat tissue kallikrein family in two gene clusters by analysis of P1 bacteriophage clones. *Genomics* **22**. 404-417.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**. 503-517.

Sternberg, N. (1990). Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. *Proc. Natl. Acad. Sci.* **87**. 103-107.

Stewart, A.A., Campbell, J.M., McGadey, J., Clarke, D.J., Gilmore, D.P. and Payne, A.P. (1994). Serotonergic deficits in the raphe neclei of the AS/AGU rat. *J. Anat.* **185**. 702.

Stubbs, L., Carver, E.A., Shannon, M.E., Kim, J., Geisler, J., Generoso, E.E., Stanford, B.G., Dunn, W.C., Mohrenweiser, H., Zimmermann, W., Watt, S.M. and Ashworth, L.K. (1996). *Genomics* **35**. 499-508.

Takahashi, H., Ohama, E., Suzuki, S., Horikawa, Y., Ishikawa, A., Morita, T., Tsuji, S. and Ikuta, F. (1994). Familial juvenile parkinsonism: clinical and pathologic study in a family. *Neurology* **44**. 437-441.

Tatton, W.G. and Greenwood, C.E. Rescue of dying neurons: a new action for deprenyl in MPTP Parkinsonism. *J. Neurosci. Res.* **30**. 666-672.

Teglund, S., Olsen, A., Khan, W.N., Frangsmyr, L. and Hammarstrom, S. (1994). The pregnancy-specific glycoprotein (PSG) gene cluster on human chromosome 19: fine structure of the 11 PSG genes and identification of 6 new genes forming a third subgroup within the carcinoembryonic antigen (CEA) family. *Genomics* **23**. 669-684.

Thompson, J., Zimmermann, W., Osthus-Bugat, P., Schleussner, C., Eades-Perner, A.M., Barnert, S., Von Kleist, S., Willcocks, T., Craig, I., Tynan, K., Olsen, A. and Mohrenweiser, H. (1992). Long-range chromosomal mapping of the carcinoembryonic antigen (CEA) gene family cluster. *Genomics* **12**. 761-772.

Tomac, A., Lindqvist, E., Lin, L-F.H., Ögren, S.O., Young, D., Hoffer, B.J. and Olsen, L. (1995). Protection and repair of the nigrostriatal dopaminergic system by GDNF *in vivo*. *Nature* **373**. 335-339.

Tong, Y., Wei, J., Zhang, S., Strong, J.A., Dlouhy, S.R., Hodes, M.E., Ghetti, B. amd Yu, L. (1996). The weaver mutation changes the ion selectivity of the affected inwardly rectifying potassium channel GIRK2. *FEBS Lett.* **390**. 63-68.

Turbide, C., Rojas, M., Stanners, C.P. and Beauchemin, N. (1991). A mouse carcinoembryonic antigen gene family member is a calcium-dependent cell adhesion molecule. *J. Biol. Chem.* **266**. 309-315.

Tynan, K., Olsen, A., Trask, B., de Jong, P., Thompson, J., Zimmermann, W., Carrano, A. and Mohrenweiser, H. (1992). Assembly and analysis of cosmid contigs in the CEA-gene family region of human chromosome 19. *Nucleic Acids Res.* **20**. 1629-1636.

Whilhelmsen, K., Mirel, D., Marder, K., Bernstein, M., Naini, A., Leal, S.M., Core, L.J., Tang, M.X., Freyer, G., Graziano, J. and Mayeux, R. (1997). Is there a genetic susceptibility locus for Parkinson's disease on chromosome 22q13? *Ann. Neurol.* **41**. 813-817.

Williams, J.L., Olsaker, I. and Teres, V.M. Using SSCP to facilitate mapping microsatellite loci. *Mamm. Genome* **8**. 79-80.

Wullner, U., Kupsch, A., Arnold, G., Renner, P., Scheid, C., Scheid R., Oertel, W. and Klockgether T. (1992). The competitive NMDA antagonist CGP40,116 enhances L-dopa response in MPTP-treated marmosets. *Neuropharmacology* **31**. 713-715.

Yamada, J., Kuramoto, T. and Serikawa, T. (1994). A rat genetic linkage map and comparative maps for mouse or human homologous rat genes. *Mamm. Genome* **5**. 63-83.

Yannish-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains. Nucleotide sequence of the M13mp18 and pUC19 vectors. *Gene* **33**. 103-119.

Yim, M.B., Chock, P.B. and Stadtman, E.R. (1990). Copper-zinc superoxide dismutase catalyses radical production from hydrogen peroxide. *Proc. Natl. Acad. Sci.* **87**. 5006-5010.

Yong, V.W., Perry, T.L. and Krisman, A.A. Depletion of glutathione in brainstem of mice caused by N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine is prevented by antioxidant pretreatment. (1986). *Neurosci. Lett.* **63**. 56-60.

Zuddas, A., Fascetti, F. & Corsini, G. *et al.* (1994). In Brown Norway Rats, MPP+ is accumulated in the nigrostratial dopaminergic terminals but it is not neurotoxic: A model of natural resistance to MPTP toxicity. *Exp. Neurol.* **127**. 54-61.

Abbreviations

''	seconds
AMPA	a-amino-3-hydroxy-5-methyl-4-isoxasole-proprionic acid
AS	Albino Swiss
AS/AGU	Albino Swiss, Anatony Glasgow University
ATP	Adenosine triphophate
BN	Brown Norway
bp	base pairs
°C	degrees centigrade
CEA	Carcinoembryonic antigen
CGM	CEA family member
Ci	Curies
cM	centiMorgan
cpm	counts per minute
Cu/ZnSOD	Copper/Zinc superoxide dismutase
ddH_2O	deionised and distilled water
DNA	2' deoxyribonucleic acid
dNTP	2' deoxy nucleotide
EDTA	ethylene diamine tetra acetic acid
EtBr	Ethidium bromide
F344	Fischer 344
hr	hour
IPTG	isopropyl-ß-D-thiogalactoside
Kbp	Kilobase pairs
M	Molar
mg	miligrammes
min	minutes
ml	mililitres
mol	Moles
MPTP	N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
MPP+	1-methyl-4-phenyl pyridine
µCi	microCuries
µl	microlitres
NMDA	N-methyl-D-aspartate
PD	Parkinson's disease

rpm	revolutions per minute
SNc	Substantia nigra pars compacta
TAE	Tris acetate EDTA
TBE	Tris borate EDTA
TE	Tris EDTA
Tris	Tris (hydroxymethyl) aminomethane
V	volts
W	watts
w/v	weight per volume
X-gal	5-bromo-4-4chloro-3-indolyl-ß-D-galactopyranoside
YAC	yeast artificial chromosome